# 组合最优化

**(第四章)**

中国科学院大学
数学科学学院

郭田德

**电话：88256412,**

**Email: tdguo@ucas.ac.cn**

# 第四章

# Polytope, polyhedra, Fakars' lemma, and linear programming

## 4.1 仿射集

### 4..1.1 仿射集

对于 $n$ 维欧氏空间中的任意两个点 $x, y \in R^n, \lambda \in R$，若 $x \neq y$，记 $\Delta x = y - x \neq 0$，则

$$(1-\lambda)x + \lambda y = x + \lambda(y - x) = x + \lambda \Delta x$$

所以通过 $x, y$ 的直线可以表示为：$line(x, y) = \{\lambda x + (1-\lambda)y : \lambda \in R\}$.

仿射组合：$\sum_{i=1}^{p} \lambda_i x^i$，其中，$x^1, x^2, \cdots, x^p \in R^n$，$\lambda_1, \lambda_2, \cdots, \lambda_p \in R$，$\sum_{i=1}^{p} \lambda_i = 1$。

**定义 4.1** 若集合 $S \subset R^n, \forall x^1, x^2 \in S$，$S$ 必包含 $x^1$ 和 $x^2$ 两点的所有仿射组合。（$S$ 包含两个点，$S$ 必包含经过这两个点的**直线**），则称 $S$ 为一个仿射集（affine set）。仿射集也称为仿射流形（affine manifold）。

注：若集合 $S \subset R^n$ 是一个仿射集，$a \in R^n$，则

（1）$S$ 关于 $a$ 的平移 $S + a = \{x + a : x \in S\}$ 也是仿射集。

（2）若 $a \in S$，则 $S$ 关于 $-a$ 的平移 $S - a = \{x - a : x \in S\}$ 是仿射集，并且包含原点，它是 $R^n$ 的一个子空间。

（3）反之，$R^n$ 的子空间是包含原点的仿射集。

（4）对于一个非空的仿射集 $S \subset R^n$，存在 $R^n$ 的唯一的子空间 $L$ 和 $a \in R^n$，使得 $S = L + a = \{x + a : x \in L\}$。显然 $a \in S$，并且 $L = S - S = \{x - y : x, y \in S\}$。**称 $L$ 为平行于仿射集 $S$ 的子空间。**

**定义 4.2** 仿射集 $S$ 的维数是指平行于该仿射集 $S$ 的子空间的维数。

注：单点集的维数为 0，直线的维数为 1，平面的维数为 2。

### 4.1.2 超平面

**定义 4.3** $R^n$ 的维数为 $n-1$ 的仿射集 $H$ 称为 $R^n$ 中的超平面( hyperplane )。

对于超平面 $H$，维数为 $n-1$，子空间 $L$ 平行于 $H$，子空间 $L$ 的维数为 $n-1$，则 $L$ 的正交补空间 $L^{\perp}$ 维数为 1，假设 $L^{\perp}$ 的基为 $\alpha \in R^n$，$\alpha \neq 0$，则子空间 $L=\{x \in R^n : \alpha^T x = 0\}$。于是，对于 $a \in H$，超平面（仿射集）$H = L + a$ 可以表示成 $H=\{x+a \in R^n : \alpha^T x = 0\} = \{y \in R^n : \alpha^T y = \alpha^T a\}$

所以，对于 $\alpha \in R^n, \beta \in R^1$，可以定义一个超平面：

$$H=\{x \in R^n : \alpha^T x = \beta\} = \{x \in R^n : \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \beta\}。$$

反之，$R^n$ 中每一个超平面都可以表示成上述形式。，并且在相差一个非零因子的意义下 $(\alpha, \beta)$ 是唯一的。

$\alpha$ 是超平面 $H$ 的法线。$\forall z \in H, y \in H$，有

$\alpha^T(y-z) = \alpha^T y - \alpha^T z = \beta - \beta = 0$，即 $\alpha \perp (y-z)$

即：法向量 $\alpha$ 与所有平行于超平面 $H$ 的向量垂直。

对于线性规划单纯性形算法的标准型：

$$\begin{aligned} &\min c^{\mathrm{T}} x \\ &s.t.\begin{cases} Ax = b \\ x \geq 0 \end{cases} \end{aligned} \qquad \text{(LP)}$$

其中 $x \in R^n, c \in R^n, b \in R^m, A \in R^{m \times n}$。超平面 $H=\{x \in R^n : c^T x = \beta\}$ 是目标函数 $c^T x$ 的一个等值面，价格向量 $c$ 是等值面的法线。

**定理 4.1** 给定矩阵 $A \in R^{m \times n}$，向量 $b \in R^m$，则 $S=\{x \in R^n : Ax = b\}$ 是 $R^n$ 中的一个仿射集（$m$ 个超平面仿射集的交集）。反之，如果 $S$ 是 $R^n$ 中的一个仿射集，则存在矩阵 $A \in R^{m \times n}$ 和向量 $b \in R^m$ 使得 $S=\{x \in R^n : Ax = b\}$，即 $R^n$ 中的任一个仿射集必是一组超平面的交集。

**定义 4.4** 给定 $R^n$ 中的集合 $S$，包含 $S$ 的所有仿射集的交集，即包含 $S$ 的最小

仿射集，称为$S$的仿射包(affine hull)，记为$\operatorname{aff} S$。事实上，

$$\operatorname{aff} S=\left\{\sum_{i=1}^{k}\lambda_i x^i : \sum_{i=1}^{k}\lambda_i=1, \lambda_i\in R, x^i\in S, i=12,\cdots,k, k\in N\right\}$$

**定义 4.5** $R^n$中的集合$S$的维数定义为它的仿射包 $\operatorname{aff} S$的维数，即包含$S$的仿射集的最小维数。

**定义 4.6** 由$m+1$个向量组成的向量组$x^0,x^1,\cdots,x^m$称为仿射无关的（affinely independent），是指集合$\{x^0,x^1,\cdots,x^m\}$的维数为$m$，即仿射包$\operatorname{aff}\{x^0,x^1,\cdots,x^m\}$的维数为$m$。

事实上，$\operatorname{aff}\{x^0,x^1,\cdots,x^m\}=L+x^0$，其中$L=\{0,x^1-x^0,\cdots,x^m-x^0\}$是包含$\{x^1-x^0,\cdots,x^m-x^0\}$的最小子空间。

显然，$L$的维数为$m$的充分必要条件是$x^1-x^0,\cdots,x^m-x^0$线性无关。

**命题 4.1** 下面三个断言是等价的：

（1） 在$R^n$中的向量组$\{x^0,x^1,\cdots,x^m\}$是仿射无关的；

（2） 在$R^n$中的向量组$\{x^1-x^0,\cdots,x^m-x^0\}$是线性无关的；

（3） 在$R^{n+1}$中的向量组$\{(x^0,1),(x^1,1),\cdots,(x^m,1)\}$是线性无关的。

## 4.2 凸集

对于$n$维欧氏空间中的任意两个点$x,y\in R^n,\lambda\in R, 0\le\lambda\le 1$，若$x\ne y$，记$\Delta x=y-x\ne 0$，则

$$(1-\lambda)x+\lambda y=x+\lambda(y-x)=x+\lambda\Delta x$$

所以$x,y$之间的线段可以表示为：

$$\operatorname{sec} tion(x,y)=\{\lambda x+(1-\lambda)y : 0\le\lambda\le 1\}.$$

凸组合：$\sum_{i=1}^{p}\lambda_i x^i$，其中，$x^1,x^2,\cdots,x^p\in R^n$，$\lambda_1,\lambda_2,\cdots,\lambda_p\in R$，$\sum_{i=1}^{p}\lambda_i=1$，

$$1\ge\lambda_i\ge 0 (i=1,2,\cdots,p)。$$

**定义 4.7** 集合 $S \subset R^n$，若 $\forall x^1, x^2 \in S$，则 $S$ 必包含 $x^1$ 和 $x^2$ 的所有凸组合，则 $S$ 为称凸集（即若 $S$ 包含两个点，则 $S$ 必包含连接这两个点的**线段上的所有点**）。

显然，$R^n$ 中的超平面 $H = \{x \in R^n : \alpha^T x = \beta\}$ 是凸集。

**半空间：**

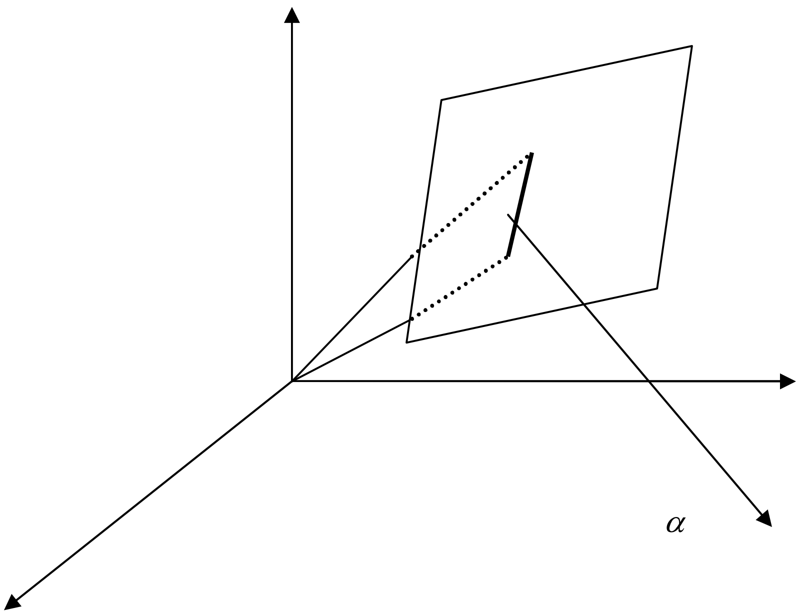
$$\left.\begin{array}{l} H_L = \{x \in R^n : \alpha^T x \le \beta\} \\ H_U = \{x \in R^n : \alpha^T x \ge \beta\} \end{array}\right. \text{——两个闭半空间}$$

$$\left.\begin{array}{l} H_L^i = \{x \in R^n : \alpha^T x < \beta\} \\ H_U^i = \{x \in R^n : \alpha^T x > \beta\} \end{array}\right. \text{——两个不相交的开半空间}$$

