MATH 3011 运筹学 **秋季 2023**

Lecture 4: 单纯形法和对偶理论

Lecturer: 陈士祥 *Scribes:* 陈士祥、*ChatGPT*

1 单纯形方法简介

单纯形法（Simplex Method）是解决线性规划问题的一种算法，由美国数学家乔治 · 丹齐格（George Dantzig）于 1947 年发明。它在数学优化和运筹学领域中占有极其重要的地位。以下是单纯形法发展历史的简述：

二战期间，丹齐格在美国空军工作，负责计划和决策问题。他发现现有的线性规划方法无法高效解决这些问题。1947 年，丹齐格发明了单纯形法，这是第一个实用且高效的线性规划求解算法。该方法使用了几何上的单纯形概念，通过从一个顶点到另一个顶点的迭代过程，寻找最优解。

初期应用：单纯形法很快被用于军事和政府项目中，尤其是在军事物资分配和经济计划方面。

计算机时代：随着计算机的发展，单纯形法的计算变得更加快速和可行。在 20 世纪 50 年代和 60 年代，它成为了计算机上执行最频繁的算法之一。

单纯形法是线性规划领域的一个里程碑，对 20 世纪的科学和工程产生了深远的影响。尽管后续出现了一些新的算法，单纯形法仍然是解决实际线性规划问题的一种重要和常用方法。

2 单纯形方法

由于标准形式 LP 的最优解若存在，则必在某个顶点取到。单纯形的**基本思想** 如下：从一个可行基解出发，求下一个使目标函数值有所改善的可行基解；通过不断迭代改进可行基解力图达到最优可行基解。单纯形法的**主要步骤：**

- 最优判定 (optimality)
- 转轴运算 (pivoting)

2.1 最优判定

首先假定对线性规划标准形问题

$$\{\min \mathbf{c}^T\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$$

已得到一个可行基的划分 $A=(B,N)$, $\mathbf{x}=\begin{pmatrix}\mathbf{x}_B\\ \mathbf{x}_N\end{pmatrix}$, $\mathbf{c}=\begin{pmatrix}\mathbf{c}_B\\ \mathbf{c}_N\end{pmatrix}$.

从可行基解 $\mathbf{x}=\begin{pmatrix}\mathbf{x}_B\\ 0\end{pmatrix}$ 开始，考虑沿着方向 $d=\begin{pmatrix}\mathbf{d}_B\\ d_N\end{pmatrix}$ 移动到 $x+\theta d, \theta>0$. 其中，$j\in N$, $d_j=1, d_i=0, i\in N, i\neq j$, d_B 待确定. 即做如下移动：对于非基变量，

$$\tilde{x}_j=x_j+\theta, \theta>0.$$

$$\tilde{x}_i=0, i\neq j, i\in N.$$

移动后的点，首先需要保证 $A(x+\theta d)=b$。

那么由 $A(x+\theta d)=b$, $Ax=b$，我们有

$$0=Ad=Bd_B+A_j,$$

所以

$$d_B=-B^{-1}A_j.$$

A_j 表示矩阵 A 的第 j 列向量。在点 $x+\theta d$ 的目标函数值可表示为

$$\begin{aligned} z &= \mathbf{c}_B^T(\mathbf{x}_B+\theta d_B)+\theta\mathbf{c}_N^T d_N \\ &= \mathbf{c}_B^T(\mathbf{x}_B-\theta B^{-1}A_j)+\theta\mathbf{c}_j \end{aligned} \tag{4.1}$$

所以，目标值相对于原来 $c^\top x$ 变化量为

$$z-c^\top x=\theta(\mathbf{c}_j-c_B^\top B^{-1}A_j).$$

我们将

$$\bar{c}_j=\mathbf{c}_j-c_B^\top B^{-1}A_j \tag{4.2}$$

称为非基方向 x_j 的减少量。

Theorem 4.1 (最优判定条件) 标准形式线性规划问题中，某个可行基解为 $\mathbf{x}=\begin{pmatrix}B^{-1}b\\ 0\end{pmatrix}$，若由式(4.2)所决定的 $\bar{c}$ 满足

$$\bar{c}_j\geq 0, \forall j\in N, \tag{4.3}$$

那么 x 是一个最优解.

Proof: 记任意可行解 y（满足 $y_j\geq 0, j\in N$），令 $d=y-x$. 那么 $Ad=0$, 可知 d 满足

$$d_B=-B^{-1}Nd_N.$$

进一步，计算

$$c^\top y-c^\top x=c^\top d=\sum_{j\in N}(c_j-c_B^\top B^{-1}A_j)d_j=\sum_{j\in N}\bar{c}_j d_j.$$

由于 y 可行，故 $d_j\geq 0, j\in N$. 因此，$c^\top x$ 为最小值。 ■

2.2 转轴运算

上述变换过程 $x+\theta d$，并未考虑 $x+\theta d$ 的关于约束 $x \geq 0$ 的可行性。也就是说，这里需要考虑是否存在合适的 $\theta > 0$ 使得 $x+\theta d \geq 0$.

