18.433 组合最优化

单纯形算法

October 16, 23 *授课教师：Luis Rademacher*

已经证明了引理：

引理 1（Farkas）令$A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$，则下面两个条件有且仅有其一成立：

1. $\exists x \in \mathbb{R}^n \ : \ Ax = b \ and \ x \geq 0.$
2. $\exists y \in \mathbb{R}^m \ : \ b^T y < 0 \ and \ A^T b \geq 0.$

考虑下面标准形式的线性规划及其对偶规划：

$$(P) \qquad \text{opt}(P) = \max\ c^T x \quad \text{s.t.} \quad Ax = b, \quad x \geq 0$$

$$\text{opt}(D) = \min\ b^T y \quad \text{s.t.} \quad A^T y \geq c \qquad (D)$$

我们称一个问题是有界的，如果其最优值是有限数。我们已经证明了弱对偶定理：

定理 2. 如果（P）和（D）有可行解，那么$\text{opt}(P) \leq \text{opt}(D)$且均为有限数。特别地，（P）是有界的。

由此马上可以得出下面的推论：

推论 3. 如果（D）有可行解，那么（P）是有界的或者无解。如果（P）有可行解，那么（D）是有界的或者无解。

1 强对偶性

求解原始及对偶规划时可能出现下面几种情况：有可行解（F），有界（B），无界（U），无解（I），但前面的推论说明，上述几种情况的某些组合是不可能出现的。例如，由推论可知原始问题有可行解，对偶问题有可行解且无界是不可能的。强对偶定理会告诉我们到底哪些组合是可能的，而且，它还将证明在某种情形下原始问题存在有界可行解且$\text{opt}(P) = \text{opt}(D)$, 这是一个最重要的结论。

定理 4（强对偶定理）对一个线性规划（P）和其对偶规划（D），仅有下面组合是可能的：

1、（P）B F（D）B F ，这种情况下 opt（P）= opt（D）。
2、（P）I ，（D）U F 。
3、（P）U F ，（D）I 。
4、（P）I ，（D）I 。

证明. 已知，有 9 种可能的组合。弱对偶性已经排除了下面三种：

- (P) B F, (D) U F,
- (P) U F, (D) B F, and
- (P) U F, (D) U F.

接下来我们要排除(P) B F, (D) I的情况，根据对偶性又可以排除(P) I, (D) B F的情况。仅有定理中给出的 4 种情况可能存在。

假设 (P) B F。令$z > \mathrm{opt}(P)$. 对 $A_0 = \binom{A}{c^T}$, $b_0 = \binom{b}{z}$应用 Farkas 引理。我们知道对所有可行解x，也就是满足$Ax = b$且$x \geq 0$的x，均有$c^T x < z$成立。换句话说，对所有的$x \geq 0$，蕴涵着$A_0 x = \binom{Ax}{c^T x} \neq \binom{b}{z} = b_0$，也就是说 Farkas 引理中的条件（1）是不满足的。从而，引理中的（2）是成立的：即存在 $y \in \mathbb{R}^m$和$\alpha \in \mathbb{R}$满足 $b^T y + z\alpha < 0$ 和$A^T y + \alpha c \geq 0.$

接下来我们要说明$\alpha < 0.$ 否则，设$x^* \in \mathbb{R}^n$是原始问题的最优解，也就是说，x^*是满足$c^T x^* = \mathrm{opt}(P)$ 的一个原始可行点。那么y和α满足的条件就可化为：

$$x^{*T} A^T y + \alpha c^T x^* \geq 0$$

如果$\alpha \geq 0$我们可以得出

$$b^T y + \alpha z \geq 0$$

这就与前面的$b^T y + z\alpha < 0$ 相矛盾。

所以，$\alpha < 0$且$y_0 = -y/\alpha$满足$b^T y_0 < z$和$A^T y_0 \geq c$。也就是说，对偶问题有可行解（且有界）。而且，$z \geq \mathrm{opt}(P)$是任意的，即对任意

$$z > \max_{x \geq 0, Ax = b} c^T x$$

存在 y_0为对偶问题的可行解且

$$\max_{x \geq 0, Ax = b} c^T x \leq \min_{y\,:\,A^T y \geq c} b^T y \leq b^T y_0 < z$$