上述四个半空间都是凸集。

$H$ 是 $H_L$ 和 $H_L^i$（当然也是 $H_U$ 和 $H_U^i$）的边界超平面。

对于 $R^n$ 中的超平面 $H = \{x \in R^n : \alpha^T x = \beta\}$，$\alpha$ 是超平面 $H$ 的法线。



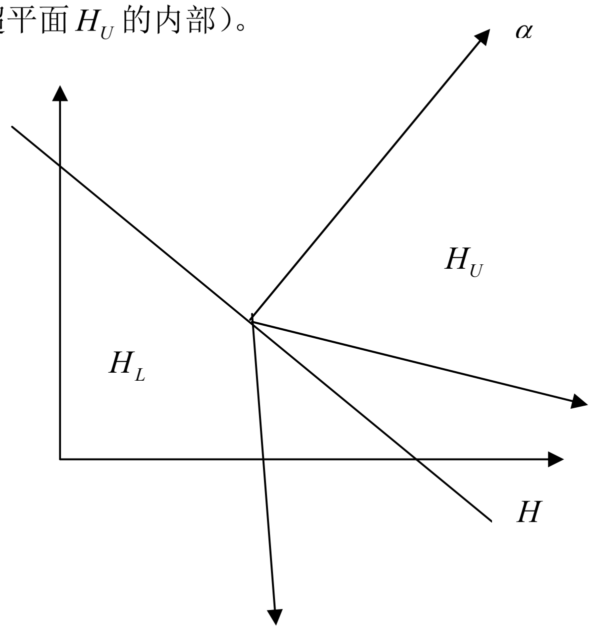

又：$\forall z \in H, w \in H_L^i$，有

$$\alpha^T (w - z) = \alpha^T w - \alpha^T z < \beta - \beta = 0$$

或者 $\forall z \in H, v \in H_U^i$，有

$\alpha^T(v-z)=\alpha^T v-\alpha^T z>\beta-\beta=0$

即：法向量$\alpha$与由超平面指向$H_L$内部的任意向量构成钝角（或者法向量$\alpha$与由超平面指向$H_U$内部的任意向量构成锐角），也即$\alpha$指向超平面$H_L$的外部（或者$\alpha$指向超平面$H_U$的内部）。




注：仿射集必是凸集，凸集不一定是仿射集。

显然：

- 超平面是仿射集（也是凸集）；
- 线性流形$\{x\in R^n : Ax=b\}$是仿射集（也是凸集）；
- 闭半空间是凸集但不是仿射集；
- 线性规划单纯性形算法的标准型的可行域：

  $P=\{x\in R^n : Ax=b, x\geq 0\}$是凸集但不是仿射集；

**定义 4.8** 内点与边界点：给定集合$S\subset R^n, x\in S, \exists\varepsilon>0$，使得

$$B=\{y\in R^n :\| x-y\|<\varepsilon\}\subset S$$

则称$x$是$S$的一个内点。反之，$x$是$S$的一个边界点。

# 4.3 凸包与凸集分离定理

## 4.3.1 Convex Hulls

Consider the line segment $L$ joint two points $u$ and $v$ in $R^n$. For any given $w \in R^n$, the mathematical programming problem

$$\max w^T x$$
$$s.t.\ x \in L$$

is particularly easy to solve: Either $u$ or $v$ is an optimal solution.

Suppose $w^T u \geq w^T v$ and let $x \in L$. We have $x = \lambda u + (1-\lambda)v$ for some $\lambda \in [0,1]$.Thus, for any $x \in L$

$$w^T x = \lambda w^T u + (1-\lambda) w^T v \leq \lambda w^T u + (1-\lambda) w^T u = w^T u$$

and so, $u$ is an optimal solution.

上述结论可以扩展多个向量：如果

$$x = \lambda_1 v_1 + \lambda_2 v_2 + \cdots \lambda_k v_k$$

其中 $v_i \in R^n, \lambda_i \in [0,1], i = 1,2,\cdots,k, \sum_{i=1}^{k} \lambda_i = 1$，则

$$w^T x = \lambda_1 w^T v_1 + \lambda_2 w^T v_2 + \cdots \lambda_k w^T v_k \leq \max\{w^T v_i : i = 1,2,\cdots,k\}$$

$x = \lambda_1 v_1 + \lambda_2 v_2 + \cdots \lambda_k v_k$ 是向量组 $v_i \in R^n, i = 1,2,\cdots,k$ 的凸组合。

**有限集合的凸包：**

**定义 4.9** The *convex hull* of a finite set $S$ (denoted by $conv.hull(S)$ is the set of all vectors that can be written as a convex combination of $S$.

$$conv.hull(S) = \{x \in R^n : x = \sum_{i=1}^{k} \lambda_i v_i, v_i \in S, \lambda_i \in [0,1], \sum_{i=1}^{k} \lambda_i = 1\}$$

显然，$conv.hull(S)$ 是包含有限集合 $S$ 的最小凸集。

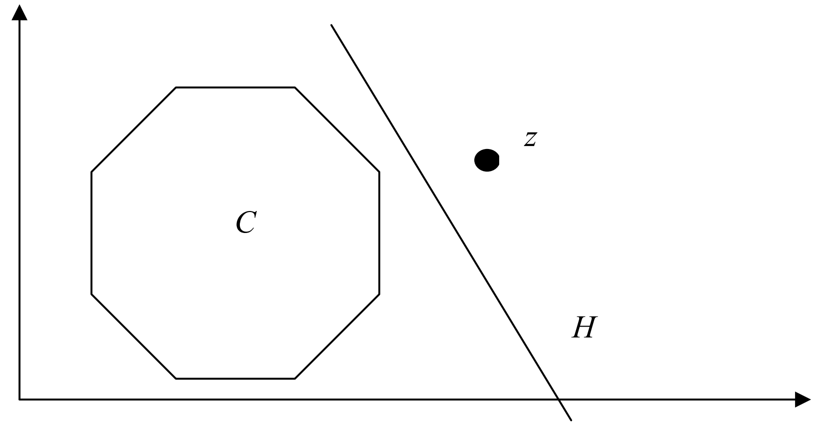
**任意集合的凸包：**

The *convex hull* of any subset $X$ of $R^n$

Clearly, the intersection of any number of convex sets is again a convex set. So, for any subset $X$ of $\mathbb{R}^n$, the smallest convex set containing $X$ exists. This set is called the *convex hull* of $X$ and is denoted by conv.hull($X$). One easily proves:

(1) $$\text{conv.hull}(X) = \{x \mid \exists t \in \mathbb{N}, \exists x_1, \ldots, x_t \in X, \exists \lambda_1, \ldots, \lambda_t \geq 0 : x = \lambda_1 x_1 + \cdots + \lambda_t x_t, \lambda_1 + \cdots + \lambda_t = 1\}.$$

对于一个闭凸集，一个关键的几何特性是下述的分割定理。




A basic property of closed convex sets is that any point not in $C$ can be separated from $C$ by a 'hyperplane'. Here a subset $H$ of $\mathbb{R}^n$ is called a *hyperplane* (or an *affine hyperplane*) if there exist a vector $c \in \mathbb{R}^n$ with $c \neq 0$ and a $\delta \in \mathbb{R}$ such that:

(2) $$H = \{x \in \mathbb{R}^n \mid c^T x = \delta\}.$$

我们称 $H$ 分离 $z$ 和 $C$，如果 $z$ 和 $C$ 分别在 $H$ 将 $R^n$ 分开的两个半空间中。

**定义 4.10** 给定非空集合 $S_1, S_2 \subset R^n$ 和超平面 $H = \{x \in R^n : \alpha^T x = \beta\}$。若 $S_1 \subseteq H_L, S_2 \subseteq H_U$ 成立，即 $\forall x^i \in S_i, i = 1,2$，有 $\alpha^T x^1 \le \beta \le \alpha^T x^2\}$，则称超平面 $H = \{x \in R^n : \alpha^T x = \beta\}$ 分离集合 $S_1$ 和 $S_2$。此外，若 $S_1 \cup S_2 \not\subset H$，则称 $H$ 正常分离 $S_1$ 和 $S_2$。若 $S_1 \subseteq H_L^i, S_2 \subseteq H_U^i$，则称 $H$ 严格分离 $S_1$ 和 $S_2$。若存在 $\varepsilon > 0$ 使得 $S_1 \subseteq H(\varepsilon) = \{x \in R^n : \alpha^T x \le \beta - \varepsilon\}, S_2 \subseteq H_U$，则称 $H$ 强分离 $S_1$ 和 $S_2$。

**点与凸集分离：**

**定理 4.2** 设 $C$ 是 $R^n$ 的一个闭凸集，$z \notin C$。则存在一个超平面分离 $z$ 和 $C$。

**Proof.** Since the theorem is trivial if $C = \emptyset$, we assume $C \neq \emptyset$. Then there exists a vector $y$ in $C$ that is nearest to $z$, i.e., that minimizes $\|z - y\|$.

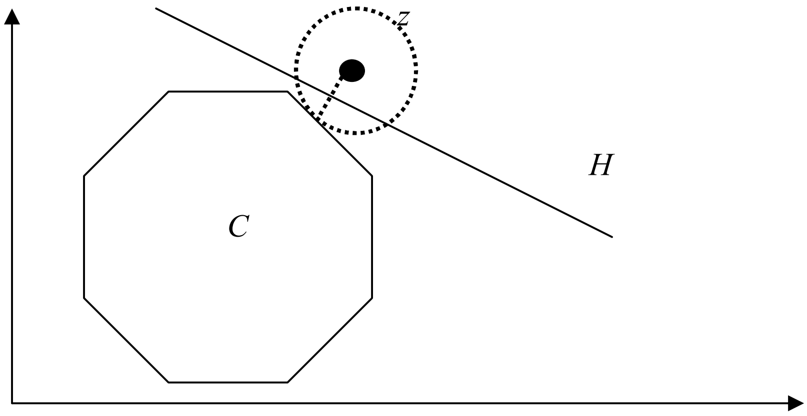
(The fact that such a $y$ exists, can be seen as follows. Since $C \neq \emptyset$, there exists an $r > 0$ such that $B(z, r) \cap C \neq \emptyset$. Here $B(z, r)$ denotes the closed ball with center $z$ and radius $r$. Then $y$ minimizes the continuous function $\|z - y\|$ over the compact set $B(z, r) \cap C$.)

