

1.1 引言

如何将现实需求转化为线性规划问题没有一定之规。美国著名数学教育家 G. Polya 教授在《How to Solve It: A New Aspect of Mathematical Method》一书中提出的如下一般数学问题求解技巧也同样适合：

- ① 求解过程中需要注意的问题：
- ② 求解什么未知数？已知什么？
- ③ 条件是什么？条件充不充分？
- ④ 但凡能画图，一定要画，把条件分解成各个部分，把自己的问题用自己的话重新讲，反复讲。

通常来说，数学规划的建模要点在于：

1. 对现实问题要看透，抓住问题里面的最重要的因素（因问题而异，无一般方法）；
2. 找到最合适的数学语言和它对应（简单而优雅）；
3. 建立模型，要尽量容易解；
4. 尽量套用已有的现成模型，方便使用成熟的求解方法和计算工具；
5. 把求得的结果从数学语言翻译成我们能看懂的语言。

1.2 常见的运筹规划问题分类

1.3 数学符号约定

在本书中, 仅涉及实数域。其中, 矩阵用大写字母 $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$ 表示。矩阵 \mathbf{A} 的转置表示为 \mathbf{A}^T 。矩阵 \mathbf{A} 的共轭转置表示为 \mathbf{A}^H 。矩阵 \mathbf{A} 的逆表示为 \mathbf{A}^{-1} 。矩阵 \mathbf{A} 的行列式表示为 $\det(\mathbf{A})$ 。矩阵 \mathbf{A} 的迹表示为 $Tr(\mathbf{A})$ 。

多维向量 (变量, 列向量) 用粗体小号字母 $\mathbf{a}, \mathbf{b}, \mathbf{c}, \dots, \mathbf{x}, \mathbf{y}, \mathbf{z}$ 表示。向量 \mathbf{a} 的转置表示为行向量 \mathbf{a}^T 。

一维变量用小写字母 a, b, c, \dots, x, y, z 表示。

集合通常用大写字母表示 Γ, \dots, Ω 。 \mathbb{R}^n 表示 n 维实数空间 (向量空间), \mathbb{R}_+^n 表示所有元素均非负的 n 维实数空间 (向量空间)。 \mathbb{N} 表示自然数空间。

$\mathbf{0}$ 是具有适当维数的所有元素都是 0 的列向量, $\mathbf{1}$ 是具有适当维数的所有元素都是 1 的列向量。 \mathbf{I} 通常表示具有适当维数的单位矩阵 (对角线元素全为 1, 其余元素全为 0)。

$\|\cdot\|$ 表示多维向量的范数。 $|\cdot|$ 表示一维变量的绝对值。

向量 $\mathbf{x} = [x_1, \dots, x_n] \in \mathbb{R}^n$ 的 0 范数 (0-norm, 本质上并非严格意义上的范数, 因为不满足三角不等式) 定义为向量 \mathbf{x} 的所有非零元素的个数。

向量 $\mathbf{x} \in \mathbb{R}^n$ 的 l_1 范数 (l_1 -norm) 定义为

$$\|\mathbf{x}\|_1 = \sum_{k=1}^n |x_k| \quad (1.1)$$

向量 $\mathbf{x} \in \mathbb{R}^n$ 的 l_2 范数 (l_2 -norm) 定义为

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{k=1}^n x_k^2} \quad (1.2)$$

显然, 我们有 $\|\mathbf{x}\|_2^2 = \mathbf{x}^T \mathbf{x}$ 。

向量 $\mathbf{x} \in \mathbb{R}^n$ 的 l_p 范数 (l_p -norm, $p \in [1, +\infty)$) 定义为

$$\|\mathbf{x}\|_p = \left(\sum_{k=1}^n |x_k|^p \right)^{1/p} \quad (1.3)$$

向量 $\mathbf{x} \in \mathbb{R}^n$ 的 l_∞ 范数 (l_∞ -norm) 定义为

$$\|\mathbf{x}\|_\infty = \max_{k=1, \dots, n} |x_k| \quad (1.4)$$

1.4 线性代数基础

一个线性方程组 (Linear Equation Set) 或者线性系统 (Linear System) 记为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \quad (1.5)$$

其中 $x_i, a_{ij}, b_i, c_i \in \mathbb{R}$, $i = 1, \dots, m$, $j = 1, \dots, n$ 。 x_i 为待定的决策变量。

我们也可以矩阵向量的形式将该方程组记为

$$\mathbf{Ax} = \mathbf{b} \quad (1.6)$$

其中向量 $\mathbf{x} = [x_1, \dots, x_n]^T \in \mathbb{R}^n$, $\mathbf{b} = [b_1, \dots, b_m]^T \in \mathbb{R}^m$, $\mathbf{A} = [a_{ij}]^T \in \mathbb{R}^{m \times n}$ 。

令矩阵 \mathbf{A} 的各列向量依次表示为 $\alpha_1, \dots, \alpha_n$, 我们也可以矩阵向量的形式将该方程组记为

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{bmatrix} = \sum_{j=1}^n x_j \alpha_j = \mathbf{b} \quad (1.7)$$

一般来说, 这个线性方程组的解有三种可能: 1) 无解; 2) 具有唯一解; 3) 具有无数多个可行解。我们可以通过如下的增广矩阵的秩来判断

$$\mathbf{A}_b = \begin{bmatrix} a_{11} & \dots & a_{1n} & b_1 \\ a_{21} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & b_m \end{bmatrix} = [\mathbf{A} \quad \mathbf{b}] \quad (1.8)$$

1) 如果 $\text{rank}(\mathbf{A}) < \text{rank}(\mathbf{A}_b)$, 这个线性方程组无解。

2) 如果 $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}_b)$, 这个线性方程组有解。进一步,

2.1) 若 $m = n = \text{rank}(\mathbf{A})$, 这个线性方程组有唯一解 $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ 。

2.2) 若 $m < n = \text{rank}(\mathbf{A})$, 这个线性方程组有无穷多个解。若 $\mathbf{x}_1 \neq \mathbf{x}_2$ 都是该线性方程组的解, 对于任意实数 $\lambda \in (0, 1)$, $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ 亦是该方程组的解。因 $\mathbf{A}[\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2] = \lambda\mathbf{Ax}_1 + (1-\lambda)\mathbf{Ax}_2 = \lambda\mathbf{b} + (1-\lambda)\mathbf{b} = \mathbf{b}$ 。

定义 1.1. 考虑 $m < n = \text{rank}(\mathbf{A})$ 的线性方程组 $\mathbf{Ax} = \mathbf{b}$ 。假设 \mathbf{A}_B 是通过从

\mathbf{A} 的 n 列中选择 m 列而形成的方阵。 $\mathbf{x}_B \in \mathbb{R}^m$ 是从 \mathbf{x} 中选择的相应变量所组成

的列向量, $\mathbf{x}_N \in \mathbb{R}^{n-m}$ 是代表 \mathbf{x} 的其余变量所组成的列向量。如果 \mathbf{A}_B 是非奇异的, 任意给定一个 \mathbf{x}_N , 我们可确定唯一解 $\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix} = \begin{bmatrix} \mathbf{A}_B^{-1}\mathbf{b} - \mathbf{A}_B^{-1}\mathbf{A}_N\mathbf{x}_N \\ \mathbf{x}_N \end{bmatrix}$ 。

进一步, 我们将 $\mathbf{x}_N = \mathbf{0}$ 时确定的 $\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_B^{-1}\mathbf{b} \\ \mathbf{0} \end{bmatrix}$ 称为 $\mathbf{Ax} = \mathbf{b}$ 的一个基

本解 (Basic Solution), 并将 $\mathbf{x}_B = \mathbf{A}_B^{-1}\mathbf{b}$ 对应的这 m 个变量称为一组基变量

(Basic Variables), 而剩下的其他变量则称为一组非基变量 (Non-Basic Variables)。 \mathbf{x}_B 中存在元素为 0, 则称之为一个退化基本解 (Degenerate Basic Solution); 否则称之为一个非退化基本解 (Non-Degenerate Basic Solution)。

定义 1.2. 对称矩阵 $\mathbf{A} \in \mathbb{R}^{n \times n}$ 是半正定矩阵, 如果矩阵 \mathbf{A} 的所有特征值都是非负的。我们用集合 $S_+^{n \times n}$ 表示所有 $n \times n$ 维的半正定矩阵构成集合。

对称矩阵 $\mathbf{A} \in \mathbb{R}^{n \times n}$ 是正定矩阵, 如果矩阵 \mathbf{A} 的所有特征值都是正的。我们用集合 $S_{++}^{n \times n}$ 表示所有 $n \times n$ 维的正定矩阵构成集合。

1.5 凸分析基础

为了更好地对于线性规划问题和非线性规划问题进行分析, 我们需要掌握一定的凸分析知识。本节中, 我们将依次介绍凸集的定义, 球、超平面、半平面、多面体、锥等常见的凸集, 极点和极线的定义, 凸集分隔定理, 支撑超平面定理, 下有界的凸集支撑超平面必有极点定理, Farkas 引理, 凸函数的定义, 光滑凸函数的一阶和二阶条件。

定义 1.3. 凸集 (Convex Set): 如果某个集合中任意两点连起来的直线段都属于该集合, 则称其为凸集, 否则为非凸集。或者等价的, 如果对于集合 $\Omega \in \mathbb{R}^n$, 任取 $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$, 对于任意实数 $\lambda \in [0, 1]$ 存在

$$\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \in \Omega \quad (1.9)$$

则集合 Ω 为凸集。

定理 1.1. 有限个凸集的交集是凸集。

证明：假设我们有 m 个凸集 $\Omega_1, \Omega_2, \dots, \Omega_m$, $m \in \mathbb{N}$ 。这 m 个凸集的交集为 $\Omega = \bigcap_{k=1}^m \Omega_k$ ，对于任意两个属于 Ω 的点 $\mathbf{x}_1, \mathbf{x}_2$ ，根据交集的定义可知满足 $\mathbf{x}_1, \mathbf{x}_2 \in \Omega_1, \Omega_2, \dots, \Omega_m$ 。由于这些集合都是凸集，我们直接可知，对于任意实数 $\lambda \in [0, 1]$ 存在

$$\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \in \Omega_k, \quad k = 1, \dots, m$$

则根据凸集的定义，集合 Ω 显然为凸集。□

定义 1.4. 圆心位于 $\mathbf{x}^* \in \mathbb{R}^n$ ，半径为 $r \in \mathbb{R}$ 的开球 B (Open Ball) 为

$$B_r(\mathbf{x}^*) = \{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}^*\|_2 < r, \mathbf{x} \in \mathbb{R}^n\} \quad (1.10)$$

圆心位于 $\mathbf{x}^* \in \mathbb{R}^n$ ，半径为 $r \in \mathbb{R}$ 的闭球 \bar{B} (Closed Ball) 为

$$\bar{B}_r(\mathbf{x}^*) = \{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}^*\|_2 \leq r, \mathbf{x} \in \mathbb{R}^n\} \quad (1.11)$$

定理 1.2. 任意一个开球和闭球都是凸集。

证明：我们仅证明开球是凸集，闭球证明与之类似。

对于任意两个点 $\mathbf{x}_1, \mathbf{x}_2 \in B_r(\mathbf{x}^*)$ 和任意实数 $\lambda \in [0, 1]$ ，我们可以得到

$$\begin{aligned} \|\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 - \mathbf{x}^*\|_2 &= \|\lambda(\mathbf{x}_1 - \mathbf{x}^*) + (1 - \lambda)(\mathbf{x}_2 - \mathbf{x}^*)\|_2 \\ &\leq \lambda \|\mathbf{x}_1 - \mathbf{x}^*\|_2 + \|(1 - \lambda)(\mathbf{x}_2 - \mathbf{x}^*)\|_2 < \lambda r + (1 - \lambda)r = r \end{aligned}$$

其中第一个不等号根据范数的三角不等式关系获得，第二个不等号根据开球定义获得。所以，根据凸集的定义，开球 $B_r(\mathbf{x}^*)$ 是凸集。□

定义 1.5. 超平面 (Hyperplane) 定义为

$$H = \{\mathbf{x} \mid \mathbf{h}^T \mathbf{x} = z, \mathbf{x} \in \mathbb{R}^n\} \quad (1.12)$$

其中 $\mathbf{h} \in \mathbb{R}^n / \{\mathbf{0}\}$ 是 \mathbb{R}^n 中的给定的非零 n 维向量，也被称为超平面 H 的法线方向。 $z \in \mathbb{R}$ 为给定实数。

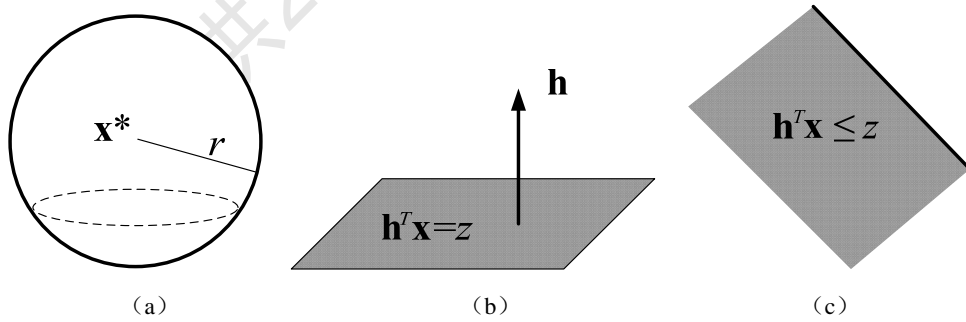


图 1.1 典型凸集示意图：(a) 球；(b) 超平面；(c) 半空间。

定理 1.3. 法线方向 \mathbf{h} 和超平面 H 中所有直线正交。

证明: 对于超平面 H 中任意两点 $\mathbf{x}_1, \mathbf{x}_2 \in H$, 过这两点的直线方向为 $\mathbf{x}_1 - \mathbf{x}_2$, 必然满足 $\mathbf{h}^T (\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{h}^T \mathbf{x}_1 - \mathbf{h}^T \mathbf{x}_2 = z - z = 0$, 也即该直线必然与法线方向 \mathbf{h} 正交。则上述结论显然。□

定义 1.6. 开半空间 (Open Half-Space) 定义为

$$S = \{\mathbf{x} | \mathbf{h}^T \mathbf{x} < z, \mathbf{x} \in \mathbb{R}^n\} \quad (1.13)$$

闭半空间 (Closed Half-Space) 定义为

$$S = \{\mathbf{x} | \mathbf{h}^T \mathbf{x} \leq z, \mathbf{x} \in \mathbb{R}^n\} \quad (1.14)$$

其中 $\mathbf{h} \in \mathbb{R}^n / \{\mathbf{0}\}$ 是 \mathbb{R}^n 中的给定的非零 n 维向量, $z \in \mathbb{R}$ 为给定实数。

定理 1.4. 任意一个超平面, 半空间都是凸集。

证明: 我们仅证明超平面是凸集, 半空间证明与之类似。

对于任意两点 $\mathbf{x}_1, \mathbf{x}_2 \in H$ 和任意实数 $\lambda \in [0, 1]$, 我们可以得到

$$\mathbf{h}^T [\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2] = \lambda \mathbf{h}^T \mathbf{x}_1 + (1 - \lambda) \mathbf{h}^T \mathbf{x}_2 = \lambda z + (1 - \lambda) z = z$$

所以根据凸集的定义, 超平面 H 是凸集。□

定义 1.7. 多面体 (Polyhedron) 定义为同时满足有限个仿射不等式的 n 维向量所构成的集合

$$P = \{\mathbf{x} | \mathbf{a}_i^T \mathbf{x} \leq b_i, i = 1, \dots, m, \mathbf{x} \in \mathbb{R}^n\} \quad (1.15)$$

其中 $\mathbf{a}_i \in \mathbb{R}^n$ 为给定的 n 维向量, $b_i \in \mathbb{R}$ 为给定实数。很显然, 一个多面体可以视为有限个闭半空间的交集。我们也可以将其记为 $P = \{\mathbf{x} | \mathbf{A} \mathbf{x} \leq \mathbf{b}, i = 1, \dots, m, \mathbf{x} \in \mathbb{R}^n\}$, 其中 $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ 。

有界多面体 (Polytope) 定义为同时满足有限个仿射不等式的有界 n 维向量所构成的集合

$$P = \{\mathbf{x} | \mathbf{a}_i^T \mathbf{x} \leq b_i, i = 1, \dots, m, l_j \leq x_j \leq u_j, j = 1, \dots, n, \mathbf{x} \in \mathbb{R}^n\} \quad (1.16)$$

其中 $l_j, u_j \in \mathbb{R}$ 为给定实数, 表示 x_j 的上下界。

定理 1.5. 任意一个多面体是凸集。

证明: 注意到半平面是凸集, 结合定理 1 和多面体的定义, 结论显然。□

定义 1.8. 假设集合 $\Omega \in \mathbb{R}^n$, 如果以点 \mathbf{x} 为中心的任意一个开球都包含集合 Ω 中的一个点和集合 $\mathbb{R}^n - \Omega$ 中的一个点, 则点 \mathbf{x} 为集合 Ω 的边界点 (Boundary Point)。 Ω 的所有边界点构成的集合记为 $\partial\Omega$, 称为 Ω 的边界 (Boundary)。

如果 $\partial\Omega \subset \Omega$ ，则 Ω 为闭集 (Closed Set)。如果 $\mathbb{R}^n - \partial\Omega$ 为闭集，则 Ω 为开集 (Open Set)。

集合 Ω 的闭包 (Closure) 定义为 $\bar{\Omega} = \Omega \cup \partial\Omega$ 。集合 Ω 的内部 (Interior) 定义为 $\Omega^\circ = \Omega - \partial\Omega$ 。 Ω 为闭集，当且仅当 $\bar{\Omega} = \Omega$ 。 Ω 为开集，当且仅当 $\Omega = \Omega^\circ$ 。

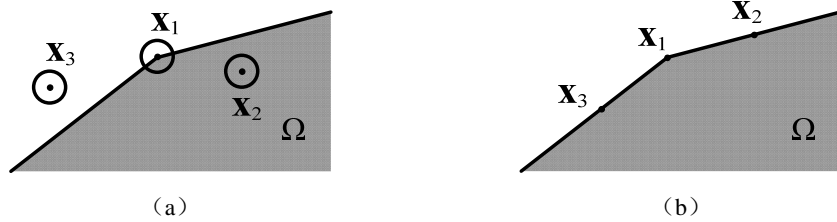


图 1.2 示意图: (a) \mathbf{x}_1 是集合 Ω 的边界点, \mathbf{x}_2 是集合 Ω 的内部点, \mathbf{x}_3 是集合 Ω 之外的点; (b) \mathbf{x}_1 是集合 Ω 的极点, \mathbf{x}_2 和 \mathbf{x}_3 是集合 Ω 的边界点但非极点。

定义 1.9. 点 \mathbf{x} 属于凸集 Ω ，如果不存在两个不同的点 $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ 和实数 $\lambda \in [0, 1]$ ，使得点 $\mathbf{x} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ ，则点 \mathbf{x} 为凸集 Ω 的极点 (Extreme Point)。

定义 1.10. 如果一个 C 中的任意一个点 $\mathbf{x} \in C$ ，满足对于任意的 $\lambda > 0$ ， $\lambda\mathbf{x} \in C$ ，则该集合 C 称为锥 (Cone)。如果集合 C 同时是凸集，则称为凸锥 (Convex Cone)。显然，多面体 $P = \{\mathbf{x} | \mathbf{A}\mathbf{x} \geq \mathbf{0}, \mathbf{x} \in \mathbb{R}^n\}$ 满足锥的定义，我们称这样的多面体为多面体锥 (Polyhedral Cone)。

定义 1.11. 考虑一个非空多面体 $P = \{\mathbf{x} | \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \in \mathbb{R}^n\}$ ，如果固定某个点 $\mathbf{y} \in P$ ，我们将点 \mathbf{y} 处的回收锥 (Recession Cone) 定义为所有方向 $\mathbf{d} \in \mathbb{R}^n$ 的集合，沿着这些方向 \mathbf{d} 我们可以无限远离点 \mathbf{y} ，而不会离开集合 P 。换句话说，回收锥表示为集合

$$R = \{\mathbf{d} | \mathbf{A}[\mathbf{y} + \lambda\mathbf{d}] \geq \mathbf{b}, \text{ for all } \lambda \geq 0, \mathbf{d} \in \mathbb{R}^n\} \quad (1.17)$$

实际上，很容易证明回收锥的定义等价于 $R = \{\mathbf{d} | \mathbf{A}\mathbf{d} \geq \mathbf{0}, \mathbf{d} \in \mathbb{R}^n\}$ 。因此，回收锥实际上是与 \mathbf{y} 无关的一个多面体锥。