又因为z是任意的，所以

$$\max_{x \geq 0, Ax = b} c^T x = \min_{y\,:\,A^T y \geq c} b^T y.$$

$\square$

2　线性规划

回顾标准形式的线性规划：

$$\text{maximize } z = \mathbf{c}^t\mathbf{x}, \text{ subject to } \begin{array}{l} \mathbf{A}\mathbf{x} \leq \mathbf{b}, \\ \mathbf{x} \geq 0. \end{array} \tag{1}$$

让我们来看一个具体的实例：

$$\text{maximize } z = 5x_1 + 4x_2 + 3x_3, \text{ subject to } \begin{array}{l} 2x_1 + 3x_2 + x_3 \leq 5, \\ 4x_1 + x_2 + 2x_3 \leq 11, \\ 3x_1 + 4x_2 + 2x_3 \leq 8, \\ x_1, x_2, x_3 \geq 0. \end{array} \tag{2}$$

3　单纯形算法

3.1 引入松弛变量

为了用单纯形法来求解上面的实例，首先引入**松弛变量**，把含有$\leq$的不等式约束转化为等式约束。在每个不等式约束$\sum_{j=1}^{n} a_{ij}x_j \leq b_i$中引入一个松弛变量$x_{n+i}$，使之变为等式约束$\sum_{j=1}^{n} a_{ij}x_j + x_{n+i} = b_i$，并增加约束条件：松弛变量是非负的，即 $x_{n+i} \geq 0$.

用上面的实例来说，引入松弛变量：

$$\text{maximize } z = 5x_1 + 4x_2 + 3x_3, \text{ subject to } \begin{array}{l} 2x_1 + 3x_2 + x_3 + x_4 = 5, \\ 4x_1 + x_2 + 2x_3 + x_5 = 11, \\ 3x_1 + 4x_2 + 2x_3 + x_6 = 8, \\ x_1, x_2, x_3, x_4, x_5, x_6 \geq 0. \end{array} \tag{3}$$

3.2 通过一次转轴操作来增加目标函数值

还是回到上面的实例。把松弛变量移到等式的一边，可得

$$\begin{aligned} x_4 &= 5 - 2x_1 - 3x_2 - x_3 \\ x_5 &= 11 - 4x_1 - x_2 - 2x_3 \\ x_6 &= 8 - 3x_1 - 4x_2 - 2x_3 \end{aligned} \tag{4}$$

这就给出了一个可行解：

$$\begin{aligned} x_1, x_2, x_3 &= 0, \\ x_4 &= 5, \\ x_5 &= 11, \\ x_6 &= 8. \end{aligned} \tag{5}$$

此时目标函数值为：

$$z = 5x_1 + 4x_2 + 3x_3 = 0, \tag{6}$$

这个值看起来似乎有些小。如何才能得到较大的目标函数值呢？或许我们可以尝试一下增加

变量x_1的值。

如果我们增加x_1而保持x_2和x_3为零，由(4)可以计算出相应的x_4, x_5, x_6的值。例如：

$$\begin{array}{llllllll} x_1 = 1, & x_2, x_3 = 0, & \Rightarrow & z = 5, & x_4 = 3, & x_5 = 7, & x_6 = 5, & \\ x_1 = 2, & x_2, x_3 = 0, & \Rightarrow & z = 10, & x_4 = 1, & x_5 = 3, & x_6 = 2, & (7) \\ x_1 = 3, & x_2, x_3 = 0, & \Rightarrow & z = 15, & x_4 = -1, & x_5 = -1, & x_6 = -1. & \end{array}$$

增加x_1可以增加目标函数值，但我们不能过多地增加x_1，否则松弛变量会变为负值，例如上面$x_1 = 3$的情况。

在保持松弛变量非负的前提下，由(4)可以计算出x_1相应的值。

$$\begin{aligned} & x_4 \geq 0 \Rightarrow x_1 \leq \tfrac{5}{2}, \\ & x_5 \geq 0 \Rightarrow x_1 \leq \tfrac{11}{4}, \\ & x_6 \geq 0 \Rightarrow x_1 \leq \tfrac{8}{3}, \end{aligned} \tag{8}$$