Definition 4.1 (可行方向) *令 $x \in P$ 是一个可行点，我们称 d 是 x 的可行方向，如果存在正数 $\theta > 0$，使得 $x+\theta d \in P$.*

Definition 4.2 (退化基解) *基解 x 称为退化基解，如果 x 有大于 n 个积极约束。*

对于标准形式的多面集，基解 x 总有 $n-m$ 个坐标为 0。如果 x 有大于 n 个主动约束，那么有大于 $n-m$ 个主动约束是 $x_i = 0$, 也就是有大于 $n-m$ 个分量是 0。反之，如果 x 是一个基解，并且 x 有大于 $n-m$ 个分量是 0，那么 x 处有大于 n 个主动约束（因为等式约束有 m 个）。所以，对于标准形式，我们有如下定义:

Definition 4.3 (退化基解) *对于标准形式的多面集，基解 x 称为退化基解，如果 x 有大于 $n-m$ 个分量是 0。*

假设 x 是一个可行基解，对 x 做上述移动：$x+\theta d$:

1. 如果 x 非退化，那么 $x_B > 0$, 存在充分小 $\theta > 0$，使得 $x_B + \theta d_B \geq 0$ 成立；
2. 如果 x 退化，因为有某个 $x_i = 0, i \in B$, 故方向 d 可能非可行方向。

Example 4.1 *考虑 $P = \{x = (x_1, \ldots, x_7) | Ax = b, x \geq 0\}$，其中*

$$A = \begin{pmatrix} 1 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 6 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 8 \\ 12 \\ 4 \\ 6 \end{pmatrix},$$

考虑由线性独立列 A_1, A_2, A_3, A_7 组成的基。为了计算相应的基解，我们首先将非基变量 x_4, x_5, 和 x_6 设为零，然后解系统 $Ax = b$ 以获得其余变量，得到 $x = (4,0,2,0,0,0,6)$。这是一个退化的基可行解，因为我们有四个变量为零，而 $n-m = 3$. 我们也可以选线性独立列 A_1, A_3, A_4, A_7 组成基，同样得到 $x = (4,0,2,0,0,0,6)$.

我们令 $B = (A_1, A_2, A_3, A_7)$. 对于非基变量 x_4, $d_B = -B^{-1}A_4 = (0, -1.5, 0.25, 1.5)$, 所以 $d = (0, -1.5, 0.25, 1, 0, 0, 1.5)$, 不是可行方向。
对于非基变量 x_5, 我们有 $d_B = -B^{-1}A_5 = (0, 0.5, -0.25, -0.5)$, $d = (0, 0.5, -0.25, 0, 1, 0, -0.5)$。只要

$\theta \leq 8$, $x + \theta d$ 可行。目标值的减少量为 $c5 - c_B^\top B^{-1} A_5 = c_5 + c_B^\top d_B$。最后，$x + 8d = (4, 4, 0, 0, 8, 0, 2)$ 也是一个可行基解。

由上例可知，对于非退化可行基解，我们可以适当的选择非基变量 x_j 和 $\theta > 0$, 到达新的可行基解。单纯形法中的转轴运算便是这种想法。

我们先假设任意可行基解都非退化，后面我们会看到，单纯形法可以巧妙的处理退化解。

回顾之前的做法，为了保证解的可行性，有必要进一步分析 θ 的取值及基变量 $\mathbf{x}_B$ 的变化情况。记 $\tilde{\mathbf{x}} = \mathbf{x} + \theta d$. 若取 $x_j = \theta$, $x_i = 0 (i, j \in N, i \neq j)$, 则由线性方程组 $A\mathbf{x} = \mathbf{b}$ 得 $\tilde{\mathbf{x}}_B$ 的值为

$$\tilde{\mathbf{x}}_B = x_B - \theta B^{-1} \mathbf{A}_j.$$

记 $d_B = -B^{-1}\mathbf{A}_j$, 上式可写成

$$\begin{pmatrix} \tilde{x}_{B_1} \\ \tilde{x}_{B_2} \\ \vdots \\ \tilde{x}_{B_m} \end{pmatrix} = \begin{pmatrix} x_{B_1} \\ x_{B_2} \\ \vdots \\ x_{B_m} \end{pmatrix} + \theta \begin{pmatrix} d_{B_1} \\ d_{B_2} \\ \vdots \\ d_{B_m} \end{pmatrix}. \tag{4.4}$$

1. 如果 $d_B \geq 0$, $\theta = +\infty$. 这时若 $\bar{c}_j < 0$，线性规划的最优值为 $-\infty$
2. 如果存在 $i \in \{1, 2, \ldots, m\}$, $d_{B_i} < 0$. θ 必须满足 $\theta \leq -x_{B_i}/d_{B_i}$. 我们取最大可能的 θ, 记作

   $$\theta^* = \min_{\{i | d_{B_i} < 0\}} (-\frac{x_{B_i}}{d_{B_i}}). \tag{4.5}$$

   得到新的点

   $$\tilde{\mathbf{x}} = \mathbf{x} + \theta^* d$$

   记 r 是取到(4.5)中最小值的下标 (可能在多个下标取到最小值，可以任取一个下标. 这种情况下，新得到的点是退化的。)，那么 $\tilde{\mathbf{x}}_{B_r} = 0$.

   注意到，$\tilde{\mathbf{x}}_j = \theta^* > 0$, $\tilde{\mathbf{x}}_i = 0, i \in N, i \neq j$.

至此，x_{B_r} 和 $x_i (i \in N, i \neq j)$ 合起来不少于 $n-m$ 个变量取值为 0; 取值为正的只能是 $\{x_{B_1}, \cdots, x_{B_{r-1}}, x_j, x_{B_{r+1}}, \cdots, x_{B_m}\}$. 把这 m 个变量看成新的基变量，另外 $n - m$ 个看成非基变量，就得到了新的 “划分”。

简单来说，我们将 A_{B_r} 从 B 中踢出，让 A_j 进入，得到新的基矩阵 $\tilde{B}$.