定义 1.12. 回收锥 $R = \{\mathbf{d} | \mathbf{A}\mathbf{d} \geq \mathbf{0}, \mathbf{d} \in \mathbb{R}^n\}$ 的任意一个非零成员为多面体 $P = \{\mathbf{x} | \mathbf{a}_i^T \mathbf{x} \geq b_i, i = 1, \dots, m, \mathbf{x} \in \mathbb{R}^n\}$ 的射线 (Ray)。而如果该非零成员使得上述不等式组中恰有 $n-1$ 个线性独立的约束条件 $\mathbf{a}_i^T \mathbf{x} \leq b_i$ 起作用 (Active)，也即 $n-1$ 个线性独立的不等式约束条件恰好取等号满足，则该非零成员为多面体 $P = \{\mathbf{x} | \mathbf{a}_i^T \mathbf{x} \geq b_i, i = 1, \dots, m, \mathbf{x} \in \mathbb{R}^n\}$ 的极线 (Extreme Ray)。

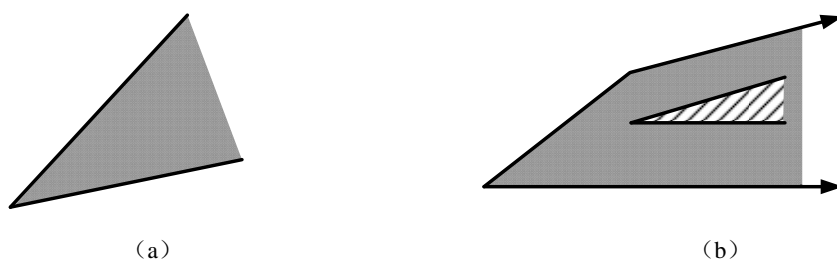


图 1.3 典型凸集示意图：(a) 多面体锥；(b) 多面体内的回收锥（斜纹阴影所示）和极线（箭头射线所示）。

定理 1.6. [凸集分隔定理] 如果某个点 $\mathbf{y} \in \mathbb{R}^n$ 位于闭凸集 $\Omega \subset \mathbb{R}^n$ 之外，则必然存在一个超平面 $H = \{\mathbf{x} \mid \mathbf{h}^T \mathbf{x} = z, \mathbf{x} \in \mathbb{R}^n\}$ 通过点 \mathbf{y} ，且满足 $\Omega \subset \{\mathbf{x} \mid \mathbf{h}^T \mathbf{x} > z\}$ ； $\mathbf{h} \in \mathbb{R}^n / \{0\}$ 是 \mathbb{R}^n 中的某个非零 n 维向量， $z \in \mathbb{R}$ 为某个实数。

更进一步，同样必然存在另外一个超平面 $H = \{\mathbf{x} \mid \mathbf{h}^T \mathbf{x} = z, \mathbf{x} \in \mathbb{R}^n\}$ 同时满足 $\Omega \subset \{\mathbf{x} \mid \mathbf{h}^T \mathbf{x} > z\}$ 和 $\mathbf{h}^T \mathbf{y} < z$ 。我们称之为存在超平面分隔点 \mathbf{y} 和闭凸集 Ω 。

证明：我们可以定义点 \mathbf{y} 到集合 Ω 的距离为 $\inf_{\mathbf{x} \in \Omega} \|\mathbf{x} - \mathbf{y}\|_2 > 0$ 。由于 Ω 是闭凸集，因此必然存在一个点 \mathbf{x}_0 使得连续函数 $\|\mathbf{x} - \mathbf{y}\|_2$ 取得最小值，参见图 1.4。

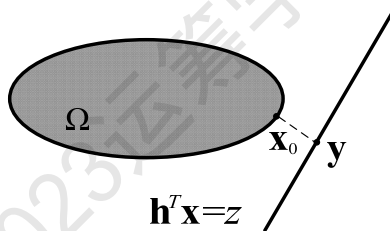


图 1.4 凸集分隔定理证明示意图。

下面我们使用反证法证明这个点 \mathbf{x}_0 一定是个边界点。如果 $\mathbf{x}_0 \in \Omega^\circ$ ，则必然存在 $2\varepsilon > 0$ ，使得以 \mathbf{x}_0 为球心， 2ε 为半径的开球位于 Ω 内部，也即 $B_{2\varepsilon}(\mathbf{x}_0) \subset \Omega$ 。我们可以由此确定一个新点

$$\mathbf{x}_1 = \mathbf{x}_0 + \varepsilon \frac{\mathbf{y} - \mathbf{x}_0}{\|\mathbf{x}_0 - \mathbf{y}\|_2} \in B_{2\varepsilon}(\mathbf{x}_0) \subset \Omega$$

不难发现，该点 \mathbf{x}_1 到点 \mathbf{y} 的距离比点 \mathbf{x}_0 到点 \mathbf{y} 的距离更小

$$\|\mathbf{x}_1 - \mathbf{y}\|_2 = \left\| \mathbf{x}_0 + \varepsilon \frac{\mathbf{y} - \mathbf{x}_0}{\|\mathbf{y} - \mathbf{x}_0\|_2} - \mathbf{y} \right\| = \left(1 - \frac{\varepsilon}{\|\mathbf{y} - \mathbf{x}_0\|_2} \right) \|\mathbf{x}_0 - \mathbf{y}\|_2 < \|\mathbf{x}_0 - \mathbf{y}\|_2$$

这和我们假设的点 \mathbf{x}_0 到点 \mathbf{y} 的距离最小矛盾，因此点 \mathbf{x}_0 一定是个边界点。

下面，我们定义 $\mathbf{h} = \mathbf{x}_0 - \mathbf{y}$ 和 $z = \mathbf{h}^T \mathbf{y}$ ，然后证明 $\Omega \subset \{\mathbf{x} | \mathbf{h}^T \mathbf{x} > z\}$ 。

对于所有的 $\mathbf{x} \in \Omega$ 和任意的实数 $\lambda \in (0, 1)$ ，由于 Ω 是凸集，我们有

$$\bar{\mathbf{x}} = \mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0) \in \Omega, \text{ 且 } \|\bar{\mathbf{x}} - \mathbf{y}\|_2 \geq \|\mathbf{x}_0 - \mathbf{y}\|_2$$

也即有

$$\begin{aligned} \|\bar{\mathbf{x}} - \mathbf{y}\|_2^2 &\geq \|\mathbf{x}_0 - \mathbf{y}\|_2^2 \Rightarrow (\bar{\mathbf{x}} - \mathbf{y})^T (\bar{\mathbf{x}} - \mathbf{y}) \geq (\mathbf{x}_0 - \mathbf{y})^T (\mathbf{x}_0 - \mathbf{y}) \\ &\Rightarrow 2\lambda(\mathbf{x}_0 - \mathbf{y})^T (\mathbf{x} - \mathbf{x}_0) + \lambda^2(\mathbf{x} - \mathbf{x}_0)^T (\mathbf{x} - \mathbf{x}_0) \geq 0 \end{aligned}$$

注意到 $\lambda > 0$ ，可以从上式两端同时除以 λ 保持不等式不变号成立，再带入 $\mathbf{h} = \mathbf{x}_0 - \mathbf{y}$ ，整理可得

$$\mathbf{h}^T (\mathbf{x} - \mathbf{x}_0) + \lambda(\mathbf{x} - \mathbf{x}_0)^T (\mathbf{x} - \mathbf{x}_0) \geq 0$$

接下来，我们使用反证法证明 $\mathbf{h}^T (\mathbf{x} - \mathbf{x}_0) \geq 0$ 。假设 $\mathbf{h}^T (\mathbf{x} - \mathbf{x}_0) < 0$ ，我们总能够找到一个足够小的 $\lambda \rightarrow 0$ 使得 $\lambda(\mathbf{x} - \mathbf{x}_0)^T (\mathbf{x} - \mathbf{x}_0) \rightarrow 0$ ，从而使得 $\mathbf{h}^T (\mathbf{x} - \mathbf{x}_0) + \lambda(\mathbf{x} - \mathbf{x}_0)^T (\mathbf{x} - \mathbf{x}_0) < 0$ ，但这和我们上面推出的不等式矛盾。

因此，我们可以得到

$$\mathbf{h}^T \mathbf{x} = \mathbf{h}^T \mathbf{x}(\mathbf{x} - \mathbf{x}_0) + \mathbf{h}^T (\mathbf{x}_0 - \mathbf{y}) + \mathbf{h}^T \mathbf{y} \geq \mathbf{h}^T \mathbf{h} + z > z$$

因此， $\Omega \subset \{\mathbf{x} | \mathbf{h}^T \mathbf{x} > z\}$ 。

更进一步，如果我们定义 $\mathbf{h} = \mathbf{x}_0 - \mathbf{y}$ 和 $z_1 = \mathbf{h}^T \left(\frac{1}{2} \mathbf{y} + \frac{1}{2} \mathbf{x}_0 \right)$ ，则可以发现超

平面 $\mathbf{h}^T \mathbf{y} < z_1$ 分隔点 \mathbf{y} 和闭凸集 Ω ，同时满足 $\Omega \subset \{\mathbf{x} | \mathbf{h}^T \mathbf{x} > z_1\}$ 和 $\mathbf{h}^T \mathbf{x} = z_1$ 。□

定义 1.13. 考虑闭凸集 $\Omega \subset \mathbb{R}^n$ 的边界点 \mathbf{x} ，如果存在一个超平面 $H = \{\mathbf{x} | \mathbf{h}^T \mathbf{x} = z, \mathbf{x} \in \mathbb{R}^n\}$ 通过点 \mathbf{x} ，且满足 $\Omega \subseteq \{\mathbf{x} | \mathbf{h}^T \mathbf{x} \geq z\}$ ； $\mathbf{h} \in \mathbb{R}^n / \{\mathbf{0}\}$ 是 \mathbb{R}^n 中的某个非零 n 维向量， $z \in \mathbb{R}$ 为某个实数。我们称这样的超平面为闭凸集 Ω 的支撑超平面。

定理 1.7. [支撑超平面定理] 对于闭凸集 $\Omega \subset \mathbb{R}^n$ 的任意一个边界点 \mathbf{y} ，必然存在一个支撑超平面 $H = \{\mathbf{x} | \mathbf{h}^T \mathbf{x} = z, \mathbf{x} \in \mathbb{R}^n\}$ ； $\mathbf{h} \in \mathbb{R}^n / \{\mathbf{0}\}$ 是 \mathbb{R}^n 中的某个非零 n 维向量， $z \in \mathbb{R}$ 为某个实数。

证明：由于 \mathbf{y} 为闭凸集 Ω 的边界点，则对于任意 $n \in \mathbb{N}$ ，以 \mathbf{x} 为球心， $1/n$ 为半径的球中必然含有一个点 \mathbf{y}_n 不属于集合 Ω 的闭包，也即

$$\mathbf{y}_n \notin \bar{\Omega}, \text{ 但 } \lim_{n \rightarrow +\infty} \mathbf{y}_n = \mathbf{y}$$

根据定理 1.6，可得到一系列分隔超平面 $\Omega \subseteq H = \{\mathbf{x} \mid \mathbf{h}_n^T \mathbf{x} > \mathbf{h}_n^T \mathbf{y}_n, \mathbf{x} \in \mathbb{R}^n\}$ ， $n \in \mathbb{N}$ ，将闭凸集 Ω 和点 \mathbf{y}_n 分隔开来。

我们可以进一步要求 $\|\mathbf{h}_n\|_2 = 1$ ，则 $\{\mathbf{h}_n\}$ 是紧空间中的有界点列，因此其中必

有一个子列 $\{\mathbf{h}_{n_k}, k \in \mathbb{N}\}$ 收敛于 $\mathbf{h} = \lim_{k \rightarrow +\infty} \mathbf{h}_{n_k}$ 。

我们定义 $z = \mathbf{h}^T \mathbf{y}$ ，则对于任意点 $\mathbf{x} \in \Omega$ ，我们有

$$\mathbf{h}^T \mathbf{x} = \lim_{n \rightarrow +\infty} \mathbf{h}_{n_k}^T \mathbf{x} \geq \lim_{n \rightarrow +\infty} \mathbf{h}_{n_k}^T \mathbf{y}_{n_k} = \mathbf{h}^T \mathbf{y} = z$$

因此，超平面 $H = \{\mathbf{x} \mid \mathbf{h}^T \mathbf{x} = z, \mathbf{x} \in \mathbb{R}^n\}$ 是闭凸集 Ω 的支撑超平面。□

定义 1.14. 一个集合 $\Omega \subset \mathbb{R}^n$ 被称为下有界 (Bounded from Below)，如果对于任意一个点 $\mathbf{x} = [x_1, \dots, x_n] \in \Omega$ ，存在

$$\inf \{x_i\} > -\infty, \quad i = 1, \dots, n \quad (1.18)$$

定理 1.8. 如果闭凸集 $\Omega \subset \mathbb{R}^n$ 下有界，则集合 Ω 的任意一个支撑超平面上至少含有一个集合 Ω 的极点。

证明：假设点 \mathbf{x}_0 是闭凸集 Ω 的一个边界点，而 $H = \{\mathbf{x} \mid \mathbf{h}^T \mathbf{x} = z, \mathbf{x} \in \mathbb{R}^n\}$ 是通过点 \mathbf{x}_0 的相对于集合 Ω 的一个支撑超平面，满足 $\Omega \subseteq \{\mathbf{x} \mid \mathbf{h}^T \mathbf{x} \geq z\}$ 。我们定义该支撑超平面和集合 Ω 的交集为 $\Delta = \Omega \cap H$ 。显然，由于 $\mathbf{x}_0 \in \Delta$ ， Δ 不是空集。

接下来，我们先证明任意一个集合 Δ 的极点同时也是集合 Ω 的极点。实际上，我们转而证明该命题的逆否命题，也即如果一个点 \mathbf{t} 不是集合 Ω 的极点，则该点 \mathbf{t} 也一定不是集合 Δ 的极点。具体而言，此时有两种情况。

第一种情况，如果 $\mathbf{t} \notin \Omega$ ，则结论显然。

第二种情况， $\mathbf{t} \in \Omega$ 。如果 \mathbf{t} 不是集合 Ω 的极点，则存在两个不同的点 $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ 和某个实数 $\lambda \in (0, 1)$ ，满足 $\mathbf{t} = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2$ 。而根据支撑超平面的性质，我们有

$$\mathbf{h}^T \mathbf{t} = \lambda \mathbf{h}^T \mathbf{x}_1 + (1 - \lambda) \mathbf{h}^T \mathbf{x}_2 = z, \text{ 同时 } \mathbf{h}^T \mathbf{x}_1 \geq z, \mathbf{h}^T \mathbf{x}_2 \geq z$$

上述关系同时成立，则必然有 $\mathbf{h}^T \mathbf{x}_1 = \mathbf{h}^T \mathbf{x}_2 = z$ ，所以点 $\mathbf{x}_1, \mathbf{x}_2 \in \Delta$ ，故而 \mathbf{t} 在集合 Δ 内不同两点 $\mathbf{x}_1, \mathbf{x}_2$ 的连线段上，因此 \mathbf{t} 不是集合 Δ 的极点。该命题的逆否命题得证。根据原命题和逆否命题的关系，原命题亦得证。

然后，我们证明集合 Δ 至少存在一个极点。显然，如果集合 Δ 仅含一个元素

\mathbf{x}_0 , 则 \mathbf{x}_0 必然为一个极点, 我们已经找到一个满足要求的极点, 可以停止。如果集合 Δ 含有多个元素, 我们使用如下有限递降法进行证明。

我们构造新的集合 $\Delta^1 = \arg \min \{x_1 | \mathbf{x} \in \Omega\}$ 。因为 $\Delta \subset \Omega$ 且集合 Ω 下有界, 闭集合 Δ^1 实际上定义为集合 Ω 中, 所有 x_1 取得最小值的点 \mathbf{x} 所构成的集合。

如果 Δ^1 只含一个元素, 则我们已经找到一个满足要求的极点, 可以停止。否则, 我们进一步构造新的集合 $\Delta^j = \arg \min \{x_j | \mathbf{x} \in \Delta^{j-1}\}$, $j=2, \dots, n$ 。如此进行下去, 这一构造过程必然在最多 n 次之后停止; 因为 $\Delta \subset \mathbb{R}^n$, 构造过程执行 n 次之后将得到唯一的一个点。假设我们在 k 次之后停止, 得到单点集合 $\Delta^k = \{\mathbf{x}^k\}$, 此时 $\Delta^k \subseteq \Delta^{k-1} \subseteq \dots \subseteq \Delta^1 \subseteq \Delta$ 。

最后, 我们使用反证法证明点 \mathbf{x}^k 必为集合 Δ 的一个极点。假设存在两个不同的点 $\mathbf{x}^1, \mathbf{x}^2 \in \Delta$ 和某个实数 $\lambda \in (0, 1)$, 满足 $\mathbf{x}^k = \lambda \mathbf{x}^1 + (1-\lambda) \mathbf{x}^2$ 。则对于任意 $i \leq k$, 存在 $x_i^k = \lambda x_i^1 + (1-\lambda) x_i^2$ 。根据集合 Δ^1 的构造方法, x_1^k 为 x_1 能取到的最小值, 则必然有 $x_1^k = x_1^1 = x_1^2$, 也即 $\mathbf{x}^1, \mathbf{x}^2 \in \Delta^1$ 。类似的, 我们可以依次得到 $x_j^k = x_j^1 = x_j^2$, 也即 $\mathbf{x}^1, \mathbf{x}^2 \in \Delta^j$, $j=2, \dots, k$, 并最终得出 $\mathbf{x}^1 = \mathbf{x}^2 = \mathbf{x}^k \in \Delta^k$, 引出矛盾。因此, 点 \mathbf{x}^k 必为集合 Δ 的一个极点。□

定理 1.9. [Farkas 引理] 给定矩阵 $\mathbf{A} \in \mathbb{R}^{m \times n}$, 向量 $\mathbf{b} \in \mathbb{R}^m$, 下述两个线性系统有且仅有一个存在可行解 [8]:

$$\text{线性系统 I: 对于 } \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \quad (1.19)$$

$$\text{线性系统 II: 对于 } \mathbf{y} \in \mathbb{R}^m, \mathbf{A}^T \mathbf{y} \geq \mathbf{0}, \mathbf{b}^T \mathbf{y} < 0 \quad (1.20)$$

证明: 根据多面体的定义, 上述两个线性系统的可行解构成的集合 Ω_1, Ω_2 必然是凸集。为了应用凸集分隔定理, 我们还需要先证明集合 Ω_1 是闭集。这个证明有两种思路。其一是引入锥包 (Conic Hull) 的概念, 接着证明锥包是闭集, 最后证明 Ω_1 为锥包。其二是将 $\mathbf{A} \in \mathbb{R}^{m \times n}$ 看成是 \mathbb{R}^n 到 \mathbf{A} 的列空间 $\text{span}(\mathbf{A})$ 的一个映射, 接着分析这个映射的性质, 考察闭集 \mathbb{R}^n 在映射下, 原像和像之间的关系证明 Ω_1 为锥包。因为这些证明较为繁琐且和本书主线相距较远, 留给大家作为练习。

接下来, 我们应用凸集分隔定理证明 Farkas 引理。

首先我们使用反证法证明上述两个线性系统不可能同时有可行解。如果集合 Ω_1, Ω_2 同时非空, 则存在可行解 $\mathbf{y} \in \mathbb{R}^m$ 满足 $\mathbf{b}^T \mathbf{y} < 0$, 这表明 $\mathbf{b}^T \neq \mathbf{0}$ 。而 $\mathbf{b}^T \neq \mathbf{0}$, 则如果有可行解 $\mathbf{x} \in \mathbb{R}^n$ 满足 $\mathbf{Ax} = \mathbf{b}$, 则可行解 $\mathbf{x} \neq \mathbf{0}$ 。考虑到 $\mathbf{x} \geq \mathbf{0}, \mathbf{A}^T \mathbf{y} \geq \mathbf{0}$, 我们得到

$$\mathbf{b}^T \mathbf{y} = (\mathbf{Ax})^T \mathbf{y} = \mathbf{x}^T (\mathbf{Ay}) \geq 0$$

这与 $\mathbf{b}^T \mathbf{y} < 0$ 的假设矛盾, 因此集合 Ω_1, Ω_2 不能同时非空。

最后, 我们证明上述两个线性系统必然有一个有解。具体而言, 此时有两种情况。

第一种情况, 线性系统 I 有解, 则结论显然。

第二种情况, 线性系统 I 无解, 这说明点 $\mathbf{b} \notin \Omega = \{\mathbf{Ax}, \mathbf{x} \geq \mathbf{0}\}$, 而我们前面证明了 $\Omega = \{\mathbf{Ax}, \mathbf{x} \geq \mathbf{0}\}$ 是闭凸集, 则根据凸集分隔定理, 存在一个超平面 $H = \{\mathbf{x} | \mathbf{y}^T \mathbf{x} = z, \mathbf{s} \in \mathbb{R}^m\}$ 分隔点 \mathbf{b} 和闭凸集 Ω 。也即 $\mathbf{y}^T \mathbf{b} < z$, 同时任取一点 $\mathbf{s} \in \Omega$, 满足 $\mathbf{y}^T \mathbf{s} \geq z$ 。

