最优化

Optimization

學院 數據科學學院 (SDS)

2025 年 11 月 2 日

Abstract

这是对课程 MAT3007 期中以前的重要内容分析。

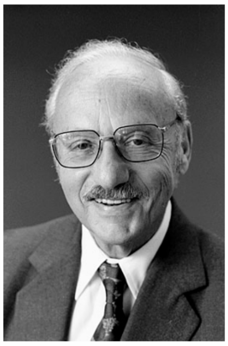
线性模型 LP

1 线性规划的定义与概述

线性模型是最优化理论中的基础组成部分，在科学、工程和经济学等多个领域有着广泛应用。线性规划 (Linear Programming, LP) 是研究线性目标函数在线性约束条件下极值问题的数学理论和方法。它是最优化理论中发展最为成熟的分支之一，具有完善的理论体系和高效的求解算法。

LP 由乔治 · B · 丹齐格（*George B. Dantzig*）于 1947 年左右构思而成，其中第一个用于求解线性规划模型的算法（单纯形法，(*Simplex method*)）是由丹齐格（1949 年）发表的。

Dantzig



1.1 线性模型的定义

Definition

线性规划（linear programming）(或线性模型) 可以定义为在一组线性约束条件下，优化 (最大化或最小化) 一个线性目标函数的问题。其一般形式可以表示为:

$$
\begin{aligned}
&\text{优化目标} \quad \min(\text{或}\max) \quad c_1x_1 + c_2x_2 + \cdots + c_nx_n \\
&\text{约束条件} \quad a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \le (\text{或} =, \ge) b_1 \\
&\qquad\qquad\quad \vdots \\
&\qquad\qquad\quad a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \le (\text{或} =, \ge) b_m \\
&\text{变量限制} \quad x_1, x_2, \cdots, x_n \ge 0 \ (\text{或无符号限制})
\end{aligned}
$$

其中，$x_1, x_2, \cdots, x_n$ 是决策变量，$c_1, c_2, \cdots, c_n$ 是目标函数系数，a_{ij} 是约束条件系数，$b_1, b_2, \cdots, b_m$ 是约束条件的右端项。

从几何角度看，线性规划问题的可行解区域是一个凸多面体；当可行域有界时，也称为凸多胞形

目标函数是一个线性函数，**其等值面是一组平行的超平面**。优化过程就是在这个凸多面体上寻找使目标函数达到极值的点。

1.2 线性规划的应用领域

线性规划在资源分配、生产计划、物流配送、投资组合优化等领域有着广泛的应用。例如：

1. 资源分配问题：在有限资源条件下，如何分配资源使得产出最大化或成本最小化

2. 运输问题：如何安排货物运输路线，使得总运输成本最低

3. 生产计划问题：如何安排不同产品的生产量，使得总利润最大且不超过生产能力限制

4. 投资组合问题：如何选择不同投资产品的比例，在风险一定的情况下最大化收益。

2 划归标准 LP

2.1 逐分量比较

在数学优化中，" **逐分量比较**（coordinate-wise comparison）" 指的是对向量的每个分量分别进行比较，而不是作为整体进行比较。例如：

* $x \geq 0$ 表示向量 x 的每个分量都非负，即对于所有的 i，$x_i \geq 0$;

* $Ax \leq b$ 表示矩阵 A 与向量 x 的乘积的每个分量都小于或等于向量 b 的对应分量

这种逐分量比较在优化问题中非常重要，因为约束条件通常是对变量的各个分量或其线性组合进行限制，而不是对整个向量进行整体限制。

于是我们可以进行对标准 LP 格式的制定：

2.2 线性规划的标准形式

为了便于理论分析和算法设计，通常将线性规划问题转化为标准形式。标准形式的线性规划问题具有以下特点：

1. 目标函数是最小化形式

2. 所有约束条件都是等式

3. 所有变量都是非负的

$$\begin{cases} \underset{x}{\text{minimize}} & c^\top x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{cases}$$

非标准型转标准型的技巧

1. 最大化问题：若目标是最大化，用 $-c$ 代替 c，并转换为最小化问题。

2. 消除不等式约束：$Ax \leq b$ 或 $Ax \geq b$

- 对于 $Ax \leq b$，将其改写为 $Ax + s = b$，其中 $s \geq 0$；

- 对于 $Ax \geq b$，将其改写为 $Ax - s = b$，其中 $s \geq 0$；

称 s 为**松弛变量（slack variables）**。

3. 处理 $x_i \leq 0$ 的变量：若存在变量 $x_i \leq 0$，将 x_i 替换为 $-y_i$，并添加约束 $y_i \geq 0$。

4. 消除"自由"变量（无约束的变量 x_i）：若变量 x_i 无约束（称为"自由变量"），将 x_i 替换为 $x_i^+ - x_i^-$，其中 $x_i^+ \geq 0$，$x_i^- \geq 0$。

Example

将以下线性规划问题转化为标准形式：

$$\begin{aligned} \max \quad & 3x_1 + 2x_2 \\ \text{subject to} \quad & x_1 + x_2 \leq 5 \\ & 2x_1 - x_2 \geq 1 \\ & x_1 \geq 0 \end{aligned}$$

转换过程：

1. 将最大化目标函数转换为最小化：$\min \quad -3x_1 - 2x_2$

2. 处理第一个不等式约束 $x_1 + x_2 \leq 5$，引入松弛变量（slack variables）$s_1 \geq 0$，得到等式约束：$x_1 + x_2 + s_1 = 5$

3. 处理第二个不等式约束 $2x_1 - x_2 \geq 1$，引入松弛变量 $s_2 \geq 0$，得到等式约束：$2x_1 - x_2 - s_2 = 1$

4. 变量 x_2 没有非负限制，用 $x_2 = x_2^+ - x_2^-$ 代替，其中 $x_2^+ \geq 0$ 和 $x_2^- \geq 0$

最终的标准形式为：

$$\begin{aligned} \min \quad & -3x_1 - 2x_2^+ + 2x_2^- \\ \text{subject to} \quad & x_1 + x_2^+ - x_2^- + s_1 = 5 \\ & 2x_1 - x_2^+ + x_2^- - s_2 = 1 \\ & x_1, x_2^+, x_2^-, s_1, s_2 \geq 0 \end{aligned}$$

2.3 转换过程的数学证明

Theorem

任何线性规划问题都可以转换为标准形式。

证明. 考虑一般的线性规划问题：

$$\begin{aligned} \max \quad & c^T x \\ \text{subject to} \quad & A_1 x \leq b_1 \\ & A_2 x \geq b_2 \\ & A_3 x = b_3 \\ & x \geq 0 \end{aligned}$$

其中，A_1, A_2, A_3 是维度匹配的矩阵，b_1, b_2, b_3 是相应的向量。

我们需要将其转换为标准形式：

1. 目标函数转换：将最大化转换为最小化，得到 $\min -c^T x$

2. 不等式约束转换：

- 对于 $A_1 x \leq b_1$，引入松弛变量 $s_1 \geq 0$，得到 $A_1 x + s_1 = b_1$
- 对于 $A_2 x \geq b_2$，引入松弛变量 $s_2 \geq 0$，得到 $A_2 x - s_2 = b_2$

3. 等式约束保留：$A_3 x = b_3$ 保持不变

4. 变量非负约束保留：$x \geq 0$

合并所有约束，得到标准形式：

$$\begin{aligned} \min \quad & -c^T x \\ \text{subject to} \quad & \begin{bmatrix} A_1 & I & 0 \\ A_2 & 0 & -I \\ A_3 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \\ & \begin{bmatrix} x \\ s_1 \\ s_2 \end{bmatrix} \geq 0 \end{aligned}$$

其中，I 是适当维度的单位矩阵。

对于存在无约束变量的情况，假设变量 x_j 无符号限制，可以用 $x_j = x_j^+ - x_j^-$ 代替，其中 $x_j^+ \geq 0$ 和 $x_j^- \geq 0$。这样，所有变量都满足非负约束。

因此，任何线性规划问题都可以转换为标准形式。

□

2.4 对标准形代数结构的理解

进一步讨论：我们知道，LP 的标准形式数学表达式为：

$$\begin{aligned} \min \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

假设元素 x 有 n 个维度，那么：

如果 A 满秩，则一共 n 个不同的约束条件

如果 A 不满秩，那么要么有可以被抵消的约束条件（redundant constraints）；要么有不成立的约束条件（constraints are not consistent）。

3 把 NLP 转换为 LP 的经典例子

3.1 绝对值问题

Example

问题描述：最小化一个非光滑函数 $\max_i(c_i^T x + d_i)$，该函数是多个线性函数的最大值。

转换为 LP：

1. 引入变量 z 表示最大值，即 $z = \max_i(c_i^T x + d_i)$

2. 最小化 z，同时确保对于所有 i，$z \geq c_i^T x + d_i$

最终的线性规划形式为：

$$\begin{aligned} \min \quad & z \\ \text{subject to} \quad & z \geq c_i^T x + d_i, \quad \forall i \\ & \text{其他约束条件} \end{aligned}$$

示例：

考虑以下非光滑优化问题：

$$\min_x \max(3x + 4, -2x + 1)$$

转换为线性规划：

$$\begin{aligned} \min \quad & z \\ \text{subject to} \quad & z \geq 3x + 4 \\ & z \geq -2x + 1 \\ & x \geq 0 \end{aligned}$$

最终标准型为: $\begin{cases} \min & z \\ \text{subject to} & z - 3x - s_1 = 4 \\ & z + 2x - s_2 = 1 \\ & x, s_1, s_2 \geq 0 \end{cases}$

最优解为 $x = \frac{3}{5}$, $z = \frac{29}{5}$。

3.2 最优化函数和

问题描述：最小化多个函数的和，每个函数 $f_i(x)$ 可能是非线性或非光滑的。

- 考虑优化问题：

$$\underset{x}{\text{minimize}} \sum_{i=1}^{m} f_i(x) \quad \text{s.t.} \quad x \in \Omega \tag{1}$$

- 以及另一个优化问题：

$$\underset{x,t}{\text{minimize}} \sum_{i=1}^{m} t_i \quad \text{s.t.} \quad x \in \Omega,\ f_i(x) \leq t_i,\ \forall i \tag{2}$$

其中 $f_i : \mathbb{R}^n \to \mathbb{R}$ 为给定函数，$\Omega \subset \mathbb{R}^n$ 是可行集。

问题 2 的本质是让 t 尽可能小，并且贴近 $f(x)$)

我们将给出如下 Claim（建模工具 Modeling tools）：

Claim

问题 (1) 和 (2) 具有如下等价性：

- 若 x^* 是 (1) 的最优解，则 $(x^*, f_1(x^*), \ldots, f_m(x^*))$（记 $t^* = (f_1(x^*), \ldots, f_m(x^*))$）是 (2) 的最优解。
- 若 (x^*, t^*) 是 (2) 的最优解，则 x^* 是 (1) 的最优解。
- 此时，两个问题的最优值相同。

证明. 为简化分析，假设 $m = 1$（证明思想可推广至一般 m 的情形）。

- 设 x^* 是 (1) 的最优解，令 $t^* = f(x^*)$，则 (x^*, t^*) 是 (2) 的可行解（满足 $f(x^*) \leq t^*$）。
- 设 $(\bar{x}, \bar{t})$ 是 (2) 的最优解，则显然有 $\bar{t} = f(\bar{x})$（原因：由约束 $f(\bar{x}) \leq \bar{t}$，且 (2) 要最小化 $\bar{t}$，故等号必须成立）；同时，$\bar{x}$ 是 (1) 的可行解（因 $\bar{x} \in \Omega$）。
- 对问题 (1)，由于 $\bar{x}$ 是可行解，且 x^* 是 (1) 的最优解，故 $f(\bar{x}) \geq f(x^*)$。

- 对问题 (2)，由于 (x^*, t^*) 是可行解，且 $(\bar{x}, \bar{t})$ 是 (2) 的最优解，故 $f(x^*) = t^* \geq \bar{t} = f(\bar{x})$。
- 结合上述两点，得 $f(x^*) = t^* = \bar{t} = f(\bar{x})$。再结合可行性可知：$(x^*, t^*)$ 是 (2) 的最优解，且 $\bar{x}$ 是 (1) 的最优解。

$\square$

转换为 LP 的具体方法：

该问题能否转换为线性规划取决于函数 $f_i(x)$ 的具体形式。以下是几种常见情况：

1. 线性函数的绝对值： 如果 $f_i(x) = |a_i^T x + b_i|$，可以引入变量 y_i 使得 $y_i \geq a_i^T x + b_i$ 和 $y_i \geq -a_i^T x - b_i$，从而将目标函数转换为 $\sum y_i$，这是一个线性函数。

2. 分段线性函数： 如果 $f_i(x)$ 是分段线性的，可以将其表示为多个线性函数的最大值或最小值，然后应用类似上述的方法。

3. 指示函数： 如果 $f_i(x)$ 是指示函数（即当满足某些条件时为 0，否则为无穷大），可以通过引入适当的约束条件来处理。

Example

考虑以下优化问题：

$$\min_x \sum_{i=1}^{3} |x_i| \quad \text{subject to} \quad Ax = b, \quad x \in \mathbb{R}^3$$

转换为线性规划：

1. 引入变量 $y_i \geq |x_i|$，即 $y_i \geq x_i$ 和 $y_i \geq -x_i$
2. 目标函数转换为 $\min \sum_{i=1}^{3} y_i$
3. 约束条件保持为 $Ax = b$

最终的线性规划形式为：

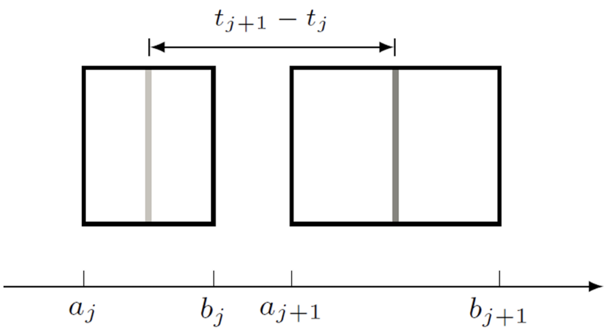
$$\begin{aligned}
\min \quad & y_1 + y_2 + y_3 \\
\text{subject to} \quad & Ax = b \\
& y_i \geq x_i, \quad i = 1, 2, 3 \\
& y_i \geq -x_i, \quad i = 1, 2, 3 \\
& y_i \geq 0, \quad i = 1, 2, 3
\end{aligned}$$

3.3 最大化最小问题

我们以一个例子导入：

Example

问题描述：最大化最小间隔问题，即在一系列变量 $t_1, t_2, \cdots, t_n$ 中，找到最大的最小间隔 $t_{j+1} - t_j$。



转换为 LP：

1. 引入变量 t 表示最小间隔，即 $t = \min_{j=1,\cdots,n-1}(t_{j+1} - t_j)$

2. 最大化 t，同时确保对于所有 j，$t_{j+1} - t_j \geq t$

3. 可以引入额外的约束条件来限制变量 t_j 的范围，例如 $t_1 \geq 0$ 或其他边界条件（如：$Ax = b$）

最终的线性规划形式为：

$$
\begin{aligned}
\max \quad & t \\
\text{subject to} \quad & t_{j+1} - t_j \geq t, \quad j = 1, 2, \cdots, n-1 \\
& Ax = b \\
& t \geq 0
\end{aligned}
$$

数学解释：

假设原问题的最优解为 t^*，则对于所有 j，有 $t_{j+1} - t_j \geq t^*$，并且存在某个 k 使得 $t_{k+1} - t_k = t^*$。在转换后的线性规划中，t 被最大化，同时所有 $t_{j+1} - t_j \geq t$ 的约束确保 t 不超过任何间隔 $t_{j+1} - t_j$，因此 t 的最大值即为原问题中的最大最小间隔。

下面给出一个示例：

Example

考虑 $n = 3$ 的情况，在区间 $[0, 10]$ 上放置 3 个点 t_1, t_2, t_3（满足 $0 \leq t_1 \leq t_2 \leq t_3 \leq 10$），目标是最大化这三个点中最小的相邻间隔，即最大化 $\min\{t_2 - t_1, t_3 - t_2\}$。用数学表达：

$$
\max \quad \min(t_2 - t_1, t_3 - t_2)
$$

转换为线性规划：

$$
\begin{aligned}
\max \quad & t \\
\text{subject to} \quad & t_2 - t_1 \geq t \\
& t_3 - t_2 \geq t \\
& t_1 \geq 0 \\
& t_3 \leq 10 \\
& t \geq 0
\end{aligned}
$$

最终标准型为：$\begin{cases} \min & -t \\ \text{subject to} & -t_1 + t_2 - t + s_1 = 0 \\ & -t_2 + t_3 - t + s_2 = 0 \\ & t_3 + s_3 = 10 \\ & t_1, t_2, t_3, t, s_1, s_2, s_3 \geq 0 \end{cases}$

最优解为 $t = \frac{10}{2} = 5$，对应的变量值为 $t_1 = 0, t_2 = 5, t_3 = 10$。

3.4 线性分式规划

- 原始线性分式规划问题：

  最小化目标函数：
  $$\frac{\boldsymbol{c}^\top \boldsymbol{x} + d}{\boldsymbol{e}^\top \boldsymbol{x} + a}$$
  约束条件：
  $$A\boldsymbol{x} \leq \boldsymbol{b}$$
  假设：对所有满足 $A\boldsymbol{x} \leq \boldsymbol{b}$ 的 $\boldsymbol{x}$，有 $\boldsymbol{e}^\top \boldsymbol{x} + a > 0$（保证分母为正）。

- **变量替换思路**

  定义新变量：
  $$\boldsymbol{y} = \frac{\boldsymbol{x}}{\boldsymbol{e}^\top \boldsymbol{x} + a}, \quad z = \frac{1}{\boldsymbol{e}^\top \boldsymbol{x} + a}$$
  目的是将分式结构转化为线性结构。

- 转化后的线性规划（LP）

  最小化目标函数：
  $$\boldsymbol{c}^\top \boldsymbol{y} + dz$$

约束条件:

$$\begin{cases} A\boldsymbol{y} - \boldsymbol{b}z \leq \mathbf{0}, \\ \boldsymbol{e}^\top \boldsymbol{y} + az = 1, \\ z \geq 0. \end{cases}$$

4 基解与基可行解概述

4.1 可行解与基解

Definition

若向量 $\mathbf{x}$ 满足所有约束条件（即 $A\mathbf{x} = \mathbf{b}$ 且 $\mathbf{x} \geq \mathbf{0}$），则称 $\mathbf{x}$ 为该线性规划问题的**可行解** (Feasible solution)。

所有可行解的集合称为**可行域**（Feasible Region）。

简言之，可行解是“满足线性规划全部约束（等式约束与变量非负约束）的解向量”，是后续分析“基本可行解”“最优解”的基础。

但是线性规划标准形中，如果约束 $\mathbf{A}\mathbf{x} = \mathbf{b}$ 是 m 个等式、n 个变量（$n > m$）的**欠定方程组**，那么可行解有无穷多个。

为找到“有限的关键解”，需引入**基（Basis）**的概念：

Definition

从系数矩阵 $\mathbf{A}$ 的 n 个列向量中，选出 m 个**线性无关**的列向量，构成 m 阶**可逆**矩阵 $\mathbf{B}$，称为**基矩阵（Basis Matrix）**。

对应地：

基矩阵 $\mathbf{B}$ 对应的 m 个变量，称为**基变量（Basic Variables）**，记为 $\mathbf{x}_B = (x_{B_1}, x_{B_2}, \ldots, x_{B_m})^T$；

剩余 $n-m$ 个变量，称为**非基变量(Non-Basic Variables)**，记为 $\mathbf{x}_N = (x_{N_1}, x_{N_2}, \ldots, x_{N_{n-m}})^T$。

由此，我们给出基解的概念：

Definition

对于标准形式线性规划，向量 $\boldsymbol{x}$ 为**基解（Basic Solution, BS）**需满足：

1. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$（满足约束方程）；
2. 存在指标 $B(1), \cdots, B(m)$，使 $\boldsymbol{A}$ 的列 $\boldsymbol{A}_{B(1)}, \cdots, \boldsymbol{A}_{B(m)}$ 线性无关，且 $i \neq B(1), \cdots, B(m)$ 时，$x_i = 0$。

由于“从 n 列中选 m 个线性无关列”的组合数是有限的，**因此基解的数量也是有限的**。这就把“无穷多解”的问题转化为“有限个关键解”的分析问题。

基解的数量最多为组合数 $C(n,m)=\frac{n!}{m!(n-m)!}$。由排列组合，容易证明。

同时，根据定义，我们可以观察到：

Fact or Background

1. **基解存在且唯一**

证明. 基解由选 m 个线性无关列（构成可逆基矩阵 $\boldsymbol{A}_B$）、非基变量置 0，再解 $\boldsymbol{A}_B\boldsymbol{x}_B=\boldsymbol{b}$（$\boldsymbol{x}_B$ 为基变量）得到，并且因为基矩阵可逆，故**基解存在且唯一**。 □