Now define:

(3) $$c := z - y, \delta := \frac{1}{2}(\|z\|^2 - \|y\|^2).$$

We show

(4) (i) $c^T z > \delta$,
(ii) $c^T x < \delta$ for each $x \in C$.




Indeed, $c^T z = (z - y)^T z > (z - y)^T z - \frac{1}{2}\|z - y\|^2 = \delta$. This shows (4)(i).

If (4)(ii) would not hold, there exists an $x$ in $C$ such that $c^T x \geq \delta$. Since $c^T y < c^T y + \frac{1}{2}\|c\|^2 = \delta$, we know $c^T(x - y) > 0$. Hence there exists a $\lambda$ with $0 < \lambda \leq 1$ and

(5) $$\lambda < \frac{2c^T(x - y)}{\|x - y\|^2}.$$

Define

(6) $$w := \lambda x + (1 - \lambda)y.$$

So $w$ belongs to $C$. Moreover,

(7) $$\|w - z\|^2 = \|\lambda(x - y) + (y - z)\|^2 = \|\lambda(x - y) - c\|^2$$
$$= \lambda^2\|x - y\|^2 - 2\lambda c^T(x - y) + \|c\|^2 < \|c\|^2 = \|y - z\|^2.$$

Here $<$ follows from (5).

However, (7) contradicts the fact that $y$ is a point in $C$ nearest to $z$. ∎

Call a subset $H$ of $\mathbb{R}^n$ a *halfspace* (or an *affine halfspace*) if there exist a vector $c \in \mathbb{R}^n$ with $c \neq 0$ and a $\delta \in \mathbb{R}$ such that

$$H = \{x \in \mathbb{R}^n \mid c^T x \leq \delta\}. \tag{8}$$

Clearly, each affine halfspace is a closed convex set.

Theorem 4.2 implies that if $C$ is a closed convex set and $z \notin C$ then there exists an affine halfspace $H$ so that $C \subseteq H$ and $z \notin H$

对于一个一般凸集，另一个关键的几何特性是下述**过边界上点的分割定理**：

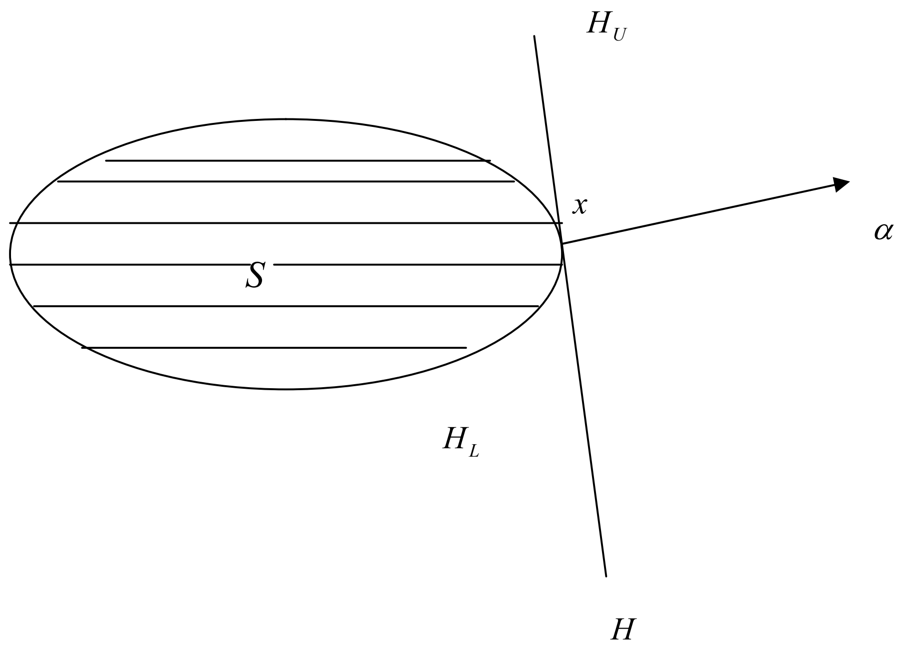
**定理 4.3** 令 $S$ 是 $R^n$ 中的凸子集，且 $x$ 是 $S$ 的一个边界点，则存在一个包含 $x$ 的超平面 $H$，使得 $S$ 或包含在 $H_L$ 中，或包含在 $H_U$ 中。

证明：思考题。

因此，我们可以定义一个支撑超平面 $H$：

1). $H$ 和 $S$ 的交是非空的；

2). $H_L$ 包含 $S$。




**定理 4.3** 令 $S$ 是 $R^n$ 中的凸子集，$x \notin S$，则存在超平面 $H$，使得 $S$ 包含在 $H_L$ 中，$x$ 包含在 $H_U$ 中。。

证明：思考题。

**凸集与凸集分离：**

**定理 4.4** （凸集分离定理）设 $S_1, S_2 \subset R^n$ 是两个非空凸集，$S_1 \cap S_2 \neq \phi$，则存在超平面 $H$，使得 $S_1$ 包含在 $H_L$ 中，$S_2$ 包含在 $H_U$ 中。

证明　令 $S = S_1 - S_2$，则 $S$ 为非空凸集且 $0 \notin S$。

## 4.4 多面体(Polyhedron)、多胞形(Polytopes)和多面锥（polyhedral cone ）

Special classes of closed convex sets are formed by the polytopes and the polyhedra. In the previous section we saw that each closed convex set is the intersection of affine halfspaces, possibly infinitely many. If it is the intersection of a *finite* number of affine halfspaces, the convex set is called a *polyhedron*.

So a subset $P$ of $\mathbb{R}^n$ is a polyhedron if and only if there exists an $m \times n$ matrix $A$ and a vector $b \in \mathbb{R}^m$ such that

$$P = \{x \in R^n : Ax \leq b\}$$

**多面体**（polyhedron）：由有限个闭半空间的交集形成的一个集合。

**多胞形**（polytope）：非空有界多面体。

给定多面体 $P$ 及其支撑超平面 $H$，称 $F = P \cap H$ 为 $P$ 的一个面。

- 若 $\dim(F) = 0$：就有 $P$ 一个顶点；
- 若 $\dim(F) = 1$：就有 $P$ 一条边；
- 若 $\dim(F) = \dim(P) - 1$：就有 $P$ 一个面；

**锥**：非空集合 $C \subset R^n$，若 $\forall x \in C, \lambda \geq 0$ 总有 $\lambda x \in C$，则称 $C$ 是一个锥。

**凸锥**：非空集合 $C \subset R^n$，若 $\forall x, y \in C, \lambda \geq 0, \mu \geq 0,$ 总有 $\lambda x + \mu y \in C$，则称 $C$ 是一个凸锥。

凸锥与线性半空间：

Let $C \subseteq \mathbb{R}^n$. Then $C$ is a closed convex cone if and only if $C = \bigcap \mathcal{F}$ for some collection $\mathcal{F}$ of linear halfspaces.

(A subset $H$ of $\mathbb{R}^n$ is called a *linear halfspace* if $H = \{x \in \mathbb{R}^n \mid c^T x \leq 0\}$ for some nonzero vector $c$.)

（证明：见作业题 2.3）

显然，凸锥一定是锥。

**生成锥**：一个锥被称之为由 $x_1, x_2, \cdots, x_k$ 生成的，如果 $x_1, x_2, \cdots, x_k \in C$，并且对 $\forall x \in C$, 存在 $\lambda_1, \lambda_2, \cdots, \lambda_k \ge 0$，使得 $x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k$。

**有限生成锥**：一个锥被称之为有限生成锥，如果它是由一个有限向量集合生成的锥。

**多面锥**（polyhedral cone ）：$C = \{x : Ax \le 0\}$（有限个线性半空间的交）

**定理 4. 5** Minkowski 1896, Steiniz 1916, Weyl 1935) A set $P$ is a polytope if and only if it is the convex hull of a finite set of points.

证明：见阅读材料。。

**定理 4. 6** (Minkowski 1896, Weyl 1935) A cone is polyhedral if and only if it is finitely generated.

证明：思考题。

**Proposition 4.4** Let $S \subseteq R^n$ be a finite set and let $w \in R^n$. Then

$$\max\{w^T x : x \in S\} = \max\{w^T x : x \in conv.hull(S)\}$$

粗看起来，这个结论似乎对我们用处不大，因为我们本来是在一个有限集合上求最大值，反而变成了在在一个无限集合上求最大值。但是，The *convex hull* of a finite set $S$ 有非常好的集合性质。

$|S| = 2$：$conv.hull(S)$ 是线段

$|S| = 3$：$conv.hull(S)$ 是三角形

（其中 $|S|$ 表示集合 $S$ 中元素个数）

一般地，$conv.hull(S)$ 是包含 $S$ 的最小多面体（多面体的顶点全部属于 $S$）。

给定一个集合 $S = \{s_1, s_2, \cdots, s_k\} \subset R^n$，对于一个给定的向量 $v \in R^n$，我们很容易判断它是否包含在 $conv.hull(S)$：我们只要判断是否存在 $\lambda_i \in R^1$，$i = 1,2,\cdots,k$，使得：

$$\sum_{i=1}^{k} \lambda_i = 1$$

$$\sum_{i=1}^{k} \lambda_i s_i = v$$

$$\lambda_i \geq 0, i = 1,2,\cdots,k$$

上述系统是一个线性系统，变量是 $\lambda_i, i = 1,2,\cdots,k$。判断是否包含在 $con.hull(S)$ 中就是判断上述系统是否有解。可以用 Farkas 引理。

**Theorem 4.7** (Farkas' lemma). The system $Ax = b$ has a nonnegative solution if and only if there is no vector $y$ satisfying $y^T A \geq 0$ and $y^T b < 0$.