考虑到 $\mathbf{0} \in \Omega$, 必然有 $z < 0$, 因此 $\mathbf{y}^T \mathbf{b} < z < 0$ 。

另一方面, 对于任意一点 $\mathbf{x} \geq \mathbf{0}$, 满足 $\mathbf{y}^T \mathbf{Ax} > 0$, 则一定有 $\mathbf{y}^T \mathbf{A} > \mathbf{0}$ 。我们使用反证法证明此命题。若向量 $\mathbf{y}^T \mathbf{A}$ 的第 i 个元素是负值, $i = 1, \dots, m$, 我们可以选择除了对应的第 i 个元素是正值, 其余元素全为 0 的某个特定的 $\hat{\mathbf{x}}$ 使得 $\mathbf{y}^T \mathbf{Ax} < 0$, 产生矛盾。

综上所述, Farkas 引理成立。□

实际上, 根据公式 (1.7), 我们可以 $\{\mathbf{Ax} | \mathbf{x} \geq \mathbf{0}\}$ 视为矩阵 \mathbf{A} 的列向量的所有非负加权组合所构成的点集。不难证明这个集合实际是一个凸锥。如图 1.5 所示, 如果向量 \mathbf{b} 落在该凸锥内, 线性系统 I 一定有解。反之, 如果向量 \mathbf{b} 未落在该凸锥内, 线性系统 I 一定无解; 并且此时, 根据凸集分隔定理, 一定能够找到一个过原点的超平面 $\mathbf{y}^T \mathbf{x} = 0$, 使得向量 \mathbf{b} 在超平面 $\mathbf{y}^T \mathbf{x} = 0$ 的一边, 凸锥 $\{\mathbf{Ax} | \mathbf{x} \geq \mathbf{0}\}$ 在超平面 $\mathbf{y}^T \mathbf{x} = 0$ 的另外一边。因此, 线性系统 II 一定有解。

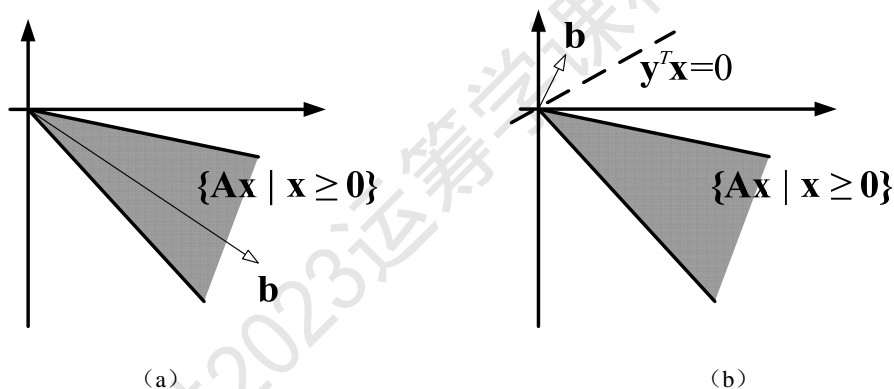


图 1.5 Farkas 引理示意图: (a) 向量 \mathbf{b} 落在凸锥 $\{\mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ 内, 线性系统 I 一定有解; (b) 向量 \mathbf{b} 未落在凸锥 $\{\mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ 内, 线性系统 I 一定无解。

定义 1.15. $f: \Omega \rightarrow \mathbb{R}$ 是定义在凸集 $\Omega \subset \mathbb{R}^n$ 上函数,

如果对于任意的 $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$, 和任意实数 $\lambda \in [0, 1]$ 存在

$$f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda) f(\mathbf{x}_2) \quad (1.21)$$

则函数 f 为凸集 Ω 上的凸函数 (Convex Function)。

如果对于任意的 $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$, 和任意实数 $\lambda \in [0, 1]$ 存在

$$f(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) < \lambda f(\mathbf{x}_1) + (1-\lambda)f(\mathbf{x}_2) \quad (1.22)$$

则函数 f 为凸集 Ω 上的严格凸函数 (Strictly Convex Function)。

如果对于任意的 $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$, 和任意实数 $\lambda \in [0, 1]$ 存在

$$f(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \geq \lambda f(\mathbf{x}_1) + (1-\lambda)f(\mathbf{x}_2) \quad (1.23)$$

则函数 f 为凸集 Ω 上的凹函数 (Concave Function)。

如果对于任意的 $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$, 和任意实数 $\lambda \in [0, 1]$ 存在

$$f(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) > \lambda f(\mathbf{x}_1) + (1-\lambda)f(\mathbf{x}_2) \quad (1.24)$$

则函数 f 为凸集 Ω 上的严格凹函数 (Strictly Concave Function)。

定理 1.10. [Jensen 不等式 (Jensen's Inequality)] 若 $f: \Omega \rightarrow \mathbb{R}$ 是定义在凸集 $\Omega \subset \mathbb{R}^n$ 上凸函数, 对于任意的 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m \in \Omega$, 任意实数 $\lambda_i \in [0, 1]$ 且满足

$\sum_{i=1}^m \lambda_i = 1, m \in \mathbb{N}$, 则成立

$$f(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \dots + \lambda_m \mathbf{x}_m) \leq \lambda_1 f(\mathbf{x}_1) + \lambda_2 f(\mathbf{x}_2) + \dots + \lambda_m f(\mathbf{x}_m) \quad (1.25)$$

反之, 如果成立 Jensen 不等式, $f: \Omega \rightarrow \mathbb{R}$ 是定义凸集 Ω 上的凸函数。

证明: 我们使用数学归纳法来证明正向论断。显然, $m = 2$ 时, 根据凸函数的定义, Jensen 不等式成立。假设 $m = k$ 时, Jensen 不等式成立, 则当 $m = k + 1$ 时, 我们利用凸函数的性质, 可以得到

$$\begin{aligned} & f(\lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k + \lambda_{k+1} \mathbf{x}_{k+1}) \\ & \leq (1 - \lambda_{k+1}) f\left(\frac{\lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k}{1 - \lambda_{k+1}}\right) + \lambda_{k+1} f(\mathbf{x}_{k+1}) \\ & \leq (1 - \lambda_{k+1}) \left[\frac{\lambda_1 f(\mathbf{x}_1)}{1 - \lambda_{k+1}} + \dots + \frac{\lambda_k f(\mathbf{x}_k)}{1 - \lambda_{k+1}} \right] + \lambda_{k+1} f(\mathbf{x}_{k+1}) \\ & = \lambda_1 f(\mathbf{x}_1) + \dots + \lambda_k f(\mathbf{x}_k) + \lambda_{k+1} f(\mathbf{x}_{k+1}) \end{aligned}$$

故当 $m = k + 1$ 时, Jensen 不等式也成立。根据数学归纳法, 正向命题成立。注意到 $m = 2$ 时, Jensen 不等式即为凸函数定义, 因此反向命题成立。□

定理 1.11. 定义在凸集 $\Omega \subset \mathbb{R}^n$ 上的仿射函数 (Affine Function) $f = \mathbf{a}^T \mathbf{x} + b$, $\mathbf{x} \in \mathbb{R}^n$, 既是凸函数, 又是凹函数; 其中 $\mathbf{a} \in \mathbb{R}^n$ 为给定的 n 维向量, $b \in \mathbb{R}$ 为给定实数。

证明: 根据凸函数和凹函数的定义, 结论显然。□

定理 1.12. [一元凸函数和多元凸函数的关系] 假设 $f: \Omega \rightarrow \mathbb{R}$ 是定义在凸集 $\Omega \subset \mathbb{R}^n$ 上的函数；对于任意点 $\mathbf{x} \in \Omega$ 和方向 $\mathbf{d} \in \mathbb{R}^n$ ，我们定义一族一元函数 $\varphi(t | \mathbf{x}, \mathbf{d}): \mathbb{R} \rightarrow \mathbb{R}$ 为

$$\varphi(t | \mathbf{x}, \mathbf{d}) = f(\mathbf{x} + t\mathbf{d}), \quad \forall \mathbf{x} + t\mathbf{d} \in \Omega \quad (1.26)$$

则函数 f 为凸集 Ω 上的凸函数的充分必要条件是，这一族函数 φ 均为凸函数。

证明：我们先证明必要性。对于给定的点 $\mathbf{x} \in \Omega$ 和方向 $\mathbf{d} \in \mathbb{R}^n$ ，如果实数 t_1, t_2 满足 $\mathbf{x} + t_1\mathbf{d}, \mathbf{x} + t_2\mathbf{d} \in \Omega$ ，则根据凸集的定义，对于任意的实数 $\lambda \in [0, 1]$ ，满足 $\lambda(\mathbf{x} + t_1\mathbf{d}) + (1-\lambda)(\mathbf{x} + t_2\mathbf{d}) \in \Omega$ 。进一步根据 f 是凸函数的性质，有

$$\begin{aligned} \varphi(\lambda t_1 + (1-\lambda)t_2 | \mathbf{x}, \mathbf{d}) &= f(\lambda(\mathbf{x} + t_1\mathbf{d}) + (1-\lambda)(\mathbf{x} + t_2\mathbf{d})) \\ &\leq \lambda f(\mathbf{x} + t_1\mathbf{d}) + (1-\lambda)f(\mathbf{x} + t_2\mathbf{d}) = \lambda\varphi(t_1 | \mathbf{x}, \mathbf{d}) + (1-\lambda)\varphi(t_2 | \mathbf{x}, \mathbf{d}) \end{aligned}$$

显然根据定义，对于任意点 $\mathbf{x} \in \Omega$ 和方向 $\mathbf{d} \in \mathbb{R}^n$ ，函数 φ 均为凸函数。

我们接着证明充分性。给定 $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ ，实数 $\lambda \in [0, 1]$ ，考虑函数 φ 是凸函数的性质，我们可以得到

$$\begin{aligned} f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) &= f(\mathbf{x}_1 + (1-\lambda)(\mathbf{x}_2 - \mathbf{x}_1)) \\ &= \varphi(\lambda \cdot 0 + (1-\lambda) \cdot 1 | \mathbf{x}_1, \mathbf{x}_2 - \mathbf{x}_1) \\ &\leq \lambda\varphi(0 | \mathbf{x}_1, \mathbf{x}_2 - \mathbf{x}_1) + (1-\lambda)\varphi(1 | \mathbf{x}_1, \mathbf{x}_2 - \mathbf{x}_1) \\ &= \lambda f(\mathbf{x}_1) + (1-\lambda)f(\mathbf{x}_2) \end{aligned}$$

显然根据定义，函数 f 均为凸函数。□

定理 1.13. [可微凸函数的一阶条件 (First-Order Condition)] 假设 $f: \Omega \rightarrow \mathbb{R}$ 是定义在凸集 $\Omega \subset \mathbb{R}^n$ 上可微函数；换言之，函数 f 的梯度 (Gradient) ∇f 在凸集 Ω 中的每个点都存在。函数 f 为凸集 Ω 上的凸函数的充分必要条件是，对于任意的 $\mathbf{x}, \mathbf{y} \in \Omega$ ，存在

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}) \quad (1.27)$$

函数 f 为凸集 Ω 上的严格凸函数的充分必要条件是，对于任意的 $\mathbf{x}, \mathbf{y} \in \Omega$ ， $\mathbf{x} \neq \mathbf{y}$ ，存在

$$f(\mathbf{y}) > f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}) \quad (1.28)$$

证明：我们仅证明凸函数的结论，严格凸函数结论的证明与之类似。

首先证明必要性。给定 $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ ，对于任意的实数 $\lambda \in (0, 1)$ ，根据凸集的性质，有 $\mathbf{x} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in \Omega$ ，应用一阶条件，得到

$$\begin{aligned} f(\mathbf{x}_1) &\geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{x}_1 - \mathbf{x}) = f(\mathbf{x}) + (1-\lambda)\nabla f(\mathbf{x})^T (\mathbf{x}_1 - \mathbf{x}_2) \\ f(\mathbf{x}_2) &\geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{x}_2 - \mathbf{x}) = f(\mathbf{x}) - \lambda\nabla f(\mathbf{x})^T (\mathbf{x}_1 - \mathbf{x}_2) \end{aligned}$$

移项整理有

$$f(\mathbf{x}_1) - f(\mathbf{x}) \geq (1 - \lambda) \nabla f(\mathbf{x})^T (\mathbf{x}_1 - \mathbf{x}_2)$$

$$f(\mathbf{x}_2) - f(\mathbf{x}) \geq -\lambda \nabla f(\mathbf{x})^T (\mathbf{x}_1 - \mathbf{x}_2)$$

加权合并这两个不等式有

$$\begin{aligned} & \lambda [f(\mathbf{x}_1) - f(\mathbf{x})] + (1 - \lambda) [f(\mathbf{x}_2) - f(\mathbf{x})] \\ & \geq \lambda (1 - \lambda) \nabla f(\mathbf{x})^T (\mathbf{x}_1 - \mathbf{x}_2) - \lambda (1 - \lambda) \nabla f(\mathbf{x})^T (\mathbf{x}_1 - \mathbf{x}_2) = 0 \end{aligned}$$

也即

$$\lambda f(\mathbf{x}_1) + (1 - \lambda) f(\mathbf{x}_2) \geq f(\mathbf{x}) = f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2)$$

则根据定义, 函数 f 为凸集 Ω 上的凸函数。

接着证明充分性。给定 $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$, $t \in (0, 1)$, 有

$$tf(\mathbf{x}_2) + (1 - t)f(\mathbf{x}_1) \geq f(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1))$$

将上式移项, 两边同时除以 t , 注意到 $t > 0$, 有

$$f(\mathbf{x}_2) - f(\mathbf{x}_1) \geq \frac{f(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1)) - f(\mathbf{x}_1)}{t}$$

令 $t \rightarrow 0$, 由极限保号性和方向导数的性质可得

$$f(\mathbf{x}_2) - f(\mathbf{x}_1) \geq \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1)) - f(\mathbf{x}_1)}{t} = \nabla f(\mathbf{x}_1)^T (\mathbf{x}_2 - \mathbf{x}_1)$$

则函数 f 满足一阶条件。□

定理 1.14. [二次可微凸函数的二阶条件 (Second-Order Condition)] 假设 $f: \Omega \rightarrow \mathbb{R}$ 是定义在凸集 $\Omega \subset \mathbb{R}^n$ 上二次可微函数; 换言之, 函数 f 的二阶导 (Second-Order Derivative, 或 Hessian) $\nabla^2 f$ 在凸集 Ω 中的每个点都存在。函数 f 为凸集 Ω 上的凸函数的充分必要条件是, 对于任意的 $\mathbf{x} \in \Omega$, $\nabla^2 f(\mathbf{x})$ 是半正定矩阵

$$\nabla^2 f(\mathbf{x}) \geq 0 \quad (1.29)$$

函数 f 为凸集 Ω 上的严格凸函数的充分必要条件是, 对于任意的 $\mathbf{x} \in \Omega$, $\nabla^2 f(\mathbf{x})$ 是正定矩阵

$$\nabla^2 f(\mathbf{x}) > 0 \quad (1.30)$$

证明: 我们仅证明凸函数的结论, 严格凸函数结论的证明与之类似。

我们首先使用反证法证明必要性。假设在点 $\mathbf{x} \in \Omega$, 存在 $\nabla^2 f(\mathbf{x})$ 的某个特征值为负。则一定存在非零向量 $\mathbf{d} \in \mathbb{R}^n$ 使得 $\mathbf{d}^T \nabla^2 f(\mathbf{x}) \mathbf{d} < 0$ 。考察函数 f 在点 \mathbf{x} 附近的带 Peano 余项的二次 Taylor 展开, 我们有

$$f(\mathbf{x} + t\mathbf{d}) = f(\mathbf{x}) + t\nabla f(\mathbf{x})^T \mathbf{d} + \frac{t^2}{2} \mathbf{d}^T \nabla^2 f(\mathbf{x}) \mathbf{d} + o(t^2)$$

移项后等式两边同时除以 t^2 , 有

$$\frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x}) - t\nabla f(\mathbf{x})^T \mathbf{d}}{t^2} = \frac{1}{2} \mathbf{d}^T \nabla^2 f(\mathbf{x}) \mathbf{d} + o(1)$$

显然, 当 t 充分小时, 有

$$\frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x}) - t\nabla f(\mathbf{x})^T \mathbf{d}}{t^2} < 0$$

也即 $f(\mathbf{x} + t\mathbf{d}) < f(\mathbf{x}) + t\nabla f(\mathbf{x})^T \mathbf{d}$ 。这和定理 1.13 的一阶条件矛盾, 因此必有 $\nabla^2 f(\mathbf{x}) \geq 0$ 。

我们接着证明充分性。给定 $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$, 考察函数 f 在点 \mathbf{x} 附近的二次 Taylor 展开, 我们有

$$f(\mathbf{x}_2) = f(\mathbf{x}_1) + \nabla f(\mathbf{x}_1)^T (\mathbf{x}_2 - \mathbf{x}_1) + \frac{1}{2} (\mathbf{x}_2 - \mathbf{x}_1)^T \nabla^2 f(\mathbf{x} + t(\mathbf{x}_2 - \mathbf{x}_1)) (\mathbf{x}_2 - \mathbf{x}_1)$$

其中 $t \in (0, 1)$ 是和 $\mathbf{x}_1, \mathbf{x}_2$ 有关的常数。

如果 $\nabla^2 f(\mathbf{x}) \geq 0$, 则

$$f(\mathbf{x}_2) \geq f(\mathbf{x}_1) + \nabla f(\mathbf{x}_1)^T (\mathbf{x}_2 - \mathbf{x}_1)$$

由定理 1.13 可知函数 f 为凸集 Ω 上的凸函数。□

1.6 计算复杂性

多项式复杂程度的非确定性问题 (Non-Deterministic Polynomial-Complete Problems, NP-C), 我们尚未找到此类问题的多项式时间算法 [10]。

1.7 作业题

1.1 假设矩阵 $\mathbf{A} \in \mathbb{R}^{m \times n}$ 是满秩的, $\text{rank}(\mathbf{A}) = n$ 。请证明或证否: 可以通过改变 \mathbf{A} 的至多 m^2 个元素的值, $m \leq n$, 使得新得到的矩阵 $\bar{\mathbf{A}}$ 满足 $\text{rank}(\bar{\mathbf{A}}) = n - m$ 。

1.2 如果一个集合 $S \in \mathbb{R}^n$, 且 S 和 \mathbb{R}^n 空间中任意一条直线的交集为凸集, 试证明集合 S 为凸集。

1.3 试证明一个多面体是有界多面体, 当且仅当该多面体不包含任何射线。

1.4 假设以下集合均为非空集合, 请判断哪些集合一定有极点, 并给出理由:

- a) $\Omega = \{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \in \mathbb{R}_+^n, \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m\}$ 。
- b) $\Omega = \{\mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b}, \mathbf{x} \in \mathbb{R}^n, \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m\}$ ，其中 \mathbf{A} 为行满秩矩阵。
- c) $\Omega = \{\mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b}, \mathbf{x} \in \mathbb{R}^n, \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m\}$ ，其中 \mathbf{A} 为列满秩矩阵。
- d) $\Omega = \{\mathbf{x} \mid \mathbf{Ax} \geq \mathbf{0}, \mathbf{c}^T \mathbf{x} = -1, \mathbf{x} \in \mathbb{R}^n, \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m, \mathbf{c} \in \mathbb{R}^n\}$ ，其中 \mathbf{A} 为

列满秩矩阵。

1.5 试证明多面体可行域 $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{b}\}$ 存在至少一个极点，必有 $\text{rank}(\mathbf{A}) = n$

1.6 试证明 $\Omega = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}, \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m\}$ 为闭集。

1.7 [Gordon 定理] 试证明给定矩阵 $\mathbf{A} \in \mathbb{R}^{m \times n}$ ，下述两个线性系统有且仅有一个存在可行解：

线性系统 I：对于 $\mathbf{x} \in \mathbb{R}^n$ ， $\mathbf{Ax} < \mathbf{0}$

线性系统 II：对于 $\mathbf{y} \in \mathbb{R}^m$ ， $\mathbf{A}^T \mathbf{y} = \mathbf{0}$ ， $\mathbf{y} \geq \mathbf{0}$ 但 $\mathbf{y} \neq \mathbf{0}$

1.7 [Stiemke 定理] 试证明给定矩阵 $\mathbf{A} \in \mathbb{R}^{m \times n}$ ，下述两个线性系统有且仅有一个存在可行解：

线性系统 I：对于 $\mathbf{x} \in \mathbb{R}^n$ ， $\mathbf{Ax} \geq \mathbf{0}$ 但 $\mathbf{Ax} \neq \mathbf{0}$

线性系统 II：对于 $\mathbf{y} \in \mathbb{R}^m$ ， $\mathbf{A}^T \mathbf{y} \geq \mathbf{0}$ ， $\mathbf{y} > \mathbf{0}$

1.8 参考文献

- [1] G. Polya, *How to Solve It: A New Aspect of Mathematical Method*, Princeton University Press, 1945.
- [2] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley, 1986.
- [3] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1996.
- [4] D. Bertsimas, J. N. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific Press, 1997.
- [5] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, Springer, 2003.
- [6] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [7] D. P. Bertsekas, *Convex Optimization Algorithms*, Athena Scientific, 2015.
- [8] D. G. Luenberger, Y. Ye, *Linear and Nonlinear Programming*, 4th edition, Springer, 2016.
- [9] G. Farkas, "A Fourier-féle mechanikai elv alkalmazásának algebrai alapjai," *Mathematikai Természettudományi Értesítő*, vol. 16, pp. 361-364, 1898.
- [10] R. M. Karp, "Reducibility among combinatorial problems," in *Complexity of Computer Computation*, R. E. Miller, J. W. Thatcher, eds., Plenum Press, pp. 85-103, 1972.