Fact or Background

2. 基解最多 m 个非零值；

【Remark】：

基解最多有 m 个非零值；但是这不意味着在每一个最优解中，最多有 m 个变量可以为正数（因为最优解可以说多个基（可行）解的凸组合）

Example

考虑下面这个反例：

minimize 0。那么，任何可行的 x 都是最优的，无论它有多少个正分量。

4.2 相邻基解 (Neighboring Basic Solutions)

Definition

若两个基解（Basic Solution, BS）的 **基指标集合**（由约束矩阵中线性无关列的索引组成）仅相差一个元素，则称这两个基解是 **相邻** 的。

即从一个基解的基矩阵（Basis Matrix, $\boldsymbol{A}_B$）中替换一列（保持列的线性无关性），得到的新基解与原基解相邻；若两个基解的基指标集合相差超过一个元素，则它们不相邻。

Example

如：（1，4，5）与（1，2，5）、（1，2，4）都相邻

但是又例如，基指标集合 $B_3 = \{1,3,4\}$ 对应的基解与 $B_1 = \{1,2,3\}$ 对应的基解，基指标相差两个元素（2 和 4），因此不相邻。

还如：（1，2，3）与（1，2，5）、（1，2，4）、（1，3，5）、（1，3，4）均相邻；而与（1，4，5）不相邻

4.3 基可行解

Definition

若基解 $\boldsymbol{x}$ 满足 $\boldsymbol{x} \geq 0$（变量非负），则为**基可行解**（Basic Feasible Solution, BFS）。

我们可以认为：

Fact or Background

在规范的 LP 规划问题中，**基可行解对应线性规划可行域的极点**，即 corner feasible point（CFP）与 extreme point 这两个概念是等价的。

或者说，标准形式多面体 $P := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0\}$ 中，$\boldsymbol{x}$ 是极点**当且仅当**为基可行解。

Theorem

BFS 与极点一一对应：给定一个基 B，对应的 BFS 是唯一的极点。

即：给定一个基 B 后，不存在另一个不同的极点，在"相同非基变量全为 0"的条件下与它共存。

证明. 假设"存在极点 $y \neq x$，且对所有 $i \notin B$（非基变量），$x_i = y_i = 0$"。我们可以通过以下步骤推导矛盾，进而理解其意义：

由"$i \notin B$ 时 $x_i = y_i = 0$"，可知 $x_N = y_N = 0$（N 是非基变量集合）；因为 $x, y \in P$，需满足 $Ax = b$ 和 $Ay = b$。代入 $x_N = y_N = 0$，得 $A_B x_B = b$ 和 $A_B y_B = b$（A_B 是基矩阵 B 对应的列）；

由于 A 行满秩，A_B 可逆，因此 $x_B = A_B^{-1} b$，$y_B = A_B^{-1} b$，即 $x_B = y_B$；

结合 $x_N = y_N = 0$，可得 $x = y$，与"$x \neq y$"矛盾。

□

基本可行解由其基变量唯一确定，不存在另一个共享相同零非基变量的不同极值点。

首先，为了导引下节，我们不加证明的得出一条引理与一个实例：

Lemma

Optimality Test: 在标准 LP 中，如果一个候选可行解没有具有更优目标值的**相邻候选可行解**，那么它**一定是最优解**。

我们以一个实际的生产计划作为示例:

Example

原问题最大化 $x_1 + 2x_2$，约束:
$x_1 \le 100$，$2x_2 \le 200$，$x_1 + x_2 \le 150$，$x_1, x_2 \ge 0$

转标准形式（最小化 $-x_1 - 2x_2$）后，约束为 $x_1 + s_1 = 100$，$2x_2 + s_2 = 200$，$x_1 + x_2 + s_3 = 150$（$x_1, x_2, s_1, s_2, s_3 \ge 0$），其中 $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}$，$\boldsymbol{b} = \begin{bmatrix} 100 \\ 200 \\ 150 \end{bmatrix}$。

选 $\boldsymbol{A}$ 前 3 列（基指标 $B = \{1, 2, 3\}$），基矩阵 $\boldsymbol{A}_B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix}$，解得 $\boldsymbol{x}_B = \boldsymbol{A}_B^{-1}\boldsymbol{b} = \begin{bmatrix} 50 \\ 100 \\ 50 \end{bmatrix}$，即解 $(50, 100, 50, 0, 0)$，满足 $\boldsymbol{x} \ge 0$，为基可行解。

通过选不同 $m = 3$ 个线性无关列，可得到多个基解，部分为基可行解（如上述解），部分因变量负非可行。

表 1: 基（可行）解列举

基指标集合	解向量	状态
$\{1,2,3\}$	$(50, 100, 50, 0, 0)$	BFS
$\{1,2,4\}$	$(100, 50, 0, 100, 0)$	BFS
$\{1,2,5\}$	$(100, 100, 0, 0, -50)$	BS（非可行）
$\{1,3,4\}$	$(150, 0, -50, 200, 0)$	BS（非可行）
$\{1,4,5\}$	$(100, 0, 0, 200, 50)$	BFS
$\{2,3,4\}$	$(0, 150, 100, -100, 0)$	BS（非可行）
$\{2,3,5\}$	$(0, 100, 100, 0, 50)$	BFS
$\{3,4,5\}$	$(0, 0, 100, 200, 150)$	BFS

基可行解对应可行域极点，如可行域由约束围成的粉色区域的角落点。

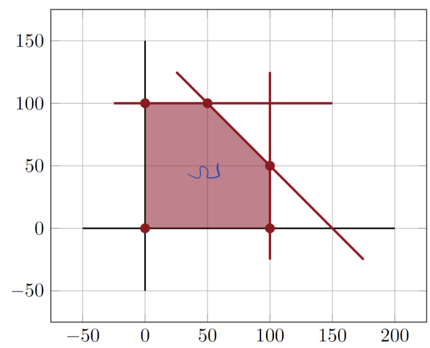


图 1: BFS 与 corner points

例如:

基可行解 $(50, 100, 50, 0, 0)$ 对应可行域中 $x_1 = 50$，$x_2 = 100$ 的点，该点是 $x_2 = 100$（由 $2x_2 \leq 200$ 取等得到）和 $x_1 + x_2 = 150$（取等）的交点。

基可行解 $(100, 50, 0, 100, 0)$ 对应 $x_1 = 100$（由 $x_1 \leq 100$ 取等得到）和 $x_1 + x_2 = 150$（取等）的交点，此时 $x_2 = 50$。

4.4 线性模型的最优解

我们这里给出**线性规划**存在最优解的**充要条件**:

> **Theorem**
>
> **线性规划**存在最优解的**充要条件**是: **问题可行**（存在满足所有约束的解）且**目标函数有界**（最大化时不会无限大，最小化时不会无限小）

后续我们可以观察到: 线性规划（LP）与非线性规划的 2 个核心区别:

线性规划的最优解如果存在，必然是全局最优，不存在 “局部最优但非全局最优” 的情况。即目标函数的最优值（如最大值、最小值）是唯一的;

但是最优解的**解向量**（决策变量的取值组合）有两种可能:

- **唯一最优解:** 可行域的顶点中，只有一个顶点（一个 BFS）能使目标函数取到最优值。此时，只有 1 组决策变量的取值是最优解。
- **无穷多最优解:** 当目标函数的 “等高线”（如二维下的直线）与可行域的某条边界平行时，这条边界上的所有点都是最优解（因为这些点的目标函数值都等于最优值）。此时，目标函数值唯一，但决策变量的取值组合有无穷多组。

所以我们可以得到 2 条重要的性质:

在求最小值的标准线性优化问题中，我们去掉了一个决策变量的非负约束（可选的解集变大了）。

假设原始问题和修正后的问题都存在最优解，那么去掉该约束后，**最优值一定不会增加**。

证明. 原问题的最优解 $\mathbf{x}^*$ 必然也是修正后问题 P' 的可行解。

由线性规划 LP 的唯一最优值，可知，最优值依然不变 □

无限的情境：

我们考虑一个无界线性规划问题。那么，如果向该问题中添加一个新变量，**这个线性规划问题仍然是无界的**。

证明. 将新变量 x_{new} 设为零，将能沿着相同的无界方向迭化目标函数。 □

接下来介绍：**线性规划基本定理**（Fundamental LP Theorem），考虑标准形式的线性规划问题:

Theorem

线性规划基本定理（Fundamental LP Theorem）：假设约束矩阵 $\boldsymbol{A}$ 的行满秩为 m：

1. 若可行域非空，那么存在基可行解（BFS）；
2. 若存在最优解，则存在一个同时也是基本可行解的最优解。

这意味着, 只要线性规划有可行解，就一定能找到基可行解这种特殊的可行解。若线性规划存在最优解，那么存在一个同时也是基可行解的最优解。

所以线性规划的最优解一定能在基可行解中找到, 为了找到最优解，我们只需要在基可行解的范围内进行搜索。

我们可以总结出最优解，基解和基可行解的逻辑关系：

Fact or Background

对于 $m \times n$ 的矩阵（即 m 个约束，n 个变量的优化问题）：

- 最优解一定在可行解中找，但是可行解往往是无限的，难以寻觅；
- 基解有限，并且是构建最优解的“基本单元”，且基解最多 m 个非零值，但是基解不一定“可行”
- 由此我们需要让基解“可行”，即得到基可行解（特殊的基解）。
- 而我们知道：最优解可能是基可行解，也可能是多个基可行解的凸组合；

故最优解的正变量个数不受基解“最多 m 个非零值”的限制。

这里结合基可行解与最优解的关系，给出一条推论：

Corollary & Secondary Conclusion

若一个具有 m 个约束的标准型线性规划（LP）存在最优解，则必**存在**一个正分量个数不超过 m 的最优解。

不过，直接列举所有基可行解并比较目标函数值的方法并不可行，因为基可行解的数量可能非常多。对于有 m 个约束、n 个变量的线性规划，基可行解的数量**最多**为组合数 $C(n,m)$（即基解的个数）。

因此，需要更巧妙的方法，即**单纯形法**（Simplex method）来高效地在基可行解中寻找最优解。

单纯形法从**一个基可行解**（可行域的一个极点）出发，转移到相邻的基可行解，持续改进目标函数的值，直至达到最优（类似于爬山思想）。所以我们只需要展示如何进行其中一步。然后，我们就可以简单地**递归**相同的方法（直到达到最优状态）。

为此我们需要：

1. 明确“相邻基可行解”的定义。
2. 设计一种高效方法，用于找到并转移到能改进目标函数值的相邻基可行解。
3. 设计一个有效的停止准则，用于判断是否已达到最优解。

5 单纯形法（Simplex Method）

单纯形法通过从一个基可行解迭代到 **相邻基可行解**，逐步寻找最优解。

5.1 确定基方向（Basic Directions）

基方向是从当前基可行解到相邻基可行解的移动方向。通过选择一个 **非基变量**（Non-basic Variable，当前基可行解中取值为 0 的变量）增加，即进基（Entering the Basis），并调整基变量以保持约束可行性，得到变量的变化方向。

首先，假设我们已经找到一个基本可行解（BFS），其基为 $B = \{B(1), \ldots, B(m)\}$。

定义：

$$A_B = \left[\; A_{B(1)} \;\middle|\; A_{B(2)} \;\middle|\; \cdots \;\middle|\; A_{B(m)} \;\right]$$

其中 A_N 是由矩阵 A 的非基列组成的矩阵。在后续中，N 表示非基索引集，满足：

- B 和 N 不相交；
- $|B| = m$，$|N| = n - m$；
- $B \cup N = \{1, 2, \ldots, n\}$。

重新排列变量后，可将 $A = [A_B, A_N]$，$\boldsymbol{x} = [\boldsymbol{x}_B; \boldsymbol{x}_N]$，其中 $\boldsymbol{x}_B$ 是基变量，$\boldsymbol{x}_N$ 是非基变量。

根据定义，有：

$$\boldsymbol{x}_B = A_B^{-1}\boldsymbol{b}, \quad \boldsymbol{x}_N = 0$$

接着我们需要找到一种寻找相邻 BFS 的方法:

- 寻找邻居意味着改变一个基索引。
- 我们要选择一个非基变量 $x_j, j \in N$ 进入基。
- 这意味着要从当前 BFS 中增加 x_j。

考虑将当前 BFS $\boldsymbol{x}$ 移动到相邻的 $\boldsymbol{y}$，形式为：

$$\boldsymbol{y} = \boldsymbol{x} + \theta \boldsymbol{d}, \quad \theta \geq 0$$

其中：

$d_j = 1$；并且对于所有其他**非基索引** $N \ni j' \neq j$，有 $d_{j'} = 0$。

那么 $\boldsymbol{d}$ 需要满足什么约束？

我们需要保证得到的步长 $\boldsymbol{y} = \boldsymbol{x} + \theta\boldsymbol{d}$ 仍然可行，即：

$$A(\boldsymbol{x} + \theta\boldsymbol{d}) = \boldsymbol{b} = A\boldsymbol{x}$$

也就是 $A\boldsymbol{d} = 0$（因为 θ 可以非零）。

- 现在，我们将 $\boldsymbol{d} = [\boldsymbol{d}_B; \boldsymbol{d}_N]$ 写出。由于 $d_j = 1$，且对于所有其他非基索引 j' 有 $d_{j'} = 0$，因此有：
$$0 = \begin{bmatrix} \boldsymbol{A}_B & \boldsymbol{A}_N \end{bmatrix} \begin{bmatrix} \boldsymbol{d}_B \\ \boldsymbol{d}_N \end{bmatrix} = \boldsymbol{A}_B\boldsymbol{d}_B + \boldsymbol{A}_j$$
- 因此：
$$\boldsymbol{d}_B = -\boldsymbol{A}_B^{-1}\boldsymbol{A}_j$$
- 这意味着，**一旦选择了 j，方向 $\boldsymbol{d}$ 就被唯一确定**：
$$\boldsymbol{d} = [\boldsymbol{d}_B; \boldsymbol{d}_N] = [-\boldsymbol{A}_B^{-1}\boldsymbol{A}_j; 0; \ldots; 1; \ldots; 0]$$

  其中, 1 位于第 j 个位置。**我们将这样的 $\boldsymbol{d}$ 称为第 j 个基方向**（jth basic direction）。

沿着第 j 个基本方向移动，目前仅能保证等式约束 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 仍然成立。

我们还需要考虑约束 $\boldsymbol{x} \geq 0$。需要验证该约束仍然成立（即 $\boldsymbol{y} = \boldsymbol{x} + \theta\boldsymbol{d} \geq 0$）：

- 对于任意非基索引 $j \in N$，x_j 原本为 0，且 d_j 非负（为 1 或 0），因此 $x_j + \theta d_j$ 仍然非负。
- 对于任意基变量 x_i（其中 $i \in B$），只要它们原本都严格为正，就**一定存在一个小**的 θ，使得 $x_i + \theta d_i \geq 0$。

【Remark】:

- 通常，基本可行解（BFS）的基变量全为正（即有 m 个正的分量）。
- 存在基本可行解中某些基变量等于 0 的情况（退化，degeneracy），我们之后会讨论这些情况。

5.2 目标函数变化 (Change in the Objective Value)

回忆原线性规划（LP）的目标函数为 $\boldsymbol{c}^\top \boldsymbol{x}$。

我们可以类似地将 $\boldsymbol{c}$ 分解为基部分和非基部分，对应基 / 非基索引，即：$\boldsymbol{c} = [\boldsymbol{c}_B; \boldsymbol{c}_N]$

现在，我们研究当从基本可行解（BFS）$\boldsymbol{x}$ 移动到 $\boldsymbol{y} = \boldsymbol{x} + \theta \boldsymbol{d}$ 时，目标函数的变化情况。

变化量为：$\theta \boldsymbol{c}^\top \boldsymbol{d}$

Claim

如果 $\boldsymbol{d}$ 是第 j 个基本方向，那么我们有：$\boldsymbol{c}^\top \boldsymbol{d} = c_j - \boldsymbol{c}_B^\top \boldsymbol{A}_B^{-1} \boldsymbol{A}_j =: \bar{c}_j$

证明. 根据“基本方向”的定义：非基变量 x_j 对应的分量 $d_j = 1$；其他非基变量对应的分量 $d_{j'} = 0$（$j' \in N, j' \neq j$）。

将目标函数的系数向量 $\boldsymbol{c}$ 也按“基变量系数 $\boldsymbol{c}_B$”和“非基变量系数 $\boldsymbol{c}_N$”分解，即 $\boldsymbol{c} = [\boldsymbol{c}_B; \boldsymbol{c}_N]$。

目标函数的变化由 $\boldsymbol{c}^\top \boldsymbol{d}$ 表示，其中 $\boldsymbol{d} = [\boldsymbol{d}_B; \boldsymbol{d}_N]$。

展开计算：$\boldsymbol{c}^\top \boldsymbol{d} = \begin{bmatrix} \boldsymbol{c}_B^\top & \boldsymbol{c}_N^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{d}_B \\ \boldsymbol{d}_N \end{bmatrix} = \boldsymbol{c}_B^\top \boldsymbol{d}_B + \boldsymbol{c}_N^\top \boldsymbol{d}_N$

而：

$\boldsymbol{c}_N^\top \boldsymbol{d}_N = c_j \cdot d_j + \sum_{\substack{j' \in N \\ j' \neq j}} c_{j'} \cdot d_{j'} = c_j \cdot 1 + \sum_{\substack{j' \in N \\ j' \neq j}} c_{j'} \cdot 0 = c_j$

我们已得到基变量部分的方向向量满足：$\boldsymbol{d}_B = -\boldsymbol{A}_B^{-1} \boldsymbol{A}_j$

将其代入 $\boldsymbol{c}_B^\top \boldsymbol{d}_B$：$\boldsymbol{c}_B^\top \boldsymbol{d}_B = \boldsymbol{c}_B^\top \left(-\boldsymbol{A}_B^{-1} \boldsymbol{A}_j \right) = -\boldsymbol{c}_B^\top \boldsymbol{A}_B^{-1} \boldsymbol{A}_j$

将 $\boldsymbol{c}_B^\top \boldsymbol{d}_B$ 和 $\boldsymbol{c}_N^\top \boldsymbol{d}_N$ 的结果代入 $\boldsymbol{c}^\top \boldsymbol{d}$ 的展开式：

$\boldsymbol{c}^\top \boldsymbol{d} = \boldsymbol{c}_B^\top \boldsymbol{d}_B + \boldsymbol{c}_N^\top \boldsymbol{d}_N = -\boldsymbol{c}_B^\top \boldsymbol{A}_B^{-1} \boldsymbol{A}_j + c_j$

按照定义，将其记为 $\bar{c}_j$ □

Definition

我们称 $c_j - \boldsymbol{c}_B^\top \boldsymbol{A}_B^{-1} \boldsymbol{A}_j =: \bar{c}_j$ 为变量 x_j 的**约化成本**（reduced cost）。

约化成本是单纯形法中一个非常重要的概念，是我们想要移动方向的指示器。它对应着当尝试改变基（将 $x_j, j \in N$ 纳入基变量）时，目标函数值的变化。

给定当前的基以及想要加入基的索引 (j)，约化成本可以很容易地计算:

- 正的约化成本意味着将 j 纳入当前基会增加目标函数。
- 负的约化成本意味着将 j 纳入基会降低目标函数。

于是，对于不同的问题，我们可以给出进动的不同判别规则:

最大化问题：若 $\bar{c}_j > 0$，则 x_j 进基能改进目标函数；

最小化问题：若 $\bar{c}_j < 0$，则 x_j 进基能改进目标函数。

Fact or Background

在线性规划中，**基变量**的约化成本恒为 0

证明. 展开转置并整理：$\bar{c}_i = \tilde{c}_i - \tilde{\mathbf{c}}_B^T \mathbf{A}_B^{-1} \mathbf{A}_i$

由于 $\mathbf{A}_i$ 是 $\mathbf{A}_B$ 的第 i 列，因此 $\mathbf{A}_B^{-1}\mathbf{A}_i$ 是单位矩阵的第 i 列（记为 $\mathbf{e}_i$，仅第 i 个元素为 1，其余为 0)。

因此：$\tilde{\mathbf{c}}_B^T \mathbf{e}_i = \tilde{c}_i$ （因为 $\tilde{\mathbf{c}}_B$的第 i 个元素就是基变量 x_i 的目标系数 $\tilde{c}_i$) 代入约化成本公式得：$\bar{c}_i = \tilde{c}_i - \tilde{c}_i = 0$ □

Example

最大化 $z = x_1 + 2x_2$ 的延续：

$$
\begin{aligned}
&\underset{x\in\mathbb{R}^5}{\text{maxmize}} \quad f(x) = x_1 + 2x_2 \\
&\text{subject to} \\
&x_1 + s_1 - 100 = 0 \\
&2x_2 + s_2 - 200 = 0 \\
&x_1 + x_2 + s_3 - 150 = 0 \\
&x_1, x_2, s_1, s_2, s_3 \geq 0
\end{aligned}
$$

当前基可行解 $\boldsymbol{x}_1 = (50, 100, 50, 0, 0)^T$，基变量为 x_1, x_2, s_1.