**Proof.** *Necessity.* Suppose $Ax = b$ has a solution $x_0 \geq 0$ and suppose there exists a vector $y_0$ satisfying $y_0^T A \geq 0$ and $y_0^T b < 0$. Then we obtain the contradiction

$$0 > y_0^T b = y_0^T (Ax_0) = (y_0^T A) x_0 \geq 0. \tag{33}$$

*Sufficiency.* Suppose $Ax = b$ has no solution $x \geq 0$. Let $a_1, \ldots, a_n$ be the columns of $A$. So

$$b \notin C := \text{cone}\{a_1, \ldots, a_n\}. \tag{34}$$

So by Exercise 2.3 there exists a linear halfspace $H$ containing $C$ and not containing $b$. That is, there exists a vector $c$ such that $c^T b < 0$ while $c^T x \geq 0$ for each $x$ in $C$. In particular, $c^T a_j \geq 0$ for $j = 1, \ldots, n$. So $y := c$ satisfies $y^T A \geq 0$ and $y^T b < 0$. ∎

So Farkas' lemma states that exactly one of the following two assertions is true:

(35) (i) $\exists x \geq 0 : Ax = b$,
(ii) $\exists y : y^T A \geq 0$ and $y^T b < 0$.

（线性系统 $Ax = b, x \geq 0$ 与线性系统 $y^T A \geq 0, y^T b < 0$ 有且仅有一个有解）

There exist several variants of Farkas' lemma, that can be easily derived from Theorem 4.5.

**Corollary 4.7a**. The system $Ax \leq b$ has a solution $x$ if and only if there is no vector $y$ satisfying $y \geq 0$, $y^T A = 0$ and $y^T b < 0$.

**Proof.** Let $A'$ be the matrix

$$A' := [A \quad -A \quad I], \tag{36}$$

where $I$ denotes the $m \times m$ identity matrix.

Then $Ax \leq b$ has a solution $x$ if and only if the system $A'x' = b$ has a nonnegative solution $x'$. Applying Theorem 2.5 to $A'x' = b$ gives the corollary. ∎

**Corollary 4.7b**. Suppose that system $Ax \leq b$ has at least one solution. Then for every solution $x$ of $Ax \leq b$ one has $c^T x \leq \delta$ if and only if there exists a vector $y$ satisfying $y \geq 0$, such that $y^T A = c^T$ and $y^T b < \delta$.

**Proof.** *Sufficiency.* If such a vector $y$ exists, then for every vector $x$ one has

$$Ax \leq b \Longrightarrow y^T Ax \leq y^T b \Longrightarrow c^T x \leq y^T b \Longrightarrow c^T x \leq \delta. \tag{37}$$

*Necessity.* Suppose that such a vector $y$ does not exist. It means that the following system of linear inequalities in the variables $y$ and $\lambda$ has no solution $(y^T \ \lambda) \geq (0 \ 0)$:

$$(y^T \quad \lambda)\begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} = (c^T \quad \delta). \tag{38}$$

According to Farkas' lemma this implies that there exists a vector $\begin{pmatrix} z \\ \mu \end{pmatrix}$ so that

$$\begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}\begin{pmatrix} z \\ \mu \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ and } (c^T \quad \delta)\begin{pmatrix} z \\ \mu \end{pmatrix} < 0. \tag{39}$$

We distinguish two cases.

**Case 1:** $\mu = 0$. Then $Az \geq 0$ and $c^T z < 0$. However, by assumption, $Ax \leq b$ has a solution $x_0$. Then, for $\tau$ large enough:

$$A(x_0 - \tau z) \leq b \text{ and } c^T(x_0 - \tau z) > \delta. \tag{40}$$

This contradicts the fact that $Ax \leq b$ implies $c^T x \leq \delta$.

**Case 2:** $\mu > 0$. As (39) is homogeneous, we may assume that $\mu = 1$. Then for $x := -z$ one has:

$$Ax \leq b \text{ and } c^T x > \delta. \tag{41}$$

Again this contradicts the fact that $Ax \leq b$ implies $c^T x \leq \delta$. ∎

**Proposition 4.7c** Let $S = \{s_1, s_2, \cdots, s_k\} \subset R^n$ and let $v \in R^n \setminus conv.hull(S)$ .Then there exists an inequality $w^T x \leq t$ that separates $v$ from $conv.hull(S)$ , that is, $w^T s \leq t$ for all $s \in conv.hull(S)$ but $w^T v > t$ .

## 4.5 线性规划

线性规划问题最早是由 G.B.Dantzig 在 1947 年以前设想出来的。他当时作为联邦空军审计员的一名数学顾问，需要开发一个数学规划的工具，用于制定布置、训练、后勤保障的方案。

由于这项工作，他于 1948 年出版了《线性结构的规划》一书。

1948 年夏天：T.C. Koopmans&G.B.Dantzig 提出了“线性规划”的名称；

1949 年：G.B.Dantzig 提出了单纯形方法。

在此之前：Fourier, W.Karush, L.V. Kantorovich 等人的工作都曾涉及到线性规划的有关工作。

1950－1960：线性规划的理论得到了进一步的发展；

1975 年：L.V. Kantorovich 和 T.C. Koopmans 获得诺贝尔经济奖－对资源最优分配理论的贡献；

1979 年：L.G.. Khanchian 的椭球算法；

1984 年：N. Karmarkar 的投影尺度算法

### 4.5.1 线性规划问题的几何解释

一、线性规划问题的标准型

$$\min c^{\mathrm{T}} x \qquad s.t. \begin{cases} Ax = b \\ x \geq 0 \end{cases} \qquad \text{(LP)}$$

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

另外，总假设：$b \geq 0.$ $(A, b, c)$ 的元素都为整数，$\operatorname{rank}(A) = m$ 记：

可行域：$P = \{x \in R^n : Ax = b, x \geq 0\}$

最优解集：$P^* = \{x \in P : x\text{是(LP)的最优解}\}$

1.凸多面体的顶点和线性规划的基可行解

凸多面体的顶点：几何实体；

线性等式与不等式组的基可行解：代数上的定义。

在线性规划的理论中，将这两个概念联系在一起，用几何直觉导引出代数工具。

顶点：在凸集$C$中的一个点$x$，如果$x$不是$C$另外两个不同点的凸组合，则称它是$C$的一个顶点。

即：一个顶点是这样一个点，它不能位于凸集中另外两个点的连线的线段之中——凸多面体的“端点”。

线性规划问题

$$\min c^{\mathrm{T}} x \\ s.t.\begin{cases} Ax = b \\ x \geq 0 \end{cases} \qquad \text{(LP)}$$

其中$x \in R^n, c \in R^n, b \in R^m, A \in R^{m\times n}$， $m \leq n$。

可行域$P = \{x \in R^n : Ax = b, x \geq 0\}$

令$A = (A_1, A_2, \cdots, A_n)$，对于$x \in P$有：

$$(A_1, A_2, \cdots, A_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = b$$

即：$x_1 A_1 + x_2 A_2 + \cdots + x_n A_n = b$。称$A_j$为变量$x_j$对应的列。

**定理4.8** $x \in P$，则$x$是$P$的一个顶点$\Leftrightarrow$ $x$的正分量对应的$A$中的各列是线性独立的。

**证明：**不仿设$x = (x_1, x_2, \cdots, x_p, 0, 0, \cdots, 0)^T = (x_B, 0)^T$，其中$x_i > 0 (i = 1, 2, \cdots, p)$，记$A = (B, N)$， $x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}$，其中$B$是$A$的前$p$列。则， $Ax = b \Leftrightarrow Bx_B = b$。

"$\Rightarrow$" 用反证法，假设$x$是$p$的一个顶点，但$B$的各列不线性独立，则存在一个非零向量$w$，使得$Bw = 0$。令$x_B^1 = x_B + \delta w, x_B^2 = x_B - \delta w$，由于$x_B > 0$，所以对于充分小的$\delta > 0$，有$x_B^1 \geq 0, x_B^2 \geq 0$。显然$Bx_B^1 = Bx_B^2 = b$。定义：$x^1 = (x_B^1, 0)^T, x^2 = (x_B^2, 0)^T$，则$Ax^1 = Ax^2 = b$，所以$x^1, x^2 \in P$，且$x = \frac{1}{2}x^1 + \frac{1}{2}x^2$，与$x$是$p$的一个顶点矛盾。

"$\Leftarrow$"用反证法，假设$x$不是$p$的一个顶点，则存在$y^1, y^2 \in P, 0<\lambda<1,$使得$x=\lambda y^1+(1-\lambda)y^2$。因为$x_N=0$，所以$y_N^1=y_N^2=0$。令$w=x-y^1$，则$w$为非零向量，且$Bw_B=Bx_B-By_B^1=b-b=0$，所以$B$中各列线性相关，与假设矛盾。

证毕

令$A=(B,N)$，$B\in R^{m\times m}$为满秩矩阵，是(LP)的一个基，$x=\begin{pmatrix} x_B \\ x_N \end{pmatrix}$，

$Ax=b \quad \Leftrightarrow \quad Bx_B+Nx_N=b \quad \Leftrightarrow \quad x_B+B^{-1}Nx_N=B^{-1}b$，$\Leftrightarrow x_B=B^{-1}b-B^{-1}Nx_N$，令$x_N$=0，若$x=\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}\geq 0$，称之为(LP)的一个基可行解。