2.1 引言

线性规划 (Linear Programming, LP) 是运筹学中的一个重要分支, 研究线性约束条件下线性目标函数的实数极值规划问题。线性规划有着研究较早、应用广泛、方法较成熟的特点, 常被作为运筹学的基础。随着计算机的逐渐普及, 我们已经可以轻松求解数十万变量乃至更大的线性规划问题。目前, 线性规划在工程技术、商业活动、军事行动和工农业生产中发挥着越来越广泛的作用。

本章先介绍线性规划的基本概念, 然后通过几何和代数的观点分析线性规划问题, 之后介绍线性规划问题中广泛使用的对偶理论和单纯型法, 最后讨论一些相关问题。

2.2 线性规划基础

2.2.1 线性规划的历史

1939 年, 苏联学者 Kantorovich 关注到具有线性约束的极值优化问题, 并提出了解乘数法, 对资源最优分配理论做出了贡献。该方法是线性规划的雏形。1975 年, Kantorovich 获得诺贝尔经济学奖。

与此同时, 美国研究者们也开始关注线性规划问题。1941 年, Hitchcock 提出运输问题 [4]; 1945 年, Koopmans 同样对运输问题进行了研究 [5]。1947 年, Dantzig 担任美国空军数学顾问, 为找到解决问题的机制化工具, 完整提出了在一组线性方程或不等式约束下, 求某一线性形式极小值问题的数学模型, 也即线性规划问题; 并提出了单纯型法 (Simplex Method)。1948 年, Koopmans 和 Dantzig 在海滩散步时共同提出使用“线性规划”作为该类问题的名称 [6]。

1971 年, Klee 和 Minty 构造出一个例子, 证明在最坏情况下, 单纯形法是一个指数时间算法 [7]。

1979 年, Khachiyan 提出了第一个解决线性规划问题的多项式时间算法——椭球法 (Ellipsoid Method) [8]。这个算法证明了线性规划问题存在多项式时间算法, 但椭球法实际计算效率不高。

1984 年, Karmarkar 发明了内点法 (Interior Point Method), 这是线性规划第一个实际可用的多项式时间算法 [9]。

2004 年, Spielman 和 Teng 提出算法复杂性平滑分析法 (Smoothed Analysis of Algorithms), 证明单纯型法在很多线性规划问题的计算中只需多项式时间即可收敛到最优解 [10]。

现在有多种成熟的软件包来计算求解大规模线性规划问题。这些软件包大多同时集成实现了单纯型法和内点法。

2.2.2 典型问题和图解法

在各类经济活动中, 经常遇到这样的问题: 在特定的生产条件下, 如何通过统筹安排, 合理安排人力、物力资源, 组织生产过程, 使总的经济效益最好。这类问题常常可以化成或近似地化成线性规划问题。本节举例介绍线性规划的典型问题, 然后给出解决线性规划最基本的图解法。

例 2.1. 某公司可以生产 A、B 两种产品, 已知每生产 1 吨这两种产品的价格, 必须的原料 1、2、3 的多少及原料每日供及量的上限如下表所示, 试问该公司应如何制定生产计划, 使得在满足要求的情况下总的产品价格最多。

	每吨产品 A	每吨产品 B	每日该原料最高需求量
原料 1	5 吨	0 吨	15 吨
原料 2	6 吨	2 吨	24 吨
原料 3	1 吨	1 吨	5 吨
产品价格	2 万元	1 万元	

解: 首先, 建立数学规划模型:

① 确定决策变量:

设 x_1 为计划生产产品 A 的量, x_2 为计划生产产品 B 的量

② 确定目标函数:

目标函数为: $\max_{x_1, x_2} 2x_1 + x_2$

③ 确定约束条件:

$$\begin{cases} 5x_1 \leq 15 \\ 6x_1 + 2x_2 \leq 24 \\ x_1 + x_2 \leq 5 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

④ 建立线性规划模型:

$$\begin{aligned} \max_{x_1, x_2} \quad & 2x_1 + x_2 \\ \text{s.t.} \quad & 5x_1 \leq 15 \\ & 6x_1 + 2x_2 \leq 24 \\ & x_1 + x_2 \leq 5 \\ & x_1 \geq 0, x_2 \geq 0 \end{aligned}$$

其次, 利用图解法求得具体数值解:

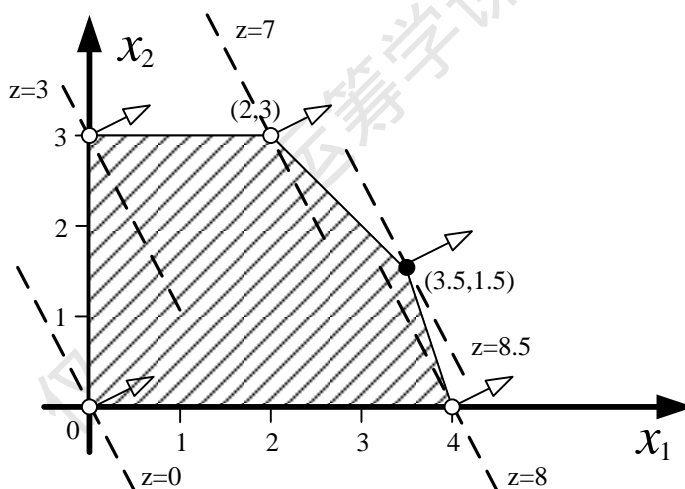


图 2.1 生产问题的图解法示例

对于少于 4 个变量的线性规划模型可以建立图解表示。如图 2.1 所示, 生产问题所有可行解围成的区域为斜纹阴影区域, 也即点 $(0,0)$, $(0,3)$, $(2,3)$, $(3.5,1.5)$, $(4,0)$ 围成的五边形, 这些点同时是该五边形的 5 个极点。虚线是目标函数的若干条等值线, 空心箭头是目标函数的梯度方向。点 $(0,0)$, $(0,3)$, $(2,3)$, $(3.5,1.5)$, $(4,0)$ 处的目标函数分别为 0, 3, 7, 8.5, 8。不难观察得知, 实心点 $(3.5,1.5)$ 是该问题的最优解。

例 2.2. 农民特德有 500 英亩的土地种植小麦、玉米或甜菜。他需要 200 吨小麦和 240 吨玉米来养牛。这些农作物可以在他的土地上种植，也可以从批发商那里购买。任何超过这些数量的产品都可以出售：小麦 170 美元/吨,玉米 150 美元/吨。任何短缺必须以小麦 238 美元/吨,玉米 210 美元/吨的价格从批发商处购买。特德也能种甜菜。少于 6000 吨甜菜可以卖 36 美元/吨。但由于甜菜生产的经济配额，超过 6000 吨的甜菜只能以每吨 10 美元的价格出售。亩产量：小麦 2.5 吨/英亩,玉米 3 吨/英亩,甜菜 20 吨/英亩。种植花费：小麦 150 美元/英亩,玉米 230 美元/英亩,甜菜 260 美元/英亩。请问特德如何种地才能获得最大收益？

解：① 确定决策变量：

小麦、玉米和甜菜种植的亩数分别为： x_W , x_C , x_B

小麦、玉米和甜菜（高价）卖出的吨数： w_W , w_C , w_B

甜菜（低价）卖出的吨数： e_B

从批发商处购买的小麦和玉米吨数： y_W , y_C

② 确定目标函数：

$$\max -150x_W - 230x_C - 260x_B - 238y_W + 170w_W - 210y_C + 150w_C + 36w_B + 10e_B$$

③ 约束条件：

$$x_W + x_C + x_B \leq 500$$

$$2.5x_W + y_W - w_W = 200$$

$$3x_C + y_C - w_C = 240$$

$$20x_B - w_B - e_B = 0$$

$$w_B \leq 6000$$

$$x_W, x_C, x_B, y_W, y_C, e_B, w_W, w_C, w_B \geq 0$$

由于该问题的特殊性，少于 6000 吨时甜菜卖价要小于多于 6000 吨时甜菜卖价，这使得我们可以简单设计如上的决策变量和线性约束关系。请读者思考，如果少于 6000 吨时甜菜卖价要大于多于 6000 吨时甜菜卖价，还是否能设计如此简单的决策变量和线性约束关系。

④ 数值求解：

可利用单纯型法求出最优解为：

	小麦	玉米	甜菜
种植亩数/亩	120	80	300
产量/吨	300	240	6000
销售量/吨	100	0	6000
购买量/吨	0	0	0

最终的利润为：118600 美元。

2.2.3 线性规划问题的定义和表达形式转换

研究者从实际问题中抽象得到了线性规划问题的一般定义,进而给出了不同线性规划问题表达形式的转换方法,以便针对标准形式的线性规划问题进行分析和求解。

定义 2.1. 线性规划问题的标准形式 (Standard Form) 为如下形式

$$\begin{aligned}
 \min_{x_1, x_2, \dots, x_n} \quad & z = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \\
 \text{s.t.} \quad & a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n = b_1 \\
 & a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n = b_2 \\
 & \dots \\
 & a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n = b_m \\
 & x_1 \geq 0, \dots, x_n \geq 0
 \end{aligned} \tag{2.1}$$

其中 $x_i, a_{ij}, b_i, c_i \in \mathbb{R}$, $i = 1, \dots, m$, $j = 1, \dots, n$ 。 x_i 为待定的决策变量。一般的, 我们有 $m < n$, 也即约束条件个数小于决策变量的个数。

我们也可以矩阵向量的形式将此规范形式记为

$$\begin{aligned}
 \min_{\mathbf{x}} \quad & z = \mathbf{c}^T \mathbf{x} \\
 \text{s.t.} \quad & \mathbf{A} \mathbf{x} = \mathbf{b} \\
 & \mathbf{x} \geq \mathbf{0}
 \end{aligned} \tag{2.2}$$

其中向量 $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{c} \in \mathbb{R}^n$ 。这里, 向量 $\mathbf{c} = [c_1 \ \dots \ c_n]^T$ 称为价值向量, $\mathbf{c}^T \mathbf{x}$ 称为目标函数。向量 $\mathbf{b} = [b_1 \ \dots \ b_m]^T$ 称为右侧向量。

需要说明的是, 有些教材将线性规划问题的标准形式设为最大化目标函数 $z = \mathbf{c}^T \mathbf{x}$ 。这在分析和求解时没有太大影响, 仅需要在后面谈到的单纯型算法条件数检验的时候考虑相反的符号设置即可。我们这里设置线性规划问题的标准形式设为最小化目标函数 $z = \mathbf{c}^T \mathbf{x}$, 是为了和后续非线性规划问题的分析保持一致。这也是国际教材中常见的定义方式。

定义 2.2. 线性规划问题的规范形式 (Canonical Form) 为如下形式

$$\begin{aligned}
& \min_{x_1, x_2, \dots, x_n} \quad z = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \\
& \text{s.t.} \quad a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n \geq b_1 \\
& \quad \quad a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n \geq b_2 \\
& \quad \quad \dots \\
& \quad \quad a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n \geq b_m \\
& \quad \quad x_1 \geq 0, \dots, x_n \geq 0
\end{aligned} \tag{2.3}$$

其中 $x_i, a_{ij}, b_i, c_i \in \mathbb{R}$, $i = 1, \dots, m$, $j = 1, \dots, n$ 。 x_i 为待定的决策变量。

我们也可以矩阵向量的形式将此规范形式记为

$$\begin{aligned}
& \min_{\mathbf{x}} \quad z = \mathbf{c}^T \mathbf{x} \\
& \text{s.t.} \quad \mathbf{A} \mathbf{x} \geq \mathbf{b} \\
& \quad \quad \mathbf{x} \geq \mathbf{0}
\end{aligned} \tag{2.4}$$

其中向量 $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{c} \in \mathbb{R}^n$ 。

实际上, 线性规划问题的标准形式和规范形式是等价的。我们也可以将任意的线性规划问题通过变换归为标准形式的线性规划问题。变换规则如下:

1) 如果一个线性规划问题的目标函数是求价值函数 $\mathbf{c}^T \mathbf{x}$ 的最大值, 则我们可以等价的求新的价值函数 $-\mathbf{c}^T \mathbf{x}$ 的最小值; 反之亦然。

2) 对于一个线性规划问题的等式约束条件

$$\sum_{j=1}^n a_{ij} x_j = b_i \tag{2.5}$$

我们可以用如下两个不等式约束条件等价替换

$$\sum_{j=1}^n a_{ij} x_j \geq b_i, \quad -\sum_{j=1}^n a_{ij} x_j \geq -b_i \tag{2.6}$$

3) 对于一个线性规划问题的不等式约束条件

$$\sum_{j=1}^n a_{ij} x_j \geq b_i \tag{2.7}$$

我们可以引入一个新的决策变量 x_{n+1} , 将上述不等式约束条件等价写作

$$\sum_{j=1}^n a_{ij} x_j - x_{n+1} = b_i, \quad x_{n+1} \geq 0 \tag{2.8}$$

4) 对于一个无非负约束的变量 x_j , 我们可以引入两个新的决策变量 $x_j^+ \geq 0$ 和 $x_j^- \geq 0$, 将变量 x_j 替换为 $x_j = x_j^+ - x_j^-$ 。

2.3 线性规划问题的有解性和几何解释

定义 2.3. 满足线性规划问题所有约束条件的解被称为可行解 (Feasible Solution)。线性规划问题所有可行解所组成的集合称为可行区域 (Feasible Region, 或者 Feasible Set)。其中最小化目标函数的可行解被称为最优解。

定理 2.1. 标准形式线性规划问题 (2.3) 或者规范形式线性规划问题 (2.4) 的可行区域构成一个下有界的闭凸集。

证明: 注意到标准形式线性规划问题 (2.3) 的约束条件分为两个部分, 一部分是有限个超平面的交集, 另一部分是闭凸集 \mathbb{R}_+^n 。这两个部分的交集构成最终的可行区域。因为闭凸集 \mathbb{R}_+^n 下有界。而且闭集合的交集保持为闭集。结论显然。

类似的, 注意到标准形式线性规划问题 (2.4) 的约束条件分为两个部分, 一部分是有限个半平面的交集, 另一部分是闭凸集 \mathbb{R}_+^n 。这两个部分的交集构成最终的可行区域。同样, 结论显然。□

而根据等价转换方式 (2.6), 标准线性规划问题也可以等价转换为规范形式线性规划问题, 其对应的可行区域同样可以视为一个下有界的多面体。

一般而言, 求解线性规划问题有三种可能的情况。

第一种情况, 线性规划问题无解 (Infeasible)。例如如下问题

$$\begin{aligned} \min_{x_1, x_2} \quad & z = x_1 + x_2 \\ \text{s.t.} \quad & x_1 + x_2 \geq 2 \\ & -x_1 - x_2 \geq -1 \\ & x_1 \geq 0, x_2 \geq 0 \end{aligned}$$

第二种情况, 线性规划问题存在无界解 (Unbounded), 也即随着决策变量趋向于无穷, 目标函数值也会趋向于无穷。例如如下问题

$$\begin{aligned} \min_{x_1, x_2} \quad & z = x_1 - x_2 \\ \text{s.t.} \quad & x_1 + x_2 \geq 1 \\ & x_1 \geq 0, x_2 \geq 0 \end{aligned}$$

第三种情况, 线性规划问题存在有界的最优解。从 2.2.2 节介绍的图解法可以观察到, 例 2.1 中的线性规划问题最优解在对应的可行区域多面体的一个极点上取得。下面, 我们将这一观察结果推广到一般线性规划问题。

定理 2.2. 如果标准形式线性规划问题 (2.3) 或者规范形式线性规划问题 (2.4) 存在有界的最优解, 则一定在其可行区域多面体的一个极点上取得最优解。

证明：我们首先使用反证法证明，该最优解必然在可行域的一个边界点取得。假设最优解在可行域内部的一个点 \mathbf{x}_0 处取得，则存在 $2\varepsilon > 0$ ，使得以 \mathbf{x}_0 为球心， 2ε 为半径的开球位于 Ω 内部，也即 $B_{2\varepsilon}(\mathbf{x}_0) \subset \Omega$ 。我们可以由此确定一个新点

$$\mathbf{x}_1 = \mathbf{x}_0 - \varepsilon \frac{\mathbf{c}}{\|\mathbf{c}\|_2} \in B_{2\varepsilon}(\mathbf{x}_0) \subset \Omega$$

不难发现，该点对应的目标函数值更小

$$\mathbf{c}^T \mathbf{x}_1 = \mathbf{c}^T \mathbf{x}_0 - \varepsilon \mathbf{c}^T \frac{\mathbf{c}}{\|\mathbf{c}\|_2^2} = \mathbf{c}^T \mathbf{x}_0 - \varepsilon > \mathbf{c}^T \mathbf{x}_0$$

这和我们假设的点 \mathbf{x}_0 处取得最优解的假设矛盾，因此最优解不可能在可行域的一个内部点取得。

假设该问题有界的最优解值为 z 。对于任意可行的 \mathbf{x} ， $\mathbf{c}^T \mathbf{x} \leq z$ ，则根据定理 1.7， $H = \{\mathbf{x} \mid \mathbf{c}^T \mathbf{x} = z, \mathbf{x} \in \mathbb{R}^n\}$ 为可行域下有界多面体的一个支撑超平面。而根据定理 1.8，该支撑超平面上一定存在一个极点，该极点同时也是可行域多面体的一个极点；在这个极点上，线性规划问题取得最优解。□

进一步细分的话，存在有界最优解的线性规划问题还可以分为有唯一最优解和有无穷多个最优解的情况。例如下述问题就存在无穷多个最优解

$$\begin{aligned} \min_{x_1, x_2} \quad & z = x_1 + x_2 \\ \text{s.t.} \quad & x_1 + x_2 \geq 1 \\ & x_1 \geq 0, x_2 \geq 0 \end{aligned}$$

例 2.3. 如果某个线性规划问题在两个不同的点取得最优解，则该问题存在无穷多个最优解。

解：假设某个线性规划问题的可行域为 Ω ，在集合 Ω 内两个不同的点 $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ 取得最优解 $\mathbf{c}^T \mathbf{x}_1 = \mathbf{c}^T \mathbf{x}_2 = z$ 。则任意取实数 $\lambda \in (0, 1)$ ，可以得到

$$\mathbf{c}^T [\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2] = \lambda \mathbf{c}^T \mathbf{x}_1 + (1 - \lambda) \mathbf{c}^T \mathbf{x}_2 = \lambda z + (1 - \lambda) z = z$$

显然，该两点连线段上的任意一点同样取得最优解；故这样点有无穷个。□

从几何角度来看，我们可以搜索线性规划问题的所有极点，其中对应目标函数值最低的极点可能就是该线性规划问题的最优解（相关讨论作为作业题 2.1-2.2 留给大家练习）。不过在具体计算求解线性规划问题时，我们还需要利用代数方法来刻画极点搜索的过程。

2.4 线性规划的代数解释

定义 2.4. 如果标准形式线性规划问题 (2.3) 的一个可行解 \mathbf{x} , 根据定义 1.1 同时是方程组 $\mathbf{Ax} = \mathbf{b}$ 的基本解, 则 \mathbf{x} 是标准形式线性规划问题 (2.3) 的一个基本可行解 (Basic Feasible Solution)。如果一个基本可行解是非退化的, 则称之为一个非退化基本可行解 (Non-Degenerate Basic Feasible Solution), 否则称之为一个退化基本可行解 (Degenerate Basic Feasible Solution)。

