第三章 线性规划及其原始—对偶算法（II）

3.5 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：

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

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

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

也可以用单纯形方法求得最优解为：$y_1{*} = \frac{3}{2}, y_2{*} = \frac{1}{8}, y_3{*} = 0$，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 3.13 （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：

$$u_i=y_i(a_i^Tx-b_i)=0$$
$$v_j=(c_j-A_j^Ty)x_j=0$$

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$u+v=c^Tx-b^Ty$

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

$\Leftrightarrow u+v=0\Leftrightarrow c^Tx-b^Ty=0\Leftrightarrow x$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**。

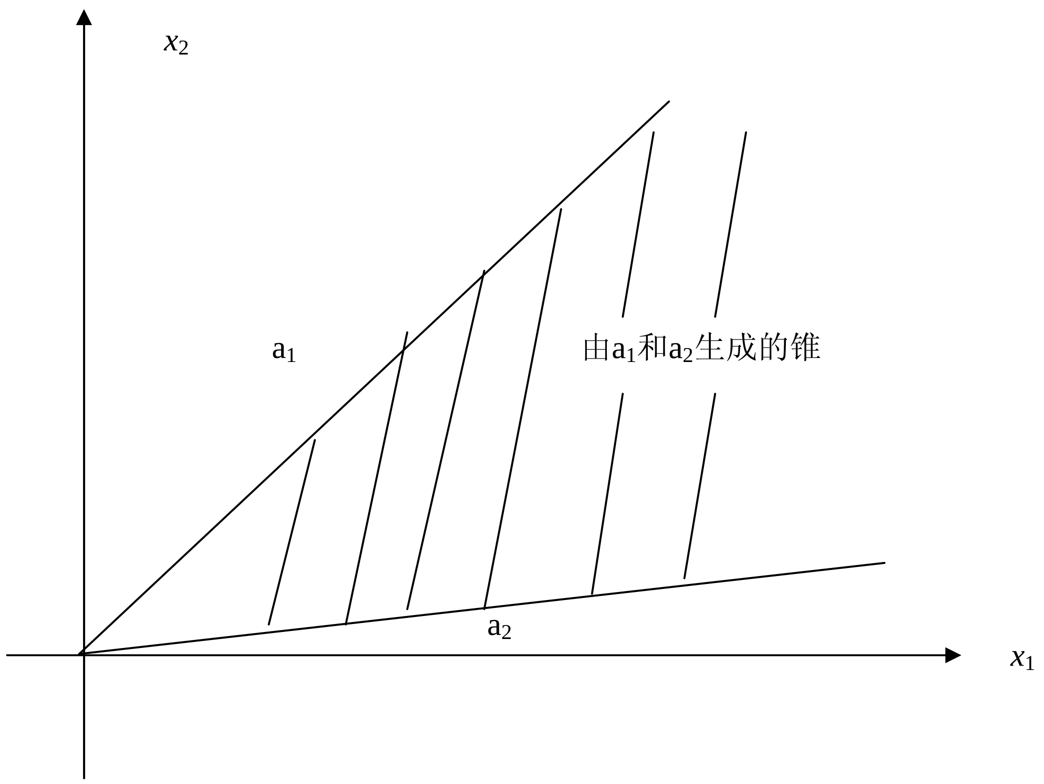
Farkas 引理

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 $C(a_i)$：

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

即 $\{a_i\}$ 非负线性组合。

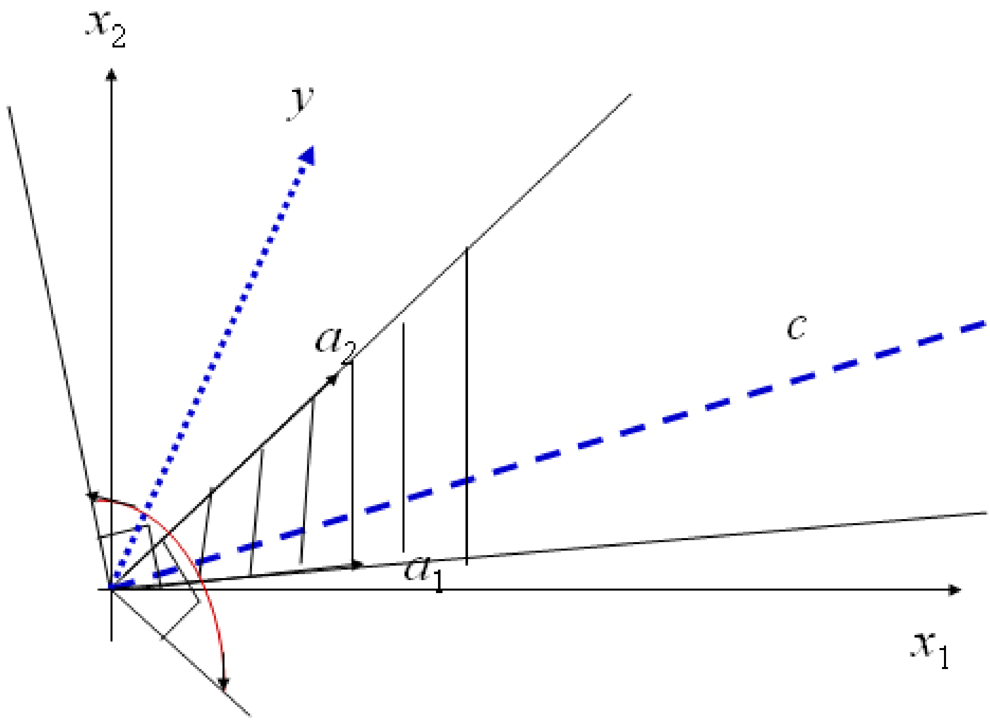


给定向量的一个集合$\{a_i\}_{i=1,2,\cdots,k}$及另外一个向量$c \in R^n$，“如果对一切向量$y \in R^n$，若 y 在$\{a_i\}_{i=1,2,\cdots,k}$有非负投影，那么 y 在 c 上也有非负投影”

Farkas 引理断言

$\Longleftrightarrow$

C 在锥$C(a_i)$里



定理 3.14（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：

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

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$a_i^T y \geq 0, i=1,2,\cdots m$，则必有$c^T y \geq 0 \Leftrightarrow c \in C(a_i)$。事实上,

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

则$c^T y = \sum_{i=1}^{m} y_i a_i^T y \geq 0$。

“$\Rightarrow$” 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 $a_i^T y \geq 0, i=1,2,\cdots m$，则一定有 $c^T y \geq 0 \Leftrightarrow c \in C(a_i)$，要证明必有：$c \in C(a_i)$。事实上，可考察下述线性规划问题：

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

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

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

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

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

所以

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