**推论 4.8a** $x$是(LP)的一个基可行解$\Leftrightarrow$ $x\in P$是$P$的一个顶点。

**推论 4.8b** (LP)的可行域至多有$C_n^m$个顶点。

由于假设$(A,b,c)$的元素都为整数，所以任一个基解其分量的绝对值是有界的。

**定理 4.9** 令$x$是一个基解，则有$|x_j|\leq m!\alpha^{m-1}\beta$，其中

$$\alpha=\max_{i,j}\{|a_{ij}|\}，\quad \beta=\max_{1\leq j\leq m}\{|b_j|\}$$

注：该结论及其证明的思想非常有用。

**证明：** 因为$B^{-1}=\dfrac{B^*}{\det B}$，而$\det B\neq 0$为整数，则必有$|\det B|\geq 1$，所以分母绝对值大于或等于 1，而$B^*$每个元素等于$B$的$(m-1)$阶子式的行列式，而$B$的$(m-1)$阶子式的行列式是$A$中的$(m-1)!$个$(m-1)$个元素连乘积之和，其绝对值不大于$(m-1)!\alpha^{m-1}$。由于每一个$x_j$是$B^{-1}$中的$m$个元素与$b$中的$m$个元素对应乘积之和，所以$|x_j|\leq m!\alpha^{m-1}\beta$。

**定理 4.10** 假定标准的线性规划问题满足(i) $\mathrm{rank}(A)=m$, (ii) $P\neq\phi$, (iii) 目标函数$c^Tx$有下界，则在最优值相等的意义下，它与下述线性规划等价：

$$
\begin{aligned}
&\min c^T x \\
&s.t. \quad Ax = b \\
&\qquad\quad x \geq 0 \\
&\qquad\quad x_i \leq M, i = 1,2,\cdots,n
\end{aligned}
$$

其中 $M = (m+1)!\alpha^m \beta, \alpha = \max\{|a_{ij}|,|c_i|\}, \beta = \max\{|b_j|,|z|\}$，$z$ 是集合 $\{c^T x \mid Ax = b, x \geq 0\}$ 的最大下界。

由该定理：我们总可以假定可行域 $F$ 是有界区域。

证明：

五、非退化与相邻性

(1) 基可行解和顶点不是一一对应：

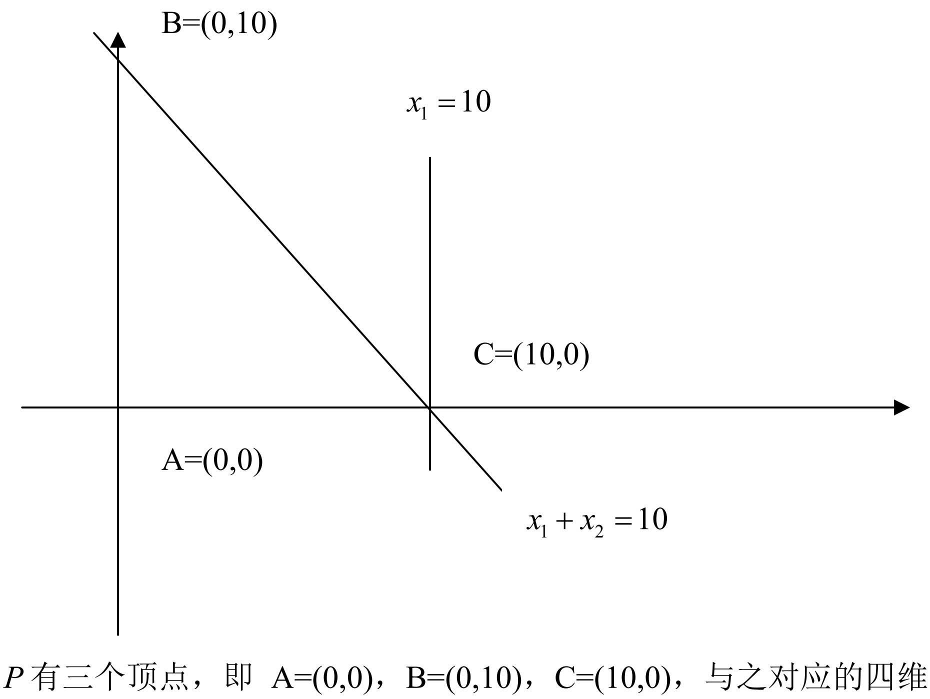
- 任给一个基可行解，存在唯一的一个顶点与之对应；
- 对于 $P$ 中的一个顶点，可能有多个基可行解与之对应。

例如：设

$P = \{x \in R^4 : x_1 + x_2 + x_3 = 10, x_1 + x_4 = 10, x_1, x_2, x_3, x_4 \geq 0\}$

即 $P$ 等价于下述图形：

$P = \{x \in R^2 : x_1 + x_2 \leq 10, x_1 \leq 10, x_1, x_2 \geq 0\}$




$P$ 有三个顶点，即 A=(0,0)，B=(0,10)，C=(10,0)，与之对应的四维空间的坐标为：A=(0,0,10,10)，B=(0,10,0,10)，C=(10,0,0,0)，

显然：

- 顶点 A 是对应基变量为 $x_3, x_4$ 的一个基可行解；
- 顶点 B 是对应基变量为 $x_2, x_4$ 的一个基可行解；

● 顶点 C 对应三个基可行解：基变量为 $x_1, x_2$

基变量为 $x_1, x_3$

基变量为 $x_1, x_4$

在这三个基可行解中，都有一个基变量取值为零！

(2) 一个基可行解 $\begin{pmatrix} x_B \\ x_N \end{pmatrix}$ 非退化：如果 $x_B > 0$， $x_N = 0$。

(3) 一个线性规划问题非退化：如果它对应的所有基可行解都是非退化的。

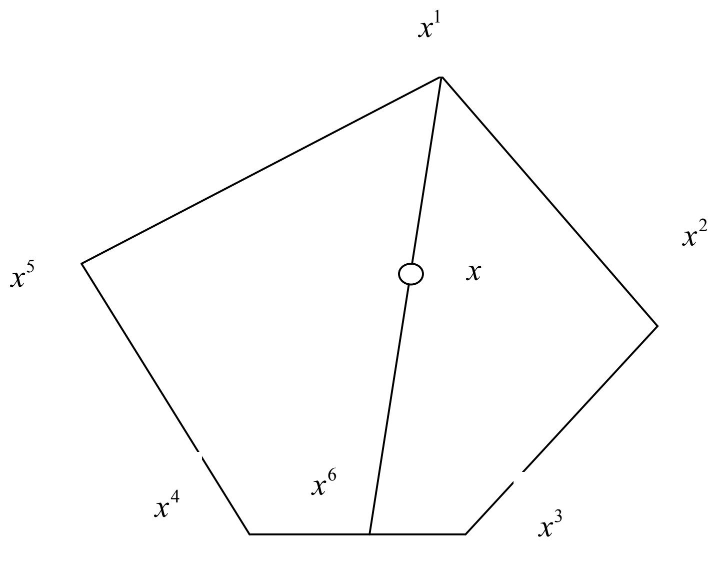
(4) 基可行解（顶点）相邻：只有一个基变量不相同（即共用 m-1 个基变量）的基可行解（顶点），称之为是相邻的。

注 1：每一个基可行解（顶点）都有 n-m 个相邻的基可行解（顶点）；

2：每一个相邻的基可行解，都可以通过下述方式达到：将一个非基变量的值由零增加到正，而同时将一个正基变量由正值减为零，并保持可行性。

六、非退化与相邻性

设可行域 $P$ 有界，即 $P$ 是一个多胞形，如：




显然：$\forall x \in P$， $x$ 可以表示成 $P$ 的顶点的凸组合

(1) 多面体的顶点方向：设非零向量 $d \in R^n$，若任给 $x^0 \in P$，射线

$$\{x \in R^n : x = x^0 + \lambda d, \lambda \geq 0\} \subset P$$

则称非零向量 $d$ 为多面体 $P$ 的顶点方向。

显然：$d$ 为多面体 $P$ 的顶点方向 $\Leftrightarrow$ $Ad = 0, d \geq 0$。

多面体 $P$ 无界 $\Leftrightarrow$ $P$ 至少有一个顶点方向。

**定理 4.11**（分解定理）：令 $V=\{v^i \in R^n : i\in I\}$ 是 $P$ 的所有顶点集合，则 d $\forall x\in P$，有

$$x=\sum_{i\in I}\lambda_i v^i+\lambda d$$

其中 $\sum_{i\in I}\lambda_i=1, \lambda_i\geq 0, i\in I$，$\lambda\geq 0$ 且 $d$ 或者是零向量，或者是 $P$ 的一个顶点方向。

证明：数学归纳法，对于 $x\in P$ 的正分量的个数归纳。

**推论 4.11a**：若 $P$ 是多胞形，则 $\forall x\in P$，$x$ 可以表示成 $P$ 的顶点的凸组合。

**推论 4.11b**：若 $P\neq\phi$，则 $P$ 至少存在一个顶点。

### 4.5.2 线性规划的基本定理

**定理 4.12** 对于标准的线性规划问题

$$\min z=c^{\mathrm{T}}x$$
$$s.t.\begin{cases}Ax=b\\ x\geq 0\end{cases} \qquad \text{(LP)}$$

若 $P\neq\phi$，则 $z$ 或者无下界，或者至少存在 $P$ 的一个顶点，$z$ 其上达到最小值。

证明：令 $V=\{v^i\in P : i\in I\}$ 是 $P$ 的所有顶点。

$\because P\neq\phi$，$\therefore V\neq\phi$。

根据分解定理，$\forall x\in P$ 有 $x=\sum_{i\in I}\lambda_i v^i+\lambda d$，其中 $\sum_{i\in I}\lambda_i=1, \lambda_i\geq 0, i\in I$，$\lambda\geq 0$ 且 $d$ 或者是零向量，或者是 $P$ 的一个顶点方向。

(1) 若 $P$ 存在一个顶点方向 $d$ 满足 $c^T d<0$，则 $z$ 必无界。事实上，任给 $x^0\in P$，$\{x\in R^n : x=x^0+\lambda d, \lambda\geq 0\}\subset P$，当 $\lambda\to+\infty$ 时，$c^T x\to-\infty$，所以 $z$ 无下界。