不失一般性, 我们可以通过从 \mathbf{A} 的 n 列中选择线性无关的 m 列而形成的可逆方阵 $\mathbf{A}_B \in \mathbb{R}^{m \times m}$, 称之为基矩阵。约束条件写为 $\mathbf{b} = \mathbf{Ax} = [\mathbf{A}_B \quad \mathbf{A}_N] \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix}$ 。

$\mathbf{x}_B \in \mathbb{R}^m$ 是从 \mathbf{x} 中选择的相应变量所组成的列向量。如果 $\mathbf{x}_B \geq \mathbf{0}$, 则 \mathbf{A}_B 被称为可行基矩阵, \mathbf{x}_B 被称为可行基或基变量。 $\mathbf{x}_N \in \mathbb{R}^{n-m}$ 是代表 \mathbf{x} 的其余变量所组成的列向量, 被称为非可行基或非基变量。对应的, 我们可以把目标函数也记为 $z = \mathbf{c}_B^T \mathbf{x}_B + \mathbf{c}_N^T \mathbf{x}_N$, $\mathbf{c}_B \in \mathbb{R}^m$ 是可行基对应的价值向量, $\mathbf{c}_N \in \mathbb{R}^{n-m}$ 是非可行基对应的价值向量。

特殊的, 一个基本可行解 \mathbf{x} 可写为 $\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ 0 \end{bmatrix}$, 此时 $\mathbf{x}_B = \mathbf{A}_B^{-1} \mathbf{b}$, 目标函数值为 $z = \mathbf{c}_B^T \mathbf{A}_B^{-1} \mathbf{b}$ 。

我们可以进一步建立线性规划问题的基本可行解和对应可行域的极点之间的关系, 得到如下几个重要的定理, 将线性规划的几何解释和代数解释联系起来, 为后续的单纯型法分析提供理论基础。

定理 2.3. 标准形式线性规划问题 (2.3) 的一个基本可行解必然对应可行域的一个极点。

证明: 下面使用反证法来证明。

假设 \mathbf{x} 是一个基本可行解。不失一般性, 我们假设 \mathbf{x} 可以写为 $\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ 0 \end{bmatrix}$, $\mathbf{b} = \mathbf{Ax} = [\mathbf{A}_B \quad \mathbf{A}_N] \begin{bmatrix} \mathbf{x}_B \\ 0 \end{bmatrix}$, \mathbf{A}_B 为对应的可逆基矩阵。

若 \mathbf{x} 不是一个极点, 则存在两个不同的可行解 $\mathbf{x}_1 = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{bmatrix}$ 以及实数 $\lambda \in (0, 1)$ 使得 $\mathbf{x} = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2$, 其中 $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{R}^m$ 和 $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^{n-m}$ 分别是 m

维和 $n-m$ 维列向量。

显然, 此时有 $\mathbf{0} = \lambda \mathbf{v}_1 + (1-\lambda) \mathbf{v}_2$, 由于 $\mathbf{x}_1, \mathbf{x}_2$ 是可行解, $\mathbf{v}_1, \mathbf{v}_2 \geq \mathbf{0}$, 所以只能是 $\mathbf{v}_1 = \mathbf{v}_2 = \mathbf{0}$ 。故

$$\mathbf{b} = \mathbf{Ax} = [\mathbf{A}_B \quad \mathbf{A}_N] \begin{bmatrix} \mathbf{x}_B \\ \mathbf{0} \end{bmatrix} = \mathbf{A}_B \mathbf{x}_B = [\mathbf{A}_B \quad \mathbf{A}_N] \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{0} \end{bmatrix} = [\mathbf{A}_B \quad \mathbf{A}_N] \begin{bmatrix} \mathbf{u}_2 \\ \mathbf{0} \end{bmatrix}$$

由于 \mathbf{A}_B 可逆, 有 $\mathbf{x}_B = \mathbf{A}_B^{-1} \mathbf{b} = \mathbf{u}_1 = \mathbf{u}_2$, 则 $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}$, 推出矛盾。故原命题成立。□

定理 2.4. 标准形式线性规划问题 (2.3) 可行域的一个极点必然对应一个基本可行解。

证明: 假设 \mathbf{x} 是一个极点, 有 p 个正元素。为了方便起见, 假设 \mathbf{x} 的前 p 个元素是正的, 令矩阵 \mathbf{A} 的各列向量依次表示为 $\alpha_1, \dots, \alpha_n$, 则有

$$\mathbf{x} = [x_1, \dots, x_p, 0, \dots, 0]^T, \quad \sum_{j=1}^p x_j \alpha_j = \mathbf{b}$$

我们使用反证法证明 $\alpha_1, \dots, \alpha_p$ 是线性独立的。假设这些列向量线性相关, 则存在一组不全为 0 的实数 $\{\lambda_j, j=1, \dots, p\}$ 使得 $\sum_{j=1}^p \lambda_j \alpha_j = \mathbf{0}$ 。进一步取

$$\varepsilon \in \left(0, \min_{\lambda_j \neq 0} \frac{x_j}{|\lambda_j|}\right), \quad \text{我们有 } x_j \pm \varepsilon \lambda_j > 0, \quad j=1, \dots, p。$$

记 $\mathbf{y} = [\lambda_1, \dots, \lambda_p, 0, \dots, 0]^T$, 我们可以得到两个新的可行解 $\mathbf{x}_1 = \mathbf{x} + \varepsilon \mathbf{y} \geq \mathbf{0}$, $\mathbf{x}_2 = \mathbf{x} - \varepsilon \mathbf{y} \geq \mathbf{0}$, $\mathbf{Ax}_1 = \mathbf{Ax} + \varepsilon \mathbf{Ay} = \mathbf{b} + \mathbf{0} = \mathbf{b}$, $\mathbf{Ax}_2 = \mathbf{Ax} - \varepsilon \mathbf{Ay} = \mathbf{b} - \mathbf{0} = \mathbf{b}$ 。注意到 $\mathbf{x} = \frac{1}{2} \mathbf{x}_1 + \frac{1}{2} \mathbf{x}_2$ 。这与我们假设的 \mathbf{x} 是一个极点矛盾。

因此, $\alpha_1, \dots, \alpha_p$ 是线性独立的, \mathbf{x} 是一个基本可行解。□

定理 2.5. [线性规划基本定理] 如果标准形式线性规划问题 (2.3) 存在可行解, 则必然存在基本可行解。

如果标准形式线性规划问题 (2.3) 存在有界最优解, 则必然存在一个基本可行解处取得最优值。

证明: 不妨设矩阵 \mathbf{A} 是行满秩的, 不然的话, 我们可以消除多余的约束再进一步讨论。同时, 我们设增广矩阵 $\mathbf{A}_b = [\mathbf{A} \quad \mathbf{b}]$ 和矩阵 \mathbf{A} 的秩相当, 否则线性规划问题没有可行解。若 $m = n = \text{rank}(\mathbf{A})$, 则该问题有唯一解, 且这个解必然是基本可行解。我们下面几种讨论 $m < n$ 的情况。

如果 $\mathbf{b} = \mathbf{0}$ ，则 $\mathbf{x} = \mathbf{0}$ 显然是一个没有正元素的基本可行解。以下讨论 $\mathbf{b} \neq \mathbf{0}$ 的情况。假设 \mathbf{x} 是一个可行解，有 p 个正元素。为了方便起见，假设 \mathbf{x} 的前 p 个元素是正的，令矩阵 \mathbf{A} 的各列向量依次表示为 $\alpha_1, \dots, \alpha_n$ ，则有

$$\mathbf{x} = [x_1, \dots, x_p, 0, \dots, 0]^T, \quad \sum_{j=1}^p x_j \alpha_j = \mathbf{b}$$

接下来我们分两种情况讨论。

第一种情况，如果 $\{\alpha_j, j = 1, \dots, p\}$ 是线性独立的一组向量，则我们有 $p \leq \text{rank}(\mathbf{A}) = m$ 。如果进一步 $p = m$ ，则根据定义， \mathbf{x} 是一个非退化的基本可行解。如果 $p < m$ ，那么一定还存在 $m - p$ 个向量 $\{\alpha_j, j = p+1, \dots, m\}$ 使得 $\{\alpha_j, j = 1, \dots, m\}$ 构成线性独立的一组向量。由于 \mathbf{x} 的第 $p+1$ 个到第 m 个元素都是 0，因此有

$$\sum_{j=1}^m x_j \alpha_j = \sum_{j=1}^p x_j \alpha_j = \mathbf{b}$$

那么根据定义， \mathbf{x} 是一个退化的基本可行解。

第二种情况，如果 $\{\alpha_j, j = 1, \dots, p\}$ 是线性相关的一组向量，则存在一组不全为 0 的实数 $\{\lambda_j, j = 1, \dots, m\}$ 使得 $\sum_{j=1}^p \lambda_j \alpha_j = \mathbf{0}$ 。显然，这组实数中一定有正数。

对于任意一个 $\lambda_r > 0$ ，则可以有 $\lambda_r = - \sum_{j=1, j \neq r}^p \frac{\lambda_j}{\lambda_r} \alpha_j$ 。将此式代入约束条件，

可得 $\sum_{j=1, j \neq r}^p \left(\mathbf{x}_j - \mathbf{x}_r \frac{\lambda_j}{\lambda_r} \right) \alpha_j = \mathbf{b}$ 。由此，我们可以得到一个新的向量

$$\hat{\mathbf{x}} = \left[x_1 - x_r \frac{\lambda_1}{\lambda_r}, \dots, x_{r-1} - x_r \frac{\lambda_{r-1}}{\lambda_r}, 0, x_{r+1} - x_r \frac{\lambda_{r+1}}{\lambda_r}, \dots, x_p - x_r \frac{\lambda_p}{\lambda_r}, 0, \dots, 0 \right]^T$$

显然，这个 $\hat{\mathbf{x}}$ 中非零元素的个数不超过 $p-1$ 。但任意选择 $\lambda_r > 0$ 进行上述操作不能保证 $\hat{\mathbf{x}}$ 中的每个元素都非负。

因此，我们可以选择 $r = \arg \min_j \left\{ \frac{x_j}{\lambda_j} : \lambda_j > 0, j = 1, \dots, m \right\}$ ，保证 $\hat{\mathbf{x}}$ 中的每个

元素都非负。因为,如此选择的话,若 $\lambda_j > 0$, 上述取最小的操作保证 $\frac{x_j}{\lambda_j} - \frac{x_r}{\lambda_r} \geq 0$,

也即 $x_j - x_r \frac{\lambda_j}{\lambda_r} \geq 0$, $j=1, \dots, p$ 成立; 若 $\lambda_j < 0$, 则 $\frac{x_j}{\lambda_j} - \frac{x_r}{\lambda_r} \leq 0$ 自然成立, 同

样有 $x_j - x_r \frac{\lambda_j}{\lambda_r} \geq 0$, $j=1, \dots, p$ 。

如果 $\hat{\mathbf{x}}$ 的正元素对应的矩阵 \mathbf{A} 的列向量是线性独立的, 那么如前所述, 这个 $\hat{\mathbf{x}}$ 是一个基本可行解。不然的话, 我们可以重复上述过程, 得到一个另一个可行解, 其正元素的个数不超过 $p-2$ 。由于 p 是有限的, 上述过程最多可能执行 $p-1$ 次, 最终我们得到一个仅有一个非零元素的可行解, 其对应的 \mathbf{A} 的列自然是线性独立的。因此, 执行上述过程一定可以找到一个基本可行解。

进一步, 根据定理 2.2, 标准形式线性规划问题 (2.3) 存在有界最优解, 则一定在对应的可行区域多面体的一个极点上取得最优解。根据定理 2.4, 这个极点一定对应一个基本可行解。□

定理 2.6. 多面体 $P = \{\mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b}, \mathbf{x} \in \mathbb{R}_+^n, \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m\}$ 仅有有限个极点和/或极线。

证明: 根据定义 1.9, 每一条极线由 $n-1$ 个线性独立的约束条件确定, 则我们最多有 $C_m^{n-1} = \frac{m!}{(n-1)!(m-n+1)!}$ 条极线。

而根据转换公式 (2.8), 我们可以通过增加 m 个新的决策变量, 将上述多面体的 m 个线性不等式约束条件等价转换为含有 $n+m$ 个变量的 m 个线性等式约束条件。根据定理 2.4, 每个多面体的极点对应这 m 个线性等式的至少一个基本可行解。因此, 多面体的极点最多有 $C_{n+m}^n = \frac{(n+m)!}{n!m!}$ 个。□

标准形式线性规划问题 (2.3) 的一个基本可行解, 一定对应对应的可行区域多面体的一个极点。但是可行区域多面体的一个极点可能对应多个基本可行解; 这多个基本可行解一定对应退化基本可行解。

2.5 线性规划的对偶理论

每一个线性规划问题（Primal Problem）都有一个伴生的线性规划问题，我们称之为对偶问题（Dual Problem）。这两个线性规划问题存在密切的联系，对于分析和求解线性规划问题有着重要的意义。

定义 2.5. 标准形式线性规划问题（2.3）的对偶问题写作

$$\begin{aligned} \max_{\mathbf{u}} \quad & \mathbf{b}^T \mathbf{u} \\ \text{s.t.} \quad & \mathbf{A}^T \mathbf{u} \leq \mathbf{c} \end{aligned} \quad (2.9)$$

其中向量 $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{c} \in \mathbb{R}^n$ 。

规范形式线性规划问题（2.4）的对偶问题写作

$$\begin{aligned} \max_{\mathbf{u}} \quad & \mathbf{b}^T \mathbf{u} \\ \text{s.t.} \quad & \mathbf{A}^T \mathbf{u} \leq \mathbf{c} \\ & \mathbf{u} \geq \mathbf{0} \end{aligned} \quad (2.10)$$

其中向量 $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{c} \in \mathbb{R}^n$ 。

其它形式的线性规划问题可以先化为标准形式线性或者规范形式规划问题，根据上述定义推出对偶问题，再进行化简。或者根据如下的对应关系进行推导

- 1) 每个对偶变量对应原问题的一个约束
- 2) 原问题是等式约束则对偶变量是自由变量（无非负约束）
- 3) 原问题是不等式约束则对偶变量有非负约束
- 4) 原问题变量和对偶问题的约束同样符合上述规律

上述关系也可以总结为对应表。

表 2.1. 线性规划原问题和对偶问题的对应关系

原问题	对偶问题
$\min_{\mathbf{x}} \mathbf{c}^T \mathbf{x}$	$\max_{\mathbf{u}} \mathbf{b}^T \mathbf{u}$
s.t. $a_{i1}x_1 + \dots + a_{in}x_n = b_i \quad i = 1, \dots, p$	s.t. u_i 是自由变量
$a_{i1}x_1 + \dots + a_{in}x_n \leq b_i \quad i = p+1, \dots, m$	$u_i \geq 0$
$x_j \geq 0 \quad j = 1, \dots, q$	$\alpha_j^T \mathbf{u} \geq c_j$
x_j 是自由变量 $j = q+1, \dots, n$	$\alpha_j^T \mathbf{u} = c_j$

其中矩阵 \mathbf{A} 的各列向量依次表示为 $\alpha_1, \dots, \alpha_n$ 。

例 2.4. 根据规范形式线性规划问题（2.3）的对偶问题定义推导标准形式线性规划问题（2.4）的对偶问题。

解：先将标准形式线性规划问题（2.4）改写为规范形式线性规划问题

$$\begin{aligned}
\min_{\mathbf{x}} \quad & z = \mathbf{c}^T \mathbf{x} \\
\text{s.t.} \quad & \mathbf{Ax} \geq \mathbf{b} \\
& -\mathbf{Ax} \geq -\mathbf{b} \\
& \mathbf{x} \geq \mathbf{0}
\end{aligned}$$

则根据定义 2.5，其对应的对偶问题为

$$\begin{aligned}
\max_{\mathbf{u}_1, \mathbf{u}_2} \quad & \mathbf{b}^T \mathbf{u}_1 - \mathbf{b}^T \mathbf{u}_2 \\
\text{s.t.} \quad & \mathbf{A}^T \mathbf{u}_1 - \mathbf{A}^T \mathbf{u}_2 \leq \mathbf{c} \\
& \mathbf{u}_1, \mathbf{u}_2 \geq \mathbf{0}
\end{aligned}$$

令 $\mathbf{u} = \mathbf{u}_1 - \mathbf{u}_2$ ，化简可得线性规划问题 (2.9)。

定理 2.7. 一个线性规划问题的对偶问题的对偶问题是原问题。

证明：我们仅对规范形式线性规划问题 (2.4) 及其对偶问题 (2.10)，其它形式的结论可转化为规范形式之后类似得到。

先将对偶问题 (2.10) 写成规范形式

$$\begin{aligned}
\min_{\mathbf{u}} \quad & -\mathbf{b}^T \mathbf{u} \\
\text{s.t.} \quad & -\mathbf{A}^T \mathbf{u} \geq -\mathbf{c} \\
& \mathbf{u} \geq \mathbf{0}
\end{aligned}$$

根据定义 2.5，其对应的对偶问题为

$$\begin{aligned}
\max_{\mathbf{x}} \quad & -\mathbf{c}^T \mathbf{x} \\
\text{s.t.} \quad & -\mathbf{Ax} \leq -\mathbf{b} \\
& \mathbf{x} \geq \mathbf{0}
\end{aligned}$$

而这显然就是规范形式线性规划问题 (2.4)。□

线性规划原问题和对偶问题之间的联系由下面三个重要的定理给出。

定理 2.8. [弱对偶性定理 (Weak Duality Theorem)] 如果 $\mathbf{x} \in \mathbb{R}^n$ 是规范形式线性规划问题 (2.4) 的一个可行解，是 $\mathbf{u} \in \mathbb{R}^m$ 是其对偶问题 (2.10) 的一个可行解，则

$$\mathbf{c}^T \mathbf{x} \geq \mathbf{b}^T \mathbf{u} \quad (2.11)$$

证明： $\mathbf{x} \in \mathbb{R}^n$ 是规范形式线性规划问题 (2.4) 的一个可行解，则因为其满足约束条件，有 $\mathbf{Ax} \geq \mathbf{b}$ 。故对于任何 $\mathbf{u} \geq \mathbf{0}$ ，我们有

$$\mathbf{u}^T \mathbf{Ax} = \mathbf{u}^T (\mathbf{Ax}) \geq \mathbf{u}^T \mathbf{b} = \mathbf{b}^T \mathbf{u}$$

$\mathbf{u} \in \mathbb{R}^m$ 是其对偶问题 (2.10) 的一个可行解，则因为其满足约束条件，有

$\mathbf{A}^T \mathbf{u} \leq \mathbf{c}$ 。故对于任何 $\mathbf{x} \geq \mathbf{0}$ ，我们有

$$\mathbf{x}^T \mathbf{A} \mathbf{u} = \mathbf{x}^T (\mathbf{A} \mathbf{u}) \leq \mathbf{x}^T \mathbf{c} = \mathbf{c}^T \mathbf{x}$$

结合上述两个不等式有 $\mathbf{c}^T \mathbf{x} \geq \mathbf{x}^T \mathbf{A} \mathbf{u} = \mathbf{u}^T \mathbf{A} \mathbf{x} \geq \mathbf{b}^T \mathbf{u}$ 。□

推论 2.1. 如果规范形式线性规划问题 (2.4) 的一个可行解 $\mathbf{x} \in \mathbb{R}^n$ 和对偶问题 (2.10) 的一个可行解 $\mathbf{u} \in \mathbb{R}^m$ ，满足 $\mathbf{c}^T \mathbf{x} = \mathbf{b}^T \mathbf{u}$ ，则 \mathbf{x} 是原问题的最优解， \mathbf{u} 是对偶问题的最优解。

定理 2.9. [强对偶性定理 (Strong Duality Theorem)] 如果规范形式线性规划问题 (2.4) 存在有界的最优解 $\mathbf{x}^* \in \mathbb{R}^n$ ，当且仅当对偶问题 (2.10) 也存在有界的最优解 $\mathbf{u}^* \in \mathbb{R}^m$ ，且满足 $\mathbf{c}^T \mathbf{x}^* = \mathbf{b}^T \mathbf{u}^*$ 。反之亦然。

证明：充分条件也就是推论 2.1 很容易证明。根据定理 2.8，对于规范形式线性规划问题 (2.4) 的任意一个可行解 $\mathbf{x} \in \mathbb{R}^n$ ，总有 $\mathbf{c}^T \mathbf{x} \geq \mathbf{b}^T \mathbf{u}^* = \mathbf{c}^T \mathbf{x}^*$ ，因此 \mathbf{x}^* 必然是原问题的最优解。类似的，对于对偶问题的任意一个可行解 $\mathbf{u} \in \mathbb{R}^m$ ，总有 $\mathbf{b}^T \mathbf{u} = \mathbf{c}^T \mathbf{x}^* \geq \mathbf{b}^T \mathbf{u}$ ，因此 \mathbf{u}^* 必然是对偶问题的最优解。