故 $\boldsymbol{c}_B = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$；

基矩阵 $\boldsymbol{A}_B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix}$。

计算基矩阵的逆 $\boldsymbol{A}_B^{-1} = \begin{bmatrix} 0 & 0.5 & -1 \\ 0 & 0.5 & 0 \\ 1 & -0.5 & 1 \end{bmatrix}$。

非基变量 s_2 的约化成本

其对应列 $\boldsymbol{A}_{s_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$，则：

$$\boldsymbol{c}_B^T \boldsymbol{A}_B^{-1} \boldsymbol{A}_{s_2} = \begin{bmatrix} 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0.5 & -1 \\ 0 & 0.5 & 0 \\ 1 & -0.5 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 1.5$$

约化成本 $\bar{c}_{s_2} = c_{s_2} - 1.5 = 0 - 1.5 = -1$。最大化问题中 $\bar{c}_{s_2} < 0$，故 s_2 进基会使目标函数减少，不适合进基。

非基变量 s_3 的约化成本

其对应列 $\boldsymbol{A}_{s_3} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$，则：

$$\boldsymbol{c}_B^T \boldsymbol{A}_B^{-1} \boldsymbol{A}_{s_3} = \begin{bmatrix} 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0.5 & -1 \\ 0 & 0.5 & 0 \\ 1 & -0.5 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = -1$$

约化成本 $\bar{c}_{s_3} = c_{s_3} - (-1) = 0 + 1 = 1 > 0$。最大化问题中 $\bar{c}_{s_3} > 0$，故 s_3 进基理论上能使目标函数增加，但实际受约束限制（如 $x_2 \leq 100$）。

所以，我们需要一条：Stopping Criterion（停止准则）

Fact or Background

最大化问题：若当前基可行解的 **所有非基变量** 的约化成本 $\bar{c}_j \leq 0$，则当前基可行解是最优解；

最小化问题：若当前基可行解的 **所有非基变量** 的约化成本 $\bar{c}_j \geq 0$，则当前基可行解是最优解。

5.3 确定步长与出基方向确定 (Moving steps)

在线性规划的迭代优化（如单纯形法）中，当找到能改进目标函数的基方向 d 后，需确定沿该方向移动的步长 θ。

Definition

步长 θ^* 是满足“移动后解仍非负”的最大非负数，即：

$$\theta^* = \max\{\theta \geq 0 : x + \theta d \geq 0\}$$

核心原则是：**在保持解可行的前提下，尽可能最大化步长，以充分改进目标函数。**

基方向 d 的分量符号决定了步长的有限性，分两种情况：

情况 1： 基方向 $d \geq 0$ 此时所有分量随 θ 增大不会变为负数，因此 $\theta^* = \infty$。同时 $\mathbf{c}^\top \mathbf{d} < 0$(对于最小化问题而言). 这说明原问题**无界**（目标函数可无限向优化方向改进）。

情况 2： 基方向存在负分量（即存在 $i \in B$，B 为基变量集合，使得 $d_i < 0$）

基变量中，若 $d_i < 0$，则 $x_i + \theta d_i$ 会随 θ 增大而减小，直到为 0。为保证解非负，需找到 **“第一个使 $x_i + \theta d_i = 0$ 的 θ”**：对所有满足 $d_i < 0$ 的基变量 i，计算 $-\frac{x_i}{d_i}$，并取其中的**最小值**作为 θ^*：

$$\theta^* = \min_{\{i \in B : d_i < 0\}} \left(-\frac{x_i}{d_i} \right)$$

当步长 θ^* 有限时（即问题有界），沿基方向移动会生成新的可行解，并伴随基变量与非基变量的替换（基的更新），推动算法迭代。具体逻辑如下：

1. 新可行解的生成

设当前可行解为 x，基方向为 d（由某个非基变量的单位向量扩展而来，满足 $Ad = 0$ 以保证可行性）。

新解定义为：

$$y = x + \theta^* \cdot d$$

其中，步长 θ^* 是确保新解非负的最大有限值，即：

$$\theta^* = \min \left\{ -\frac{x_{B(i)}}{d_{B(i)}} \mid d_{B(i)} < 0,\, i = 1, \ldots, m \right\}$$

（仅考虑基变量方向分量为负的情况，因方向分量非负时变量值随步长增大而增大，不会违反非负约束）。

2. 基变量的替换逻辑

由 θ^* 的定义，存在某个基变量索引 ℓ，使得：

$$\theta^* = -\frac{x_{B(\ell)}}{d_{B(\ell)}}$$

此时，新解中该基变量的值为：

$$y_{B(\ell)} = x_{B(\ell)} + \theta^* \cdot d_{B(\ell)} = 0$$

即原基变量 $x_{B(\ell)}$ 在新解中变为 0，失去“基变量”的本质（基变量需对应线性无关的列向量且通常取正值以保证解的非退化性），因此从基集合中移除，称为**出基**。

另一方面，引发基方向 d 的原非基变量 x_j（其方向分量 $d_j = 1$），在新解中的值为：

$$y_j = x_j + \theta^* \cdot d_j = 0 + \theta^* \cdot 1 = \theta^* > 0$$

由于其值从 0 变为正，且对应列向量与原基中其余列向量保持线性无关（基方向的构造保证了这一点），因此加入基集合，称为**入基**。

3. 新基的构成

新的基集合由“原基中除 $B(\ell)$ 外的所有基变量”与“入基变量 j”组成，即：

$$B_{\text{新}} = \{B(1), \ldots, B(\ell - 1), j, B(\ell + 1), \ldots, B(m)\}$$

其规模仍为约束维数 m，且因入基变量的列向量与原基其余列向量线性无关，新基矩阵依然可逆，可支撑下一轮迭代。

退化情形的补充说明

若当前解 x 存在基变量 $x_{B(i)} = 0$（退化解），则可能出现 $\theta^* = 0$（因 $x_{B(i)} = 0$ 时，即使 $d_{B(i)} < 0$，$-\frac{x_{B(i)}}{d_{B(i)}} = 0$）。

此时新解 $y = x$，与原解完全相同，基变量的替换（出基与入基）仅改变基的构成，但目标函数值不变。若多次迭代均出现此类情况，可能导致算法循环（始终在相同解之间切换），这是退化情形下需要特殊处理的核心问题。

Exercise

T or F? : 考虑应用于标准型线性规划（LP）问题的单纯形法，假设矩阵 $\mathbf{A}$ 的行线性无关。单纯形法的一次迭代可能在将基本可行解（BFS）移动一段正距离的同时，保持成本不变。

解答：错误。（单纯形法）进入变量 x_j 的检验数始终满足 $\bar{c}_j < 0$。此处最小比值 $\theta^* > 0$，成本的净变化为 $\bar{c}_j\theta^* < 0$，因此成本无法保持不变。

5.4 退化与退化问题解

Definition

退化性（Degeneracy）：如果某些基变量 x_B 为 0，则我们称这个基本可行解 x 是**退化的**；否则，称其为**非退化的**。

Example

考虑如下线性规划的约束与变量非负条件：

$$
\begin{cases}
x_1 + 2x_2 + s_1 \qquad\qquad = 8 \\
x_1 - x_2 \qquad + s_2 \qquad = 4 \\
-x_1 + x_2 \qquad\qquad + s_3 = 4 \\
x_1,\ x_2,\ s_1,\ s_2,\ s_3 \geq 0
\end{cases}
$$

若选择基指标为 $\{1,2,4\}$（这里 s_2 对应第 4 个变量），则对应的基变量为 $x_B = (0,4,8)$。此时，该基本可行解是**退化的**

【Remark】：

考虑标准型多面体 $\{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$，且 $\mathbf{A}$ 的行线性无关。当一个基本解是退化的时，**不一定**存在一个相邻的退化基本解。

（几何角度）例如，考虑如下标准型的多面体

$$\{(x,y) \in \mathbb{R}^2 \mid x+y \geq 0, x-y \geq 0, \quad x,y \geq 0\}.$$

该多面体仅包含一个基本解（0，0），因此不存在相邻的基本解。

对于退化的问题，再实际操作中可能会陷入“循环”，即：在迭代中不断出现重复的基。

那么为什么会形成循环？

其原因大致有 2 点：

Fact or Background

- 退化解可能使得目标函数的值无法变更；
- 不同的基可能对应同一个基本可行解

【Remark】：

该 fact 论述的是：**“不同基，但是同一 BFS”**，**而非：“同一基故唯一 BFS”**，二者的逻辑前提不同，因此可以同时成立。

我们知道：对所有满足 $d_i < 0$ 的基变量 i，计算 $-\frac{x_i}{d_i}$，并取其中的最小值作为 θ^*：

$$\theta^* = \min_{\{i \in B: d_i < 0\}} \left(-\frac{x_i}{d_i}\right)$$

由于出基变量 $x_{Bk}=0$，此时 $\theta=0$（因为 $x_{Bk}/a_{kj}=0/a_{kj}=0$，且是最小值），因此进基变量 x_j 的新取值为 $\theta=0$；

进基变量 x_j 因为步长为 0，使用即使加入了也没有前进，对目标函数的贡献是 0，加入它也不会增加目标函数值。因此，替换前后目标函数值都是 $\sum c_{Bi}x_{Bi}$（只是去掉了 0 贡献的 x_{Bk}，加入了 0 贡献的 x_j），总和不变。

而两个基 $\{x_1,x_2,x_3\}$ 和 $\{x_1,x_4,x_3\}$ 不同，但对应的解向量 $(x_1,x_2,x_3,x_4,x_5)=(5,0,3,0,0)$ 完全相同。

这种“基不同但解相同”的现象，**会让单纯形法误以为每次迭代都在探索新的解**（因为基变了），但实际上解的数值没变，目标函数也没变，**最终可能在多个基之间循环**。

Bland 规则的应用步骤：为避免循环，我们使用 Bland 规则，规定：

Method

进基变量选择：在所有符合改进条件的非基变量中（如检验数 $\bar{c}_j<0$ 的最大化问题），选择 $\min\{j\mid\bar{c}_j<0\}$。
出基变量选择：在计算最小比值 $\theta=\min\{x_{B(i)}/d_i\mid d_i>0\}$ 时，若多个基变量比值相同，选择 $\min\{i\mid x_{B(i)}/d_i=\theta\}$。

即进基变量和出基变量的选择都遵循“**最小下标优先**”的约束：

迭代过程对比（有无 Bland 规则）

无 Bland 规则时：

若随意选择进/出基变量，可能出现如下循环：

$$\{x_4,x_5,x_6\}\to\{x_1,x_5,x_6\}\to\{x_1,x_2,x_6\}\to\{x_4,x_2,x_6\}\to\{x_4,x_5,x_6\}$$

目标函数值始终不变，陷入死循环。

有 Bland 规则时：

后续迭代始终按“最小下标”原则选择，基序列严格递增：

$$\{x_4,x_5,x_6\}\to\{x_1,x_5,x_6\}\to\{x_1,x_2,x_6\}\to\{x_1,x_2,x_3\}$$

最终收敛到最优解，无循环。

适用场景： Bland 规则适用于所有存在退化基本可行解的线性规划问题，尤其在目标函数值迭代中不改进时。

5.5 初始化

在我们了解了其运行过程后，我们又面临一个问题：在大量的基解中，如何选择初始的基可行解？

我们提出了一种普适的构造法: **两阶段单纯形法**

核心目的

单纯形法的迭代必须从**基本可行解（BFS）**开始，但很多时候初始 BFS 无法直接获得（如约束右端项有负数、无明显松弛变量构成基等）。两阶段单纯形法通过“分阶段处理”，先解决 **“初始 BFS 获取”** 和 **“原问题可行性判定”**，再求解原问题，是处理复杂线性规划的通用方法。

若 $b > 0$，则非常容易找到一组 BFS,

Example

$$\begin{cases} \max\ 2x_1 + 3x_2 \\ \text{约束条件：}\quad x_1 + x_2 + x_3 = 4 \quad (\text{松弛变量}x_3 \geq 0) \\ \qquad x_1 + x_4 = 3 \quad (\text{松弛变量}x_4 \geq 0) \\ \qquad x_1, x_2, x_3, x_4 \geq 0 \end{cases}$$

这里:

$\mathbf{b} = \begin{bmatrix} 4 \\ 3 \end{bmatrix} \geq \mathbf{0}$（向量 $\mathbf{b}$ 的元素都是非负的）。

x_3、x_4 对应的列组成了单位矩阵 $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$（这是一组有效的基）。

要找一个基本可行解（BFS）:

1. 令非基变量 $x_1 = 0$，$x_2 = 0$。

2. 求解基变量 x_3、x_4:

- 由 $0 + 0 + x_3 = 4$，得 $x_3 = 4$;

- 由 $0 + x_4 = 3$，得 $x_4 = 3$。

结果: $\mathbf{x} = (0, 0, 4, 3)^\top$，它是可行的（所有变量都非负），也是基本的（基由 x_3、x_4 组成）。

这种情况能顺利找到初始 BFS，是因为 $\mathbf{b} \geq \mathbf{0}$，所以基变量 $\mathbf{x}_B = \mathbf{b}$ 天生就满足非负性。

但是如果 $b_i < 0$ 呢?

Method

第一阶段：构造辅助问题，判定可行性并找初始 BFS

当原问题无法直接找到初始 BFS 时，如 $b_i < 0$, 或约束系数矩阵 A 无现成单位矩阵作为基，需构造**辅助问题**，通过求解辅助问题完成两件事:

1. 判断原问题是否可行;

2. 若可行，得到原问题的初始 BFS。

Step 1：处理负右端项：若原问题中存在 $b_i < 0$，对该约束两端乘 -1，使所有 $b_i \geq 0$（保证后续人工变量能构成可行解）。

Step 2：引入人工变量（auxiliary variables）：对**每个约束**添加一个**人工变量** $y_i \geq 0$（共 m 个），构造辅助问题的约束：

$$a_{i1}x_1 + \cdots + a_{in}x_n + y_i = b_i \quad (i = 1, 2, \ldots, m)$$

此时，辅助问题的系数矩阵中，人工变量 $y_1, \ldots, y_m$ 对应的列向量构成单位矩阵:（y_i 列只有第 i 行为 1，其余为 0），可直接作为初始基变量。

Step 3：辅助问题的目标函数：辅助问题以“最小化人工变量之和”为目标（目的是让人工变量尽可能为 0，若能全为 0，说明原问题有可行解）：

$$\min \quad w = y_1 + y_2 + \cdots + y_m$$

辅助问题的初始 BFS 与求解

辅助问题的初始 BFS：直接取 $x_1 = \cdots = x_n = 0$（非基变量），$y_i = b_i$（基变量），因 $b_i \geq 0$，故该解可行。

求解方法：用单纯形法求解辅助问题（按最小化目标迭代，检验数判断是否最优）。

原问题可行性的判定：

辅助问题的最优值 w^* 决定原问题是否可行：

若 $w^* > 0$：说明无论如何调整 $x_1, \ldots, x_n$，人工变量都无法全为 0（至少有一个 $y_i > 0$），原问题**无可行解**，算法终止。

若 $w^* = 0$：说明存在 $x_1, \ldots, x_n \geq 0$ 使得 $y_1 = \cdots = y_m = 0$（此时 x 满足原问题所有约束），原问题**可行**。此时，从辅助问题的最优解中提取 x 的值，作为原问题的初始可行解。

从辅助问题到原问题的初始 BFS

当 $w^* = 0$ 时，辅助问题的最优解中 $y_1 = \cdots = y_m = 0$，但需进一步判断 x 是否为原问题的 BFS：

情况 1：人工变量均为非基变量

此时基变量全是原问题的变量（x 中的一部分），且基变量个数为 m（满秩），直接作为原问题的初始 BFS。

情况 2：仍有部分人工变量是基变量（但值为 0，因 $w^* = 0$）

这是退化情形（基变量值为 0），需替换这些人工基变量：从原问题的非基变量中选一个，其对应列向量与当前基中其他向量线性无关，将人工变量换出基，直至基变量全为原问题的变量，得到初始 BFS。

第一阶段的核心是“通过最小化人工变量和”判断原问题是否有可行解，本质是用单纯形法求解一个“可行性问题”。

Method

第二阶段：用初始 BFS 求解原问题

若第一阶段判定原问题可行，且已获得初始 BFS，则进入第二阶段，用单纯形法求解原问题（按原目标函数迭代）。

【Remark】:

人工变量仅用于第一阶段“临时构造基”，第二阶段必须剔除（或替换）。

与此同时，我们还有一种方法——**Big-M 方法（大 M 法）：**

为解决线性规划（LP）**无初始基本可行解（BFS）**时的求解问题，通过引入“惩罚因子 M”构造辅助问题，**在单阶段内同时完成“可行性判定”与“原问题求解”**。

Method

辅助问题构造: 对原线性规划问题（标准型，若 b 含负分量，对对应约束行乘 -1 使 $b \geq 0$），构造辅助问题：

$$\begin{cases} \min \quad c^\top x + M\sum_{i=1}^{m} y_i \\ \text{subject to} \quad Ax + y = b \\ x, y \geq 0 \end{cases}$$

其中：

$y = (y_1, y_2, \ldots, y_m)^\top$：**人工变量**，用于构造初始 BFS；

M：“足够大的正数”（惩罚因子），对人工变量出现在基中施加“惩罚”，迫使最优解中人工变量尽可能为 0。

辅助问题的初始 BFS 可直接取 $x = 0$，$y = b$。由于假设 $b \geq 0$，故 $y \geq 0$，满足可行性。

求解逻辑（单纯形法迭代与可行性判定）：

- **原问题可行的情况**: 若原问题存在可行解 $\hat{x} \geq 0$ 满足 $A\hat{x} = b$，则辅助问题中 $\hat{y} = 0$ 可满足约束 $A\hat{x} + \hat{y} = b$。由于 M 是极大正数，单纯形迭代会“极力避免”人工变量 y_i

留在基中（否则目标函数值会因 M 被大幅拉高，无法最小化）。

因此，迭代过程中人工变量会被逐步换出基，最终最优解中 $\hat{y}=0$。此时提取 $\hat{x}$，它既满足原问题约束 $A\hat{x}=b$ 且 $\hat{x}\geq 0$，可作为原问题的初始可行解，继续用单纯形法迭代即可求原问题最优解。

- **原问题不可行的情况**：若原问题无可行解，则不存在任何 $x\geq 0$ 使得 $Ax=b$。此时约束 $Ax+y=b$ 必须依赖人工变量 y 的正分量才能满足，即最优解中 $\hat{y}\neq 0$（因无法让所有 $y_i=0$ 同时满足约束）。由此可判定：原问题无可行解。

【Remark】：

M 需取“足够大的正数”（远大于原目标函数中其他系数的量级），但“足够大”的程度依赖经验。若 M 取值不当，可能因计算机浮点数精度限制引入数值误差，影响迭代结果。

与两阶段单纯形法的对比

Method	Big-M 法	两阶段法
人工变量处理	用大正数 M 惩罚人工变量，迫使为 0	最小化人工变量和，消去人工变量
计算复杂度	单阶段但 M 取值易引发数值问题	两阶段数值稳定性好
可行性判定	最优解人工变量是否为 0	一阶段辅助问题最优值是否为 0
应用场景	教学易理解，实际因数值问题少用	更通用，单纯形表实现中常用
对退化的敏感性	M 可能放大退化导致的循环风险	结合 Bland 规则可有效避免循环

总结：Big-M 法通过“惩罚因子”单阶段解决问题，概念简洁但存在数值缺陷；两阶段法分阶段处理，逻辑更清晰、数值稳定性更强，是实际应用中更常用的方法。

6 单纯形表（The Simplex Tableau）

6.1 单纯性表的结构

理解单纯形表的每个部分：

$\boldsymbol{c}^\top-\boldsymbol{c}_B^\top\boldsymbol{A}_B^{-1}\boldsymbol{A}$	$-\boldsymbol{c}_B^\top\boldsymbol{A}_B^{-1}\boldsymbol{b}$
$\boldsymbol{A}_B^{-1}\boldsymbol{A}$	$\boldsymbol{A}_B^{-1}\boldsymbol{b}$

- 左上角是检验数向量。
- 右上角 $\boldsymbol{c}_B^\top\boldsymbol{A}_B^{-1}\boldsymbol{b}=\boldsymbol{c}_B^\top x_B$ 是目标函数值的**相反数**，在标准 LP 下。
- 下方部分是变换后的约束及对应的基变量。
- 结合后续知识我们还知道：对偶问题的最优解就是原松弛变量的 reduced cost