于是，我们获得了新的可行解

$$\tilde{\mathbf{x}} = (x_{B_1}, \cdots, x_{B_{r-1}}, 0, x_{B_{r+1}}, \cdots, x_{B_m}, 0, \cdots, x_j, \cdots, 0)^T.$$

Theorem 4.2 上述新得到的 $\tilde{\mathbf{x}}$ 是可行基解, 即 $\tilde{B} = (A_{B_1}, \cdots, A_{B_{r-1}}, A_j, A_{B_{r+1}}, \cdots, A_{B_m})$ 的秩为 m.

Proof: 因为 $B^{-1}A_j = -d_B$, d_B 的第 r 个分量非零。并且 $B^{-1}A_{B_i}, i \neq r$ 是单位矩阵 I_m 的第 i 列。所以他们线性无关。故 $A_{B_1}, \cdots, A_{B_{r-1}}, A_j, A_{B_{r+1}}, \cdots, A_{B_m}$ 线性无关。 ■

以上操作是通过替换基变量 x_{B_r} 与非基变量 x_j 来确定新的可行基解，我们称之为**转轴运算** (pivoting).

另外，称这时的 x_{B_r} 为出基变量，x_j 为入基变量。这样得到的新可行基解，利用式(4.3)进行最优判定，若未达到最优解，则重复上述操作。

单纯形算法 (SIMPLEX)

1. 第一步（初始化）:

   确定一个可行基划分 $A = (B, N)$, 并计算 $\mathbf{x}_B = \bar{\mathbf{b}} := B^{-1}\mathbf{b}$.

2. 第二步（最优判定）:

   计算向量 $\mathbf{w} = (B^T)^{-1}\mathbf{c}_B$, 对所有的非基分量 $j \in N$ 求出 $z_j = \mathbf{w}^T\mathbf{A}_j$.

   IF $\bar{c}_j = c_j - z_j \geq 0 (j \in N)$, 则当前的可行基解 $\mathbf{x} = \begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{b}} \\ \mathbf{0} \end{pmatrix}$ 已是最优解 [**stop!**]。

   ELSE 选取一个满足 $c_j - z_j < 0$ 的 $j \in N$ 进入下一步。(**注：** 选取 j 的方式有多种。如：下标从小到大计算 $\bar{c}_j$，遇到小于 0 时直接选取。或者，计算出所有的情况，取 $\theta_j^*\bar{c}_j$ 中最小的那个。但是，后者在实际中并未提升求解效率。)

3. 第三步（转轴运算）:

   计算向量 $\mathbf{d} = -B^{-1}\mathbf{A}_j$.

   IF $\mathbf{d} \geq \mathbf{0}$, 则问题无有界解 [**stop!!**]。

   ELSE

   找出 r 使得 $\theta^* = \min_{\{i|d_{B_i}<0\}}(-\frac{x_{B_i}}{d_{B_i}})$, 将基矩阵 B 的列向量 $\mathbf{A}_{B_r}$ 用 $\mathbf{A}_j$ 替换，得到新的基矩阵

   $$B = (\mathbf{A}_{B_1}, \cdots, \mathbf{A}_{B_{r-1}}, \mathbf{A}_j, \mathbf{A}_{B_{r+1}}, \cdots, \mathbf{A}_{B_m}).$$

   进而，记新基变量的值为

   $$x_{B_i} = x_{B_i} + \theta^* d_i (i = 1, \cdots, m, i \neq r), x_j = \theta^*.$$

   令基变量的下标集合与非基变量的下标集合分别为

   $$B := (B \cup \{j\}) - \{B_r\}, \quad N := (N \cup \{B_r\}) - \{j\}.$$

   回到第二步。

2.3 单纯形法的收敛结论

如前所述，经过转轴运算，对应于 $\bar{c}_j < 0$ 的非基变量 x_j 入基得到新的可行基解时，目标函数值将减少 $\theta^*|\bar{c}_j|$.

对于非退化情况，每次转轴后，目标值严格减小，所以每个可行基解不会被重复到达 2 次及以上。另外，可行基解的个数是有限的，迭代不会无限重复下去，必在有限次迭代后结束计算。

Theorem 4.3（单纯形法的有限收敛性定理（非退化情形）） *若所给的线性规划问题可以求出初始可行基解，而且转轴运算过程中的所有可行基解都是非退化的，则利用单纯形法在有限次迭代后，要么找到最优解，要么识别出问题无有界解，从而结束计算。*

3 单纯形表

在上述算法中，如果我们每一步直接计算新的基矩阵的逆矩阵 B^{-1}, 浮点数乘法复杂度为 $\mathcal{O}(m^3)$. 其余计算为矩阵向量乘法，例如 $B^{-1}A_j, C_B^TB^{-1}$，复杂度为 $\mathcal{O}(mn+m^2)$. 故总体复杂度为 $\mathcal{O}(mn+m^3)$.

实际上，如果上一步的基矩阵的逆 B^{-1} 已经存在，那么计算新的基矩阵 $\tilde{B}$ 的逆，仅需要 $\mathcal{O}(m^2)$. 这是因为，

$$B^{-1}\tilde{B} = [e_1, e_2, \ldots, e_{r-1}, u_r, e_{r+1}, \ldots, e_m],$$

这里的 e_i 是 $m \times m$ 单位矩阵 I_m 的第 i 列，$u_r = -d_B = B^{-1}A_j$.