接下来用 Farkas 引理和构造法证明必要性。假设原问题存在有界的最优解 $\mathbf{x}^* \in \mathbb{R}^n$ ，我们需要找到对应的 $\mathbf{u}^* \in \mathbb{R}^m$ ，满足对偶问题的最优解。假设 \mathbf{x}^* 使得约束条件 $\mathbf{A} \mathbf{x} \geq \mathbf{b}$ 中取等号那些约束的下标们构成集合 I ，则我们有

$$\mathbf{a}_i^T \mathbf{x}^* = b_i, i \in I, \text{ 以及 } \mathbf{a}_i^T \mathbf{x}^* > b_i, i \notin I$$

\mathbf{x}^* 是最优解，则集合 $\{\mathbf{d} \mid \mathbf{a}_i^T \mathbf{d} \geq 0, \mathbf{c}^T \mathbf{d} < 0, i \in I\}$ 必为空集。否则，总可以找到一个足够小的 $\varepsilon > 0$ 使得 $\mathbf{a}_i^T (\mathbf{x}^* + \varepsilon \mathbf{d}) \geq b_i, \mathbf{c}^T (\mathbf{x}^* + \varepsilon \mathbf{d}) < \mathbf{c}^T \mathbf{x}^*, i = 1, \dots, m$ 。

根据定理 1.9 所述的 Farkas 引理，我们知道集合 $\{\mathbf{d} \mid \mathbf{a}_i^T \mathbf{d} \geq 0, \mathbf{c}^T \mathbf{d} < 0, i \in I\}$ 为空集，则必然存在一组实数 $\lambda_i, i \in I$ 满足

$$\lambda_i \geq 0, \sum_{i \in I} \lambda_i \mathbf{a}_i^T = \mathbf{c},$$

则我们可以根据

$$\mathbf{u}_i^* = \lambda_i, i \in I, \text{ 以及 } \mathbf{u}_i^* = 0, i \notin I$$

来构造对偶问题的可行解 $\mathbf{u}^* \in \mathbb{R}^m$ 。此时成立

$$\mathbf{b}^T \mathbf{u}_i^* = \sum_{i \in I} b_i \mathbf{u}_i^* = \sum_{i \in I} (\mathbf{a}_i^T \mathbf{x}^*) \mathbf{u}_i^* = \mathbf{u}^* \mathbf{A} \mathbf{x}^* = \mathbf{c}^T \mathbf{x}^*$$

则根据弱对偶性定理 2.8， \mathbf{u}^* 是对偶问题的最优解，所以对于原问题，强对偶性得证。类似的可以证明对于对偶问题亦然。□

定理 2.10. [互补松弛性定理 (Complementary Slackness)] 如果规范形式线性规划问题 (2.4) 的一个可行解 $\mathbf{x} \in \mathbb{R}^n$ 和对偶问题 (2.10) 的一个可行解 $\mathbf{u} \in \mathbb{R}^m$ ，

则 \mathbf{x} 是原问题的最优解且 \mathbf{u} 是对偶问题的最优解的充分必要条件是同时满足

$$u_i (\mathbf{a}_i^T \mathbf{x} - b_i) = 0 \text{ 和 } (c_j - \mathbf{u}^T \alpha_j) x_j = 0 \quad (2.12)$$

其中 $\{\alpha_j \in \mathbb{R}^m, j=1, \dots, n\}$ 是矩阵 \mathbf{A} 的各列向量。

证明：我们首先证明充分性。如果 (2.12) 成立，则有 $(\mathbf{u}^T \mathbf{A} - \mathbf{c}^T) \mathbf{x} = 0$ ，也即 $\mathbf{u}^T \mathbf{A} \mathbf{x} = \mathbf{u}^T \mathbf{b} = \mathbf{c}^T \mathbf{x}$ 。则根据强对偶性定理， \mathbf{x} 是原问题的最优解且 \mathbf{u} 是对偶问题的最优解。

我们接下来证明必要性。如果 \mathbf{x} 是原问题的最优解且 \mathbf{u} 是对偶问题的最优解，则根据强对偶性定理，我们得到 $(\mathbf{u}^T \mathbf{A} - \mathbf{c}^T) \mathbf{x} = 0$ 且 $\mathbf{u}^T (\mathbf{A} \mathbf{x} - \mathbf{b}) = 0$ 。考虑到 $\mathbf{x} \geq \mathbf{0}$ 且 $\mathbf{u}^T \mathbf{A} - \mathbf{c}^T \geq \mathbf{0}$ ， $\mathbf{u} \geq \mathbf{0}$ 且 $\mathbf{A} \mathbf{x} - \mathbf{b} \geq \mathbf{0}$ ，则 (2.12) 必然成立。□

根据上述上个定理，我们可以进一步分析线性规划原问题和对偶问题有解性之间的联系。这些联系将在我们的后续分析证明中反复使用。

定理 2.11. 原问题存在有界的最优解，则对偶问题存在有界的最优解。原问题存在无界解，则对偶问题无解。原问题无解，则对偶问题无解或者存在无界解。

对偶问题存在有界的最优解，则原问题存在有界的最优解。对偶问题存在无界解，则原问题无解。对偶问题无解，则原问题无解或者存在无界的解。

证明：如果原始问题是无界的，这意味着 $\mathbf{c}^T \mathbf{x}$ 可以趋向 $-\infty$ ，根据弱对偶性定理 2.8，我们有 $\mathbf{b}^T \mathbf{u} \leq \mathbf{c}^T \mathbf{x}$ ，这是不可能的，这意味着对偶问题一定无解。

如果原始问题存在有界的最优解，根据强对偶定理 2.9，则对偶问题存在有界的最优解。

如果原问题无解，根据弱对偶定理 2.8，对偶问题可能是无界的。我们提出下面的例子来证明对偶问题也可能是无解。考虑原问题

$$\begin{aligned} \min_{x_1, x_2} \quad & \begin{bmatrix} -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ \text{s.t.} \quad & \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \geq \begin{bmatrix} 1 \\ -0.5 \end{bmatrix} \\ & x_1, x_2 \geq 0 \end{aligned}$$

其对偶问题为

$$\begin{aligned}
 & \max_{u_1, u_2} \quad [1 \quad -0.5] \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\
 & \text{s.t.} \quad \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \leq \begin{bmatrix} -2 \\ 1 \end{bmatrix} \\
 & \quad \quad \quad u_1, u_2 \geq 0
 \end{aligned}$$

很显然，这两个线性规划都无解。

对于对偶问题，结论类似可以证明。□

总之，原问题和对偶问题的解的对应关系如下表。

表 2.2. 线性规划原问题和对偶问题的解的对应关系

原始 \ 对偶	存在有界最优解	存在无界解	无解
存在有界最优解	√	×	×
存在无界解	×	×	√
无解	×	√	√

接下来，我们顺势给出线性规划多面体可行域的极点—极线表示法，为第 4 章的算法分析提供理论基础。

定理 2.12. $\mathbf{0}$ 是多面体锥 $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{0}\}$ 的极点，当且仅当 $\text{rank}(\mathbf{A}) = n$ 。

证明：首先，我们使用反证法证明必要性。

若 $\mathbf{0}$ 是该多面体锥的极点。 $\text{rank}(\mathbf{A}) < n$ ，必然存在一个点 $\mathbf{y} \neq \mathbf{0}$ 满足 $\mathbf{Ay} = \mathbf{0}$ 。进一步，点 $-\mathbf{y} \neq \mathbf{0}$ 也满足 $\mathbf{A}(-\mathbf{y}) = \mathbf{0}$ 。所以点 $\mathbf{y} \neq \mathbf{0}$ 和 $-\mathbf{y} \neq \mathbf{0}$ 都在这个多面体锥中，且 $\mathbf{y} + (-\mathbf{y}) = \mathbf{0}$ ，这和 $\mathbf{0}$ 是该多面体锥的极点矛盾。因此，若 $\mathbf{0}$ 是该多面体锥的极点，则 $\text{rank}(\mathbf{A}) = n$ 。

接着，我们使用反证法证明充分性。

若 $\text{rank}(\mathbf{A}) = n$ ， $\mathbf{y} = \mathbf{0}$ 是 $\mathbf{Ay} = \mathbf{0}$ 的唯一解。假设存在 $\mathbf{x}_1, \mathbf{x}_2 \neq \mathbf{y}$ 以及实数 $\lambda \in (0, 1)$ 使得 $\mathbf{0} = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2$ 。则

$$\mathbf{A}[\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2] = \lambda \mathbf{Ax}_1 + (1 - \lambda) \mathbf{Ax}_2 = \mathbf{0}$$

但 $\mathbf{Ax}_1 \geq \mathbf{0}$ ， $\mathbf{Ax}_2 \geq \mathbf{0}$ ， $\lambda \geq 0$ ， $1 - \lambda \geq 0$ 。必有 $\mathbf{Ax}_1 = \mathbf{Ax}_2 = \mathbf{0}$ ，这和 $\mathbf{y} = \mathbf{0}$ 是 $\mathbf{Ay} = \mathbf{0}$ 的唯一解矛盾。因此，若 $\text{rank}(\mathbf{A}) = n$ ， $\mathbf{0}$ 是该多面体锥的极点。□

定理 2.13. 如果线性规划问题 $\left\{ \min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} \geq \mathbf{0} \right\}$ 存在无界解，则锥可行域

$\{\mathbf{x} \mid \mathbf{Ax} \geq \mathbf{0}\}$ 有且必有一条射线 \mathbf{r} 满足 $\mathbf{c}^T \mathbf{r} \leq 0$ 。

证明：根据定义，上述充分性命题显而易见。我们着重分析必要性。

如果原始问题是无界的，则根据强对偶性定理，它的对偶问题一定是无解的。换句话说，对偶问题的可行域集合 $\{\mathbf{u} \in \mathbb{R}^m \mid \mathbf{A}^T \mathbf{u} = \mathbf{c}, \mathbf{u} \geq \mathbf{0}\}$ 一定为空集。根据定理 1.9 所述的 Farkas 引理，存在向量 \mathbf{r} 使得 $\mathbf{Ar} \geq \mathbf{0}$ 和 $\mathbf{c}^T \mathbf{r} \leq 0$ 。显然， \mathbf{r} 是满足要求的锥多面体 $\Omega = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{0}\}$ 的一条射线。□

定理 2.14. 如果线性规划问题 $\left\{ \min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b} \right\}$ 存在无界解，且其定义域中存在至少一个极点，则多面体可行域 $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{b}\}$ 有且必有一条极线 \mathbf{d} 满足 $\mathbf{c}^T \mathbf{d} \leq 0$ 。

证明：根据定义，上述充分性命题显而易见。我们着重分析必要性。

首先，考虑该线性规划问题 $\left\{ \min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b} \right\}$ 的对偶问题

$$\begin{aligned} \max_{\mathbf{u}} \quad & z = \mathbf{b}^T \mathbf{u} \\ \text{s.t.} \quad & \mathbf{A}^T \mathbf{u} = \mathbf{c}, \mathbf{u} \geq \mathbf{0} \end{aligned}$$

如果原问题存在无界解，则对偶问题无解。因此可行域 $\{\mathbf{u} \in \mathbb{R}^m \mid \mathbf{A}^T \mathbf{u} = \mathbf{c}, \mathbf{u} \geq \mathbf{0}\}$ 是空集，因此下述线性规划问题也无解

$$\begin{aligned} \max_{\mathbf{u}} \quad & \mathbf{0}^T \mathbf{u} = 0 \\ \text{s.t.} \quad & \mathbf{A}^T \mathbf{u} = \mathbf{c}, \mathbf{u} \geq \mathbf{0} \end{aligned}$$

而这个新的线性规划问题的对偶问题

$$\begin{aligned} \min_{\mathbf{x}} \quad & z = \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} \geq \mathbf{0} \end{aligned}$$

要么存在无界解，要么无解。但 $\mathbf{x} = \mathbf{0}$ 显然是该问题的一个平凡解，则线性规划问题 $\left\{ \min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} \geq \mathbf{0} \right\}$ 显然存在无界解，根据定理 2.13，锥多面体

$\Omega = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{0}\}$ 存在一条射线 \mathbf{r} 满足 $\mathbf{Ar} \geq \mathbf{0}$ 和 $\mathbf{c}^T \mathbf{r} \leq 0$ 。

接着，考虑如下的线性规划问题

$$\begin{aligned} \min_{\mathbf{x}} \quad & z = \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} \geq \mathbf{0}, \mathbf{c}^T \mathbf{x} = -1 \end{aligned}$$

显然, 约束 $\mathbf{c}^T \mathbf{x} = -1$ 确保该线性规划问题存在有界的最优解。而且, 这个最优解显然可以取到。例如, 上面证明存在一个向量 \mathbf{r} 满足使得 $\mathbf{Ar} \geq \mathbf{0}$ 和 $\mathbf{c}^T \mathbf{r} \leq \mathbf{0}$ 。

令向量 $\mathbf{x}_1 = -\frac{\mathbf{r}}{\mathbf{c}^T \mathbf{r}}$, 我们可以看到 $\mathbf{Ax}_1 = -\frac{\mathbf{Ar}}{\mathbf{c}^T \mathbf{r}} \geq \mathbf{0}$, $\mathbf{c}^T \mathbf{x}_1 = -\frac{\mathbf{c}^T \mathbf{r}}{\mathbf{c}^T \mathbf{r}} = -1$, 则线性规划问题 $\left\{ \min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} \geq \mathbf{0}, \mathbf{c}^T \mathbf{x} = -1 \right\}$ 在点 \mathbf{x}_1 处取得最优解 $z = \mathbf{c}^T \mathbf{x}_1 = -1$ 。

然后, 由于多面体可行域 $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{b}\}$ 存在至少一个极点, 根据作业题 1.5 的结论, 必有 $\text{rank}(\mathbf{A}) = n$ 。根据作业题 1.4 的结论, 若 $\text{rank}(\mathbf{A}) = n$, 多面体 $\Gamma = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{0}, \mathbf{c}^T \mathbf{x} = -1\}$ 必然存在极点。根据定理 2.3, 线性规划问题 $\left\{ \min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} \geq \mathbf{0}, \mathbf{c}^T \mathbf{x} = -1 \right\}$ 存在有界的最优解, 且定义域存在极点, 则最优解必然在某一个极点处取得。不妨设在多面体可行域 $\Gamma = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{0}, \mathbf{c}^T \mathbf{x} = -1\}$ 的极点 \mathbf{d} 处取得最优解。

如果在点 \mathbf{d} 处存在 n 个线性独立的约束条件取等号 $\mathbf{a}_i^T \mathbf{d} = 0$, $i = 1, \dots, n$, 则点 \mathbf{d} 也会是锥多面体 $\Omega = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{0}\}$ 的极点。 $\text{rank}(\mathbf{A}) = n$ 。并且 $\mathbf{d} \neq \mathbf{0}$, 因为 $\mathbf{c}^T \mathbf{d} = -1$ 。而这会与定理 2.12 矛盾。

如果在点 \mathbf{d} 处存在 $n-1$ 个线性独立的约束条件取等号 $\mathbf{a}_i^T \mathbf{d} = 0$, $i = 1, \dots, n-1$, 则根据定义 1.12, 点 \mathbf{d} 就是我们寻找的满足原命题要求的极线。

如果在点 \mathbf{d} 处存在少于 $n-1$ 个线性独立的约束条件取等号, 则我们可以简单的使用反证法证明点 \mathbf{d} 不可能是多面体 $\Gamma = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{0}, \mathbf{c}^T \mathbf{x} = -1\}$ 的极点。
□

定理 2.15. [弱化的 Minkowski 表示定理] 假设 $\Omega = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{b}\}$ 表示的多面体非空, 且含有至少一个极点。如果多面体 Ω 的所有极点构成集合 $\{\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n\}$, Ω 的所有极线构成集合 $\{\mathbf{d}_1, \dots, \mathbf{d}_r \in \mathbb{R}^n\}$ 。定义集合

$$\Gamma = \left\{ \sum_{i=1}^k \lambda_i \mathbf{x}_i + \sum_{j=1}^r \theta_j \mathbf{d}_j \mid \lambda_i \geq 0, \theta_j \geq 0, \sum_{i=1}^k \lambda_i = 1 \right\} \quad (2.12)$$

则 $\Gamma = \Omega$ 。

证明：我们首先证明 $\Gamma \subseteq \Omega$ 。对于任意一个 $\mathbf{x} = \sum_{i=1}^k \lambda_i \mathbf{x}_i + \sum_{j=1}^r \theta_j \mathbf{d}_j \in \Gamma$ ，有

$$\mathbf{A} \sum_{i=1}^k \lambda_i \mathbf{x}_i = \sum_{i=1}^k \lambda_i \mathbf{A} \mathbf{x}_i \geq \sum_{i=1}^k \lambda_i \mathbf{b} = \mathbf{b}, \quad \mathbf{A} \sum_{j=1}^r \theta_j \mathbf{d}_j = \sum_{j=1}^r \theta_j \mathbf{A} \mathbf{d}_j \geq \sum_{j=1}^r \theta_j \mathbf{0} = \mathbf{0}$$

因此 $\mathbf{A} \mathbf{x} \geq \mathbf{b}$ ，也即有 $\Gamma \subseteq \Omega$ 。

接着，我们使用反证法证明 $\Omega \subseteq \Gamma$ 。假设存在一个点 \mathbf{y} 使得 $\mathbf{y} \in \Omega$ 同时 $\mathbf{y} \notin \Gamma$ 。注意到多面体是闭集，则根据定理 1.6 所述的凸集分隔定理，必有一个超平面分隔点 \mathbf{y} 和集合 Γ 。

考虑如下的线性规划问题

$$\begin{aligned} \max_{\lambda, \theta} \quad & \mathbf{0}^T \lambda + \mathbf{0}^T \theta \\ \text{s.t.} \quad & \mathbf{y} = \sum_{i=1}^k \lambda_i \mathbf{x}_i + \sum_{j=1}^r \theta_j \mathbf{d}_j \\ & \lambda^T \mathbf{1} = 1 \\ & \lambda \geq 0, \theta \geq 0 \end{aligned} \quad (2.13)$$

其中 $\lambda = [\lambda_1, \dots, \lambda_k] \in \mathbb{R}^k$ ， $\theta = [\theta_1, \dots, \theta_r] \in \mathbb{R}^r$ 。

$\mathbf{y} \notin \Gamma$ ，则上述线性规划问题无解，则下述对偶问题无解或者存在无解的解。

$$\begin{aligned} \min_{\mathbf{p}, q} \quad & \mathbf{p}^T \mathbf{y} + q \\ \text{s.t.} \quad & \mathbf{p}^T \mathbf{x}_i + q \geq 0, i = 1, \dots, k \\ & \mathbf{p}^T \mathbf{d}_j \geq 0, j = 1, \dots, r \end{aligned} \quad (2.14)$$

但显然 $\mathbf{p} = \mathbf{0}$ ， $q = 0$ 是上述对偶问题的平凡解，则该对偶问题存在无解的解。则意味着存在一个解 \mathbf{p}_1 ， q_1 使得 $\mathbf{p}_1^T \mathbf{y} + q_1 < 0$ 。同时观察到 $\mathbf{p}_1^T \mathbf{x}_i + q_1 \geq 0, i = 1, \dots, k$ 。则有 $\mathbf{p}_1^T \mathbf{y} \leq \mathbf{p}_1^T \mathbf{x}_i, i = 1, \dots, k$ 。

考虑如下一个新的线性规划问题

$$\begin{aligned} \min_{\mathbf{x}} \quad & z = \mathbf{p}_1^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A} \mathbf{x} \geq \mathbf{b} \end{aligned}$$

由于我们假设多面体 Ω 非空，且含有至少一个极点；则该线性规划问题是有解的。

如果这个线性规划问题存在有界的最优解，则根据定理 2.5 表述的线性规划基本定理，该问题一定在一个极点处取得最优解，但 $\mathbf{y} \in \Omega$ ， \mathbf{y} 显然不是极点，则应该存在 $\mathbf{p}_1^T \mathbf{y} > \mathbf{p}_1^T \mathbf{x}_i, i = 1, \dots, k$ ，这导致矛盾。

如果这个线性规划问题存在无界的解；则根据定理 2.14，存在一条极线 \mathbf{d}_j 使得 $\mathbf{p}_1^T \mathbf{d}_j < 0$ ，这和 \mathbf{p}_1 ， q_1 是上述对偶问题 (2.14) 的解，满足

$\mathbf{p}^T \mathbf{d}_j \geq 0, j=1, \dots, r$ 相矛盾。因此, 只有可能 $\Omega \subseteq \Gamma$ 。

因为 $\Gamma \subseteq \Omega$ 和 $\Omega \subseteq \Gamma$ 同时成立, 则 $\Gamma = \Omega$ 。□