(2) 否则，设 $P$ 的所有顶点方向 $d$ 满足 $c^T d\geq 0$（或者不存在顶点方向）。记 $v^{\min}=\arg\min\{c^T v^i : i\in I\}$ 是目标函数值最小的顶点，且，则 $\forall x\in P$ 有

$$C^T x=\sum_{i\in I}\lambda_i C^T v^i+C^T d\geq\sum_{i\in I}\lambda_i C^T v^{\min}=C^T v^{\min}$$

所以 $z$ 在顶点 $v^{\min}$ 上达到最小值。

### 4.5.3 线性规划问题的单纯形算法

给定一个非退化的基可行解$\bar{x}$，对应的可行基为$B$则等式约束变为：

$$x_B + B^{-1}Nx_N = B^{-1}b$$

$$x_B = B^{-1}b - B^{-1}Nx_N$$

目标函数$c^{\mathrm{T}}x = c_B^{\mathrm{T}}x_B + c_N^{\mathrm{T}}x_N$

$$= c_B^{\mathrm{T}}(B^{-1}b - B^{-1}Nx_N) + c_N^{\mathrm{T}}x_N$$

$$= c_B^{\mathrm{T}}B^{-1}b - (c_B^{\mathrm{T}}B^{-1}N - c_N^{\mathrm{T}})x_N$$

规划等价于

$$\min c_B^{\mathrm{T}}B^{-1}b - (c_B^{\mathrm{T}}B^{-1}N - c_N^{\mathrm{T}})x_N$$
$$s.t.\begin{cases} x_B = B^{-1}b - B^{-1}Nx_N \\ x \ge 0 \end{cases}$$

如果令：

$$B^{-1}b = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix},\quad B^{-1}N = \begin{pmatrix} \beta_{1,m+1} & \beta_{1,m+2} & \cdots & \beta_{1,n} \\ \beta_{2,m+1} & \beta_{2,m+2} & \cdots & \beta_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{m,m+1} & \beta_{m,m+2} & \cdots & \beta_{m,n} \end{pmatrix},$$

$(\lambda_{m+1}, \lambda_{m+2}, \cdots, \lambda_n)^T = C_B^T B^{-1}N - C_N^T$，$f_0 = C_B^T B^{-1}b$

则上述线性规划问题就变成：

$$\min f_0 - (\lambda_{m+1}x_{m+1} + \lambda_{m+2}x_{m+2} + \cdots, \lambda_n x_n)$$
$$s.t.\begin{cases} x_1 = \alpha_1 - (\beta_{1,m+1}x_{m+1} + \beta_{1,m+2}x_{m+2} + \cdots + \beta_{1,n}x_n) \\ x_2 = \alpha_2 - (\beta_{2,m+1}x_{m+1} + \beta_{2,m+2}x_{m+2} + \cdots + \beta_{2,n}x_n) \\ \cdots\cdots \\ x_m = \alpha_m - (\beta_{m,m+1}x_{m+1} + \beta_{m,m+2}x_{m+2} + \cdots + \beta_{m,n}x_n) \\ x_i \ge 0 (i = 1, 2, \cdots, n) \end{cases} \qquad \text{典式(LP')}$$

对应的基可行解为$(\alpha_1, \alpha_2, \cdots, \alpha_m, 0, \cdots, 0)^T$，目标函数值为$f_0$。

若有一个$\lambda_i > 0$，不妨设$\lambda_{m+1} > 0$，则可令$x_{m+1}$从 0 上升到某个$\theta > 0$，显然，可以使得目标函数值下降。

**定理 4.13** 设$\{x_1, x_2, \cdots, x_m\}$是（LP）的一个可行基，（LP'）为其对应的典式。如果$\lambda_{m+1} \le 0, \lambda_{m+2} \le 0, \cdots, \lambda_{m+n} \le 0$，则基$\{x_1, x_2, \cdots, x_m\}$对应的基可行解

$$x^0 = (\alpha_1, \alpha_2, \cdots, \alpha_m, 0, \cdots, 0)^T$$

是最优解。

**定理 4.14** 设$\{x_1, x_2, \cdots, x_m\}$是（LP）的一个可行基，（LP'）为其对应的典式。如果目标函数有下界且存在一个检验数$\lambda_{m+k} > 0$，则非基变量$x_{m+k}$对应的系数$\beta_{1,m+k}, \beta_{2,m+k}, \cdots, \beta_{m,m+k}$中至少有一个大于零。

**定理 4.15** 设$\{x_1, x_2, \cdots, x_m\}$是（LP）的一个可行基，（LP'）为其对应的典式。如果存在一个检验数$\lambda_{m+k} > 0$，使得：

（1）非基变量$x_{m+k}$对应的系数$\beta_{1,m+k}, \beta_{2,m+k}, \cdots, \beta_{m,m+k}$中至少有一个大于零；

（2）所有$\alpha_i > 0 (i = 1, 2, \cdots, m)$，

则一定存在另一个可行基，它对应的基可行解代入目标函数所得到的值比$f_0$小（即：新的基可行解要比原来的更好）。

## 4.5.4 线性规划的对偶理论

先从一个简单的例子谈起。

例子：某工厂在计划期内要安排生产 I、II 两种产品，已知生产单位产品所需要的设备台时及 A、B 两种原材料的消耗，如下表：

| | I | II | |
|---|---|---|---|
| 设备 | 1 | 2 | 8 台时 |
| 原材料 A | 4 | 0 | 16 |
| 原材料 B | 0 | 4 | 12 |

该工厂每生产一件产品 I 可获利 2，每生产一件产品 II 可获利 3。问应该如何安排生产计划使该工厂获利最大？

设$x_1$、$x_2$分别表示 I、II 的产量，则该问题的数学模型为：

$$
\begin{aligned}
\max \quad & z = 2x_1 + 3x_2 \\
\text{s.t.} \quad & x_1 + 2x_2 \le 8 \\
& 4x_1 \qquad \le 16 \\
& \qquad 4x_2 \le 12 \\
& x_1, x_2 \ge 0
\end{aligned}
$$

用单纯形方法可以求得最优解为：$x_1^* = 4, x_2^* = 2$，最优值为$z^* = 14$。

假设：该工厂的决策者决定不生产产品 I、II，而将其所有资源出租或外售。

这时工厂的决策者就要考虑给每种资源定价的问题。设用 $y_1, y_2, y_3$ 分别表示出租单位设备台时的租金和出让单位原材料 A，B 的附加额，则

$$
\begin{aligned}
\min \quad & \omega = 8y_1 + 16y_2 + 12y_3 \\
\text{s.t.} \quad & y_1 + 4y_2 \geq 2 \\
& 2y_1 + 4y_3 \geq 3 \\
& y_i \geq 0, i = 1,2,3
\end{aligned}
$$

也可以用单纯形方法求得最优解为：$y_1* = \dfrac{3}{2}, y_2* = \dfrac{1}{8}, y_3* = 0$，最优值为 $\varpi* = 14$。

显然：

$(8 - x_1 * - 2x_2*) = 0$ 且 $y_1* > 0$，即 $(8 - x_1 * - 2x_2*)y_1* = 0$——原始约束紧，对偶变量松

$(16 - 4x_1*) = 0$ 且 $y_2* > 0$，即 $(16 - 4x_1*)y_2* = 0$——原始约束紧，对偶变量松

$(12 - 4x_2*) > 0$ 且 $y_3* = 0$，即 $(12 - 4x_2*)y_3* = 0$——原始约束松，对偶变量紧

同样，对称的，

$x_1* > 0$ 且 $(y_1 * + 4y_2 * - 2) = 0$，即 $x_1 * \cdot (y_1 * + 4y_2 * - 2) = 0$——原始变量松，对偶约束紧

$x_2* > 0$ 且 $(2y_1 * + 4y_3 * - 3) = 0$，即 $x_2 * \cdot (2y_1 * + 4y_3 * - 3) = 0$——原始变量松，对偶约束紧

**最终达到平衡，原始－对偶目标函数取值相等，得到原始－对偶最优解。这就是所谓的“互补松弛性”**

互补松弛性

原始与对偶规划之间存在者拉锯式争夺：

一个问题里的某个约束越紧，另一个问题中对应的变量就越松；最终的平衡表示式，就是 $x$ 和 $y$ 是原始－对偶问题最优解的充分必要条件，这就是所谓的**互补松弛性条件**

**定理 4.16** （互补松弛性条件）$x$ 和 $y$ 分别为原始－对偶可行解，则它们分别是

原始－对偶最优解$\Leftrightarrow$对一切$i$和$j$有：

$$u_i = y_i(a_i^T x - b_i) = 0$$
$$v_j = (c_j - A_j^T y)x_j = 0$$

证明：显然，对一切$i$和$j$有：$u_i \ge 0, v_j \ge 0$。定义

$$u = \sum_i u_i,\quad v = \sum_j v_j$$

则　$u = 0 \Leftrightarrow u_i = 0$（对一切$i$）

$v = 0 \Leftrightarrow v_j = 0$（对一切 $j$ ）

而$u + v = c^T x - b^T y$

所以，$u_i = 0$（对一切$i$）且$v_j = 0$（对一切 $j$ ）$\Leftrightarrow u = 0$且$v = 0$

$\Leftrightarrow u + v = 0 \Leftrightarrow c^T x - b^T y = 0 \Leftrightarrow x$和$y$是原始－对偶问题最优解。

注：上述定理隐含着下述事实：

- 对最优解$x$和$y$，如果对偶中一个约束取严格不等式，则原始规划中对应的变量取值必须为 0；
- 对称地，如果一个非负变量取值为正值，则其对应的约束必取等式。