因为 $\tilde{B}^{-1}\tilde{B} = I_m$, 我们只需要对 $B^{-1}\tilde{B}$ 作行变换，使得第 r 列变为 e_r 即可。这需要先将 $B^{-1}\tilde{B}$ 的第 r 行乘以 $-\frac{u_i}{u_r}$ 加到第 i 行，$i \neq r, i = 1, ..., m$. 最后让第 r 行除以 u_r.

我们可以将更新 $B^{-1}\tilde{B}$ 写入下面的单纯形表格中，该表格共有 $m+1$ 行和 $n+1$ 列：

$-\mathbf{c}_B'\mathbf{B}^{-1}\mathbf{b}$	$\mathbf{c}' - \mathbf{c}_B'\mathbf{B}^{-1}\mathbf{A}$
$\mathbf{B}^{-1}\mathbf{b}$	$\mathbf{B}^{-1}\mathbf{A}$

或者 （最上面这行叫做第 0 行）

$-\mathbf{c}_B'\mathbf{x}_B$	$\bar{c}_1$	$\ldots$	$\bar{c}_n$
$x_{B(1)}$	$\vert$		$\vert$
$\vdots$	$\mathbf{B}^{-1}\mathbf{A}_1$	$\ldots$	$\mathbf{B}^{-1}\mathbf{A}_n$
$x_{B(m)}$	$\vert$		$\vert$

单纯形表更新第 1 到 m 行的过程，就是更新逆矩阵 B^{-1} 的过程，或者说是对单纯性表作行变换。

- 注意，若 i 是可行基解的某个下标，那么 $\bar{c}_i = c_i - c_B^TB^{-1}A_i = c_i - c_i = 0$. 对于第 0 行，若 j 是新的入基坐标，我们用第 r 行乘以某个数加到第 0 行，使得 $\bar{c}_j = 0$.

- 对于非出基坐标 r 对应的行，需要先将 $B^{-1}\tilde{B}$ 的第 r 行乘以 $-\frac{u_i}{u_r}$ 加到第 i 行，使其变为 0，$i \neq r, i = 1, ..., m$. 最后让第 r 行除以 u_r，使其变为 1.

Example 4.2 考虑问题

$$\begin{aligned} \min \quad & -10x_1 - 12x_2 - 12x_3 \\ \textit{subject to} \quad & x_1 + 2x_2 + 2x_3 \leq 20 \\ & 2x_1 + x_2 + 2x_3 \leq 20 \\ & 2x_1 + 2x_2 + x_3 \leq 20 \\ & x_1, x_2, x_3 \geq 0. \end{aligned}$$

加入松弛变量，我们得到如下问题：

$$\begin{aligned} \min \quad & -10x_1 - 12x_2 - 12x_3 \\ \textit{subject to} \quad & x_1 + 2x_2 + 2x_3 + x_4 \quad = 20 \\ & 2x_1 + x_2 + 2x_3 + \quad x_5 \quad = 20 \\ & 2x_1 + 2x_2 + x_3 + x_6 \quad = 20 \\ & x_1, \ldots, x_6 \geq 0. \end{aligned}$$

$\mathbf{x} = (0, 0, 0, 20, 20, 20)$ 是一个可行基解. 所以，$B(1) = 4, B(2) = 5$, *and* $B(3) = 6$. 基矩阵为单位阵 $\mathbf{I}$. 我们有 $\mathbf{c}_B = \mathbf{0}$, $\mathbf{c}'_B\mathbf{x}_B = 0$ 以及 $\bar{\mathbf{c}} = \mathbf{c}$.

		x_1	x_2	x_3	x_4	x_5	x_6
	0	-10	-12	-12	0	0	0
$x_4=$	20	1	2	2	1	0	0
$x_5=$	20	2 *	1	2	0	1	0
$x_6=$	20	2	2	1	0	0	1

表 4.1: 初始单纯形表。由于第 0 行中，-10 对应的 x_1 表示可以转轴到 x_1 减小函数目标值。因为 20/2 最小，故将 2 标 * 号。入基变量 x_1, 出基 x_5. 将第 2 行乘以 5 加到第 0 行，将第 2 行乘以 0.5 减到第一行，再用第 3 行减去第 2 行，最后将第二行除以 2. 得到下一张表。

		x_1	x_2	x_3	x_4	x_5	x_6
	100	0	-7	-2	0	5	0
x_4	10	0	1.5	1*	1	-0.5	0
$x_1=$	10	1	0.5	1	0	0.5	0
$x_6=$	0	0	1	-1	0	-1	1

表 4.2: 入基变量 x_3, 出基 x_4

		x_1	x_2	x_3	x_4	x_5	x_6
	120	0	-4	0	2	4	0
$x_3=$	10	0	1.5	1	1	-0.5	0
$x_1=$	0	1	-1	0	-1	1	0
$x_6=$	10	0	2.5*	0	1	-1.5	1

表 4.3: 入基变量 x_2, 出基 x_6

		x_1	x_2	x_3	x_4	x_5	x_6
	136	0	0	0	3.6	1.6	1.6
$x_3=$	4	0	0	1	0.4	0.4	-0.6
$x_1=$	4	1	0	0	-0.6	0.4	0.4
$x_2=$	4	0	1	0	0.4	-0.6	0.4

表 4.4: 第 0 行全为非负，最优解得到, 最小值为-136

3.1 单纯性表：退化基解的循环

当存在退化解时，下面这个例子说明单纯形法可能陷入循环。

Example 4.3 假设我们有如下初始单纯形表：

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-3/4$	20	$-1/2$	6	0	0	0
$x_5=$	0	$1/4$*	-8	-1	9	1	0	0
$x_6=$	0	$1/2$	-12	$-1/2$	3	0	1	0
$x_7=$	1	0	0	1	0	0	0	1