定理 2.16. 如果集合 Γ 定义为 (2.12) 形式, 则集合 Γ 为多面体。

证明: 如果存在点 \mathbf{y} 满足 $\mathbf{y} \in \Gamma$, 则必然线性规划问题 (2.13) 有解。而线性问题 (2.13) 有解, 则最优解的值必为 $\mathbf{0}$ 。根据强对偶性定理, 对偶问题 (2.14) 也有存在有界的最优解, 值为 $\mathbf{0}$ 。

我们引入 $\mathbf{p}^+, \mathbf{p}^- \geq \mathbf{0}$ 和 $q^+, q^- \geq 0$, 将变量 \mathbf{p} 和 q 替换为 $\mathbf{p} = \mathbf{p}^+ - \mathbf{p}^-$ 和 $q = q^+ - q^-$, 将对偶问题 (2.14) 改写成标准形式的线性规划问题, 则任意可行解点 \mathbf{y} 对应的目标函数值要大于或者等于最优解的值 $\mathbf{0}$, 也即

$$\mathbf{p}^T \mathbf{y} + q = (\mathbf{p}^+ - \mathbf{p}^-)^T \mathbf{y} + q^+ - q^- \geq 0 \quad (2.15)$$

根据定理 2.14, 对于改写成标准形式的对偶问题 (2.14), 其可行域的所有极线都必须满足不等式 (2.15), 而根据定理 2.6, 极线的个数是有限的, 因此任意可行解 \mathbf{y} 只需要满足有限个形如 (2.15) 的不等式约束条件; 根据多面体的定义, 集合 Γ 为多面体。□

2.6 线性规划问题的单纯型法

2.6.1 单纯型法

秉持着“如果线性规划问题的最优解存在, 则一定可以在其可行区域的定点中找到”的基本理念, George Dantzig 于 1947 年提出了一个寻找最优解的思路: 先找出可行域的一个极点, 据一定规则判断其是否最优; 若否, 则转换到与之相邻的另一极点, 并使目标函数值更优; 如此往复, 直到找到某最优解为止。本节将会讨论这种“在极点中搜索最优解的算法”——单纯型法。

我们首先给出标准形式线性规划问题 (2.3) 的单纯型法的一般步骤, 然后逐一解释为什么这些步骤不可或缺, 并证明单纯型法的收敛性和复杂度。

定义 2.6. 对于标准形式线性规划问题 (2.3), 给定的可行基矩阵 \mathbf{A}_B , 非

可行基矩阵 \mathbf{A}_B , 我们定义检验数 (Condition) 为向量 $\boldsymbol{\sigma}_N = \mathbf{c}_N^T - \mathbf{c}_B^T \mathbf{A}_B^{-1} \mathbf{A}_N \in \mathbb{R}^{n-m}$ 。

值得注意的是, 有部分国内教材的检验数定义与国际大多数教材恰好反号, 则对应的判断是否达到的准备也恰好反号。本书还是按国际主流方式来定义检验

数。

算法 2.1. 单纯型法

Step 1. [确定初始可行解] 如果观察原问题即能找到明显的初始可行基矩阵 \mathbf{A}_B 和对应的可行基 $\mathbf{x}_B = \mathbf{A}_B^{-1}\mathbf{b} \geq \mathbf{0}$ ，则直接进入 Step 2。否则使用大 M 法或者两阶段法找到初始可行基矩阵 \mathbf{A}_B 和对应的可行基 $\mathbf{x}_B = \mathbf{A}_B^{-1}\mathbf{b} \geq \mathbf{0}$ ，再进入 Step 2。

Step 2. [判断最优性] 注意到标准形式线性规划问题 (2.3) 的约束条件和目标函数可以写成

$$\begin{cases} \mathbf{A}_B \mathbf{x}_B + \mathbf{A}_N \mathbf{x}_N = \mathbf{b} \\ \mathbf{c}_B^T \mathbf{x}_B + \mathbf{c}_N^T \mathbf{x}_N = z \end{cases} \quad (2.16)$$

或等价的

$$\begin{cases} \mathbf{x}_B + \mathbf{A}_B^{-1} \mathbf{A}_N \mathbf{x}_N = \mathbf{A}_B^{-1} \mathbf{b} \\ \mathbf{c}_B^T (\mathbf{A}_B^{-1} \mathbf{b} - \mathbf{A}_B^{-1} \mathbf{A}_N \mathbf{x}_N) + \mathbf{c}_N^T \mathbf{x}_N = z \end{cases}$$

移项化简为

$$\begin{cases} \mathbf{x}_B + \mathbf{A}_B^{-1} \mathbf{A}_N \mathbf{x}_N = \mathbf{A}_B^{-1} \mathbf{b} \\ (\mathbf{c}_N^T - \mathbf{c}_B^T \mathbf{A}_B^{-1} \mathbf{A}_N) \mathbf{x}_N = z - \mathbf{c}_B^T \mathbf{A}_B^{-1} \mathbf{b} \end{cases} \quad (2.17)$$

我们根据当前的可行基矩阵 \mathbf{A}_B ，构建如下的单纯型表 (Tableau)

	\mathbf{x}_B	\mathbf{x}_N	RHS
\mathbf{x}_B	\mathbf{I}	$\mathbf{A}_B^{-1} \mathbf{A}_N$	$\mathbf{A}_B^{-1} \mathbf{b}$
	$\mathbf{0}$	$\boldsymbol{\sigma}^T$	$z - \mathbf{c}_B^T \mathbf{A}_B^{-1} \mathbf{b}$

令 $\bar{\mathbf{A}} = \mathbf{A}_B^{-1} \mathbf{A}_N = [\bar{\boldsymbol{\alpha}}_{m+1}, \dots, \bar{\boldsymbol{\alpha}}_n] = [\bar{a}_{ij}] \in R^{m \times (n-m)}$ ， $\bar{\mathbf{b}} = \mathbf{A}_B^{-1} \mathbf{b} = [\bar{b}_i] \in R_+^m$ 来简

化表述上表中的对应矩阵分块，其中 $\bar{\boldsymbol{\alpha}}_j$ ， $j = 1, \dots, n-m$ ，表示矩阵 $\bar{\mathbf{A}}$ 的各个列

向量。 \bar{a}_{ij} ， $i = 1, \dots, m$ ， $j = m+1, \dots, n$ 表示矩阵 $\bar{\mathbf{A}}$ 的各个元素。 \bar{b}_i 表示向量 $\bar{\mathbf{b}}$

的各个元素。 $\boldsymbol{\sigma}_N = [\sigma_{m+1}, \dots, \sigma_n]^T = [c_{m+1} - \mathbf{c}_B^T \mathbf{A}_B^{-1} \bar{\boldsymbol{\alpha}}_{m+1}, \dots, c_n - \mathbf{c}_B^T \mathbf{A}_B^{-1} \bar{\boldsymbol{\alpha}}_n]^T$ 。我们

可以将上面的单纯型表的各列详细表述出来

\mathbf{x}_B	x_1	x_2	...	x_{m+1}	...	x_{m+k}	...	RHS
x_1	1	0	...	$\bar{a}_{1,m+1}$...	$\bar{a}_{1,m+k}$...	\bar{b}_1
x_2	0	1	...	$\bar{a}_{2,m+1}$...	$\bar{a}_{2,m+k}$...	\bar{b}_2
...
	0	0	0	σ_{m+1}	...	σ_{m+k}	...	$z - \mathbf{c}_B^T \mathbf{A}_B^{-1} \mathbf{b}$

如果 $\sigma \geq 0$ ，则当前解为最优解，停止。

如果检验数向量 σ 的第 k 个元素 $\sigma_k < 0$ ，而对应的列向量 $\bar{\alpha}_k \leq 0$ ，则原问题无界，停止。

如果检验数向量 σ 的第 k 个元素 $\sigma_k < 0$ ，而对应的列向量 $\bar{\alpha}_k$ 中至少有一个元素 \bar{a}_{jk} 为正，注意到我们可以将对应的非基变量 x_{m+k} 从 0 变为正来减少目标函数值，则进入 Step 3，通过翻转找到一个可行的基 $\hat{\mathbf{x}}_B$ ，使得 $\mathbf{c}_B^T \hat{\mathbf{x}}_B \leq \mathbf{c}_B^T \mathbf{x}_B$ 。

Step 3. [翻转 (Pivots)] 首先确定进基变量，按照 Bland 法则 (Bland's Rule) [13]，则选取检验数向量 σ 的所有非负元素中下标最小的元素对应的决策变量作为进基变量，由非基变量变为基变量。

$$k = \min \{j \mid \sigma_j < 0, j = m+1, \dots, n\} \quad (2.18)$$

此时 \mathbf{x}_N 中的第 k 个元素 x_{m+k} 即为选定的进基变量。

接着计算求解下述问题，确定可能的出基变量的下标所组成的集合

$$L = \left\{ \operatorname{arccmin}_i \left\{ \frac{\bar{b}_i}{\bar{a}_{ik}} \mid \bar{a}_{ik} > 0, i = 1, \dots, m \right\} \right\} \quad (2.19)$$

如果有多个 \bar{b}_i / \bar{a}_{ik} 同时取得最小值，则同样按照 Bland 法则，选取这其中下标最小的 i 作为出基变量下标 l

$$l = \min \{i \mid i \in L, i = 1, \dots, m\} \quad (2.20)$$

则出基变量为此时可行基 \mathbf{x}_B 中的第 l 个元素 x_l 。

根据原有可行基 \mathbf{x}_B ，进基变量 x_{m+k} 和出基变量 x_l 构造新的可行基矩阵 $\hat{\mathbf{A}}_B$ ，进而求解除新的可行基本解 $\hat{\mathbf{x}}_B = \hat{\mathbf{A}}_B^{-1} \mathbf{b} \geq 0$ 。这一步可以直接在表中，由进基变量所在列和出基变量所在行决定的约束矩阵系数通过初等变换得到，参见例 2.5 的求解过程。最后回到 Step 2 进行下一轮迭代。

如果我们使用单纯型法求解线性规划问题 $\left\{ \max_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \mid \mathbf{A} \mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0 \right\}$ ，则有两种求解选择。第一种选择是转化为标准形式线性规划 (2.3)，也即求解 $\left\{ \min_{\mathbf{x}} -\mathbf{c}^T \mathbf{x} \mid \mathbf{A} \mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0 \right\}$ 。第二种选择是在 Step 2 根据如下规则来直接确定问题的翻转方向：

如果 $\sigma \leq 0$ ，则当前解为最优解，停止。

如果检验数向量 σ 的第 k 个元素 $\sigma_k > 0$ ，而对应的列向量 $\bar{\alpha}_k \leq 0$ ，则原问题无界，停止。

如果检验数向量 σ 的第 k 个元素 $\sigma_k > 0$ ，而对应的列向量 $\bar{\alpha}_k$ 中至少有一个元素 $\bar{\alpha}_{jk}$ 为正，注意到我们可以将对应的非基变量 x_{m+k} 从 0 变为正来增加目标函数值，则进入 Step 3，通过翻转找到一个新的可行基 $\hat{\mathbf{x}}_B$ ，使得

$$\mathbf{c}_B^T \hat{\mathbf{x}}_B \geq \mathbf{c}_B^T \mathbf{x}_B。$$

接着，我们依次证明几个定理，解释上述单纯型算法几个关键点的必要性和可行性。

定理 2.17. 如果单纯型法迭代到某一步，满足 $\sigma \geq 0$ ，则当前解为最优解。

证明：若满足 $\sigma \geq 0$ 的解是 \mathbf{x} ，对任意可行解 \mathbf{y} 有 $\mathbf{A}\mathbf{x} - \mathbf{A}\mathbf{y} = \mathbf{b} - \mathbf{b} = \mathbf{0}$ 。

我们定义 $\mathbf{d} = \mathbf{y} - \mathbf{x} = \begin{bmatrix} \mathbf{d}_B \\ \mathbf{d}_N \end{bmatrix}$ ，则上式改写为 $\mathbf{A}\mathbf{d} = \mathbf{0}$ 。或者进一步考虑对应

的可行基矩阵 \mathbf{A}_B ，写为 $\mathbf{A}_B \mathbf{d}_B + \sum_{i \in N} \alpha_i d_i = \mathbf{0}$ ，其中 N 表示所有非基变量的下标

构成的集合， $\{\alpha_1, \dots, \alpha_n\}$ 表示矩阵 \mathbf{A} 的各个列向量。

因矩阵 \mathbf{A}_B 可逆，有 $\mathbf{d}_B = -\sum_{i \in N} \mathbf{A}_B^{-1} \alpha_i d_i$ 和 $\mathbf{c}^T \mathbf{d} = \mathbf{c}_B^T \mathbf{d}_B + \mathbf{c}_N^T \mathbf{d}_N = \sum_{i \in N} \sigma_i d_i$ 。

因为对于非基变量 $x_i = 0$ ，对于可行解 \mathbf{y} 有 $y_i \geq 0$ 。因此 $d_i \geq 0$ 。同时根据假

设， $\sigma \geq 0$ 。所以 $\mathbf{c}^T \mathbf{d} \geq 0$ 。故而对于任意可行解 \mathbf{y} ， $\mathbf{c}^T \mathbf{y} \geq \mathbf{c}^T \mathbf{x}$ 。□

与之相关，我们还可以得到其它一些有用的推论，这里作为作业题 2.3-2.4 留给大家练习。

定理 2.18. 如果单纯型法迭代到某一步，满足检验数向量 σ 的第 k 个元素 $\sigma_k < 0$ ，且对应的列向量 $\bar{\alpha}_k \leq 0$ ，则原问题无界，且单纯型法停止时，我们可以直接构造出标准形式线性规划问题 (2.3) 可行域的一条极线。

证明：停止时可行基矩阵为 \mathbf{A}_B ，矩阵 $\bar{\mathbf{A}} = \mathbf{A}_B^{-1} \mathbf{A}_N = [\bar{\alpha}_1, \dots, \bar{\alpha}_{n-m}] \in R^{m \times (n-m)}$ 的第 k 列 $\bar{\alpha}_k \leq 0$ ，则我们可以构造如下的方向 $\mathbf{d} = [\mathbf{d}_B, \mathbf{d}_N] \in R^n$ ，其中 $\mathbf{d}_B = -\bar{\alpha}_k \in R^m$ ，而除了 $d_{m+k} = 1$ ， $\mathbf{d}_N \in R^{n-m}$ 的其它各元素都等于 0。

我们把标准形式线性规划问题 (2.3) 的可行域 $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ 等价变换为 $\Omega = \{\mathbf{x} \in \mathbb{R}^n \mid [\mathbf{A} \quad -\mathbf{A} \quad \mathbf{I}]\mathbf{x} \geq [\mathbf{b} \quad -\mathbf{b} \quad \mathbf{0}]^T\}$ 。则该可行域对应的回归锥为 $\Gamma = \{\mathbf{x} \in \mathbb{R}^n \mid [\mathbf{A} \quad -\mathbf{A} \quad \mathbf{I}]\mathbf{x} \geq \mathbf{0}\}$ 。

不难验证 $\mathbf{d} \geq \mathbf{0}$, $\mathbf{Ad} = \mathbf{0}$, 因此 \mathbf{d} 属于标准形式线性规划问题 (2.3) 可行域的回归锥 Γ 。同时, 根据 $\sigma_k < 0$, 不难验证 $\mathbf{c}^T \mathbf{d} \leq 0$ 。

最后, 因为 \mathbf{A}_B 可逆, 我们一定可以挑出 $\mathbf{Ad} = \mathbf{0}$ 中对应可行基矩阵 \mathbf{A}_B 的 m 个线性独立的等式约束。再加上 $d_{m+j} = 0, j = 1, \dots, n-m, j \neq k$, 这 $n-m+1$ 个线性独立的等式约束。我们确认 \mathbf{d} 恰好满足 $n-1$ 个线性独立的等式约束。□

定理 2.19. 算法 2.1 因为采用 Bland 法则, 不会在退化的基本可行解处产生迭代循环 (Cycle), 一定会在有限步内收敛。如果放弃 Bland 法则, 则可能出现循环 [11]-[14]。

证明: 显然, 下标越大的决策变量元素越难被选中进基。对于决策变量 $\mathbf{x} = [x_1, \dots, x_n]^T \in \mathbb{R}^n$, x_n 最多可能被选中 1 次进基。我们容易递归证明, x_{n-i} 只可能被选中 2^i 次进基, $i = 1, \dots, n$ 。因此, 采用 Bland 法则保证会在有限步内收敛。而翻转次数最多的线性规划问题例子可以参见定理 2.21。

如果放弃 Bland 法则, Hoffman, Beale 等研究者构造了一些有趣的例子来证实会发生循环。以下的例子来自 [12], 考虑如下的线性规划问题

$$\begin{aligned} \max_{x_1, x_2, \dots, x_7} \quad & \frac{3}{4}x_4 - 20x_5 + \frac{1}{2}x_6 - 6x_7 \\ \text{s.t.} \quad & x_1 + \frac{1}{4}x_4 - 8x_5 - x_6 + 9x_7 = 0 \\ & x_2 + \frac{1}{2}x_4 - 12x_5 - \frac{1}{2}x_6 + 3x_7 = 0 \\ & x_3 + x_6 = 1 \\ & x_i \geq 0, i = 1, 2, \dots, 7 \end{aligned}$$

假设我们选择的初始基变量是 $\{x_1, x_2, x_3\}$, 则得到如下的单纯型表

\mathbf{x}_B	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_1	1	0	0	1/4	-8	-1	9	0
x_2	0	1	0	1/2	-12	-1/2	3	0
x_3	0	0	1	0	0	1	0	1
	0	0	0	3/4	-20	1/2	-6	$z-0$

第一次选择 x_4 作为进基变量, x_1 作为出基变量, 进行翻转, 基变量变为 $\{x_4, x_2, x_3\}$, 得到如下的单纯型表

\mathbf{x}_B	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_4	4	0	0	1	-32	-4	36	0
x_2	-2	1	0	0	4	3/2	-15	0
x_3	0	0	1	0	0	1	0	1
	-3	0	0	0	4	7/2	-33	$z-0$

第二次选择 x_5 作为进基变量, x_2 作为出基变量, 进行翻转, 基变量变为 $\{x_4, x_5, x_3\}$, 得到如下的单纯型表

\mathbf{x}_B	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_4	-12	8	0	1	0	8	-84	0
x_5	-1/2	1/4	0	0	1	3/8	-15/4	0
x_3	0	0	1	0	0	1	0	1
	-1	-1	0	0	0	2	-18	$z-0$

第三次选择 x_6 作为进基变量, x_4 作为出基变量, 进行翻转, 基变量变为 $\{x_6, x_5, x_3\}$, 得到如下的单纯型表

\mathbf{x}_B	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_6	-2/3	1	0	1/8	0	1	-21/2	0
x_5	1/16	-1/8	0	-3/64	1	0	3/16	0
x_3	3/2	-1	1	-1/8	0	0	21/2	1
	2	-3	0	-1/4	0	0	3	$z-0$

第四次选择 x_7 作为进基变量, x_5 作为出基变量, 进行翻转, 基变量变为 $\{x_6, x_7, x_3\}$, 得到如下的单纯型表

\mathbf{x}_B	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_6	2	-6	0	-5/2	56	1	0	0
x_7	1/3	-2/3	0	-1/4	16/3	0	1	0
x_3	-2	6	1	5/2	-56	0	0	1
	1	-1	0	1/2	-16	0	0	$z-0$

第五次选择 x_1 作为进基变量, x_6 作为出基变量, 进行翻转, 基变量变为 $\{x_1, x_7, x_3\}$, 得到如下的单纯型表

\mathbf{x}_B	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_1	1	-3	0	-5/4	28	1/2	0	0
x_7	0	1/3	0	1/6	-4	-1/6	1	0
x_3	0	0	1	0	0	1	0	1
	0	2	0	7/4	-44	-1/2	0	$z-0$

第六次选择 x_2 作为进基变量, x_7 作为出基变量, 进行翻转, 基变量变为 $\{x_1, x_2, x_3\}$, 得到如下的单纯型表

\mathbf{x}_B	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_1	1	0	0	1/4	-8	-1	9	0
x_2	0	1	0	1/2	-12	-1/2	3	0
x_3	0	0	1	0	0	1	0	1
	0	0	0	3/4	-20	1/2	-6	$z-0$

则此时循环到了初始状态。实际上在整个迭代过程中, 虽然可行基矩阵不断改变, 但对应的基本可行解始终是 $\mathbf{x} = [0, 0, 1, 0, 0, 0, 0]^T$, 没有变化。

而如果采用 Bland 法则, 则在第四次选择 x_1 作为进基变量, x_5 作为出基变量, 进行翻转, 基变量变为 $\{x_6, x_1, x_3\}$, 得到如下的单纯型表