所以，称之为**互补松弛性**。

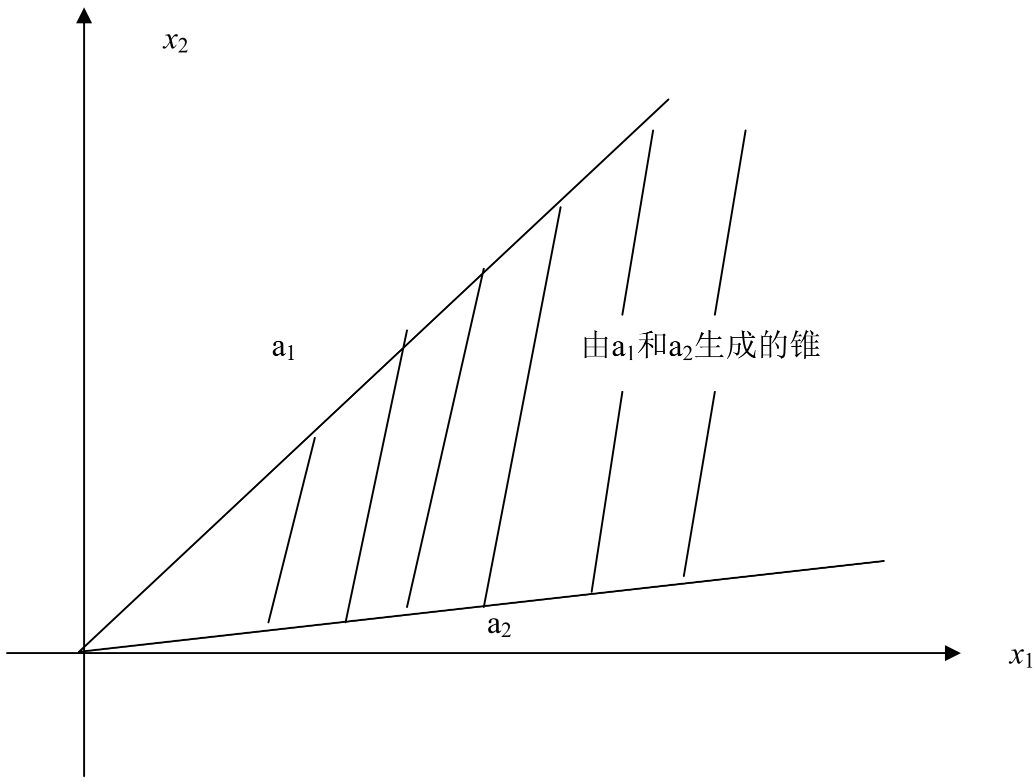
**Farkas 引理与对偶**

给定一个集合$S = \{s_1, s_2, \cdots, s_k\} \subset R^n$，对于一个给定的向量$v \in R^n$，我们可用 Farkas 引理判断它是否包含在凸包$conv.hull(S)$。同样，我们可用 Farkas 引理判断它是否包含在凸锥组合中。实际上，Farkas 引理描述了$R^n$中向量间的一种基本关系。在某种意义下，它反映了对偶的本质。

给定一组向量$a_i \in R^n (i =,2, \cdots, m)$由这组向量$\{a_i\}$生成的锥记为$C(a_i)$：

$$C(a_i) = \{x \in R^n : x = \sum_{i=1}^m y_i a_i, y_i \ge 0, i = 1,2,\cdots,m\}$$

即$\{a_i\}$非负线性组合。

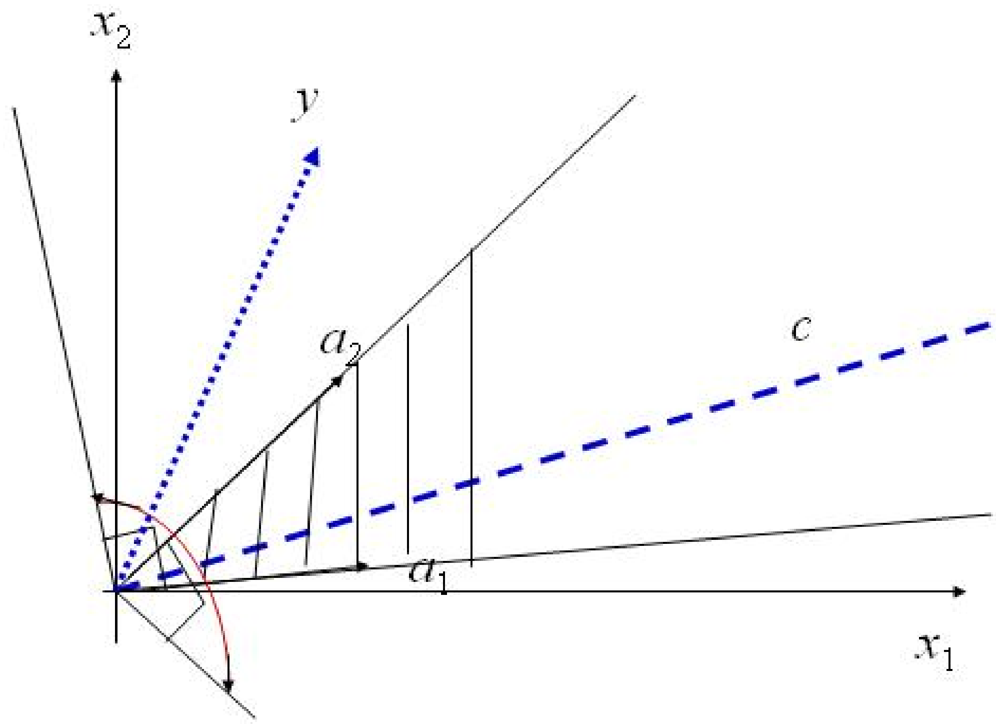



给定向量的一个集合$\{a_i\}_{i=1,2,\cdots,k}$及另外一个向量$c \in R^n$，“如果对一切向量$y \in R^n$，若$y$在$\{a_i\}_{i=1,2,\cdots,k}$有非负投影，那么$y$在$c$上也有非负投影”

***Farkas 引理断言***

$$\Longleftrightarrow$$

$c$在锥$C(a_i)$里




**定理 4.17（Farkas 引理）**给定一组向量 $a_i \in R^n (i=,2,\cdots,m)$ 及向量 $c \in R^n$，则有：

$$a_i^T y \geq 0, i=1,2,\cdots m \Rightarrow c^T y \geq 0 \Leftrightarrow c \in C(a_i)$$

**注：**给定向量的一个集合 $\{a_i\}_{i=1,2,\cdots,k}$ 及另外一个向量 $c \in R^n$，记 $A=(a_1,a_2,\cdots,a_k)$，因为：

（1） **命题**“若 $a_i^T y \geq 0, i=1,2,\cdots m$ 成立，则有 $c^T y \geq 0$ 成立”**的等价命题（逆否命题）是：**“若 $c^T y < 0$ 成立（$c^T y \geq 0$ 不成立），则存在 $i$ 使得 $a_i^T y < 0$ 成立（$a_i^T y \geq 0, i=1,2,\cdots m$ 不成立）”。**即等价于命题“**线性系统 $a_i^T y \geq 0, i=1,2,\cdots m$，$c^T y < 0$ 无解”。

（2） $c \in C(a_i)$ 的充分必要条件是线性系统 $Ax=c, x \geq 0$ 有解。

所以，**上述 Farkas 引理**可以表述成下述等价形式：**线性系统** $Ax=c, x \geq 0$ **有解充分必要条件是线性系统** $y^T A \geq 0$，$c^T y < 0$ **无解**。

若线性系统 $Ax=c, x \geq 0$ 有解，则 $c \in C(a_i)$，所以，若 $y^T A \geq 0$，则 $c^T y \geq 0$，

从而线性系统 $y^T A \ge 0, c^T y < 0$ 无解（反之，若有解，则存在 $y$ 使得 $y^T A \ge 0$ 但是 $c^T y < 0$，Farkas 引理 4.17 不成立，矛盾）；

若线性系统 $y^T A \ge 0, c^T y < 0$ 无解，则若有 $y^T A \ge 0$，则必有 $c^T y \ge 0$，从而 $c \in C(a_i)$，即线性系统 $Ax = c, x \ge 0$ 有解。

所以定理 4.17（Farkas 引理）等价形式为线性系统 $Ax = c, x \ge 0$ 与线性系统 $y^T A \ge 0, c^T y < 0$ 有且只有一个有解。**（这就是定理 4.5）**

证明:“$\Leftarrow$” 给定一组向量 $a_i \in R^n (i =,2,\cdots,m)$ 及向量 $c \in R^n$ 且 $c \in C(a_i)$，要证明：对于 $y \in R^n$，若 $a_i^T y \ge 0, i = 1,2,\cdots m$，则必有 $c^T y \ge 0$。事实上,

$\because c \in C(a_i)$，$\therefore$ $\exists y_i \ge 0 (i = 1,2,\cdots,m)$ 使得 $c = \sum_{i=1}^{m} y_i a_i$。由已知 $a_i^T y \ge 0, i = 1,2,\cdots m$，则 $c^T y = \sum_{i=1}^{m} y_i a_i^T y \ge 0$。

“$\Rightarrow$” 给定一组向量 $a_i \in R^n (i =,2,\cdots,m)$ 及向量 $c \in R^n$，满足：对于 $y \in R^n$，若 $a_i^T y \ge 0, i = 1,2,\cdots m$ 成立，则一定有 $c^T y \ge 0$ 成立。要证明必有：$c \in C(a_i)$。事实上，可考察下述线性规划问题：

$$\begin{aligned} &\min c^T y \\ &a_i^T y \ge 0, i = 1,2,\cdots,m \qquad \text{(LP)} \\ &y\text{无限制} \end{aligned}$$

$\because$ $y$=0 是一个可行解，$\therefore$ (LP)可行。又 $\because$ $a_i^T y \ge 0, i = 1,2,\cdots,m$ 及 $c^T y \ge 0$，$\therefore$ (LP)有界，$\therefore$ (LP)的对偶问题：

$$\begin{aligned} &\max 0 \\ &A_j^T x = c_j, j = 1,2,\cdots,n \qquad \text{(LP)} \\ &x \ge 0 \end{aligned}$$