- 可以在表中找到 $\boldsymbol{A}_B^{-1}\boldsymbol{A}$，也通常能得到 $\boldsymbol{A}_B^{-1}$。

我们仅仅以最小化问题为例：
从最小化的标准型（canonical form）为例，单纯形表的每次迭代操作包括：

1. 判定最优性：查看第一行（检验数）：
   - 若检验数均非负，则当前基本可行解（BFS）已最优。
   - 否则，选择含最负检验数的下标 i 使得 $\bar{c}_i < 0$ 作为入基（主元列）或者选择最小下标 j 使得 $\bar{c}_i < 0$ 作为入基。

   【Remark】：

   选最负检验数：追求更快收敛
   使用 *Bland's rule*：避免循环

2. 执行**最小比值检验**（仅在当前基变量与入基列的正元素间进行）。通过最小比值检验我们就知道进基的**限度**。
3. 选择达到最小比值的行作为**主元行（Pivot row）**，并确定出基（最小下标规则）。
4. 更新表格：
   - 将**主元行**的每个元素除以主元元素。
   - 对其他每行添加主元行的适当倍数，使得**主元列（Pivot column）**的所有其他元素（包括第一行）变为零。

在单纯形表操作过程中，我们始终维护标准型（canonical form），从开始到结束。

在上一节我们已经知道：

Claim

基变量的检验数恒为 0

所以单纯形表在对列按 $[B;N]$ **重排**后应呈现如下形式：

$\boldsymbol{0}_m^\top$	$\boldsymbol{c}_N^\top - \boldsymbol{c}_B^\top\boldsymbol{A}_B^{-1}\boldsymbol{A}_N$	$-\boldsymbol{c}_B^\top\boldsymbol{x}_B$
$\boldsymbol{I}_m$	$\boldsymbol{A}_B^{-1}\boldsymbol{A}_N$	$\boldsymbol{x}_B$

- 基变量的约束矩阵（不一定是前 m 列）是单位矩阵。
- 基变量的检验数为零。

由此可轻松判断当前基本可行解（BFS）是否最优。

6.2 单纯性表的构造及其数学原理

给出单纯形表在对列按 $[B;N]$ **重排**后应呈现如下形式：

$\mathbf{0}_m^\top$	$\boldsymbol{c}_N^\top - \boldsymbol{c}_B^\top \boldsymbol{A}_B^{-1}\boldsymbol{A}_N$	$-\boldsymbol{c}_B^\top \boldsymbol{x}_B$
$\boldsymbol{I}_m$	$\boldsymbol{A}_B^{-1}\boldsymbol{A}_N$	$\boldsymbol{x}_B$

Fact or Background

单纯形表中，除了第一行（目标函数行 / 检验数行）的 RHS（目标函数值 Z_0）可以小于 0 外，其余所有行（对应基变量的约束行）的 RHS，在有效迭代的基本可行解（BFS）阶段，都不能小于 0。

证明. 约束行 RHS = 基变量取值 → BFS 要求基变量非负 → RHS ≥0; □

Fact or Background

单纯形法的迭代核心：通过初等行变换（转轴操作）迭代更新基矩阵，使基矩阵经变换后呈现单位矩阵形式，同时通过检验数判断最优性，逐步逼近最优解。

我们知道：虽然 simplex method 给出了基变量的表达式，但直接计算 $\boldsymbol{A}_B^{-1}$ 和 $\boldsymbol{A}_B^{-1}\mathbf{A}_N$ 可能很麻烦。

单纯形表的巧妙之处在于：通过初等行变换，把基矩阵 $\boldsymbol{A}_B$ 变成单位矩阵 $\boldsymbol{I}$，这样基变量的解可以直接"读"出来。

若问题含"≤"约束且 $\boldsymbol{b} \geq \mathbf{0}$，选松弛变量构成初始基矩阵 $\boldsymbol{A}_B = \boldsymbol{I}$（单位矩阵），此时基可行解 $\boldsymbol{x}_B = \boldsymbol{b}$（非基变量 $\boldsymbol{x}_N = \mathbf{0}$）。（根据定义：当 $\boldsymbol{x}_N = 0$ 时，即约束内的非基变量为 0，我们得到基变量：$\boldsymbol{x}_B = \boldsymbol{b}$）

这就是单纯形表的初始阶段：

接下来我们展示其迭代的过程和原理：初等行变换可以用一个矩阵来表示：假设所有行变换的总效果等价于矩阵 $\boldsymbol{E}$ 左乘 $\boldsymbol{A}_B$，即：$\boldsymbol{E} \cdot \boldsymbol{A}_B = \boldsymbol{I}$

根据逆矩阵的定义：若 $\boldsymbol{E} \cdot \boldsymbol{A}_B = \boldsymbol{I}$，则 $\boldsymbol{E}$ 就是 $\boldsymbol{A}_B$ 的逆矩阵，即 $\boldsymbol{E} = \boldsymbol{A}_B^{-1}$。

【Remark】：

此时的 A_B 是最终阶段的 final table 的基向量

我们不仅要对基矩阵 $\boldsymbol{A}_B$ 做行变换，还要对整个约束矩阵 $\boldsymbol{A}$（包含 A_N 和 A_B）和右端项 $\boldsymbol{b}$ 做相同的行变换（因为约束方程是等式，两边要同时变形）。对式两边左乘 $\boldsymbol{E} = \boldsymbol{A}_B^{-1}$：$\boldsymbol{E} \cdot (\boldsymbol{A}_B\boldsymbol{x}_B + \boldsymbol{A}_N\boldsymbol{x}_N) = \boldsymbol{E} \cdot \boldsymbol{b}$

展开后：$(\boldsymbol{E}\boldsymbol{A}_B)\boldsymbol{x}_B + (\boldsymbol{E}\boldsymbol{A}_N)\boldsymbol{x}_N = \boldsymbol{E}\boldsymbol{b}$

代入 $\boldsymbol{E}\boldsymbol{A}_B = \boldsymbol{I}$ 和 $\boldsymbol{E} = \boldsymbol{A}_B^{-1}$，得到：$\boldsymbol{I}\boldsymbol{x}_B + (\boldsymbol{A}_B^{-1}\mathbf{A}_N)\boldsymbol{x}_N = \boldsymbol{A}_B^{-1}\boldsymbol{b}$

这就是行变换后的约束方程，此时**基矩阵**（final 的基）已经变成单位矩阵 $\boldsymbol{I}$，非基变量的系数矩阵变成 $\boldsymbol{A}_B^{-1}\boldsymbol{A}_N$，右端项变成 $\boldsymbol{A}_B^{-1}\boldsymbol{b}$（即基可行解）——这正是单纯形表中"约束行"的结构！

由此，我们可以捕捉到：

Fact or Background

在 $b \geq 0$ 的情况下，初始的基变量为松弛变量，构成单位阵；

- 在经过最终迭代后，得到最优解与最终基变量后，得到 $\boldsymbol{E} = \boldsymbol{A}_B^{-1}$；
- 原先由松弛变量构成的单位阵经过变换：$\boldsymbol{E} = \boldsymbol{A}_B^{-1}$ 后，变成 $\boldsymbol{A}_B^{-1}$。
- 同样的，$\boldsymbol{A}_B^{-1}\boldsymbol{A}_N$ 就是 final 表格中非基变量中读到的

所以我们可以直接从单纯形表中读出 $\boldsymbol{A}_B^{-1}$ 和 $\boldsymbol{A}_B^{-1}\boldsymbol{A}_N$。

【Remark】:

注意不要混淆，$\boldsymbol{A}_B^{-1}\boldsymbol{A}_N$ 是通过考察 final 表格的非基变量得到；而 $\boldsymbol{A}_B^{-1}$ 是考察原来松弛变量所在的单位阵位置在 final 表格周能够得到的。
并且 $\boldsymbol{A}_B^{-1}\boldsymbol{A}_N$ 不等于 $\boldsymbol{A}_B^{-1}\boldsymbol{A}$

Fact or Background

单纯形法的迭代核心：**min Ratio Test**

证明. 我们知道，类似于矩阵运算与方程组的关系，初等行变换并不会改变其核心的约束关系，只是通过恒等变换将约束关系以另一种形式呈现出来。

因此经过迭代后的单纯形表的约束就是实际约束关系，并且我们可以通过最小比值判定来确定进基的步长。 □

例如：

Example

$2x_1 + x_3 = 5, ratio = 2.5$

$2x_1 + x_4 = 4, ratio = 2$

在这个例子中，x_1 需要进基，且我们理想化当然是让他变的无穷大，但是因为其它的 x_i 都要大于 0，所以 x_1 的变大程度就会被限制，而 ratio 则是最大的步长。为了让所有都满足，我们取小的那一个 2，所以 x_4 最终变为 0，出基。

即：若 x_1 为进基变量（检验数最负），需保证 $x_3, x_4 \geq 0$。x_1 增大时，$x_3 = 5 - 2x_1 \geq 0 \implies x_1 \leq 2.5$；$x_4 = 4 - 2x_1 \geq 0 \implies x_1 \leq 2$。取最小比值 2，故 x_4 出基，x_1 进基，保证基变量非负。

相反的，如果我们遇到这种情况：

考虑某 maximization 问题的单纯形表片段，进基列（枢列）所有系数非正：

$$\begin{cases} -x_1 + x_3 = 5 \\ -2x_1 + x_4 = 3 \end{cases}$$

因为所有的 x_i 都要大于 0，在这种情况下 x_1 就可以无限增大，如果 pivot col 每一个都是负数开头（标准型下），那么就会造成原问题无界的情境。

若 x_1 为**唯一可行的进基变量**（检验数为正，需增大），则 $x_3 = 5 + x_1$（可无限增大），$x_4 = 3 + 2x_1$（可无限增大）。此时 x_1 无增长上限，原问题无界。

如果结合后续的知识，我们还知道：

Fact or Background

在标准最小化问题（约束为"≤"，引入松弛变量）中，**对偶问题的最优解$\mathbf{y}^*$** 等于原问题松弛变量的 Reduced Cost（检验数 $z_j - c_j$）。

证明. 对于松弛变量 $\mathbf{s}$，其目标系数 $c_s = 0$，且对应的系数列是单位矩阵的列（即 $\mathbf{A}_s = \mathbf{I}$）。

因此，松弛变量的检验数为：$z_s - c_s = \mathbf{y}^T\mathbf{I} - 0 = \mathbf{y}$ □

6.3 单纯形表的重要示例

我们将以三个例子为例，进行具体解读：

Example

考虑如下线性规划问题：

$$\begin{cases} \textit{maximize} & x_1+2x_2+3x_3+8x_4 \\ \text{subject to} & x_1-x_2+x_3\leq 2 \\ & x_3-x_4\leq 1 \\ & 2x_2+3x_3+4x_4\leq 8 \\ & x_1,x_2,x_3,x_4\geq 0. \end{cases}$$

使用单纯形表彻底求解该问题。

我们首先将问题转化为**标准型**：

$$\begin{cases} \textit{minimize} & -x_1-2x_2-3x_3-8x_4 \\ \text{subject to} & x_1-x_2+x_3+s_1=2 \\ & x_3-x_4+s_2=1 \\ & 2x_2+3x_3+4x_4+s_3=8 \\ & x_1,x_2,x_3,x_4,s_1,s_2,s_3\geq 0. \end{cases}$$

由于右端项非负，我们可以选择松弛变量作为基变量，令原变量为零，即点 $(x;s)=(0,0,0,0,2,1,8)^\top$ 是一个有效的初始基本可行解（BFS），基为 $B=\{5,6,7\}$（其中 s_1,s_2,s_3 分别对应下标 5,6,7）。

此时任意的 x_i 都是为 0. 只有松弛变量为基向量，那么任意的 x_ix 前面的系数就是加入这个元素后目标函数变化的值：所以我们得出第一个初始单纯形表：

B	-1	-2	-3	-8	0	0	0	0
5	1	-1	1	0	1	0	0	2
6	0	0	1	-1	0	1	0	1
7	0	2	3	4	0	0	1	8

【Remark】：

左第一行不是系数，而是约化成本，只是在无人工变量的初始表中恰好是系数。

当前目标函数值为 0。如果我们采用的是 bland rule，那么 pivot 列是 $\{1\}$；pivot 行是 $\{5\}$；pivot 元素是 1；行更新后得到新表：

基更新为 $\{1,6,7\}$（易忘）；当前 BFS 是 $(x;s)=(2,0,0,0,0,1,8)^\top$；当前目标函数值为 -2。在更新后的表中，pivot 列是 $\{2\}$；pivot 行是 $\{7\}$；pivot 元素是 2；行更新后新表为：

B	0	-3	-2	-8	1	0	0	2
1	1	-1	1	0	1	0	0	2
6	0	0	1	-1	0	1	0	1
7	0	2	3	4	0	0	1	8

B	0	0	$\frac{5}{2}$	-2	1	0	$\frac{3}{2}$	14
1	1	0	$\frac{5}{2}$	2	1	0	$\frac{1}{2}$	6
6	0	0	1	-1	0	1	0	1
2	0	1	$\frac{3}{2}$	2	0	0	$\frac{1}{2}$	4

基更新为 $\{1,6,2\}$；当前 BFS 是 $(x;s)=(6,4,0,0,0,1,0)^\top$；当前目标函数值为 -14。在新表中，pivot 列是 $\{4\}$；pivot 行是 $\{2\}$；pivot 元素是 2；行更新后新表为：

B	0	1	4	0	1	0	2	18
1	1	-1	1	0	1	0	0	2
6	0	$\frac{1}{2}$	$\frac{7}{4}$	0	0	1	$\frac{1}{4}$	3
4	0	$\frac{1}{2}$	$\frac{3}{4}$	1	0	0	$\frac{1}{4}$	2

由于所有约化成本都非负，方法停止，最优解为 $\boxed{x}^* = (2,0,0,2)^\top$，$\boxed{s}^* = (0,3,0)^\top$；最优基是 $B=\{1,4,6\}$，原问题的最优值为 18。

下面是两阶段法的单纯性表：

Example

应用两阶段单纯形法（通过单纯形表实现）求解如下线性规划问题。

$$\begin{cases} \textit{minimize} & x_1 - x_2 + 2x_3 \\ \text{subject to} & 2x_1 - x_2 + 2x_3 \leq -1 \\ & x_1 - x_2 - x_3 \leq 4 \\ & x_2 - x_4 = 0 \\ & x_1, x_2, x_3, x_4 \geq 0. \end{cases}$$

我们按照流程首先将问题转化为标准型：

$$\begin{cases} \textit{minimize} & x_1 - x_2 + 2x_3 \\ \text{subject to} & 2x_1 - x_2 + 2x_3 + s_1 = -1 \\ & x_1 - x_2 - x_3 + s_2 = 4 \\ & x_2 - x_4 = 0 \\ & x_1, x_2, x_3, x_4, s_1, s_2 \geq 0. \end{cases}$$

注意，对于该问题，可直接找到初始基本可行解（BFS）：令 $(x; s) = (0, 1, 0, 1, 0, 5)^\top$，基为 $B = \{2, 4, 6\}$（其中 x_2, x_4, s_2 分别对应下标 2,4,6）。但是为了练习两阶段法，我们还是用两阶段发单纯形表练习：

第一阶段（Phase I）

构造辅助问题以寻找原问题的初始 BFS：

$$\begin{cases} \textit{minimize} & y_1 + y_2 + y_3 \\ \text{subject to} & -2x_1 + x_2 - 2x_3 - s_1 + y_1 = 1 \\ & x_1 - x_2 - x_3 + s_2 + y_2 = 4 \\ & x_2 - x_4 + y_3 = 0 \\ & x_1, x_2, x_3, x_4, s_1, s_2, y_1, y_2, y_3 \geq 0, \end{cases}$$

【Remark】：

保证等式右端都大于 0.

其中对第一个等式约束缩放以得到非负右端项。辅助问题的初始 BFS 为 $(x; s; y) = (0, 0, 0, 0, 0, 0, 1, 4, 0)^\top$，基为 $B = \{7, 8, 9\}$（其中 y_1, y_2, y_3 分别对应下标 7,8,9）。

非基变量的约化成本计算如下：

$$(0,0,0,0,0,0) - (1,1,1)\begin{bmatrix} -2 & 1 & -2 & 0 & -1 & 0 \\ 1 & -1 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 & 0 & 0 \end{bmatrix} = (1, -1, 3, 1, 1, -1)$$

Method

当然为了让基向量的约化成本为 0（符合一般状况），我们也可以把 $y_1 + y_2 + y_3$ 用 x_i, s_i 表示出来，然后利用初始化的特殊情境，利用系数来求解其约化成本。

但是，注意常数项如果在左边，需要移到右边的 RHS。而且往往需要变号才能使用（易错）

对应的初始单纯形表为：

B	1	-1	3	1	1	-1	0	0	0	-5
7	-2	1	-2	0	-1	0	1	0	0	1
8	1	-1	-1	0	0	1	0	1	0	4
9	0	1	0	-1	0	0	0	0	1	0

- pivot 列：$\{2\}$；出基列：$\{9\}$；pivot 元素：1；当前 BFS 退化。- 行更新后新表：

B	1	0	3	0	1	-1	0	0	1	-5
7	-2	0	-2	1	-1	0	1	0	-1	1
8	1	0	-1	-1	0	1	0	1	1	4
2	0	1	0	-1	0	0	0	0	1	0

- pivot 列：$\{6\}$；出基列：$\{8\}$；pivot 元素：1；行更新后新表：

B	2	0	2	-1	1	0	0	1	2	-1
7	-2	0	-2	1	-1	0	1	0	-1	1
6	1	0	-1	-1	0	1	0	1	1	4
2	0	1	0	-1	0	0	0	0	1	0

- pivot 列：$\{4\}$；出基列：$\{7\}$；pivot 元素：1；行更新后新表：

由于所有约化成本非负，第一阶段最优值为 0，故 $(x; s) = (0, 1, 0, 1, 0, 5)^\top$ 是原问题的 BFS，基为 $B = \{2, 4, 6\}$。

第二阶段（Phase II）

利用第一阶段最终表构造第二阶段初始表，计算约化成本：

回忆单纯形表的每个部分：

$\boldsymbol{c}^\top - \boldsymbol{c}_B^\top \boldsymbol{A}_B^{-1} \boldsymbol{A}$	$-\boldsymbol{c}_B^\top \boldsymbol{A}_B^{-1} \boldsymbol{b}$
$\boldsymbol{A}_B^{-1} \boldsymbol{A}$	$\boldsymbol{A}_B^{-1} \boldsymbol{b}$

B	0	0	0	0	0	0	1	1	1	0
4	-2	0	-2	1	-1	0	1	0	-1	1
6	-1	0	-3	0	-1	1	1	1	0	5
2	-2	1	-2	0	-1	0	1	0	0	1

于是在经过对**列按照顺序的重排**后，我们有：

$$\begin{aligned}
\bar{c}^\top &= c^\top - c_B^\top A_B^{-1} A \\
&= c^\top - (-1,0,0) \begin{bmatrix} -2 & 1 & -2 & 0 & -1 & 0 \\ -2 & 0 & -2 & 1 & -1 & 0 \\ -1 & 0 & -3 & 0 & -1 & 1 \end{bmatrix} \\
&= (-1,0,0,0,-1,0)
\end{aligned}$$

初始单纯形表为：

【Remark】：

是 $\bar{c}^\top = c^\top - c_B^\top A_B^{-1} A$，而非：$\bar{c}^\top = c^\top - c_B^\top A_B^{-1} A_N$

B	-1	0	0	0	-1	0	1
2	-2	1	-2	0	-1	0	1
4	-2	0	-2	1	-1	0	1
6	-1	0	-3	0	-1	1	5

由于 pivot 列 $\{1\}$ 与 pivot 列 $\{5\}$ 中所有元素均为负，原问题无界。

我们再考察一个退化的情形：

Example

使用两阶段单纯形法（通过单纯形表实现）彻底求解如下线性优化问题。

$$\begin{cases}
\text{minimize} & x_1 + 3x_2 + x_4 - 2x_5 \\
\text{subject to} & x_1 + 2x_2 + 4x_4 + x_5 = 2 \\
& x_1 + 2x_2 - 2x_4 + x_5 = 2 \\
& -x_1 - 4x_2 + 3x_3 = 1 \\
& x_1, x_2, x_3, x_4, x_5 \geq 0.
\end{cases}$$

该问题已为标准型，直接进入第一阶段（Phase I）。

第一阶段：构造并求解辅助问题

构造辅助问题以寻找原问题的初始基本可行解（BFS）：

$$\begin{cases} \text{minimize} & y_1 + y_2 + y_3 \\ \text{subject to} & x_1 + 2x_2 + 4x_4 + x_5 + y_1 = 2 \\ & x_1 + 2x_2 - 2x_4 + x_5 + y_2 = 2 \\ & -x_1 - 4x_2 + 3x_3 + y_3 = 1 \\ & x_1, x_2, x_3, x_4, x_5, y_1, y_2, y_3 \geq 0, \end{cases}$$