因此，x_1最大可以取到$\frac{5}{2}$。

令$x_1 = \frac{5}{2}$，$x_4 = 0$，等式约束(4)可以转化为：

$$\begin{aligned} & x_1 = \tfrac{5}{2} - \tfrac{3}{2}x_2 - \tfrac{1}{2}x_3 - \tfrac{1}{2}x_4 \ , \\ & x_5 = 11 - 4(\tfrac{5}{2} - \tfrac{3}{2}x_2 - \tfrac{1}{2}x_3 - \tfrac{1}{2}x_4) - x_2 - 2x_3, \\ & x_6 = 8 - 3(\tfrac{5}{2} - \tfrac{3}{2}x_2 - \tfrac{1}{2}x_3 - \tfrac{1}{2}x_4) - 4x_2 - 2x_3. \end{aligned} \tag{9}$$

(3)式中的目标函数可以写为$z = \frac{25}{2} - \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{5}{2}x_4$.

这一系列的步骤称为一次**转轴操作**，如果用表的形式表示它，会进一步体会到它的重要性。

3.3 重复

以x_1为主元做转轴操作已经把目标函数值从 0 增加到$\frac{25}{2}$，让我们重复上述一系列运算。增加x_2和x_4的值会减小目标函数值，我们只能增加x_3。

$$\begin{aligned} & x_1 \geq 0 \Rightarrow x_3 \leq 5, \\ & x_5 \geq 0\text{，对 } x_3 \text{ 没有任何限制，} \\ & x_6 \geq 0 \Rightarrow x_3 \leq 1, \end{aligned} \tag{10}$$

因此，x_3最多可以增加到 1。
令$x_3 = 1$，$x_6 = 0$，等式约束(4)可以写为：

$$
\begin{aligned}
x_3 &= 1 + x_2 + 3x_4 - 2x_6, \\
x_5 &= 1 + 5x_2 + 2x_4, \\
x_1 &= 1 - 2x_2 - 2x_4 + x_6,
\end{aligned} \tag{11}
$$

目标函数转化为 $z = 13 - 3x_2 - x_4 - x_6$.

由目标函数的上述形式可知，增加x_2, x_4, x_6中的任何一个都会减小目标函数值。因此，当$x_2, x_4, x_6 = 0$ 时目标函数取得最大值 13。

4 单纯形算法的精确描述

在有 n个变量m个等式约束的线性规划

$$
\max\ c'x
$$

$$
\text{s.t.}\ \ Ax = b
$$

$$
x \geq 0,
$$

中，如果令其中$n - m$个变量为 0 而由m个等式约束求出其它变量的值，如果这些值是非负的，那么这个点是可行的且是一个顶点（证明留作习题），我们称其为一个基本可行解。设为 0 的变量称为非基变量，其余的称为基变量。基变量的集合称为一个基。注意，如果变量的值非 0 则其为基变量，但反之则不一定成立。 在一个基本可行解中，如果有一个基变量取值为 0，则称这个解（基）是退化的。它的几何含义是，每个基都对应一个顶点，但一个顶点可以对应几个基（在退化的情况下）。

若用 B 表示基变量的下标集，N 表示非基变量的下标集，如果 b>=0,则这种形式的规划为标准型。形如：

$$
\max\ c'_N x_N + c'_B x_B
$$

$$
\text{s.t.}\ \ A_N x_N + I x_B = b,
$$

$$
x_N, x_B \geq 0.
$$

单纯形算法从标准型的规划的一个基本可行解出发，选择一个有价值的非基变量，如增该变量的值可增加目标函数的值。接着，在不改变基变量的非负性的前提下，尽量增加该变量的值。该变量成为基变量，（至少）一个原基变量变成 0，一个基变量成为非基变量。这个步骤称为一个转轴，转换将原基下的典式换成新基下的典式。

5 再论单纯形算法（具有更好记号）

5.1 单纯形表记号

如果想在计算机上实现我们上面提到的操作，用矩阵的形式来表示系数会更方便。（3）中的线性规划表示为一个矩阵，或称**单纯形表**，如下：

z	x_1	x_2	x_3	x_4	x_5	x_6	=
	2	3	1	1			5
	4	1	2		1		11
	3	4	2			1	8
-1	5	4	3				0

(12)

5.2 表算法

利用引入的单纯形表记号，上面求解线性规划的算法归纳如下：

1、 观察表的最后一行。找到此行中的一个正元素，确定其所在的列为主元列。

z	x_1	x_2	x_3	x_4	x_5	x_6	=
	2	3	1	1			5
	4	1	2		1		11
	3	4	2			1	8
-1	5	4	3				0