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

3.6 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：

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

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，这样(LP)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**“适合互补松弛性条件(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，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

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

y 是(D)的一个可行解

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

$$\begin{aligned} \min\ & \xi = \sum_{i=1}^{m} x_i^a \\ \text{s.t. } & \sum_{j \in J} a_{ij} x_j + x_i^a = b_i, i = 1,2,\cdots,m \\ & x_j \geq 0, \quad j \in J \qquad \text{(RP)} \\ & x_j = 0, \quad j \notin J \\ & x_i^a \geq 0,\ i = 1,2,\cdots,m \end{aligned} \Longrightarrow \begin{aligned} \max\ & \varpi = b^T y \\ \text{s.t. } & A_j^T y \leq 0, \quad j \in J \qquad \text{(DRP)} \\ & y_i \leq 1,\ i = 1,2,\cdots,m \\ & y_i\text{无限制} \end{aligned}$$

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 $A_j^T \overline{y} \leq 0, j \notin J$，则(D)无上界，从而(LP)不可行，算法终止；
- 否则，取 $\theta \leq \theta_1 = \min\limits_{\substack{A_j^T \overline{y} > 0, \\ j \notin J}} \left\{ \dfrac{c_j - A_j^T y}{A_j^T \overline{y}} \right\}$， $y := y + \theta \overline{y}$

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

$k := k + 1$ 转第 1 步。

定理 3.15 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，θ 的最大取值为

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

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 $c_{n+1} = 0$：

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

显然(LP')与(LP)具有相同的最有解，而(LP')的对偶为：

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

(D')有一个可行解：

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

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有解开始求解(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，即在 J 里又在(RP)的最优基里的变量，此时不可能离开 J 。

定理 3.16 (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 3.17 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。