其中 y_1, y_2, y_3 为人工变量（对应下标 6,7,8）。

- 辅助问题的初始 BFS：$(x; y) = (0,0,0,0,0,2,2,1)^\top$，基为 $B = \{6,7,8\}$。
- 非基变量的约化成本计算：

$$(0,0,0,0,0) - (1,1,1)\begin{bmatrix} 1 & 2 & 0 & 4 & 1 \\ 1 & 2 & 0 & -2 & 1 \\ -1 & -4 & 3 & 0 & 0 \end{bmatrix} = (-1,0,-3,-2,-2)$$

- 初始单纯形表：

B	-1	0	-3	-2	-2	0	0	0	-5
6	1	2	0	4	1	1	0	0	2
7	1	2	0	-2	1	0	1	0	2
8	-1	-4	3	0	0	0	0	1	1

第一步迭代： pivot 列 $\{1\}$，出基列 $\{6\}$，pivot 元素 1；行更新后新表：

B	0	2	-3	2	-1	1	0	0	-3
1	1	2	0	4	1	1	0	0	2
7	0	0	0	-6	0	-1	1	0	0
8	0	-2	3	4	1	1	0	1	3

第二步迭代：pivot 列 $\{3\}$，出基列 $\{8\}$，pivot 元素 3；行更新后新表：

B	0	0	0	6	0	2	0	1	0
1	1	2	0	4	1	1	0	0	2
7	0	0	0	-6	0	-1	1	0	0
3	0	$-\frac{2}{3}$	1	$\frac{4}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{1}{3}$	1

由于所有约化成本非负且目标函数值为 0，第一阶段停止，得到原问题的 BFS：$(x; y) = (2,0,1,0,0,0,0,0)^\top$，基为 $B = \{1,3,7\}$。

替换人工变量（基的调整）

当前基 $B=\{1,3,7\}$ 仍含人工变量 y_2（下标 7），需用 x_4 替换。

【Remark】：

注意这里是基中含有人工变量，即在 1,3,7 中含有 7，因此退化；
而非在第一行的 reduce cost 里面的 7，8，9 三列的 reduce cost 非零。

那么如何找到新的列来替换呢？

Method

识别方法：找基列中在 $x_i=0$ 行（人工变量“7”行）有非零元素的列，进行替换。

证明. 只有在这一行非零的列，其代表的非基变量才有可能对人工变量构成约束；若为 0，则不任何构成约束。 □

在上述的例子中，仅第 4 列（元素 -6）满足（第 2、5 列该行元素为 0）。为保持标准型，对第 4 列进行行更新，将其变为单位向量（基从 $\{7\}$ 更新为 $\{4\}$），得到：

B	x_1	x_2	x_3	x_4	x_5	x_6	y_1	y_2	RHS
Z	0	0	0	0	0	1	1	1	0
1	1	1	2	0	4	1	1	0	2
3	0	$-\frac{2}{3}$	1	0	$\frac{1}{3}$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$	1
4	0	0	0	1	0	$\frac{1}{6}$	$-\frac{1}{6}$	0	0

所以：

$$A_B^{-1}A=\begin{bmatrix}1&2&0&0&1\\0&-\frac{2}{3}&1&0&\frac{1}{3}\\0&0&0&1&0\end{bmatrix},\quad c^\top-c_B^\top A_B^{-1}A=(0,1,0,0,-3)$$

【Remark】：

是 $\bar{c}^\top=c^\top-c_B^\top A_B^{-1}A$，而非：$\bar{c}^\top=c^\top-c_B^\top A_B^{-1}A_N$

第二阶段（Phase II）：求解原问题

令基 $B=\{1,3,4\}$，初始目标函数值为 $-c_B^\top x_B=-2$（因为我们把它转换成最小化问题了，所以肯定是负数）。基于第一阶段更新后的表，构造原问题初始单纯形表：

【Remark】：

初始目标函数值的计算核心是 $c_B^\top x_B=2$，而“$-c_B^\top x_B=-2$”是单纯形表表示习惯导致的符号呈现，其本质是对原目标函数值的等价变形。

迭代步骤：pivot 列 $\{5\}$，出基列 $\{1\}$，pivot 元素 1；行更新后新表：

B	0	1	0	0	-3	-2
1	1	2	0	0	1	2
4	0	0	0	1	0	0
3	0	$-\frac{2}{3}$	1	0	$\frac{1}{3}$	1

B	3	7	0	0	0	4
5	1	2	0	0	1	2
4	0	0	0	1	0	0
3	$-\frac{1}{3}$	$-\frac{4}{3}$	1	0	0	$\frac{1}{3}$

所有约化成本非负，方法停止。最优解为 $x = \left(0, 0, \frac{1}{3}, 0, 2\right)^\top$，最优值为 -4。

针对退化的情境，我们还可以举一个例子加深印象：

B	0	0	1	2/3	4/3	7/3	0	0
1	1	0	1	$-1/3$	1/3	1/3	0	4
2	0	1	1/2	1/6	$-1/6$	1/3	0	1
7	0	0	-1	$-2/3$	$-1/3$	$-4/3$	1	0

在这个例子中，我们通过执行行运算，用 $x3$ 来替换它（使用 $x4$ 也是可以的），

B	0	0	0	0	1	1	1	0
1	1	0	0	-1	0	-1	1	4
2	0	1	0	$-1/6$	$-1/3$	$-1/3$	1/2	1
3	0	0	1	2/3	1/3	4/3	-1	0

7 影子价格

我们以一个工厂生产的例子作为起手：

Plant1（门框车间）：只负责做“门框”。生产 1 扇门，必须先在这个车间花 1 小时来制作门框（没有这 1 小时，门的框架都没有，后续没法组装）。所以生产 1 扇门，必然消耗 Plant1 的 1 小时。

Plant2（窗框车间）：只负责做“窗框”。只有生产“窗”时才会用到它（比如生产 1 扇窗需要在这花 2 小时做窗框），生产“门”用不到这个车间。

Plant3（组装车间）：负责把框架组装成成品。不管是门还是窗，做好框架后，都得送到这个车间“组装成型”。生产 1 扇门，组装过程需要 3 小时；生产 1 扇窗，组装需

要 2 小时

一扇门的利润为 3 美元，窗为 5 美元：

求最大利润：

得到目标函数：$\max Z = 3x_1 + 5x_2$ （其中 x_1 为门的数量，x_2 为窗的数量，利润单位为美元）

约束条件：

$$\begin{cases} x_1 \leq 4 \quad \text{（工厂 1 工时约束）} \\ 2x_2 \leq 12 \quad \text{（工厂 2 工时约束）} \\ 3x_1 + 2x_2 \leq 18 \quad \text{（工厂 3 工时约束）} \\ x_1, x_2 \geq 0 \quad \text{（非负约束）} \end{cases}$$

经过单纯形法，得到最终目标函数形式：$Z = 36 - 0s_1 - \frac{3}{2}s_2 - s_3$ （其中 s_1, s_2, s_3 分别为三个工厂的松弛变量）

与：最优解：$x_1^* = 2$，$x_2^* = 6$，最大利润 $Z^* = 36$ 美元

其中松弛变量为：

工厂 1：$s_1^* = 2$（工时未用尽）

工厂 2：$s_2^* = 0$（工时刚好用尽）

工厂 3：$s_3^* = 0$（工时刚好用尽）

接着，我们引入影子价格的概念：

Definition

影子价格：约束条件右侧（资源量）增加 1 单位时，目标函数（利润）的增加量

1. 工厂 1 的影子价格：

原约束 $x_1 \leq 4$，增加到 $x_1 \leq 5$ 由于原最优解 $x_1^* = 2 < 4$（工时未用尽），增加工时后最优解不变，利润无增加因此，工厂 1 的影子价格为 0 美元

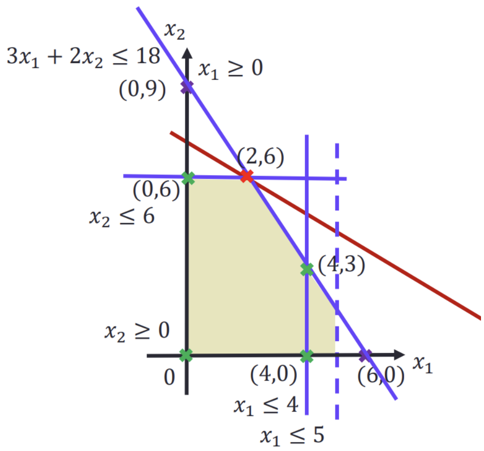


图 2: 工厂 1

2. 工厂 2 的影子价格:

原约束 $2x_2 \leq 12$，增加到 $2x_2 \leq 13$（工时 +1）新最优解：$x_1^* = \frac{5}{3}$，$x_2^* = \frac{13}{2}$ 新利润 $Z^* = 3 \times \frac{5}{3} + 5 \times \frac{13}{2} = \frac{75}{2} = 37.5$ 美元

利润增加量：$37.5 - 36 = 1.5 = \frac{3}{2}$ 美元因此，工厂 2 的影子价格为 $\frac{3}{2}$ 美元

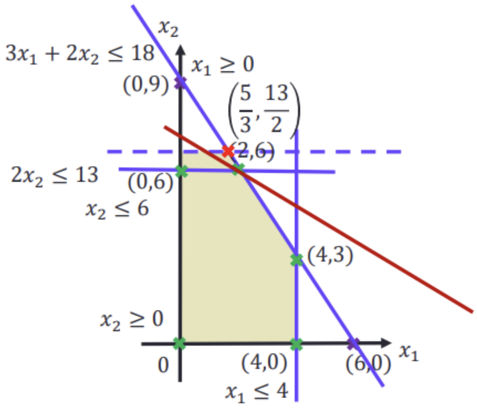


图 3: 工厂 2

3. 工厂 3 的影子价格:

原约束 $3x_1 + 2x_2 \leq 18$，增加到 $3x_1 + 2x_2 \leq 19$（工时 +1）新最优解：$x_1^* = \frac{7}{3}$，$x_2^* = 6$ 新利润 $Z^* = 3 \times \frac{7}{3} + 5 \times 6 = 37$ 美元

利润增加量：$37 - 36 = 1$ 美元因此，工厂 3 的影子价格为 1 美元.

回顾单纯形法最终目标函数形式：$Z = 36 - 0s_1 - \frac{3}{2}s_2 - s_3$（其中 s_1, s_2, s_3 分别为三个工厂的松弛变量）

观察到:

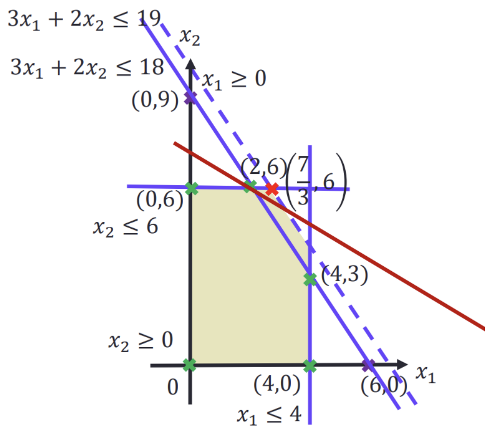


图 4: 工厂 3

Theorem

影子价格是**松弛变量系数的相反数**。

因此:

工厂 1 对应 s_1，系数为 0，影子价格 0

工厂 2 对应 s_2，系数为 $-\frac{3}{2}$，影子价格 $\frac{3}{2}$

工厂 3 对应 s_3，系数为 -1，影子价格 1

【Remark】:

影子价格只和每个约束的松弛变量有关，和其它变量无关!

松弛变量与约束松弛的关系:

松弛变量 s_i 对应”约束的剩余资源”（如 s_2 对应工厂 2 的约束 $2x_2 + s_2 = 12$）。若增加松弛变量的值（如 $s_2^{\text{new}} = s_2 + 1$），等价于约束向外松弛（如工厂 2 的工时约束从 $2x_2 \leq 12$ 变为 $2x_2 \leq 13$）。

约束松弛对利润的影响（以工厂 2 为例）:

将 $s_2 = s_2^{\text{new}} - 1$ 代入目标函数 $Z = 36 - \frac{3}{2}s_2 - s_3$，可得:

$$Z = 36 - \frac{3}{2}(s_2^{\text{new}} - 1) - s_3 = 36 + \frac{3}{2} - \frac{3}{2}s_2^{\text{new}} - s_3$$

这表明：当工厂 2 的工时约束松弛 1 单位（工时 +1），目标函数（利润）会增加 $\frac{3}{2}$。

简言之，通过单纯形法最终目标函数的结构，结合松弛变量与约束松弛的关系，推导出”约束松弛对利润的边际贡献”（即影子价格）。

8 对偶与对偶理论

接下来，我们考虑原问题的对偶问题:

原问题： 生产者的“利润最大化”逻辑生产者的核心决策是确定“门（x_1）”和“窗（x_2）”的产量，以最大化总利润

对偶问题： 从购买原问题“劳动工时”的收购者视角来看，收购者的核心决策是给 3 个工厂的“每小时工时”定价（y_1, y_2, y_3），以最小化“收购全部工时的总成本”，同时确保“收购价能覆盖生产者自己生产的利润”。

目 标：收购成本最小化收购成本由“工厂工时数 × 每小时定价”决定：

Plant1 有 4 小时，每小时定价 y_1，成本为 $4y_1$；

Plant2 有 12 小时，每小时定价 y_2，成本为 $12y_2$；

Plant3 有 18 小时，每小时定价 y_3，成本为 $18y_3$；

目标函数为： $\min W = 4y_1 + 12y_2 + 18y_3$。

约 束：收购价需覆盖生产利润收购价不能太低，否则生产者会选择“自己生产”而非“卖工时”。因此，工时的总收购价值必须 生产对应产品的利润：

针对“门（x_1）”：生产 1 扇门，需要 Plant1 的 1 小时（价值 $1 \times y_1$）和 Plant3 的 3 小时（价值 $3 \times y_3$），总收购价值为 $y_1 + 3y_3$；而生产 1 扇门的利润是 3\$，因此要求 $y_1 + 3y_3 \geq 3$；

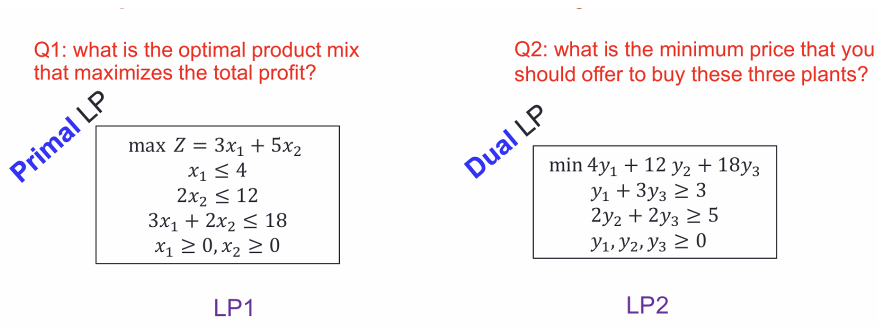
针对“窗（x_2）”：生产 1 扇窗，需要 Plant2 的 2 小时（价值 $2 \times y_2$）和 Plant3 的 2 小时（价值 $2 \times y_3$），总收购价值为 $2y_2 + 2y_3$；而生产 1 扇窗的利润是 5\$，因此要求 $2y_2 + 2y_3 \geq 5$；同时，工时的“定价”不能为负（时间有价值，收购价至少为 0）$\rightarrow y_1, y_2, y_3 \geq 0$。

由此，我们可以得到原对偶问题的表达式：

对偶问题（Dual LP，最小化购买成本）：目标：$\min 4y_1 + 12y_2 + 18y_3$（$y_1, y_2, y_3$：三个工厂工时的“影子价格”，目标系数是原始约束的右侧常数（RHS）：4、12、18）。

约束：2 个“价值限制”（对应原始变量的利润）：$y_1 + 3y_3 \geq 3$（对应 x_1 的利润系数 3）、$2y_2 + 2y_3 \geq 5$（对应 x_2 的利润系数 5）；变量非负：$y_1, y_2, y_3 \geq 0$。

即：



观察到：

Corollary & Secondary Conclusion

LP1 中的变量数量与 LP2 中的约束数量一致（不包括非负性约束）
LP1 中约束的右侧常数与 LP2 目标函数中的系数一致；
LP1 中约束左侧的列系数与 LP2 中约束左侧的行系数一致。

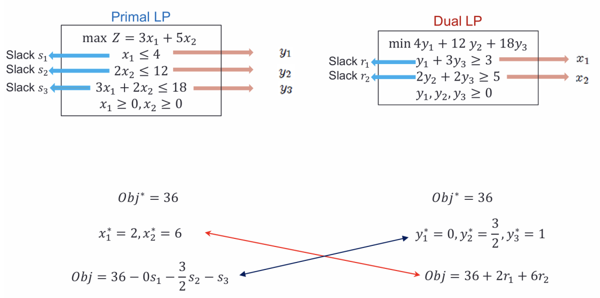
$$\begin{array}{c c} & x_1 \\ \begin{matrix} y_1 \\ y_2 \\ y_3 \end{matrix} & \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \end{array} \Longrightarrow y_1 + 3y_3 \geq 3$$

$$\begin{array}{c c} & x_2 \\ \begin{matrix} y_1 \\ y_2 \\ y_3 \end{matrix} & \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} \end{array} \Longrightarrow 2y_2 + 2y_3 \geq 5$$

【Remark】:

对偶问题的对偶还是原问题。

Primal-Dual Connection：



8.1 对偶问题的构造

接下来，我们从拉格朗日问题开始，对“对偶问题”做一个详尽的分析：

线性规划的原始问题（Primal）通常是：

$$\min_{\mathbf{x}} \mathbf{c}^\top \mathbf{x}$$

主体约束：

不等式约束：$\mathbf{A}\mathbf{x} \geq \mathbf{b}$ 或 $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ 或混合。

等式约束：$\mathbf{E}\mathbf{x} = \mathbf{d}$。

变量域：$\mathbf{x} \geq \mathbf{0}$，或部分变量 0，或自由（无限制）。

要推导对偶，我们先构造拉格朗日函数 $L(\mathbf{x}, \mathbf{y})$，其中 $\mathbf{y}$ 是对偶变量（乘子），每个约束对应一个 y_i。

$$L(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + \sum_i y_i g_i(\mathbf{x})$$

对于 LP，$f = \mathbf{c}^\top \mathbf{x}$，约束需标准化。

【Remark】：

在通用对偶中，为保持一致乘子 y_i 不一定大于 0：

如：常将　约束的乘子设为　0，等价于让 $y_i' = -y_i \geq 0$，然后项变成 $y_i'(b_i - \mathbf{a}_i^\top \mathbf{x})$。

统一形式：我们选择所有拉格朗日项为 $y_i(b_i - \mathbf{a}_i^\top \mathbf{x})$，并据约束类型调整 y_i 符号：

M_1 $(\leq)$：$y_i \leq 0$

M_2 $(\geq)$：$y_i \geq 0$

M_3 $(=)$：y_i 自由

证明. 统一用拉格朗日项 $y_i(b_i - \mathbf{a}_i^\top \mathbf{x})$，但符号由约束类型决定。

逻辑：确保当约束违反时，L 能“惩罚”，即增加值（因为原始问题是 min，所以 L 应比原始目标更大才能起到惩罚的效果）。

如果原始约束是 $\geq$ 型，$\mathbf{a}_i^\top \mathbf{x} \geq b_i$ 如果违反约束，则 $(b_i - \mathbf{a}_i^\top \mathbf{x}) \geq 0$，那么如果要增加，需要 y_i 也大于 0，正乘正还是正，可以增加。

□

这样确定完 y_i 的符号后，继续对拉格朗日函数的构造，得到拉格朗日函数为：

$$L(\mathbf{x}, \mathbf{y}) = \mathbf{c}^\top \mathbf{x} + \sum_{i=1}^m y_i(b_i - \mathbf{a}_i^\top \mathbf{x}) = \mathbf{c}^\top \mathbf{x} - (\mathbf{A}^\top \mathbf{y})^\top \mathbf{x} + \mathbf{b}^\top \mathbf{y} = (\mathbf{c} - \mathbf{A}^\top \mathbf{y})^\top \mathbf{x} + \mathbf{b}^\top \mathbf{y}$$

这里 $\mathbf{A}$ 是约束矩阵（行对应约束），$\mathbf{A}^\top \mathbf{y}$ 是对偶梯度项。

这个特殊的形式源于拉格朗日乘子定理（18 世纪拉格朗日用于等式约束，后扩展到不等式）。在凸优化中，L 是鞍点函数：在最优点，$\max_{\mathbf{y}}\min_{\mathbf{x}} L = \min_{\mathbf{x}}\max_{\mathbf{y}} L =$ 最优值。