采用如下规则选取出基入基下标：

1. 对于非基向量，选取下标 j 使得减少量 $\bar{c}_j$ 最小
2. 对于基向量，出基下标，选取所有满足条件中下标最小的。

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	0	−4	−7/2	33	3	0	0
$x_1 =$	0	1	−32	−4	36	4	0	0
$x_6 =$	0	0	4*	3/2	−15	−2	1	0
$x_7 =$	1	0	0	1	0	0	0	1

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	0	0	−2	18	1	1	0
$x_1 =$	0	1	0	8*	−84	−12	8	0
$x_2 =$	0	0	1	3/8	−15/4	−1/2	1/4	0
$x_7 =$	1	0	0	1	0	0	0	1

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	1/4	0	0	−3	−2	3	0
$x_3 =$	0	1/8	0	1	−21/2	−3/2	1	0
$x_2 =$	0	−3/64	1	0	3/16*	1/16	−1/8	0
$x_7 =$	1	−1/8	0	0	21/2	3/2	−1	1

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	−1/2	16	0	0	−1	1	0
$x_3 =$	0	−5/2	56	1	0	2*	−6	0
$x_4 =$	0	−1/4	16/3	0	1	1/3	−2/3	0
$x_7 =$	1	5/2	−56	0	0	−2	6	1

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	−7/4	44	1/2	0	0	−2	0
$x_5 =$	0	−5/4	28	1/2	0	1	−3	0
$x_4 =$	0	1/6	−4	−1/6	1	0	1/3*	0
$x_7 =$	1	0	0	1	0	0	0	1

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	−3/4	20	−1/2	6	0	0	0
$x_5 =$	0	1/4	−8	−1	9	1	0	0
$x_6 =$	0	1/2	−12	−1/2	3	0	1	0
$x_7 =$	1	0	0	1	0	0	0	1

图 4.1: 表 1-6，回到了原点。目标值一直未改变。

如何处理存在退化可行基解情况 若遇到退化基解，可采取 Bland's rule[最小下标转轴规则] 避免循环:

1. 若非基向量中，存在多个 $\bar{c}_i < 0, i \in N$. 选取 j 为最小的下标。

2. 出基过程中，即在计算下标 r 使得 $\theta^* = \min_{\{i|d_{B_i}<0\}}(-\frac{x_{B_i}}{d_{B_i}})$ 时，选取最小的下标.

感兴趣，可参考：Bland, Robert G. (May 1977). "New finite pivoting rules for the simplex method". Mathematics of Operations Research. 2 (2): 103–107. doi:10.1287/moor.2.2.103. JSTOR 3689647. MR 0459599.

作业 4.1 用单纯形表求解如下问题：

$$
\begin{aligned}
\min \quad & -4x_1 - x_2 \\
\text{s.t.} \quad & -x_1 + 2x_2 \leq 4 \\
& 2x_1 + 3x_2 \leq 12 \\
& x_1 - x_2 \leq 3 \\
& x_1, x_2 \geq 0.
\end{aligned} \tag{4.6}
$$

3.2 单纯性表初始化

初始可行基解： 前述单纯形法需要一个可行基解作为初始解。

一般问题中为给出初始可行基解，我们在这里介绍两种常用的初始化方法：

- 两阶段法

- 大 M 法

3.3 两阶段法

两阶段法我们只做简单的介绍。

两阶段法： 第一阶段求一个辅助问题的最优解；第二阶段再以这个最优解，为原问题的初始可行基解。

在**两阶段法**的第一阶段里，针对线性规划问题

$$\{\min \mathbf{c}^T\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\},$$

考虑如下的辅助问题

$$\begin{aligned} \min \quad & \mathbf{1}^T\mathbf{y} \\ \text{s.t.} \quad & A\mathbf{x}+\mathbf{y}=\mathbf{b} \\ & \mathbf{x}\geq\mathbf{0},\mathbf{y}\geq\mathbf{0}. \end{aligned} \tag{辅助问题}$$

这里，$\mathbf{y}$ 为新引进的 m 维人工变量向量，$\mathbf{1}$ 是所有分量均为 1 的常数向量.
只需要对等式进行符号处理，我们不妨可以假设 $\mathbf{b}\geq\mathbf{0}$.

在 (辅助问题)中，$\begin{pmatrix}\mathbf{x}\\\mathbf{y}\end{pmatrix}=\begin{pmatrix}\mathbf{0}\\\mathbf{b}\end{pmatrix}$ 是可行基解。把该可行基解当作初始解，就可直接用单纯形法进行求解。

(辅助问题)的目标函数 $\mathbf{1}^T\mathbf{y}$ 在可行域里取值是非负的。如果原问题存在可行解 $\mathbf{x}$，那么 $\begin{pmatrix}\mathbf{x}\\\mathbf{0}\end{pmatrix}$ 满足(辅助问题)的约束条件，且目标函数值为 0.