(13)

2、在位于主元列的元素 r 为正的所有行中，选择一行使比值 s/r 最小，其中 s 是该行在最后一列的元素。这就是主元行。

z	x_1	x_2	x_3	x_4	x_5	x_6	=
	2	3	1	1			5
	4	1	2		1		11
	3	4	2			1	8
-1	5	4	3				0

(14)

3、把主元行中的每个元素除以该行在主元列中的元素：

z	x_1	x_2	x_3	x_4	x_5	x_6	=
	1	$\frac{3}{2}$	$\frac{1}{2}$	$\frac{1}{2}$			$\frac{5}{2}$
	4	1	2		1		11
	3	4	2			1	8
-1	5	4	3				0

(15)

4、对其它行，每行减去主元行的倍数使主元列的其它元素为 0：

z	x_1	x_2	x_3	x_4	x_5	x_6	$=$
	1	$\frac{3}{2}$	$\frac{1}{2}$	$\frac{1}{2}$			$\frac{5}{2}$
		-5		-2	1		1
		$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{3}{2}$		1	$\frac{1}{2}$
-1		$-\frac{7}{2}$	$\frac{1}{2}$	$-\frac{5}{2}$			$-\frac{25}{2}$

(16)

5、重复步骤 1 到 4。如果第一步中找不到主元列，则已经达到了最优解。如果第二步中没有正的比值，则此问题无界。

5.3 注释

5.3.1 循环

在第一步中如果运气不好，算法可能会陷入循环（虽然可能性非常小）。用某种前后一致的规则来确定哪个正元素所在列作为主列可以避免循环，比如选择下标最小的正元素（Bland, Robert G., “A combinatorial abstraction of linear programming”, *J. Combinational Theory*, Ser. B, 23(1997), no. 1, 33-57）。

5.3.2 初始可行解

从上面单纯形算法的描述可以看出，算法仅能直接应用于典式 。例如我们的实例中，如果系统开始（或者可以被转化成）形式如下：

$$\begin{aligned} &\max\ c'x \\ &\text{s.t.}\ \ Ax \le b \\ &\qquad x \ge 0, \end{aligned} \tag{17}$$

且 $b \ge 0$，那么引入松弛变量后规划就变成典式。当系统形式如(17)但b有某些非负分量时，只能是非平凡情形。这时，增加一个辅助变量x_0，考虑下面的问题：

$$\begin{aligned} &\min\ x_0 \\ &\text{s.t.}\ \ Ax - x_0 \le b \\ &\qquad x, x_0 \ge 0, \end{aligned} \tag{18}$$

增加松弛变量后，我们可以以x_0为主元，以 b 的最小元素相关的行为主元行进行转轴操作。这就将系统化为典式。 然后我们就可以用单纯形法来解决(18)。求解结束后，如果$x_0 > 0$，那么原始问题(17)无可行解。否则，$x_0 = 0$。在这种情况下，如果x_0是基变量，那我们就再进行一次转换使之变为非基变量。现在$x_0 = 0$且为非基变量，那么我们就不再考虑表中的这一列，并把简化费用问题(17)的费用系数来代替。显然这就是问题(17)的典式，我们就可以用单纯形法来求解。例如：

$$
\begin{aligned}
&\max -2x_1-3x_2\\
\text{s.t.}\quad &-x_1-x_2\le -3\\
&2x_1-x_2\le -2\\
&x_1,x_2\ge 0
\end{aligned}
$$

添加辅助变量x_0和松弛变量x_3和x_4后的初始单纯形表为：

$$
\begin{array}{ccccc|c}
x_0 & x_1 & x_2 & x_3 & x_4 & \\
\hline
-1 & -1 & -1 & 1 & & -3\\
-1 & 2 & -1 & & 1 & -2\\
\hline
-1 & & & & &
\end{array}
$$

把第一行和第一列作为主元行和主元列进行转轴操作以后，可得

$$
\begin{array}{ccccc|c}
x_0 & x_1 & x_2 & x_3 & x_4 & \\
\hline
1 & 1 & 1 & -1 & & 3\\
 & 3 & & -1 & 1 & 1\\
\hline
 & 1 & 1 & -1 & & 3
\end{array}
$$