其中

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}=\begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{pmatrix}=\left(A_1, A_2, \cdots, A_n\right)$$

一定有可行解，即存在$x \geq 0$，使得 $A_j^T x = c_j, j = 1,2,\cdots,n$，即

$c_1 = a_{11}x_1 + a_{21}x_2 + \cdots + a_{m1}x_m$
$c_2 = a_{12}x_1 + a_{22}x_2 + \cdots + a_{m2}x_m$
……
$c_n = a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{mn}x_m$

所以

$$c=\begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{pmatrix}x_1+\begin{pmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2n} \end{pmatrix}x_2+\cdots+\begin{pmatrix} a_{m1} \\ a_{m2} \\ \vdots \\ a_{mn} \end{pmatrix}x_m$$

$$= x_1 a_1 + x_2 a_2 + \cdots + x_m a_m = \sum_{i=1}^{m} x_i a_i$$

即$c \in C(a_i)$。

## 4.5.5 线性规划的原始－对偶算法

### 一、算法的基本思路

线性规划的原始－对偶算法是线性规划的一个一般的算法，它实际上是由某些网络问题的一个特殊算法发展起来的，并且由它可以产生一系列与组合优化有关问题的一些特殊算法。

考虑线性规划问题及其对偶问题：

$$\begin{aligned} &\min z = c^T x \\ &\text{s.t. } Ax = b \geq 0 \quad \text{(LP)} \\ &\qquad x \geq 0 \end{aligned} \qquad\qquad \begin{aligned} &\max \varpi = b^T y \\ &\text{s.t. } \quad A^T y \leq c \qquad \text{(D)} \\ &\qquad y\text{无限制} \end{aligned}$$

互补松弛性条件：$x$ 和 $y$ 分别为原始－对偶可行解，则它们分别是原始－对偶最优解 $\Leftrightarrow$ 对一切 $i$ 和 $j$ 有：

$$y_i(a_i^T x - b_i) = 0 \tag{1}$$

$$(c_j - A_j^T y)x_j = 0 \tag{2}$$

显然，(1)对任何可行解都成立，只要讨论(2)。假定 $y$ 是(D)的可行解，我们的目标是设法找出(LP)的一个可行解 $x$，使得当 $c_j - A_j^T y > 0$ 时，有 $x_j = 0$，那么 $x$ 和 $y$ 将分别是原始－对偶最优解。

但是，由于 $y$ 不一定是(D)的一个最优解，所以，对于(D)的可行解 $y$，(LP)满足上述要求的可行解 $x$ 不一定能够找到。不过，对于给定的(D)的可行解 $y$，我们可以找出一个“**满足互补松弛性条件(2)，且最接近(LP)的可行解的向量** $x$”，**并根据该** $x$ **对(LP)的可行性的“破坏程度”调整(D)的可行解** $y$**，得到(D)的一个新的可行解** $y'$——线性规划的原始－对偶算法的基本思路。

## 二、线性规划的原始－对偶算法

线性规划的原始－对偶算法实际上是对偶算法，开始时 $y$ 是(D)的可行解，迭代过程始终保持对偶可行性。

### 1. 算法基本过程

$$\begin{aligned} &\min z = c^T x \\ &\text{s.t. } Ax = b \ge 0 \quad \text{(LP)} \\ &\qquad\quad x \ge 0 \end{aligned}$$

$$\begin{aligned} &\max \varpi = b^T y \\ &\text{s.t. } \quad A^T y \le c \qquad\qquad \text{(D)} \\ &\qquad\quad y\text{无限制} \end{aligned}$$

$y$ 是(D)的一个可行解

$J = \{j : A_j^T y = c_j\}$

$$\begin{aligned} &\min \xi = \sum_{i=1}^{m} x_i^a \\ &\text{s.t. } \sum_{j \in J} a_{ij} x_j + x_i^a = b_i, i = 1,2,\cdots,m \\ &\qquad x_j \ge 0, \quad j \in J \qquad\qquad \text{(RP)} \\ &\qquad x_j = 0, \quad j \notin J \\ &\qquad x_i^a \ge 0, \; i = 1,2,\cdots,m \end{aligned}$$

$$\begin{aligned} &\max \varpi = b^T y \\ &\text{s.t. } A_j^T y \le 0, \quad j \in J \\ &\qquad y_i \le 1, \quad i = 1,2,\cdots,m \qquad \text{(DRP)} \\ &\qquad y_i\text{无限制} \end{aligned}$$

用单纯形类方法求解(RP)， 若：

- (RP)的最优值 $\xi_{\text{opt}} = 0$，则得到(LP)和(D)的最优解，算法终止；
- (RP)的最优值 $\xi_{\text{opt}} > 0$，设 $\overline{y}$ 是(DRP)的最优解，若 $A_j^T \overline{y} \le 0, j \notin J$，则(D)无上界，从而(LP)不可行，算法终止；
- 否则，取 $\theta \le \theta_1 = \min\limits_{\substack{A_j^T \overline{y} > 0,\\ j \notin J}} \left\{ \dfrac{c_j - A_j^T y}{A_j^T \overline{y}} \right\}$， $y := y + \theta \overline{y}$

### 2. 算法

初始化：求(D)的一个可行解 $y^0$， $k := 0$；

第 1 步：求 $y^k$ 的允许列集合 $J_k = \{j : A_j^T y^k = c_j\}$；

第 2 步：构造(RP)和(DRP)，用单纯形类方法求解(RP)，得到(RP)的最优值 $\xi_{\text{opt}}$

和最有解 $x^k$ 及(DRP)的最优解 $\overline{y}$；

第 3 步：若 $\xi_{\text{opt}} = 0$，则得到(LP)和(D)的最优解 $x^k$ 和 $y^k$，算法终止；

否则，即 $\xi_{\text{opt}} > 0$，转第 4 步；

第 4 步：若 $A_j^T \overline{y} \le 0, j \notin J$，则(D)无上界，从而(LP)不可行，算法终止；

否则转第 5 步；

第 5 步：计算 $\theta_1 = \min\limits_{\substack{A_j^T \overline{y}>0, \\ j \notin J}} \left\{ \dfrac{c_j - A_j^T y}{A_j^T \overline{y}} \right\}$，取 $\theta \le \theta_1$，令 $y^{k+1} := y^k + \theta \overline{y}$，

$k := k+1$ 转第 1 步。

**定理 4.18** 设 $y$ 是(D)的一个可行解，$J = \{j : A_j^T y = c_j\}$ 是 $y$ 对应的允许列集合，$\overline{y}$ 是(DRP)的最优解，且(RP)的最优值 $\xi_{\text{opt}} > 0$：

（1） 若 $A_j^T \overline{y} \le 0, j \notin J$，则(D)无上界，从而(LP)不可行；

（2） 若 $\exists j \notin J$ 使得 $A_j^T \overline{y} > 0$，要维持 $y^* = y + \theta \overline{y}$ 的可行性，$\theta$ 的最大取值为

$$\theta_1 = \min_{\substack{A_j^T \overline{y}>0, \\ j \notin J}} \left\{ \frac{c_j - A_j^T y}{A_j^T \overline{y}} \right\}$$

并且新的费用为 $w^* = b^t y + \theta_1 b^T \overline{y} = w + \theta_1 b^T \overline{y} > w$。

**3. 算法初始可行解的求法**

(1)当 $c \ge 0$ 时，取 $y = 0$ 即可。

(2) )当 $c \ge 0$ 不成立时：在原始问题(LP)中引进变量 $x_{n+1}$ 和增加一个约束：

$$x_1 + x_2 + \cdots + x_n + x_{n+1} = b_{m+1}$$

其中 $b_{m+1}$ 大于(LP)的任意可行解 $x = (x_1, x_2, \cdots, x_n)^T$ 的分量 $x_1, x_2, \cdots, x_n$ 之和，且在目标函数中对应的费用取 $c_{n+1} = 0$：

$$
\begin{aligned}
&\min c^T x + 0 \cdot x_{n+1} \\
&\text{s.t.} \quad a_i^T x = b_i, i = 1,2,\cdots,m \\
&\qquad x_1 + x_2 + \cdots + x_n + x_{n+1} = b_{m+1} \\
&\qquad\quad x_i \ge 0 (i = 1,2,\cdots,n,n+1)
\end{aligned}
\qquad \text{(LP')}
$$

显然(LP')与(LP)具有相同的最有解，而(LP')的对偶为：

$$
\begin{aligned}
&\max \varpi = b^T y + b_{m+1} y_{m+1} \\
&\text{s.t.} \quad A_j^T y + y_{m+1} \le c_j, j = 1,2,\cdots,m \\
&\qquad y_{m+1} \le 0
\end{aligned}
\qquad \text{(D')}
$$

(D')有一个可行解：

$$
\begin{aligned}
&y_i = 0, i = 1,2,\cdots,m \\
&y_{m+1} = \min\{c_j \mid 1 \le j \le m\} < 0
\end{aligned}
$$

**4. 原始－对偶算法的几点说明**

在每一次迭代，都可以由前一次迭代得到的最优解开始求解(RP)，因此这是非常方便的。可以如此做的原因是：每次迭代结束时，既在允许列集合$J$里又在(RP)的最优基里的变量，此时不可能离开$J$。

**定理 4.19** (RP)最优基里每个允许列，在下一次迭代开始时它仍然保持是允许的。

证明：在一次迭代结束时，如果$A_j$是在(RP)的最有基里，那么它的检验数（在(RP)的相对费用）为：

$$\lambda_j = A_j^T \overline{y} = 0$$

所以：$A_j^T y^* = A_j^T y + \theta_1 A_j^T \overline{y} = A_j^T y = c_j$

所以：$j$仍保留在$J$中。

即：不仅从前一次的可行解开始迭代，而且由于不可能产生基的列变为非允许列的麻烦，所以可以用修正的单纯形算法求解。

**定理 4.20** 线性规划的上述原始－对偶算法在有限时间内能够得到(LP)的解。