- 当辅助问题的最优解 $\begin{pmatrix}\bar{\mathbf{x}}\\\bar{\mathbf{y}}\end{pmatrix}$ 成立 $\mathbf{1}^T\bar{\mathbf{y}}>0$ 时，原问题没有可行解。
- 反之，如果辅助问题的最优解为 $\begin{pmatrix}\bar{\mathbf{x}}\\0\end{pmatrix}$. 此时，$\bar{\mathbf{x}}$ 显然是原问题的一个可行解。

1. 如果辅助问题的基矩阵 B 的列全部属于 A, $\bar{\mathbf{x}}$ 是可行基解，done!
2. 否则，由于辅助问题的基矩阵 B 的某些列不在 A 中，由于 A 的秩为 m，$\begin{pmatrix}\bar{\mathbf{x}}\\0\end{pmatrix}$ 是辅助问题的退化解，$\bar{\mathbf{x}}$ 中必定有大于 $n-m$ 个坐标为 0. 我们只需要在 A 的列中，找出一些与已有的基列线性无关的列即可。

3.4 大 M 法

先求出初始可行基解后再进行第二阶段最优解的计算合并成一步作业的方法，称为**大 M 法**。

大 M 法考虑如下线性规划问题

$$\begin{aligned} \min \quad & \mathbf{c}^T\mathbf{x}+M\cdot\mathbf{1}^T\mathbf{y} \\ \text{s.t.} \quad & A\mathbf{x}+\mathbf{y}=\mathbf{b} \\ & \mathbf{x}\geq\mathbf{0},\mathbf{y}\geq\mathbf{0}. \end{aligned} \tag{4.7}$$

这里 $M>0$ 为充分大的常数，我们也假设 $\mathbf{b}\geq\mathbf{0}$, $\mathbf{y}$ 和常数向量 $\mathbf{1}$ 与两阶段法中的一样。故，我们直接得到了一个可行基解

$$\begin{pmatrix}\mathbf{x}\\\mathbf{y}\end{pmatrix}=\begin{pmatrix}0\\\mathbf{b}\end{pmatrix}.$$

在 大 M 法的问题(4.7)中，目标函数项 $M \cdot \mathbf{1}^T\mathbf{y}$ 占支配地位，所以单纯形法迭代将首先让人工变量 y_i 的值减小。其结果是，若所有人工变量都成为 0, 则目标函数 $\mathbf{c}^T\mathbf{x} + M \cdot \mathbf{1}^T\mathbf{y}$ 和原问题的目标函数 $\mathbf{c}^T\mathbf{x}$ 本质上一致，可认为在此之后单纯形法的迭代所进行的就是求原问题最优解的计算。

事实上，如果原问题可行且最优值有限，在 M 充分大时，得到大 M 法问题(4.7)的最优解 $\begin{pmatrix} \bar{\mathbf{x}} \\ \bar{\mathbf{y}} \end{pmatrix}$ 中 $\bar{\mathbf{y}} = \mathbf{0}$，且 $\bar{\mathbf{x}}$ 成为原问题的最优解。实际计算中，我们可以保留 M，默认其是一个大数。进行正常的单纯形表更新。下面的例子进一步说明大 M 单纯形法是如何工作的。

Example 4.4 用大 M 法求解如下问题：

$$
\begin{aligned}
\min \quad & x_1 + x_2 - 3x_3 \\
\text{s.t.} \quad & x_1 - 2x_2 + x_3 \le 11 \\
& 2x_1 + x_2 - 4x_3 \ge 3 \\
& x_1 - 2x_3 = 1 \\
& x_1, x_2, x_3 \ge 0.
\end{aligned} \tag{4.8}
$$

标准化模型（大 M 法），先添加松弛变量 x_4 和 x_5 标准化，将不等式变为等式。由于 x_4 已经对应了形如 $(1,0,0)^T$ 的基底，故第一个等式约束不用再添加大 M 变量，只对第 2，3 个等式约束添加 x_6 和 x_7 引入大 M 变量：

$$
\begin{aligned}
\min \quad & x_1 + x_2 - 3x_3 + M(x_6 + x_7) \\
\text{s.t.} \quad & x_1 - 2x_2 + x_3 + x_4 = 11 \\
& 2x_1 + x_2 - 4x_3 - x_5 + x_6 = 3 \\
& x_1 - 2x_3 + x_7 = 1 \\
& x_1, x_2, x_3, x_4, x_5, x_6, x_7 \ge 0.
\end{aligned} \tag{4.9}
$$

表 4.5: 迭代单纯形表-1

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	$-4M$	$1-3M$	$1-M$	$-3+6M$	0	M	0	0
x_4	11	1	-2	1	1	0	0	0
x_6	3	2	1	-4	0	-1	1	0
x_7	1	1*	0	-2	0	0	0	1

表 4.6: 迭代单纯形表-2

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	$-M-1$	0	$1-M$	-1	0	M	0	3M-1
x_4	10	0	-2	3	1	0	0	-1
x_6	1	0	1*	0	0	-1	1	-2
x_1	1	1	0	-2	0	0	0	1

表 4.7: 迭代单纯形表-3

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	-2	0	0	-1	0	1	0	M+1
x_4	12	0	0	3*	1	-2	2	-5
x_2	1	0	1	0	0	-1	1	-2
x_1	1	1	0	-2	0	0	0	1

表 4.8: 迭代单纯形表-4

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	2	0	0	0	1/3	1/3	2/3	M-2/3
x_3	4	0	0	1	1/3	-2/3	2/3	-5/3
x_2	1	0	1	0	0	-1	1	-2
x_1	9	1	0	0	2/3	-4/3	4/3	-7/3

4 总结

- 由多面集的几何性质，我们得到了线性规划标准形式的极点必存在。
- 线性规划的标准形式，如果最优解有限，必在某个可行基解处取到。
- 由大 M 法，我们可以通过单纯形表求解线性规划问题。并且，每个迭代的浮点数乘法复杂度为 $\mathcal{O}(m^2+mn)$.