现在我们给出对偶问题定义为：

Definition

$$\max_{\mathbf{y}} \inf_{\mathbf{x}\in\mathcal{X}} L(\mathbf{x},\mathbf{y})$$

其中 $\mathcal{X}$ 是 $\mathbf{x}$ 的域（基于 N_1, N_2, N_3），inf 是下确界（可能 $-\infty$）。

为什么是 max *inf* L？

$\inf_{\mathbf{x}} L$ 给出原始问题的下界：对于任何可行 $\mathbf{x}$，$L(\mathbf{x},\mathbf{y}) \le \mathbf{c}^\top\mathbf{x}$（如果 $\mathbf{y}$ 符号正确，符合约束使 L 永远小于等于原目标函数的值）。

最大化这个下界给出最好的下界，即对偶目标。

弱对偶： 对偶值小于原始值，因为 $\inf L \le L(\mathbf{x},\mathbf{y}) \le \mathbf{c}^\top\mathbf{x}$。

在 LP（凸且线性），**强对偶总是成立（无须额外假设，因为多面体）**。

现在计算 $\inf_{\mathbf{x}} L = \mathbf{b}^\top\mathbf{y} + \inf_{\mathbf{x}}(\mathbf{c}-\mathbf{A}^\top\mathbf{y})^\top\mathbf{x}$。

由于线性且分量独立，$\inf = \sum_j \inf_{x_j\in\text{domain}_j} \text{coeff}_j x_j$，其中 $\text{coeff}_j = c_j - \mathbf{A}_j^\top\mathbf{y}$（$\mathbf{A}_j$ 是第 j 列）。

对于 $j \in N_1$ $(x_j \ge 0)$:

$$\inf_{x_j\ge 0} \text{coeff}_j\, x_j = \begin{cases} 0 & \text{if coeff}_j \ge 0 \text{ (} x_j = 0) \\ -\infty & \text{if coeff}_j < 0 \; (x_j \to +\infty) \end{cases}$$

如果系数 $coeff_i \ge 0$，那么该函数单调递增且最小值为 0（因为 $x \ge 0$），则 inf 存在为 0；反之如果系数 $coeff_i \le 0$，那么该函数单调递减且最小值为负无穷（因为 $x \ge 0$，可以达到正无穷），则 inf 不存在，不符合对偶条件。

要使 inf 有限，必须 $\text{coeff}_j \ge 0$，即 $\mathbf{A}_j^\top\mathbf{y} \le c_j$。贡献值为 0。

对于 $j \in N_2$ $(x_j \le 0)$:

$$\inf_{x_j\le 0} \text{coeff}_j\, x_j = \begin{cases} 0 & \text{if coeff}_j \le 0 \text{ (} x_j = 0) \\ -\infty & \text{if coeff}_j > 0 \; (x_j \to -\infty) \end{cases}$$

如果系数 $coeff_i \ge 0$，那么该函数单调递增（因为 $x \le 0$），但是 inf 不存在，因为当 x 取到负无穷就会得到负无穷；反之如果系数 $coeff_i \le 0$，那么该函数单调递减且最小值为 0（因为 $x \le 0$，可以达到 0），则 inf 存在，符合对偶条件。

要使有限，必须 $\text{coeff}_j \le 0$，即 $\mathbf{A}_j^\top\mathbf{y} \ge c_j$。贡献值为 0。

对于 $j \in N_3$ (x_j 自由):

$$\inf_{x_j \in \mathbb{R}} \text{coeff}_j\, x_j = \begin{cases} 0 & \text{if coeff}_j = 0 \\ -\infty & \text{otherwise} \end{cases}$$

要使有限，必须 $\text{coeff}_j = 0$，即 $\mathbf{A}_j^\top \mathbf{y} = c_j$。贡献值为 0。

因此，当所有对偶约束满足时，$\inf_{\mathbf{x}} L = \mathbf{b}^\top \mathbf{y}$；否则 $-\infty$。

简言之，对偶问题是原问题的“镜像”，每个符号的选择都是为了让这面“镜子”的反射逻辑自洽，从而支撑对偶理论的核心结论。

Example

原问题:

$$\begin{aligned} \text{最小化} \quad & x_1 + 2x_2 \\ \text{受约束于} \quad & x_1 + x_2 \geq 5 \\ & x_1 - x_2 \leq 3 \\ & x_1 \geq 0,\ x_2 \in \mathbb{R}\ (\text{自由}) \end{aligned}$$

那么对偶问题:

$$\begin{aligned} \text{最大化} \quad & 5y_1 + 3y_2 \\ \text{受约束于} \quad & y_1 + y_2 \leq 1 \\ & y_1 - y_2 = 2 \\ & y_1 \geq 0,\ y_2 \leq 0 \end{aligned}$$

8.2 对偶理论

在优化理论中，原始问题（Primal）与对偶问题（Dual）是一对相互关联的优化问题。其标准形式如下：

Primal Problem	Dual Problem
min $\boldsymbol{c}^\top \boldsymbol{x}$	max $\boldsymbol{b}^\top \boldsymbol{y}$
s.t. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$	s.t. $\boldsymbol{A}^\top \boldsymbol{y} \leq \boldsymbol{c}$
$\boldsymbol{x} \geq \boldsymbol{0}$	

其中：- $\boldsymbol{x} \in \mathbb{R}^n$ 为原始问题的决策变量

- $\boldsymbol{y} \in \mathbb{R}^m$ 为对偶问题的决策变量

- $\boldsymbol{A} \in \mathbb{R}^{m\times n}$、$\boldsymbol{b} \in \mathbb{R}^m$、$\boldsymbol{c} \in \mathbb{R}^n$ 为已知参数

弱对偶定理（Weak Duality Theorem）是对偶理论的基础结论，表述如下：

Theorem

若 $\boldsymbol{x}$ 是原始问题的可行解，且 $\boldsymbol{y}$ 是对偶问题的可行解，则：

$$\boldsymbol{b}^\top \boldsymbol{y} \le \boldsymbol{c}^\top \boldsymbol{x}$$

该定理的含义包括：

1. 任何对偶可行解的目标函数值都是原始问题最优值的下界
2. 任何原始可行解的目标函数值都是对偶问题最优值的上界
3. 原始问题的最优值不小于对偶问题的最优值

证明. 已知 $\boldsymbol{x}$ 是原始可行解，满足：

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \quad \text{且} \quad \boldsymbol{x} \ge \boldsymbol{0}$$

$\boldsymbol{y}$ 是对偶可行解，满足：

$$\boldsymbol{A}^\top \boldsymbol{y} \le \boldsymbol{c}$$

对原始约束 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 两边同时左乘 $\boldsymbol{y}^\top$：

$$\boldsymbol{y}^\top \boldsymbol{b} = \boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{x}$$

根据矩阵转置性质 $(\boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{x}) = (\boldsymbol{A}^\top \boldsymbol{y})^\top \boldsymbol{x}$，可得：

$$\boldsymbol{b}^\top \boldsymbol{y} = (\boldsymbol{A}^\top \boldsymbol{y})^\top \boldsymbol{x}$$

由于 $\boldsymbol{x} \ge \boldsymbol{0}$ 且 $\boldsymbol{A}^\top \boldsymbol{y} \le \boldsymbol{c}$，两边同乘 $\boldsymbol{x}$ 后不等号方向保持不变：

$$(\boldsymbol{A}^\top \boldsymbol{y})^\top \boldsymbol{x} \le \boldsymbol{c}^\top \boldsymbol{x}$$

联立可得：

$$\boldsymbol{b}^\top \boldsymbol{y} \le \boldsymbol{c}^\top \boldsymbol{x}$$

□

以下事实：

Fact or Background

原问题状态	对偶问题状态
可行且有界	可行且有界（最优值相等）
可行但无界	不可行
不可行	不可行或无界

证明. 如果原问题无界，则对偶问题一定不可行。这是因为：

根据对偶理论，原问题无界意味着对偶问题没有可行解（即不可行）。如果对偶问题可行，那么弱对偶定理会要求原问题目标函数值有下界（对于最小化问题）或上界（对于最大化问题），这与原问题无界矛盾。 □

【Remark】：

如果原问题有唯一最优解（非最优值），但是对偶问题不一定有唯一的最优解。

由弱对偶定理可直接推出以下推论：

Corollary & Secondary Conclusion

1. 若原始问题无下界（或者无上界，即 prime 无界），则对偶（dual）问题无可行解；
2. 如果原问题不可行（不等同于无界），则对偶问题可能不可行，也可能无界。
3. 若存在可行解 $\boldsymbol{x}^*$ 和 $\boldsymbol{y}^*$ 满足 $\boldsymbol{b}^\top \boldsymbol{y}^* = \boldsymbol{c}^\top \boldsymbol{x}^*$，则二者分别为原始问题和对偶问题的最优解

由最后一条推论可得：

Theorem

具有**相同目标值**的**原始可行解**和**对偶可行解**分别是原始问题和对偶问题的**最优解**。

证明. proof: 由夹逼定理易证。 □

首先考虑线性函数 $f(y) = y^\top v$（其中 $v = b - Ax$ 是一个固定向量，$y \in \mathbb{R}^m$ 是变量）。

若 $v \neq 0$（即 $Ax \neq b$）：线性函数 $y^\top v$ 没有上界。

例如：若 v 有分量 $v_i > 0$，令 $y_i \to +\infty$、其余 $y_j = 0$，则 $y^\top v = y_i v_i \to +\infty$；

若 v 有分量 $v_i < 0$，令 $y_i \to -\infty$，则 $y^\top v = y_i v_i \to +\infty$（负负得正）。因此，$\max_{y\in\mathbb{R}^m} y^\top v = +\infty$。

若 $v = 0$（即 $Ax = b$）：此时 $y^\top v = 0$ 对任意 y 成立，故 $\max_{y\in\mathbb{R}^m} y^\top v = 0$。

接着考虑**标准型线性规划（LP）**：

$$
\begin{cases}
\text{minimize} & c^\top x \\
\text{subject to} & Ax = b \\
 & x \geq 0
\end{cases}
$$

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

隐式引入等式约束：等式约束 $Ax=b$ 可通过“最大化线性项”隐式表达。原问题等价于：

$$\begin{cases}\text{minimize} & c^\top x+\max_{y\in\mathbb{R}^m} y^\top(b-Ax)\\ \text{subject to} & x\geq 0\end{cases}$$

我们令：$y^\top$ 为变量，此时若 $Ax\neq b$，则 $y^\top(b-Ax)$ 是线性函数，最大化后会趋于 ∞（破坏最优性）；只有 $Ax=b$ 时，该项为 0，因此**隐式强制了** $Ax=b$。

交换极小与极大（假设合法）（其证明即为强对偶定理的证明）：将 “$\min_x$” 与 “$\max_y$” 交换顺序，问题变为：

$$\max_y\left(b^\top y+\min_{x\geq 0} x^\top(c-A^\top y)\right)$$

分析内层极小化项：对于 $\min_{x\geq 0} x^\top(c-A^\top y)$：

- 若 $A^\top y\leq c$：取 $x=0$，该项值为 0（最小可能值）。
- 若 $A^\top y\not\leq c$：存在分量 $(c-A^\top y)_i<0$，令 $x_i\to+\infty$，该项会趋于 $-\infty$（破坏最大化整体目标）。

因此，内层极小化有效时，约束为 $A^\top y\leq c$，此时问题等价于：

$$\begin{cases}\text{maximize} & b^\top y\\ \text{subject to} & A^\top y\leq c\end{cases}$$

即对偶问题：

这种技巧是对偶理论中常用的 **“隐式约束嵌入”**：通过线性函数的无界性，把等式 / 不等式约束转化为“目标函数的极端行为”，从而让约束“自动”被满足（不满足的情况会被目标函数的无穷大“筛掉”）。

现在考察非标准型约束的对偶（以 $Ax\geq b$ 为例）：

考虑原问题：

$$\begin{cases}\text{minimize} & c^\top x\\ \text{subject to} & Ax\geq b\end{cases}$$

隐式引入不等式约束：约束 $Ax\geq b$ 等价于 $b-Ax\leq 0$，通过“非负变量的最大化”隐式表达：原问题等价于：

$$\begin{cases}\text{minimize} & c^\top x+\max_{y\geq 0} y^\top(b-Ax)\\ \text{subject to} & \text{无额外变量约束（除隐含逻辑）}\end{cases}$$

理由：若 $Ax\not\geq b$，则 $y\geq 0$ 时 $y^\top(b-Ax)$ 最大化会趋于 ∞；只有 $Ax\geq b$ 时，该项为 0。

交换极小与极大（假设合法）：交换后得到：

$$\max_{y\geq 0}\left(b^\top y+\min_{x} x^\top(c-A^\top y)\right)$$

分析内层极小化项：对于 $\min_x x^\top(c-A^\top y)$：

- 若 $A^\top y = c$：该项值为 0（对任意 x，$x^\top(c - A^\top y) = 0$）。
- 若 $A^\top y \neq c$：存在分量使得 $x^\top(c - A^\top y)$ 可趋于 $-\infty$（破坏最大化整体目标）。

因此，约束为 $A^\top y = c$，此时对偶问题为：

约束 $Ax \geq b$ 的对偶问题：

$$\begin{cases} \text{maximize} & b^\top y \\ \text{subject to} & A^\top y = c \\ & y \geq 0 \end{cases}$$

Theorem

强对偶定理 (Strong Duality Theorem)：如果一个线性规划问题有最优解，那么其对偶问题也有最优解，且原问题与对偶问题的最优值相等。

证明. 不失一般性，我们假设原问题为标准型。

若原问题存在最优解 $\boldsymbol{x}^*$，且 $\boldsymbol{x}^*$ 是一个基可行解（可视为由单纯形法得到），则它必定与某个最优基 B 相关联，使得 $\boldsymbol{x}_B = A_B^{-1}\boldsymbol{b}$（其中 $\boldsymbol{x}_B$ 是 $\boldsymbol{x}^*$ 的基变量部分）。此外，检验数满足

$$\boldsymbol{c}^\top - \boldsymbol{c}_B^\top A_B^{-1} A \geq \boldsymbol{0} \tag{1}$$

现在我们定义 $\boldsymbol{y}^\top = \boldsymbol{c}_B^\top A_B^{-1}$。由上述等式 (1) 可知，$A^\top \boldsymbol{y} \leq \boldsymbol{c}$，即 $\boldsymbol{y}$ 是对偶问题的可行解。此外，还有

$$\boldsymbol{b}^\top \boldsymbol{y} = \boldsymbol{c}_B^\top A_B^{-1}\boldsymbol{b} = \boldsymbol{c}_B^\top \boldsymbol{x}_B = \boldsymbol{c}^\top \boldsymbol{x}^*$$

因此，根据弱对偶定理，$\boldsymbol{y}$ 必定是对偶问题的最优解，故定理得证。 □

8.3 线性规划的互补松弛性（Complementary Slackness）

原问题（Primal LP）与对偶问题（Dual LP）的标准形式：

1. 原问题（最大化型）：

$$\begin{cases} \max Z = c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \leq b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leq b_m \\ x_1, x_2, \ldots, x_n \geq 0 \end{cases}$$

2. 对偶问题（最小化型）：

$$
\begin{cases}
\min W = b_1y_1 + b_2y_2 + \cdots + b_my_m \\
a_{11}y_1 + a_{21}y_2 + \cdots + a_{m1}y_m \geq c_1 \\
a_{12}y_1 + a_{22}y_2 + \cdots + a_{m2}y_m \geq c_2 \\
\vdots \\
a_{1n}y_1 + a_{2n}y_2 + \cdots + a_{mn}y_m \geq c_n \\
y_1, y_2, \ldots, y_m \geq 0
\end{cases}
$$

松弛变量（Slack Variables）

原问题松弛变量 s_i：$s_i = b_i - (a_{i1}x_1 + \cdots + a_{in}x_n)$，满足 $s_i \geq 0$（因原约束为"≤"）。

对偶问题松弛变量 r_j：$r_j = (a_{1j}y_1 + \cdots + a_{mj}y_m) - c_j$，满足 $r_j \geq 0$（因对偶约束为"≥"）。

互补松弛性条件（Complementary Slackness Conditions）

若 $(x_1^*, \ldots, x_n^*)$ 是原问题最优解，$(y_1^*, \ldots, y_m^*)$ 是对偶问题最优解，则需满足：

1. 原约束与对偶变量的互补性：对原问题的第 i 个约束（对应对偶变量 y_i^*），有：

$$y_i^* \cdot s_i^* = 0 \quad (i = 1, 2, \ldots, m)$$

即 $y_i^* \left(b_i - \sum_{j=1}^n a_{ij}x_j^*\right) = 0$。

2. 对偶约束与原变量的互补性：对对偶问题的第 j 个约束（对应原变量 x_j^*），有：

$$x_j^* \cdot r_j^* = 0 \quad (j = 1, 2, \ldots, n)$$

即 $x_j^* \left(\sum_{i=1}^m a_{ij}y_i^* - c_j\right) = 0$。

即："非零变量对应紧约束"：

若原问题某约束的松弛变量 $s_i^* > 0$（约束"宽松"，不取等），则对应对偶变量 $y_i^* = 0$；（因为本质上对偶变量 y_i^* 是原问题的拉格朗日乘子，如果约束宽松，则说明约束"不活跃"，用 $y_i^* = 0$ 关闭该约束）

若对偶变量 $y_i^* > 0$，则原问题对应约束的松弛变量 $s_i^* = 0$（约束"紧"，则说明约束"活跃"，用 $y_i^* \neq 0$ 开启该约束）。

若对偶问题某约束的松弛变量 $r_j^* > 0$（约束"宽松"，不取等），则对应原变量 $x_j^* = 0$；（同样的道理，对偶的对偶是原问题，位置互换但思维模式不变）

若原变量 $x_j^* > 0$，则对偶问题对应约束的松弛变量 $r_j^* = 0$（约束"紧"，取等）。

由此得到：**互补松弛性的性质**（Complementary Slackness Property）

Corollary & Secondary Conclusion

若 $(\bar{x}_1,\ldots,\bar{x}_n)$ 是原问题可行解，$(\bar{y}_1,\ldots,\bar{y}_m)$ 是对偶问题可行解，且满足互补松弛性条件，则：

1. 原问题目标值 $\sum_{j=1}^{n} c_j\bar{x}_j$ 与对偶问题目标值 $\sum_{i=1}^{m} b_i\bar{y}_i$ 相等；
2. 根据“强对偶性”，这两个可行解同时是最优解。

具体示例:

1. 原问题:

$$\begin{cases} \max Z = 3x_1 + 5x_2 \\ x_1 \le 4 \\ 2x_2 \le 12 \\ 3x_1 + 2x_2 \le 18 \\ x_1, x_2 \ge 0 \end{cases}$$

2. 对偶问题:

$$\begin{cases} \min W = 4y_1 + 12y_2 + 18y_3 \\ y_1 + 3y_3 \ge 3 \\ 2y_2 + 2y_3 \ge 5 \\ y_1, y_2, y_3 \ge 0 \end{cases}$$

3. 互补松弛性条件:

$$\begin{cases} y_1^*(4 - x_1^*) = 0 \\ y_2^*(12 - 2x_2^*) = 0 \\ y_3^*(18 - 3x_1^* - 2x_2^*) = 0 \\ x_1^*(y_1^* + 3y_3^* - 3) = 0 \\ x_2^*(2y_2^* + 2y_3^* - 5) = 0 \end{cases}$$

4. 若原、对偶可行解满足上述等式，则它们都是最优解，且 $3x_1^* + 5x_2^* = 4y_1^* + 12y_2^* + 18y_3^*$。

即我们可以得到:

Theorem

设 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 分别是原问题和对偶问题的可行点。则 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是最优解当且仅当:

$$y_i \cdot (\boldsymbol{a}_i^T \boldsymbol{x} - b_i) = 0, \quad \forall i \quad 且 \quad x_j \cdot (\boldsymbol{A}_j^T \boldsymbol{y} - c_j) = 0, \quad \forall j.$$

证明. 由对偶定理可知，若 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 分别是原问题和对偶问题的最优解，则必有

$$\boldsymbol{c}^T\boldsymbol{x} = \boldsymbol{b}^T\boldsymbol{y}$$

因此有

$$0 = \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{y} = \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{y}^T\boldsymbol{A}\boldsymbol{x} = \sum_{i=1}^{n}(c_i - \boldsymbol{A}_i^T\boldsymbol{y}) \cdot x_i \tag{2}$$

由于 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 都是可行解，对每个 i，有 $c_i - \boldsymbol{A}_i^T\boldsymbol{y} \geq 0$ 且 $x_i \geq 0$。因此，要使式 (2) 成立，必须有

$$(c_i - \boldsymbol{A}_i^T\boldsymbol{y}) \cdot x_i = 0, \quad \forall i = 1, \ldots, n.$$

另一方向可通过上述论证反向推导得出。

□

我们通过一个例子来感受一下：

Example

Prime problem:

$$\begin{aligned} \text{minimize} \quad & 13x_1 + 10x_2 + 6x_3 \\ \text{s.t.} \quad & 5x_1 + x_2 + 3x_3 = 8 \\ & 3x_1 + x_2 = 3 \\ & x_1,\ x_2,\ x_3 \geq 0 \end{aligned}$$

$\rightsquigarrow$ Optimal solution $\boldsymbol{x}^* = (1, 0, 1)$.