此表即为典式。

6 强对偶定理的一个证明，基于单纯形算法

强对偶定理如下：

$$
\begin{array}{rcr}
\max\sum_{j=1}^{n} c_j x_j,\ \text{subject to} & & \min\sum_{i=1}^{m} b_i y_i,\ \text{subject to}\\
\sum_{j=1}^{n} a_{ij}x_j \le b_i & = & \sum_{i=1}^{m} a_{ij}y_i \ge c_j\\
x_j \ge 0 & & y_i \ge 0
\end{array}
\tag{19}
$$

前面我们已经用 Farkas 引理证明了它，接下来我们借用单纯形算法的思想给出它的另一个证明。

弱对偶定理证明了

$$
\sum_{j=1}^{n} c_j x_j \le \sum_{i=1}^{m} b_i y_i \tag{20}
$$

其中x_j's和y_i's取任意可行解。

要证明强对偶定理，只需要证明存在x_j's和y_i's使两边加和相等。

给定一个线性规划，

$$\text{maximize } \sum_{j=1}^{n} c_j x_j, \text{ subject to } \quad \begin{array}{l} \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \\ x_j \geq 0, \end{array} \tag{21}$$

设 $x_1^*, ..., x_n^*$ 是使目标函数达最大值的可行解，且令 $z^* = \sum_{j=1}^{n} c_j x_j^*$ 为此最大值。

引入松弛变量，

$$x_{n+i} = b_i - \sum_{j=1}^{n} a_{ij} x_j, \tag{22}$$

把标准形式的线性规划转化成单纯形算法中的等式约束形式。

因为 z^*是目标函数的最大可行值，所以，存在$\bar{c}_k \leq 0$，任意可行解对应的目标函数值可以写为：

$$z = z^* + \sum_{k=1}^{n+m} \bar{c}_k x_k \tag{23}$$

令

$$y_i^* = -\bar{c}_{n+i}. \tag{24}$$

接下来证明$z^* = \sum_{i=1}^{m} b_i y_i^*$ 且$\sum_{i=1}^{m} a_{ij} y_i^* \geq c_j$. 因为$y_i^* \geq 0$，我们将证明对偶规划的一个可行解等于原规划的一个可行解。从 z 开始

$$z = \sum_{j=1}^{n} c_j x_j = z^* + \sum_{k=1}^{m+n} \bar{c}_k x_k. \tag{25}$$

上面的和可以分解为：

$$z = z^* + \sum_{j=1}^{n} \bar{c}_j x_j + \sum_{k=n+1}^{n+m} \bar{c}_k x_k. \tag{26}$$

用(24)替代上和式中的$\bar{c}$'s 可得

$$z = z^* + \sum_{j=1}^{n} \bar{c}_j x_j + \sum_{i=1}^{m} -y_k^* x_{n+i}. \tag{27}$$

用(22)替代上式中的松弛变量可得

$$z = z^* + \sum_{j=1}^{n} \bar{c}_j x_j + \sum_{i=1}^{m} -y_k^* (b_i - \sum_{j=1}^{n} a_{ij} x_j). \tag{28}$$

整理，交换求和顺序可得

$$z = (z^* - \sum_{i=1}^{m} b_i y_i^*) + \sum_{j=1}^{n} (\bar{c}_j + \sum_{i=1}^{m} a_{ij} y_i^*) x_j. \tag{29}$$

注意前面的运算都仅仅是对目标函数的整理，因此对$x \in \mathbb{R}^n$均有效。因此，对$x_j = 0$也成立。故

$$z^* = \sum_{i=1}^{m} b_i y_i^*, \tag{30}$$

成立，这就完成了证明的一半。

为了说明y_i's是一个可行解，把(30)式代入(29)式，并与(25)式比较，可得

$$z = \sum_{j=1}^{n} (\bar{c}_j + \sum_{i=1}^{m} a_{ij} y_i^*) x_j = \sum_{j=1}^{n} c_j x_j. \tag{31}$$

上式成立仅当x_j's的所有系数相等（再次说明，因为这个等式对任意$x \in \mathbb{R}^n$成立）：

$$c_j = \bar{c}_j + \sum_{i=1}^{m} a_{ij} y_i^*. \tag{32}$$

但是，每一个 $\bar{c}_j \le 0$，所以

$$\sum_{i=1}^{m} a_{ij} y_i^* \ge c_{ij}. \tag{33}$$

这就完成了强对偶定理的证明。