补充: 对于很多问题, 通常只需要 $\mathcal{O}(m)$ 次转轴运算即可得到最优解。但是, 最差的情况需要 2^n (可参考 Chapter 3.7. in Book: Bertsimas, Dimitris, and John N. Tsitsiklis. Introduction to linear optimization.

Vol. 6. Belmont, MA: Athena scientific, 1997.)。实际实现算法时，要考虑算法的稳定性（如舍入误差），以及结合数据的稀疏性（如稀疏矩阵分解求逆）进一步加速算法。可以参考：Chapter 3.3. in Book: Bertsimas, Dimitris, and John N. Tsitsiklis. Introduction to linear optimization.）

5 对偶理论

针对任意一个最优化问题，可以定义它的一个对偶问题 (Dual Problem) 。

对偶理论将揭示原问题与对偶问题之间的内在联系，为进一步深入研究线性规划的求解算法提供理论依据。

原问题：

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

定义 Lagrange 函数 $L(x, \lambda, \mu) = c^\top x + \lambda^\top (b - Ax) - \mu^\top x, \mu \geq 0, \lambda \in \mathbb{R}^m$. 对于无约束问题

$$\min_{x \in \mathbb{R}^n} L(x, \lambda, \mu),$$

我们记 $g(\lambda, \mu) = \min_{x \in \mathbb{R}^n} L(x, \lambda, \mu)$.

对于任意原问题的可行解 $\bar{x}$，我们有

$$g(\lambda, \mu) \leq c^\top \bar{x} + \lambda^\top (b - A\bar{x}) - \mu^\top x \leq c^\top \bar{x}.$$

故，问题

$$\max_{\lambda, \mu} g(\lambda, \mu)$$

可以看作寻找原问题的最紧下界。

$$\begin{aligned} \min_{x \in \mathbb{R}^n} L(x, \lambda, \mu) &= \min_x (c - A^\top \lambda - \mu)^\top x + \lambda^\top b \\ &= \begin{cases} b^\top \lambda & \text{if } A^\top \lambda + \mu = c, \\ -\infty & \text{otherwise.} \end{cases} \end{aligned}$$

对偶问题：

$$\begin{aligned} \max \quad & \mathbf{b}^T\lambda \\ \text{s.t.} \quad & A^T\lambda \leq \mathbf{c}. \end{aligned} \tag{DP}$$

对于一般优化问题:

$$\begin{aligned} \min\, & f(x) \\ \text{s.t.} \quad & c_i(x) = 0, i \in \mathcal{E}, \\ & c_i(x) \geq 0, i \in \mathcal{I}. \end{aligned}$$

一般优化问题的 Langrange 对偶:

$$\max_{\lambda,\mu\geq 0}\min_{x} L(x,\lambda,\mu) = f(x) + \sum_{i\in\mathcal{E}}\lambda_i c_i(x) - \sum_{i\in\mathcal{I}}\mu_i c_i(x)$$

作业 4.2 推导出下面问题的对偶问题:

1.

$$\min c^\top x, \quad s.t.\ Ax \geq b.$$

2.

$$\min c^\top x, \quad s.t.\ Ax \leq b.$$

3. 并说明: *(DP)* 的对偶等价于 *(LP)*. 即, 对偶的对偶是原问题。

Theorem 4.4 (弱对偶定理) 任意线性规划问题 *(LP)* 及其对偶问题 *(DP)* 之间成立以下关系

$\mathbf{x}$ 是 *(LP)* 问题的可行解, λ 是 *(DP)* 问题的可行解, 那么 $\mathbf{c}^T\mathbf{x} \geq \mathbf{b}^T\lambda$.

推论:

(a) $\mathbf{x}$ 是 (LP) 问题的可行解 $\Rightarrow \mathbf{c}^T\mathbf{x} \geq$ 问题 (DP) 的最大值。
λ 是 (DP) 问题的可行解 $\Rightarrow \mathbf{b}^T\lambda \leq$ 问题 (LP) 的最小值。

(b) $\mathbf{x}$ 是 (LP) 问题的可行解, λ 是 (DP) 问题的可行解, 且 $\mathbf{c}^T\mathbf{x} = \mathbf{b}^T\lambda \ \Rightarrow\ \mathbf{x}$ 是 (LP) 问题的最优解, λ 是 (DP) 问题的最优解。

(c) 问题 (LP) 无界 $\Rightarrow$ 问题 (DP) 无可行解。
问题 (DP) 无界 $\Rightarrow$ 问题 (LP) 无可行解。

(d) **最优解存在性定理:** 若 (LP) 与 (DP) 都有可行解, 则它们均存在最优解。

Theorem 4.5 (强对偶定理) 原问题 *(LP)* 有最优解, 则对偶问题 *(DP)* 也有最优解, 且此时两方的最优值一致。

Proof: 设 (LP) 的最优解为某个可行基解: $\bar{x} = (x_B, 0)^\top$, 那么最优值为 $c_B^\top x_B$. 由最优判定定理得:

$$c^\top - c_B^\top B^{-1}A \geq 0^\top.$$

令 $\bar{\lambda} = B^{-T}c_B \in \mathbb{R}^m$, 那么,

$$A^\top\bar{\lambda} \leq c.$$

并且

$$\bar{\lambda}^\top b = c_B^\top B^{-1} b = c_B^\top x_B.$$

即, $\bar{\lambda}$ 是 (DP) 的最优解，且两者最优值一致。 ■

Theorem 4.6 (互补松弛定理) 设 $\bar{\mathbf{x}}$ 和 $\bar{\lambda}$ 分别是 *(LP)* 和 *(DP)* 的可行解，那么 $\bar{\mathbf{x}}$ 和 $\bar{\lambda}$ 是对应问题最优解的充要条件是:

$$\begin{cases} \bar{\mathbf{x}}^T(A^T\bar{\lambda} - \mathbf{c}) = 0 \\ \bar{\lambda}^T(A\bar{\mathbf{x}} - \mathbf{b}) = 0. \end{cases} \tag{4.10}$$