这样就跳出了循环。

□

本书仅介绍了使用 Bland 法则避免循环, 还有其它的避免循环的方法, 参见文献 [6]。

定理 2.20. 如果不按照 (2.19) 定义的方式选择出基变量, 不能保证在迭代过程中 $\mathbf{x}_B \geq \mathbf{0}$ 。

该定理很容易构造数值反例证明, 留做作业题 2.6, 给大家作为练习。

定理 2.21. 在最差情况下, 线性规划问题的单纯型法非多项式时间复杂度。

证明: Klee 和 Minty 构造了如下的特殊线性规划问题来证明这一点 [7]

$$\begin{aligned}
 \min \quad & -x_n \\
 \text{s.t.} \quad & x_1 \geq 0 \\
 & x_1 \leq 1 \\
 & x_i \geq \varepsilon x_{i-1}, i = 2, \dots, n \\
 & x_i \leq 1 - \varepsilon x_{i-1}, i = 2, \dots, n
 \end{aligned}$$

其中 $\varepsilon \in (0, 0.5)$ 。

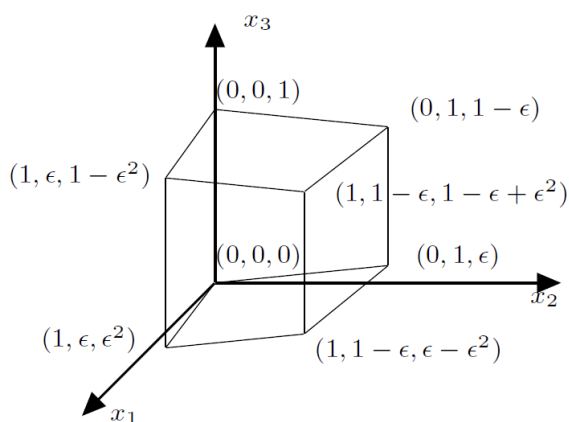


图 2.6 $n = 3$ 时, 发生变形的 3 维立方体各极点坐标示意图。

显然该问题的有界多面体可行域为一个发生变形的 n 维立方体, 参见图 2.6。不难得知该可行域有 2^n 个极点。我们需要证明, 最差的情况是初始解选择原点, 需要经过 $2^n - 1$ 次翻转, 遍历所有的极点, 才能找到最优解

$\mathbf{x} = (0, \dots, 0, 1)$ 。

首先, 我们对于 n 维立方体的极点进行 $\{0, 1\}^n$ 上的二进制编码。例如, 图 2.6 所示的立方体各极点编码为

极点	编码
$(0,0,0)$	$(0,0,0)$
$(0,\varepsilon,\varepsilon^2)$	$(1,0,0)$
$(0,1-\varepsilon,\varepsilon-\varepsilon^2)$	$(1,1,0)$
$(0,1,\varepsilon)$	$(0,1,0)$
$(0,1,1-\varepsilon)$	$(0,1,1)$
$(0,1-\varepsilon,1-\varepsilon+\varepsilon^2)$	$(1,1,1)$
$(0,\varepsilon,1-\varepsilon^2)$	$(1,0,1)$
$(0,0,1)$	$(0,0,1)$

再者，我们定义 $\{0,1\}^n$ 上的一条特殊 Hamiltonian 路径，并证明沿着该路径，目标函数不断下降。实际上这条特殊 Hamiltonian 路径可能看作一类特殊的递归定义的 Gray 码。Gray 码是一种要求相邻的两个数字仅在一个数位上相差 1 的数字编码。

$$\begin{array}{rcl}
 & & 0\dots 00 \\
 & & 1\dots 00 \\
 \Theta_{n-1} & & \dots \\
 & & 1\dots 10 \\
 \Theta_n = & & 0\dots 11 \\
 & & \dots \\
 \text{反向}\Theta_{n-1} & & \dots \\
 & & 1\dots 01 \\
 & & 0\dots 01
 \end{array}$$

接下来，我们使用数学归纳法证明对于 n 维立方体，我们沿着类似的 Hamiltonian 路径遍历所有极点时，目标函数一直是下降的。

当 $n = 2$ 时，考虑对应的线性规划问题

$$\begin{array}{ll}
 \min & -x_2 \\
 \text{s.t.} & x_1 \geq 0, x_1 \leq 1 \\
 & x_2 \geq \varepsilon x_1, x_2 \leq 1 - \varepsilon x_1
 \end{array}$$

显然，若初始解选择原点 $\mathbf{x} = (0,0)$ ，需要经过 3 次翻转 $\Theta_2 = 00 \rightarrow 10 \rightarrow 11 \rightarrow 01$ ，遍历所有的极点，才能找到最优解 $\mathbf{x} = (0,0,1)$ 。

当 $n = 3$ 时，对于图 2.6 中的 3 维立方体，若初始解选择原点 $\mathbf{x} = (0,0)$ ，易证单纯型法经过 7 次翻转，才能找到最优解 $\mathbf{x} = (0,0,1)$ 。走过的 Hamiltonian

路径为 $\Theta_3 = 000 \rightarrow 100 \rightarrow 110 \rightarrow 010 \rightarrow 011 \rightarrow 111 \rightarrow 101 \rightarrow 001$ 。而且不难发现, 这条 3 维 Hamiltonian 路径 Θ_3 可以分解为两段 2 维 Hamiltonian 路径 Θ_2 和反向 Θ_2 , 在第二段 2 维 Hamiltonian 路径和反向 Θ_2 , $x_3 = 1 - \varepsilon x_2$ 。每次翻转都保持目标函数下降。

类似的, 若 $n = k$ 时, 上述结果成立。则 $n = k + 1$ 时, 我们可以得到一条 $k + 1$ 维 Hamiltonian 路径 Θ_{k+1} , 这条 $k + 1$ 维 Hamiltonian 路径 Θ_3 可以分解为两段 k 维 Hamiltonian 路径 Θ_k 和反向 Θ_k 。在第二段 k 维 Hamiltonian 路径反向 Θ_k , $x_{k+1} = 1 - \varepsilon x_k$ 。对于每段分别应用 $n = k$ 时结论, 可以证明每次翻转都保持目标函数下降。易证单纯型法经过 $2^{k+1} - 1$ 次翻转, 才能找到最优解 $\mathbf{x} = (0, \dots, 1)$ 。

因此, 根据数学归纳法, 定理得证。□

例 2.5. 我们使用单纯型法求解例 2.1 中的线性规划问题。

解:

如果我们在一开始不能找到明显的初始可行基, 可以使用大 M 法或者两阶段法找到初始可行基。

其中大 M 法在标准形式线性规划问题 (2.3) 中添加人工变量 $\mathbf{u} \in \mathbb{R}^m$, 得到如下的线性规划问题

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{u}} \quad & z = \mathbf{c}^T \mathbf{x} + \mathbf{m}^T \mathbf{u} \\ \text{s.t.} \quad & \mathbf{Ax} + \mathbf{u} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}, \mathbf{u} \geq \mathbf{0} \end{aligned} \quad (2.21)$$

其中向量 $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{c} \in \mathbb{R}^n$ 。向量 $\mathbf{m} = [M \ \dots \ M]^T \in \mathbb{R}^m$ 的所有元素均为足够大的正数 M 。

因为 $\mathbf{u} \geq \mathbf{0}$, M 又很大, 则在线性规划问题 (2.21) 的求解过程中, 我们自然而然的会倾向于令 $\mathbf{u} \rightarrow \mathbf{0}$, 使得目标函数 $z = \mathbf{c}^T \mathbf{x} + \mathbf{m}^T \mathbf{u} \rightarrow \mathbf{c}^T \mathbf{x}$ 。如果该问题存在有界的最优解, 当应用单纯型法到停止时, 会发现 $\mathbf{u} = \mathbf{0}$, 即人工变量最终能全部出基, 最后得到的结果稍作整理就可知原问题的最优解。

但是由于 M 很大, 在应用单纯型法时会造成较大的计算舍入误差, 因此现在该方法很少使用。为了克服这一困难, 两阶段法在第一阶段先考虑问题

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{u}} \quad & z = \mathbf{1}^T \mathbf{u} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} + \mathbf{u} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}, \mathbf{u} \geq \mathbf{0} \end{aligned} \quad (2.22)$$

显然, 线性规划问题 (2.22) 如果存在有界的最优解 $\mathbf{x}_0, \mathbf{u}_0$, 必然有 $\mathbf{u}_0 = \mathbf{0}$, 即人工变量最终能全部出基。而此时得到的 \mathbf{x}_0 的解就标准形式线性规划问题 (2.3) 的初始可行解。此时, 再开始第二阶段, 正式求解原问题。

2.6.2 对偶单纯型法

考虑标准形式线性规划问题 (2.3) 及其对偶问题 (2.9)。单纯型法的本质是针对原问题, 在保证 $\mathbf{x}_B = \mathbf{A}_B^{-1}\mathbf{b} \geq \mathbf{0}$ 的前提下, 迭代寻找新的基矩阵直到满足 $\boldsymbol{\sigma}_N = \mathbf{c}_N^T - \mathbf{c}_B^T \mathbf{A}_B^{-1} \mathbf{A}_N \geq \mathbf{0}$ 。根据强对偶性, 此时对偶问题同时达到最优解。

但从对偶问题的角度来看, 我们也可以设计如下的对偶单纯型法, 在保证 $\boldsymbol{\sigma}_N = \mathbf{c}_N^T - \mathbf{c}_B^T \mathbf{A}_B^{-1} \mathbf{A}_N \geq \mathbf{0}$ 的前提下, 迭代寻找新的基矩阵直到满足 $\mathbf{x}_B = \mathbf{A}_B^{-1}\mathbf{b} \geq \mathbf{0}$ 。这样在迭代的过程中, 我们保证每次迭代找到的都是对偶问题的可行解, 直到找到对偶问题的最优解。根据强对偶性, 此时原问题同时达到最优解。

算法 2.2. 对偶单纯型法

Step 1. [确定初始对偶可行解] 如果观察原问题即能找到明显的初始基矩阵 \mathbf{A}_B , 满足 $\boldsymbol{\sigma}_N = \mathbf{c}_N^T - \mathbf{c}_B^T \mathbf{A}_B^{-1} \mathbf{A}_N \geq \mathbf{0}$, 但对应的 $\mathbf{x}_B = \mathbf{A}_B^{-1}\mathbf{b}$ 中存在小于 0 的元素, 则进入 Step 2。

Step 2. [判断最优性] 求 $r = \text{arcmmin}\{\bar{b}_i \mid i = 1, \dots, m\}$ 。

如果 $\bar{b}_r \geq 0$, 且 $\boldsymbol{\sigma} \geq \mathbf{0}$, 则已找到原始问题最优解, 停止。

如果 $\bar{b}_r < 0$, 但矩阵 \mathbf{A} 的第 r 行所有元素非负, $\mathbf{x} \geq \mathbf{0}$, 则下式不可能成立

$$x_r + \bar{a}_{r,m+1}x_{m+1} + \dots + \bar{a}_{r,n}x_n = \bar{b}_r$$

此时可以得出该线性规划问题无解, 停止。

如果 $\bar{b}_r < 0$, 但矩阵 \mathbf{A} 的第 r 行并非所有元素非负。注意到我们可以将对应的基变量 x_r 从负值变为 0 来增加解的可行性并同时减少目标函数值, 则进入 Step 3, 通过翻转找到一个新的基矩阵 \mathbf{A}_B

Step 3. [翻转 (Pivots)] 出基变量选为 x_r , 计算求解下述问题, 确定可能的进基变量的下标所组成的集合

$$L = \left\{ \operatorname{arcm}_{i} \left\{ -\frac{\sigma_i}{\bar{a}_{ir}} \mid \sigma_i < 0, i = m+1, \dots, n \right\} \right\} \quad (2.23)$$

如果有多个 $-\sigma_i / \bar{a}_{ik}$ 同时取得最小值, 则同样按照 Bland 法则, 选取这其中下标最小的 i 作为进基变量下标 l

$$l = \min \{i \mid i \in L, i = m+1, \dots, n\} \quad (2.24)$$

则进基变量为此时非基变量 x_l 。

根据原有基本解 \mathbf{x}_B , 出基变量 \mathbf{x}_r 和进基变量 \mathbf{x}_l 构造新的可行基矩阵 $\hat{\mathbf{A}}_B$, 进而求解除新的基本解 $\hat{\mathbf{x}}_B$ 。这一步可以直接在表中, 由进基变量所在列和出基变量所在行决定的约束矩阵系数通过初等变换得到。最后回到 Step 2 进行下一轮迭代。

定理 2.22. 对偶单纯型法一定能在有限时间内停止。

证明: 对偶单纯型法每次迭代都是从一个基本解前进到另外一个基本解。根据定理 2.6 和定理 2.4, 线性规划问题的基本解个数是有限的。

如果遇到非退化基本解。考虑到 $\sigma_i < 0$, $\bar{b}_r < 0$, $\bar{a}_{ik} > 0$, 不难发现迭代前后, 目标函数值单调下降

$$z' = z - \frac{\sigma_i \bar{b}_r}{\bar{a}_{ik}} < z$$

如果遇到退化基本解, 则采用 Bland 方法保证不发生循环。则对偶单纯型法一定能在有限时间内停止。□

对偶单纯型法特别适合如下类型的线性规划问题

$$\begin{aligned} \min_{\mathbf{x}} \quad & z = \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} \geq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \quad (2.25)$$

通过添加人工变量 $\mathbf{u} \in \mathbb{R}^m$, 我们可以得到如下的线性规划问题

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{u}} \quad & z = \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} + \mathbf{u} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}, \mathbf{u} \geq \mathbf{0} \end{aligned} \quad (2.26)$$

该问题直接可以观测到初始的对偶可行解 $\mathbf{x} = \mathbf{0}$, $\mathbf{u} = \mathbf{b}$; 然后开始应用对偶单纯型法求解。我们在本书 3.3 节可以看到该方法的一个典型应用。

例 2.6. 我们使用对偶单纯型法求解例 2.1 中的线性规划问题。

解:

2.7 其他

当我们考虑生产建设的时候，常可以建模得到如下的线性规划问题及其对偶问题

$$\begin{aligned} \max_{\mathbf{x}} \quad & z = \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \quad \begin{aligned} \min_{\mathbf{u}} \quad & \mathbf{b}^T \mathbf{u} \\ \text{s.t.} \quad & \mathbf{A}^T \mathbf{u} \geq \mathbf{c} \\ & \mathbf{u} \geq \mathbf{0} \end{aligned} \quad (2.27)$$

其中 b_i 表示投入的第 i 种生产资源的总量， a_{ij} 表示投入的单位数量的第 i 种生产资源产出第 j 种产品的转化率， c_j 表示单位数量的第 j 种产品的价值， x_j 表示计划产出第 j 种产品的数量； $i=1, \dots, m$ ， $j=1, \dots, n$ 。

假设对偶问题最优解为 $\hat{\mathbf{u}}$ ，由强对偶性知，原问题的目标函数最优值为

$\mathbf{b}^T \hat{\mathbf{u}} = \sum_{i=1}^m b_i \hat{u}_i$ 。所以，原问题的目标函数在最优解处关于 b_i 的偏导数分别是

\hat{u}_i 。这说明增加一个单位的第 i 种生产资源，最优目标值可望增加 \hat{u}_i ，故称 \hat{u}_i 为

b_i 的影子价格。

进一步考虑检验数 $\sigma_j = c_j - \mathbf{c}_B^T \mathbf{A}_B^{-1} \bar{\mathbf{a}}_j = c_j - \hat{\mathbf{u}}^T \bar{\mathbf{a}}_j = c_j - \sum_{i=1}^m \hat{u}_i \bar{a}_{ij}$ ，则 σ_j

等于生产一个单位第 j 种产品的利润减去按影子价格计算的资源的总成本，如

果差值 σ_j 大于零，应继续生产，所以最优解必须满足所有检验数非正。

实际上，我们可以进一步分析目标函数带有参数的线性规划问题。例如考虑目标函数中含参数 λ 的线性规划问题

$$\begin{aligned}
 \min_{\mathbf{x}} \quad & z = (\mathbf{c} + \lambda \bar{\mathbf{c}})^T \mathbf{x} \\
 \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b} \\
 & \mathbf{x} \geq \mathbf{0}
 \end{aligned} \tag{2.28}$$

我们可以首先令 $\lambda = 0$ 解无参数的问题，然后确定保持当前最优基矩阵不变的 λ 各个取值区间，最后在 λ 各个取值区间里，求解带参数的问题最优解。

对于约束条件右边常数含参数的线性规划问题

$$\begin{aligned}
 \min_{\mathbf{x}} \quad & z = \mathbf{c}^T \mathbf{x} \\
 \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b} + \lambda \bar{\mathbf{b}} \\
 & \mathbf{x} \geq \mathbf{0}
 \end{aligned} \tag{2.29}$$

我们可以转而考虑其对偶问题

$$\begin{aligned}
 \max_{\mathbf{u}} \quad & (\mathbf{b} + \lambda \bar{\mathbf{b}})^T \mathbf{u} \\
 \text{s.t.} \quad & \mathbf{A}^T \mathbf{u} \leq \mathbf{c} \\
 & \mathbf{x} \geq \mathbf{0}
 \end{aligned} \tag{2.30}$$

转换为上述目标函数带有参数的线性规划问题进行求解。显然，这一变换的想法就是保持可行域不含参数，方便我们讨论分析。

例 2.7. 我们求解如下带参数的线性规划问题。

解：

2.8 作业题

2.1 如果某个标准形式线性规划问题的一个极点比其周围的所有极点目标函数值都小，试证明或证否该极点处取得最优解。

2.2 如果某个标准形式线性规划问题的一个极点比其它所有极点目标函数值都小，试证明或证否该极点处取得最优解。

2.3 如果一个可行解是标准形式线性规划问题的最优解，则其对应的检验数满足 $\sigma \geq 0$ 。

2.4 如果从标准形式线性规划问题的约束条件矩阵 $\mathbf{A} \in \mathbb{R}^{m \times n}$ 的 n 列中挑出 m 列构成的矩阵 \mathbf{B} 可逆，剩下的 $n-m$ 列构成的矩阵 \mathbf{N} 。满足 $\mathbf{B}^{-1} \mathbf{b} \geq \mathbf{0}$,

$\mathbf{c}_N^T - \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{N} \geq \mathbf{0}$, 则其对应的解是该线性规划问题的最优解。

2.5 试证明或证否: 如果线性规划问题 $\{\min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} \geq \mathbf{0}\}$ 存在无界解, 则锥可行域 $\{\mathbf{x} \mid \mathbf{Ax} \geq \mathbf{0}\}$ 有且必有一条极线 \mathbf{d} 满足 $\mathbf{c}^T \mathbf{d} \leq \mathbf{0}$ 。

2.6 试构造反例证明, 如果不按照 (2.19) 定义的方式选择出基变量, 不能保证在迭代过程中 $\mathbf{x}_B \geq \mathbf{0}$ 。

2.9 参考文献

- [1] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley, 1986.
- [2] D. Bertsimas, J. N. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific Press, 1997.
- [3] D. G. Luenberger, Y. Ye, *Linear and Nonlinear Programming*, 4th edition, Springer, 2016.
- [4] F. L. Hitchcock, "The distribution of a product from several sources to numerous localities," *Journal of Mathematics and Physics*, vol. 20, no. 1-4, pp. 224-230 1941.
- [5] T. C. Koopmans, "Optimum utilization of the transportation system," *Proceedings of the International Statistical Conference*, Washington, DC, 1947.
- [6] G. B. Dantzig, "Linear Programming," *Operations Research*, vol. 50, no. 1, pp. 42-47, 2002.
- [7] V. Klee, G. J. Minty, "How good is the simplex algorithm?" in *Inequalities, III*, O. Shisha, ed., pp. 159-175, Academic Press, 1972.
- [8] L. G. Khachiyan, "A polynomial algorithm for linear programming," *Doklady Akademii Nauk SSSR*, vol. 244, pp. 1093-1096, 1979.
- [9] N. K. Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica*, vol. 4, pp. 373-395, 1984.
- [10] D. A. Spielman, S.-H. Teng, "Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time," *Journal of the ACM*, vol. 51, no. 3, pp. 385-463, 2004.
- [11] A. J. Hoffman, "Cycling in the simplex algorithm," Technical report 2974, National Bureau of Standards, Gaithersburg, USA, December 16th, 1953.
- [12] E. M. L. Beale, "Cycling in the dual simplex algorithm," *Naval Research Logistics Quarterly*, vol. 2, no. 4, pp. 269-275, 1955.
- [13] R. G. Bland, "New finite pivoting rules for the simplex method," *Mathematics of Operations Research*, vol. 2, no. 2, pp. 103-107, 1977.
- [14] T. Terlaky, "Lexicographic pivoting rules," in *Encyclopedia of Optimization*, C. A. Floudas, P. M. Pardalos, eds., Springer, pp. 1870-1873, 2009.