Dual problem:

$$\begin{aligned} \text{maximize} \quad & 8y_1 + 3y_2 \\ \text{s.t.} \quad & 5y_1 + 3y_2 \leq 13 \\ & y_1 + y_2 \leq 10 \\ & 3y_1 \leq 6 \end{aligned}$$

原问题是最小化线性目标函数，有 2 个等式约束与非负约束，最优解为 $x^* = (1, 0, 1)$。

其中 $x_1 = 1 > 0$、$x_3 = 1 > 0$（$x_2 = 0$）。

根据线性规划的互补松弛性，原问题中正的变量对应的对偶约束必须“紧”（即取等号）。

原变量 $x_1 > 0$，对应对偶的第一个约束 $5y_1 + 3y_2 \leq 13$，需取等：$5y_1 + 3y_2 = 13$。

原变量 $x_3 > 0$，对应对偶的第三个约束 $3y_1 \leq 6$，需取等：$3y_1 = 6$。

求解对偶变量：由 $3y_1 = 6$，得 $y_1 = 2$。将 $y_1 = 2$ 代入 $5y_1 + 3y_2 = 13$，得 $10 + 3y_2 = 13$，解得 $y_2 = 1$。

验证对偶可行性：对偶的第二个约束为 $y_1 + y_2 \leq 10$，代入 $y_1 = 2, y_2 = 1$，得 $2 + 1 = 3 \leq 10$，满足约束。

因此 $(y_1, y_2) = (2, 1)$ 是对偶可行解。验证最优性（强对偶性）：原问题最优目标值：$13 \times 1 + 10 \times 0 + 6 \times 1 = 19$。

对偶问题目标值（用 $(2, 1)$ 计算）：$8 \times 2 + 3 \times 1 = 19$。两者目标值相等，故 $(2, 1)$ 是对偶问题的最优解。

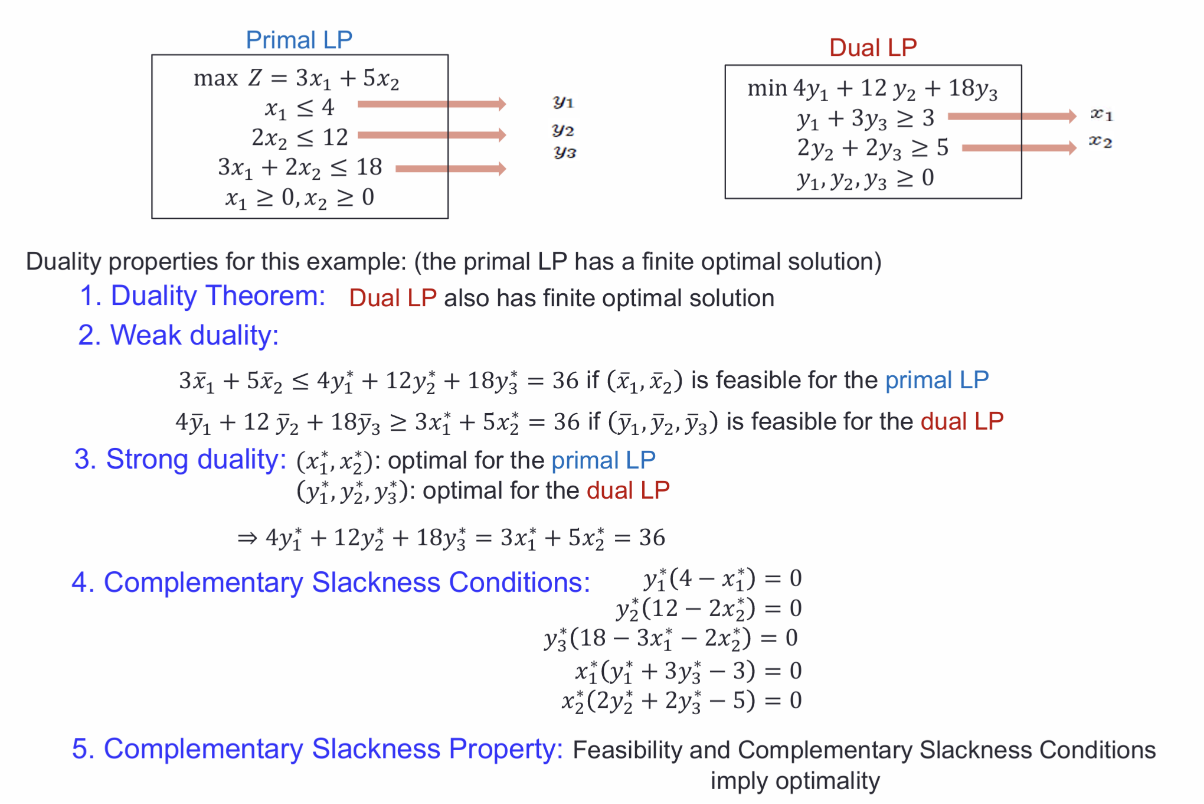
而在单纯型表中：

无论迭代到哪一步，原问题当前解的目标值 Z 与对应对偶解的目标值 $4y_1 + 5y_2$ 始终相等（这是单纯形法的数学结构决定的）。

对偶解的可行性逐步改善：中间迭代步骤中，对偶解可能不满足对偶问题的约束（如 y_2 为负、约束条件不成立），即“不可行”。

但是当原问题达到最优解时，对偶解也恰好满足所有对偶约束（可行），此时两者目标值相等（强对偶定理）。

这一节的知识可用下图表示：



9 内点法

线性规划（LP）的最优性条件

要使 (x, y) 成为线性规划的最优解，需同时满足三**个条件**：

原问题可行性：$Ax = b$ 且 $x \ge 0$（原问题变量满足约束）。

对偶可行性：$A^\top y \le c$（对偶问题变量满足约束，其中 $s = c - A^\top y$ 为对偶松弛变量）。

互补松弛性：对每个分量 i，$x_i \cdot s_i = x_i \cdot (c_i - A_i^\top y) = 0$（原变量与对偶松弛变量“互补”，即两者不能同时为正）。

内点法的核心思路

内点法旨在找到满足上述三个条件的 x, y, s，但直接求解这些非线性方程（尤其是互补松弛的非线性条件 $x_i \cdot s_i = 0$）较困难。因此采用“松弛 + 迭代”策略：

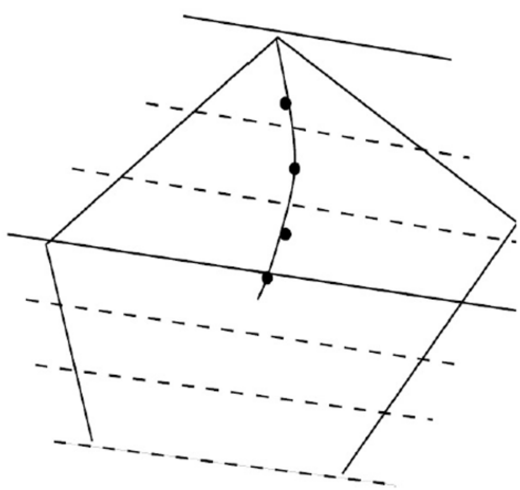
Method

松弛互补条件：将严格的互补松弛 $x_i \cdot s_i = 0$ 放宽为 $x_i \cdot s_i \le \mu$（$\mu > 0$ 称为“互补间隙”）。
迭代缩小间隙：先找到某一 μ 下的解，再逐步减小 μ，重复此过程直到 μ 趋近于 0，此时松弛后的条件近似满足原始互补松弛条件，从而得到最优解。

为何称为“内点法”？与单纯形法的区别

Fact or Background

单纯形法： 仅在可行域的**顶点**（边界点）中搜索最优解。
内点法： 从可行域的**内部**开始迭代，迭代过程中始终保持 $x > 0$ 且 $s > 0$（变量严格为正，位于可行域内部）。
最终最优解的性质：**若解唯一，则必为基可行解（BFS）；若解不唯一，可能不是基可行解**。



以下通过两个具体的线性规划例子，分别阐述解唯一和不唯一时，单纯形法和内点法的求解特性。

例子 1：线性规划问题解唯一线性规划问题：$\begin{cases} \min & z=-x_1-2x_2 \\ \text{s.t.} & x_1+x_2\le 4 \\ & x_1\ge 0 \\ & x_2\ge 0 \end{cases}$ 引入松弛变量 x_3 将不等式约束化为等式约束，得到标准型：$\begin{cases} \min & z=-x_1-2x_2+0x_3 \\ \text{s.t.} & x_1+x_2+x_3=4 \\ & x_1\ge 0 \\ & x_2\ge 0 \\ & x_3\ge 0 \end{cases}$

单纯形法求解过程：该线性规划的可行域是由 $x_1=0$，$x_2=0$，$x_1+x_2=4$ 这几条直线所围成的区域，其顶点为 $(0,0)$，$(4,0)$，$(0,4)$ 。

单纯形法从初始基本可行解（如 $(0,0,x_3=4)$ ）开始，在这些顶点之间移动，通过换基迭代来不断改进目标函数值。经过计算可知，最优解在顶点 $(0,4)$ 处取得，此时 $z=-8$ ，该最优解是基可行解。

内点法求解过程：内点法从可行域内部的一个点（比如 $(1,1)$ ）开始，在迭代过程中，始终保持 $x_1>0$，$x_2>0$ 以及松弛变量 $x_3>0$ 。通过不断调整 x_1，x_2 的值，使得目标函数值逐渐减小，最终收敛到最优解 $(0,4)$ ，该解也是基可行解。

例子 2：线性规划问题解不唯一线性规划问题：$\begin{cases} \max & z=x_1+2x_2 \\ \text{s.t.} & -x_1+x_2\le 1 \\ & x_1+x_2\le 3 \\ & x_1\ge 0 \\ & x_2\ge 0 \end{cases}$ 引入松弛变量 x_3，x_4 化为标准型：$\begin{cases} \max & z=x_1+2x_2+0x_3+0x_4 \\ \text{s.t.} & -x_1+x_2+x_3=1 \\ & x_1+x_2+x_4=3 \\ & x_1\ge 0 \\ & x_2\ge 0 \\ & x_3\ge 0 \\ & x_4\ge 0 \end{cases}$

单纯形法求解过程：可行域是由 $x_1=0$，$x_2=0$，$-x_1+x_2=1$，$x_1+x_2=3$ 所围成的多边形区域，顶点有 $(0,0)$，$(0,1)$，$(1,2)$，$(3,0)$ 。

单纯形法在顶点间迭代，会发现目标函数在边 $(1,2)$ 到 $(0,3)$ 上都能取得最大值 $z=5$ ，最优解不唯一，且这些最优解都是基可行解。

内点法求解过程：内点法从可行域内部的点（如 $(1,1)$ ）开始迭代。随着迭代进行，会发现最终收敛到的最优解可能不是基可行解。比如在迭代过程中，可能会收敛到边 $(1,2)$ 到 $(0,3)$ 上除端点外的某一点，**该点不是可行域的顶点，也就不是基可行解，但它满足最优解的条件，即目标函数值达到最大值** $z=5$ 。

实际运行效率:

Fact or Background

单纯形法：对不同问题表现差异大——部分问题迭代次数极少（速度快），但最坏情况时间复杂度为指数级。
内点法：运行时间更稳定，固定规模下不同问题的时间波动小。

内点法的解特性（与单纯形法对比）:

Fact or Background

内点法倾向于找到非零分量最多的最优解（“最高秩”解）。
示例：求解 $\min\ x_1+x_2+x_3$ s.t. $x_1+x_2+x_3=1,\ x_1,x_2,x_3\geq 0$：
单纯形法给出基可行解（BFS）（如 $(1,0,0)$，非零分量最少，“低秩”解）。
内点法给出 $(\frac{1}{3},\frac{1}{3},\frac{1}{3})$（所有分量非零，“高秩”解）。

多最优解场景下的方法选择:
若需最高秩解（非零分量最多）：选内点法。
若需低秩解（非零分量最少）：选单纯形法。

10 灵敏度分析

线性规划（LP）的灵敏度分析的**核心是研究“输入** (A,b,c) **变化时，最优解和最优值如何变化”**，大致可以从**局部灵敏度与全局灵敏度**进行分析。

10.1 Local Sensitivity

以标准型 LP 为例:

$$\begin{cases} \min & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} \\ & \boldsymbol{x}\geq \boldsymbol{0} \end{cases}$$

记最优值为 V，分析 $\boldsymbol{b}$ 或 $\boldsymbol{c}$ 微小变化时 V 的响应。

（1）约束右侧 $\boldsymbol{b}$ 变化时的灵敏度（对偶解的角色）若固定 $\boldsymbol{A},\boldsymbol{c}$，最优值 V 可视为 $\boldsymbol{b}$ 的函数 $V(\boldsymbol{b})$。

Theorem

(最优值函数的可微性)： 若对偶问题存在唯一最优解 $\boldsymbol{y}^*$，则 $V(\boldsymbol{b})$ 对 $\boldsymbol{b}$ 的梯度满足 $\nabla_b V(\boldsymbol{b}) = \boldsymbol{y}^*$。

证明. 由强对偶性，原问题与对偶问题最优值相等。对偶问题为：

$$\begin{cases} \max & \boldsymbol{b}^\top \boldsymbol{y} \\ \text{s.t.} & \boldsymbol{A}^\top \boldsymbol{y} \leq \boldsymbol{c} \end{cases}$$

因此 $V(\boldsymbol{b}) = \boldsymbol{b}^\top \boldsymbol{y}^*$。当 $\boldsymbol{b}$ 微小变化 $\Delta \boldsymbol{b}$ 且对偶最优解 $\boldsymbol{y}^*$ 不变时，**目标值变化为** $\Delta \boldsymbol{b}^\top \boldsymbol{y}^*$。 □

直观意义：若 $\boldsymbol{b}$ 的**某分量** b_i 有微小变化 Δb_i（且最优解结构不变），则目标值的变化量为 $\Delta b_i \cdot y_i^*$（y_i^* 是对偶最优解的第 i 个分量）。

【Remark】：

$y_i^* = (A_B)^{-T} C_B$，即原松弛变量的 reduced cost。

（2）目标系数 $\boldsymbol{c}$ 变化时的灵敏度（原解的角色）若固定 $\boldsymbol{A}, \boldsymbol{b}$，最优值 V 可视为 $\boldsymbol{c}$ 的函数 $V(\boldsymbol{c})$。

Theorem

(最优值函数的可微性)： 若原问题存在唯一最优解 $\boldsymbol{x}^*$，则 $V(\boldsymbol{c})$ 对 $\boldsymbol{c}$ 的梯度满足 $\nabla_c V(\boldsymbol{c}) = \boldsymbol{x}^*$。

证明. 原问题目标为 $\boldsymbol{c}^\top \boldsymbol{x}$，当 $\boldsymbol{c}$ 微小变化 $\Delta \boldsymbol{c}$ 且原最优解 $\boldsymbol{x}^*$ 不变时，**目标值变化为** $\Delta \boldsymbol{c}^\top \boldsymbol{x}^*$。

□

直观意义：若 $\boldsymbol{c}$ 的某分量 c_i 有微小变化 Δc_i（且最优解结构不变），则目标值的变化量为 $\Delta c_i \cdot x_i^*$（x_i^* 是原最优解的第 i 个分量）。

(3)推广到不等式约束/最大化问题对于更一般的 LP(如最大化问题:$\max\ \boldsymbol{c}^\top \boldsymbol{x}$ s.t. $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$，$\boldsymbol{x} \geq \boldsymbol{0}$)，上述结论仍成立：

若对偶问题有唯一最优解 $\boldsymbol{y}^*$，则 $\nabla V(\boldsymbol{b}) = \boldsymbol{y}^*$；

若原问题有唯一最优解 $\boldsymbol{x}^*$，则 $\nabla V(\boldsymbol{c}) = \boldsymbol{x}^*$。

(可通过引入松弛变量将不等式约束转化为标准型，再应用前述结论。)

这个例子是生产计划优化的线性规划问题，结合原问题、对偶问题和灵敏度分析，解释如下：

一、原问题与对偶问题

原问题（生产计划的目标与约束）

目标：最大化两种产品的总利润，即 max $x_1 + 2x_2$（x_1 是产品 1 的产量，x_2 是产品 2 的产量；产品 1 单位利润 1，产品 2 单位利润 2）。约束：

资源 1：$x_1 \leq 100$（产品 1 消耗资源 1，总量不超过 100）；

资源 2：$2x_2 \leq 200$（产品 2 消耗资源 2，总量不超过 200）；

资源 3：$x_1 + x_2 \leq 150$（两种产品共同消耗资源 3，总量不超过 150）；

非负约束：$x_1, x_2 \geq 0$。

目标函数：$\max z = x_1 + 2x_2 + 0 \cdot s_1 + 0 \cdot s_2 + 0 \cdot s_3$

约束：$\begin{cases} x_1 + s_1 = 100 \\ 2x_2 + s_2 = 200 \\ x_1 + x_2 + s_3 = 150 \\ x_1, x_2, s_1, s_2, s_3 \geq 0 \end{cases}$

原最优解：$x^* = (50, 100)$，**最优总利润为** $50 + 2 \times 100 = 250$。

对偶问题（资源的"影子价格"）

根据线性规划对偶理论，原问题的对偶问题用于描述"资源的影子价格"（即资源每增加 1 单位，总利润的边际提升）：

目标：最小化资源的"总影子成本"，即 min $100y_1 + 200y_2 + 150y_3$（y_1, y_2, y_3 分别是资源 1、2、3 的影子价格）。

约束：

产品 1 的利润约束：$y_1 + y_3 \geq 1$（资源 1 和 3 的影子成本之和，需　产品 1 的单位利润）；

产品 2 的利润约束：$2y_2 + y_3 \geq 2$（资源 2 和 3 的影子成本之和，需　产品 2 的单位利润）；非负约束：$y_1, y_2, y_3 \geq 0$。

目标函数：min　$100y_1 + 200y_2 + 150y_3$

约束条件:

$s_1, s_2 \geq 0$)：$\begin{cases} y_1 + y_3 - s_1 = 1 & \text{（产品 1 的利润约束）} \\ 2y_2 + y_3 - s_2 = 2 & \text{（产品 2 的利润约束）} \\ y_1 \geq 0,\ y_2 \geq 0,\ y_3 \geq 0,\ s_1 \geq 0,\ s_2 \geq 0 & \text{（非负约束）} \end{cases}$

对偶最优解：$y^* = (0, 0.5, 1)$，**最优总影子成本为** $100 \times 0 + 200 \times 0.5 + 150 \times 1 = 250$（与原问题最优值相等，符合强对偶性）。

二、灵敏度分析：输入变化对最优值的影响

灵敏度分析回答“当资源容量（b）或产品利润（c）变化时，最优总利润如何变化”，核心是利用原最优解 x^* 和对偶最优解 y^*。

问题 1： 资源 2 的容量从 200 增加到 202，最优值如何变?

资源 2 对应对偶解的分量 $y_2^* = 0.5$（即资源 2 的影子价格为 0.5）。资源 2 的变化量 $\Delta b_2 = 202 - 200 = 2$。根据灵敏度规则：资源 i 容量变化 Δb_i 时，最优值变化量为 $\Delta b_i \times y_i^*$。变化量：$2 \times 0.5 = 1$，因此新最优值为 $250 + 1 = 251$。

问题 2： 资源 1 的容量从 100 减少到 99，最优值如何变?

资源 1 对应对偶解的分量 $y_1^* = 0$（即资源 1 的影子价格为 0）。资源 1 的变化量 $\Delta b_1 = 99 - 100 = -1$。变化量：$-1 \times 0 = 0$，因此最优值不变，仍为 250。直观解释：资源 1 的影子价格为 0，说明它是“过剩资源”，减少少量容量不影响最优生产计划。

问题 3： 产品 1 的单位利润从 1 增加到 1.02，最优值如何变?

产品 1 的利润对应原目标系数 c_1，原最优解中 $x_1^* = 50$（产品 1 的最优产量）。产品 1 利润的变化量 $\Delta c_1 = 1.02 - 1 = 0.02$。根据灵敏度规则：产品 i 利润变化 Δc_i 时，最优值变化量为 $\Delta c_i \times x_i^*$。变化量：$0.02 \times 50 = 1$，因此新最优值为 $250 + 1 = 251$。

问题 4： 产品 2 的单位利润从 2 减少到 1.97，最优值如何变?