回顾：线性规划的 KKT 条件为:

$$(\text{KKT}) \qquad \begin{cases} A\mathbf{x} - \mathbf{b} = 0 \\ \mathbf{x} \geq \mathbf{0} \\ A^T\lambda - \mathbf{c} \leq 0 \\ \mathbf{x}^T(A^T\lambda - \mathbf{c}) = 0 \\ \lambda^T(A\mathbf{x} - \mathbf{b}) = 0 \end{cases} \tag{4.11}$$

KKT 条件为充分条件：由于约束都是线性的，故 Slater 条件成立。与此处的结论一致。

Example 4.5 利用互补松弛定理，当知道一个问题的最优解时，可求出其对偶问题的最优解。

$$\begin{aligned} \min \quad & 13x_1 + 10x_2 + 6x_3 \\ \text{s.t.} \quad & 5x_1 + x_2 + 3x_3 = 8 \\ & 3x_1 + x_2 = 3 \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{aligned} \quad \text{的最优解}\bar{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

对偶问题为:

$$\begin{aligned} \max \quad & 8\lambda_1 + 3\lambda_2 \\ \text{s.t.} \quad & 5\lambda_1 + 3\lambda_2 \leq 13 \\ & \lambda_1 + \lambda_2 \leq 10 \\ & 3\lambda_1 \leq 6. \end{aligned}$$

因为 $x_1 > 0, x_3 > 0$ 由互补松弛得,

$$5\lambda_1 + 3\lambda_2 = 13$$

$$3\lambda_1 = 6.$$

$$\text{所以对偶问题的最优解}\bar{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

5.1 对偶问题以及对偶变量的意义

灵敏度分析： 实际问题中，很多时候是基于某些采集数据来决定模型的系数。在这种情况下，势必会出现系数的扰动及引起的变化。为简单起见，考虑问题

$$\begin{aligned} \min \quad & \mathbf{c}^T\mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{4.12}$$

约束条件右边常数向量作微小变化

$$\mathbf{b} + \Delta\mathbf{b} = (b_1 + \Delta b_1, \cdots, b_m + \Delta b_m)^T,$$

得到新的线性规划问题

$$\begin{aligned} \min \quad & \mathbf{c}^T\mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} + \Delta\mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}. \end{aligned} \tag{4.13}$$

由强对偶理论，(4.13)的最优解 $\mathbf{c}^T\mathbf{x} = \lambda^T(\mathbf{b} + \Delta\mathbf{b})$, 故原问题的最优值关于约束资源 b 的变化率为

$$\partial\mathbf{c}^T\mathbf{x}/\partial\mathbf{b} = \lambda.$$

由于直接解原问题，并不知道对偶问题的解，故对偶最优解也被称为 **"影子价格"**。

Example 4.6 假设一个公司 α 生产两种产品，产品 A 和产品 B。生产这两种产品需要两种资源：劳动力和原材料。每个产品的利润和每个资源的可用量是已知的，分别用 p_A, p_B 表示。

原问题 决策变量:

- x_A, x_B - 分别表示产品 A 和产品 B 的生产量。

目标函数:

$$\text{最大化总利润} \quad P = p_A x_A + p_B x_B$$

约束条件:

$$l_A x_A + l_B x_B \leq L \quad (\text{劳动力})$$
$$m_A x_A + m_B x_B \leq M \quad (\text{原材料})$$

l_A 和 l_B 分别表示每件产品 A, B 所需要的劳动力，m_A, m_B 分别表示每件产品 A, B 所需要的材料。**对偶问题** 假设由另一个公司 β 向公司 α 购买所有的原材料和劳动力，希望总价最低。

决策变量:

- y_L, y_M - 分别对应劳动力和原材料约束的影子价格，即单位劳动力和原材料对利润的影响。

目标函数：

$$\text{最小化资源总成本} \quad C = Ly_L + My_M$$

约束条件：α 公司愿意卖给 β 的条件是，获得的收益不得低于自己制造产品 A, B 的利润。

$$l_A y_L + m_A y_M \geq p_A$$
$$l_B y_L + m_B y_M \geq p_B$$

影子价格的意义 在对偶问题中，最优解的 y_L, y_M 是影子价格，它们有以下含义：

1. **资源价值**：y_L 和 y_M 表示每单位劳动力力和原材料对总利润的贡献。
2. **成本限制**：影子价格反映了在当前生产计划下，资源的稀缺程度和价值。
3. **经济决策**：如果增加资源的成本低于其影子价格，则增加该资源是有利的。

对偶问题与原问题的关系

- **强对偶性**：如果原问题和对偶问题都有可行解，则它们的最优解值相等。
- **经济解释**：对偶问题提供了资源分配和成本效益分析的另一个视角。强对偶理论表明了两种决策下的相同利润和成本，达到了平衡。

通过对偶问题和影子价格的理解，企业和组织可以更好地洞察资源分配的经济效益，从而作出更明智的生产和投资决策。