产品 2 的利润对应原目标系数 c_2，原最优解中 $x_2^* = 100$（产品 2 的最优产量）。产品 2 利润的变化量 $\Delta c_2 = 1.97 - 2 = -0.03$。变化量：$-0.03 \times 100 = -3$，因此新最优值为 $250 - 3 = 247$。

这个例子通过生产计划场景，展示了线性规划灵敏度分析的核心逻辑：

资源的“影子价格”（对偶最优解 y^*）决定“资源容量变化对总利润的边际影响”；

产品的“最优产量”（原最优解 x^*）决定“产品利润变化对总利润的边际影响”。

【Remark】：

上述分析仅适用于**微小变化**。

微小变化意味着最优基不会改变。如果最优基发生变化，那么最优值和最优解的变化可能会大不相同，并且需要付出更多努力。

了解什么样的**变化范围**可被视为微小变化是很有用的，这样局部敏感性分析才适用。

10.2 Global Sensitivity

线性规划中右端项 b 或 c 变化的全局敏感性分析

研究线性规划中右端项向量 $\mathbf{b}$ 变为 $\mathbf{b} + \Delta\mathbf{b}$（或价值向量 $\mathbf{c}$ 变为 $\mathbf{c} + \Delta\mathbf{c}$）时，对最优解和最优基的影响，借助**单纯形表**进行分析。

回忆单纯形表：

$\boldsymbol{c}^\top - \boldsymbol{c}_B^\top \boldsymbol{A}_B^{-1}\boldsymbol{A}$	$-\boldsymbol{c}_B^\top \boldsymbol{A}_B^{-1}\boldsymbol{b}$
$\boldsymbol{A}_B^{-1}\boldsymbol{A}$	$\boldsymbol{A}_B^{-1}\boldsymbol{b}$

在最优解处：

- 检验数满足 $\boldsymbol{c}^\top - \boldsymbol{c}_B^\top \boldsymbol{A}_B^{-1}\boldsymbol{A} \geq \mathbf{0}$。
- $\boldsymbol{x}_B = \boldsymbol{A}_B^{-1}\boldsymbol{b}$ 是最优原始解的基部分，且 $\boldsymbol{y} = \boldsymbol{A}_B^{-\top}\boldsymbol{c}_B$ 是对偶最优解。

b 变化时的关键性质

检验数 $\bar{\mathbf{c}} = \mathbf{c}^\top - \mathbf{c}_B^\top A_B^{-1}A$（衡量基是否最优的指标）不依赖于 $\mathbf{b}$，因此原最优基的"最优性条件" $\bar{\mathbf{c}} \geq \mathbf{0}$ 仍成立。

所以，我们只需要考虑**基的可行性（保持最优的核心条件）：**若要原最优基 B 继续保持最优，新的基可行解（BFS）需**满足非负性**：

$$\tilde{\mathbf{x}}_B = A_B^{-1}(\mathbf{b} + \Delta\mathbf{b}) = \mathbf{x}_B^* + A_B^{-1}\Delta\mathbf{b} \geq \mathbf{0}$$

（其中 $\mathbf{x}_B^*$ 是原最优基对应的**基变量解**，而非简单点最优基，并且记得重新排序）。

新最优解与最优值：若基 B 保持最优，新的原始最优解为 $\tilde{\mathbf{x}} = [\tilde{\mathbf{x}}_B; \mathbf{0}]$（非基变量为 0）；新最优值为：

$$V(\tilde{\mathbf{b}}) = \tilde{\mathbf{b}}^\top \mathbf{y}^* = V^* + (\mathbf{y}^*)^\top \Delta\mathbf{b}$$

其中 $\mathbf{y}^* = (A_B^{-1})^\top \mathbf{c}_B$ 是对偶最优解，V^* 是原最优值。

b 的"单分量变化"分析：

若仅 $\mathbf{b}$ 的第 i 个分量变化，即 $\Delta\mathbf{b} = \lambda\mathbf{e}_i$（$\mathbf{e}_i$ 是第 i 个分量为 1 的单位向量），则只需满足：

$$\mathbf{x}_B^* + \lambda A_B^{-1}\mathbf{e}_i \geq \mathbf{0}$$

解此不等式组可得到单分量改变量 λ 的范围，在此范围内，原最优基保持最优，局部敏感性分析仍适用。

Example

问题：最大化 $x_1 + 2x_2$，约束为：

$$\begin{cases} x_1 \leq 100 \\ 2x_2 \leq 200 \\ x_1 + x_2 \leq 150 \\ x_1, x_2 \geq 0 \end{cases}$$

最优基与解：最优基为 $\{1, 2, 3\}$，最优解为 $\mathbf{x}^* = (50, 100, 50, 0, 0)^\top$.

在引入 slack variables 后：$s_1, s_2, s_3 \geq 0$:

$x_1 + s_1 = 100$ $2x_2 + s_2 = 200$ $x_1 + x_2 + s_3 = 150$

改变第三个右端项（150）：设第三个右端项变化量为 λ，即 $\Delta \mathbf{b} = \lambda \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$。

首先需要最终的单纯性表（初始基变量为 3，4，5）

Basic	x_1	x_2	s_1	s_2	s_3	RHS
B	0	0	0	1/2	1	250
1	1	0	0	-1/2	1	50
3	0	0	1	1/2	-1	50
2	0	1	0	1/2	0	100

由单纯形表得基逆矩阵 $A_B^{-1} = \begin{bmatrix} 0 & -1/2 & 1 \\ 0 & 1/2 & 0 \\ 1 & 1/2 & -1 \end{bmatrix}$，则新基变量解需满足：

$$\tilde{\mathbf{x}}_B = \begin{bmatrix} 50 \\ 100 \\ 50 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \geq \mathbf{0}$$

解不等式得 $-50 \leq \lambda \leq 50$，即第三个右端项在 $150 - 50 = 100$ 到 $150 + 50 = 200$ 之间变化时，原最优基仍保持最优。

【Remark】:

其实无需单纯形表也可以得到基逆矩阵 A_B^{-1}，方法如下：

划归标准形后的约束条件（包括松弛变量）的完整系数矩阵如下：

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

基变量为 x_1（第 1 列）、x_2（第 2 列）、s_1（第 3 列），因此：

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

然后计算逆矩阵 A_B^{-1}

但是这个方法就比较繁琐，无法直接从单纯性表中读出来。

c 变化时的关键性质

c 变化在最优范围内部的分析（当前基仍最优）

基本解的不变性：当 c 变为 $\tilde{c} = c + \Delta c$ 时，对应同一基的基本解（$x_B = B^{-1}b$，$x_N = 0$）保持不变。

证明. 原因：基本解由约束条件 $Ax = b$ 和基矩阵 B 决定，c 的变化不影响约束，因此基变量的取值 $x_B = B^{-1}b$ 不会改变。 □

为了让当前基仍为最优，需保证所有非基变量的检验数（reduced costs） 0（以最小化问题为例），即：$\bar{c}_j = c_j - c_B^T A_B^{-1} A_j$

基变量的检验数始终为 0（因为 $c_B - c_B^T B^{-1} B = 0$），因此只需验证非基变量的检验数是否满足非负性即可。

单分量 c_j 变化的范围推导：

若仅改变 c 的一个分量 c_j，变化量为 λ（即 $\tilde{c}_j = c_j + \lambda$），可通过线性不等式组推导 λ 的取值范围，需分两种情况：

- 若 $j \in B$（j 是非基变量）： 改变 c_j 直接影响**其自身的检验数**，需保证该检验数非负。

检验数推导：当基变量 j 的目标系数变为 $\mathbf{c}_B + \lambda \mathbf{e}_j$ 时（其中 $\mathbf{e}_j$ 为单位向量，第 j 个分量为 1，其余为 0），非基变量的新检验数为：

$$\begin{aligned} \mathbf{c}_N^T - (\mathbf{c}_B^T + \lambda \mathbf{e}_j^T)\mathbf{A}_B^{-1}\mathbf{A}_N &= \mathbf{c}_N^T - \mathbf{c}_B^T \mathbf{A}_B^{-1}\mathbf{A}_N - \lambda \mathbf{e}_j^T \mathbf{A}_B^{-1}\mathbf{A}_N \\ &= \mathbf{r}_N^T - \lambda \mathbf{e}_j^T \mathbf{A}_B^{-1}\mathbf{A}_N \end{aligned}$$

其中 $\mathbf{r}_N^T = \mathbf{c}_N^T - \mathbf{c}_B^T \mathbf{A}_B^{-1}\mathbf{A}_N$ 为原非基变量检验数。

最优性条件（保持当前基最优）：对于最小化问题，需所有非基变量检验数非负，即：

$$\mathbf{r}_N^T - \lambda \mathbf{e}_j^T \mathbf{A}_B^{-1}\mathbf{A}_N \geq \mathbf{0}$$

解此不等式组可得到 λ 的取值范围。

【Remark】：

要注意 C 的取值正负，尤其是最大最小问题，容易变号。

- 若 $j \in N$（j 是基变量）： 改变 c_j 会**影响所有非基变量的检验数**（因检验数公式包含 $\mathbf{c}_B$），需保证所有非基变量的检验数均非负。

检验数推导：当非基变量 j 的目标系数变为 $\mathbf{c}_N + \lambda \mathbf{e}_j$ 时（$\mathbf{e}_j$ 为非基变量集合对应的单位向量），非基变量的新检验数为：

$$
\begin{aligned}
(\mathbf{c}_N^T + \lambda \mathbf{e}_j^T) - \mathbf{c}_B^T \mathbf{A}_B^{-1} \mathbf{A}_N &= \left(\mathbf{c}_N^T - \mathbf{c}_B^T \mathbf{A}_B^{-1} \mathbf{A}_N\right) + \lambda \mathbf{e}_j^T \\
&= \mathbf{r}_N^T + \lambda \mathbf{e}_j^T
\end{aligned}
$$

此时仅第 j 个非基变量的检验数发生变化（增加 λ），其余非基变量检验数保持不变。

最优性条件（保持当前基最优）为使当前基仍为最优基（最小化问题中检验数需 0），需满足：$\mathbf{r}_N + \lambda \mathbf{e}_j \geq \mathbf{0}$ 解此不等式可得到 λ 的取值范围。

Example

考虑如下生产计划的线性规划问题：

$$
\begin{cases}
\max & x_1 + 2x_2 \\
\text{s.t.} & x_1 \leq 100 \\
& 2x_2 \leq 200 \\
& x_1 + x_2 \leq 150 \\
& x_1, x_2 \geq 0
\end{cases}
$$

其最终单纯形表为：

B	0	0	0	1/2	1	250
1	1	0	0	-1/2	1	50
3	0	0	1	1/2	-1	50
2	0	1	0	1/2	0	100

问题：第一个目标函数系数 (1) 可以在多大范围内变化，使得我们可以使用局部灵敏度分析？

我们有：

$$
\mathbf{A}_B^{-1}\mathbf{A}_N = \begin{bmatrix} -0.5 & 1 \\ 0.5 & 0 \\ 0.5 & -1 \end{bmatrix}, \quad \mathbf{r}_N = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}
$$

假设我们将第一个产品的利润从 1 变为 $1+\lambda$（即标准型中为 $-1-\lambda$）。此时，需满足：

$$
\mathbf{r}_N - (-\lambda)\mathbf{A}_N^T(\mathbf{A}_B^{-1})^T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} - \lambda \begin{bmatrix} 0.5 \\ -1 \end{bmatrix} \geq \mathbf{0}
$$

解得：$-1 \leq \lambda \leq 1$。

▷ 若第一个产品的利润系数在 0 到 2 之间，我们可以使用局部灵敏度定理，通过 $\mathbf{x}^*$ 计算最优值。

c 变化超出最优范围的处理：

Method

当 c 的变化导致当前解的检验数出现负分量时，当前基不再是最优的。此时处理方法为：继续用单纯形表迭代，直到找到新的最优解

逻辑依据：检验数为负（最大化问题）说明存在"能使目标函数值改进"的非基变量，需按照单纯形法的标准步骤（选入基变量 → 选离基变量 → 转轴操作）继续迭代，直至所有检验数非负，得到新的最优基和最优解。

Changing A

- 若变化出现在非基列（例如 A_j），则原最优解仍可行。唯一的变化是第 j 个变量的检验数。
  - 重新计算 $\bar{c}_j$。若其仍非负，则**原最优解**保持最优；否则，更新第 j 列的单纯形表以及检验数，并从该状态继续迭代。
- 若变化出现在基列，则单纯形表中几乎所有数值都会改变。一般而言，没有简单的处理方法。

Other Changes

添加变量（其余保持不变）：

- 将新增变量赋值为 0，并作为非基变量。原基本可行解（BFS）仍为 BFS，检验数不变。
- 仅需检查对应新变量的检验数。
- 若其非负，则原最优解加上新增的非基变量仍最优；否则从该状态继续使用单纯形法迭代。

添加约束

- 若原最优解满足该约束，则其仍最优。
- 若不满足，则最佳处理方式是将其理解为新增一个对偶变量，然后使用对偶问题的单纯形表继续计算。

11 博弈论与网络优化模型

11.1 博弈论 Game Theorem

意义

- 用于理解竞争局势的数学工具
- 常用于商业、经济学中，以确定最佳策略和政策

我们的主题： 双人**零和**博弈

双人零和博弈

- 双人零和博弈由以下要素确定
- $S_1, \ldots, S_m$：参与者 A 的策略，
- $T_1, \ldots, T_n$：参与者 B 的策略
- A 的收益表：显示在两种策略组合下参与者 A 的收益

参与者 A 的收益表

		B				
		T_1	$\ldots$	T_j	$\ldots$	T_n
A	S_1	p_{11}	$\ldots$	p_{1j}		p_{1n}
	$\vdots$	$\vdots$	$\ddots$	$\vdots$		$\vdots$
	S_i	p_{i1}	$\ldots$	p_{ij}		p_{in}
	$\vdots$	$\vdots$	$\ddots$	$\vdots$		$\vdots$
	S_m	p_{m1}	$\ldots$	p_{mj}		p_{mn}

博弈论中的主要假设

- 两个参与者**都是聪明**且**理性**的！

引例：

Example

两个参与者，每人有一枚硬币，各自秘密地将硬币翻为正面或反面，同时展示他们的选择

- 如果硬币相同，则参与者 A 赢得两枚硬币
- 如果不同，则参与者 B 赢得两枚硬币

$\Rightarrow$ 零和博弈

表示方法：

参与者 A 的收益表：

	$B:$	$B:$
$A:$	1	-1
$A:$	-1	1

参与者 B 的收益表：

	$B:$	$B:$
$A:$	-1	1
$A:$	1	-1

已知条件与假设

已知： 参与者 A 的收益表。

假设： 参与者都是聪明且理性的。

问题： 参与者 A 应分别以多大的概率（即比例）x_i 采用每一种策略 S_i，且参与者 B 应分别以多大的概率（即比例）y_j 采用每一种策略 T_j，**从而使得 A 最大化自身收益，同时 B 最小化自身损失**。(因为 B 在假设条件下是理性的，必然要减少损失)

参与者策略逻辑

参与者 A 知道参与者 B 是聪明的参与者，因此不会允许 A 获取越来越多的收益。所以，参与者 A 的目标是**最大化自己能获得的最小收益**，即采用**极大极小准则**（Maximin criterion）。

类似地，参与者 B 的目标是**最小化自己的最大损失**，即采用**极小极大准则**（Minimax criterion）。

博弈的价值（Value of the game）等于当双方都采用最优策略时，参与者 A 的收益。

核心概念：占优策略（Dominated Strategies）

> **Definition**
>
> 若策略 S_j 无论对手采取什么策略，都至少和策略 S_i "一样好"（对自身收益更有利，或对损失更可控），则称 S_i 被 S_j **占优**（Dominated）。

- 对参与者 A（追求收益最大化）：若 S_j 的收益 $\geq$ S_i 的收益（对所有对手策略），则 S_i 被 S_j 占优。
- 对参与者 B（追求损失最小化）：若 S_j 导致的"损失"（即 A 的收益）$\leq$ S_i 导致的"损失"（对所有对手策略），则 S_i 被 S_j 占优。

性质： **理性参与者不会选择占优策略**（因为有更优的替代策略），因此占优策略可以从策略集合中**消除**，从而简化博弈分析。

迭代消除占优策略的过程

通过" 参与者互知对方是理性的" 这一前提，逐步消除占优策略，最终得到纯策略最优解。

示例过程演绎-1

初始收益表（Player A 的收益）：

	T_1	T_2	T_3
S_1	1	2	4
S_2	1	0	5
S_3	0	1	-1

步骤 1：参与者 A 消除自身的占优策略 S_3

- 分析：无论参与者 B 选 T_1, T_2, T_3 中哪一个，A 选 S_1 的收益都高于 S_3：
  - B 选 T_1：S_1 收益 $1 > S_3$ 收益 0；
  - B 选 T_2：S_1 收益 $2 > S_3$ 收益 1；
  - B 选 T_3：S_1 收益 $4 > S_3$ 收益 -1。
- 结论：S_3 被 S_1 占优，A 会消除 S_3，博弈简化为：

	T_1	T_2	T_3
S_1	1	2	4
S_2	1	0	5

【Remark】：

B 希望的是数组越小越好，因为 A 的收益对 B 而言就是损失，B 自然趋向于损失小的方向。

步骤 2：参与者 B 消除自身的占优策略 T_3

- 前提：B 知道 A 是理性的（会消除 S_3），因此只需考虑 A 的策略 S_1, S_2。
- 分析：无论 A 选 S_1 还是 S_2，B 选 T_1 的 "损失"（即 A 的收益）都小于 T_3：
  - A 选 S_1：B 选 T_1 损失 $1 <$ 选 T_3 损失 4；
  - A 选 S_2：B 选 T_1 损失 $1 <$ 选 T_3 损失 5。
- 结论：T_3 被 T_1 占优，B 会消除 T_3，博弈简化为：

	T_1	T_2
S_1	1	2
S_2	1	0

步骤 3：参与者 A 再次消除自身的占优策略 S_2

- 前提：A 知道 B 是理性的（会消除 T_3），因此只需考虑 B 的策略 T_1, T_2。
- 分析：无论 B 选 T_1 还是 T_2，A 选 S_1 的收益都高于 S_2：
  - B 选 T_1：S_1 收益 1 = S_2 收益 1（但 S_1 在其他情况更优）；
  - B 选 T_2：S_1 收益 2 > S_2 收益 0。
- 结论：S_2 被 S_1 占优，A 会消除 S_2，博弈简化为：

	T_1	T_2
S_1	1	2

步骤 4：参与者 B 消除自身的占优策略 T_2

- 前提：B 知道 A 是理性的（会消除 S_2），因此只需考虑 A 的策略 S_1。
- 分析：当 A 固定选 S_1 时，B 选 T_1 的损失（1）小于选 T_2 的损失（2）。
- 结论：T_2 被 T_1 占优，B 会消除 T_2，博弈最终简化为：

	T_1
S_1	1

最终结论：纯策略最优解与博弈价值

- **参与者 A 的最优策略**：始终选择 S_1（通过"极大极小准则"，最大化自身的最小收益）。
- **参与者 B 的最优策略**：始终选择 T_1（通过"极小极大准则"，最小化自身的最大损失）。
- **博弈的价值**：当双方都采取最优策略时，参与者 A 的收益为 1。
- **观察**：此时双方都选择**单一策略**（而非"以概率混合多个策略"），这种最优策略称为**纯最优策略（Pure Optimal Strategy）**。

示例过程演绎-2

		B		(min profit by playing S_i)
	T_1	T_2	T_3	min
S_1	-3	-2	4	-3
S_2	2	0	2	0
S_3	5	-2	-4	-4
(max loss by playing T_i)	5	0	4	

考虑玩家 A

- 玩家 B 是聪明的，会保护自己避免大额损失, 因此 B 会选择 T_1，确保 A 遭受最大损失
- 所以如果 A 选择 S_1，她能获得的最好结果只有 -3，即行最小值
- 每一行都是如此

	B			(min profit by playing S_i)
	T_1	T_2	T_3	min
S_1	-3	-2	4	-3
S_2	2	0	2	0
S_3	5	-2	-4	-4
(max loss by playing T_i)	5	0	4	

考虑玩家 B

- 玩家 A 是聪明的，会希望赢得最多利润 (即：保护自己避免大额损失), 因此 A 会选择 S_3，确保 B 遭受最大损失
- 所以如果 B 选择 T_1，他能获得的最好结果只有 5（即损失 5），即列最大值
- 每一列都是如此

	B			(min profit by playing S_i)
	T_1	T_2	T_3	min
S_1	-3	-2	4	-3
S_2	2	0	2	0
S_3	5	-2	-4	-4
(max loss by playing T_i)	5	0	4	

但是我们知道：

玩家 A：最大化最小值准则（Maximin criterion）

玩家 B：最小化最大值准则（Minimax criterion）

观察： 行最小值中的最大值和列最大值中的最小值在同一个位置达到

- 如果玩家 A 始终使用单一策略，那一定是 S_2
- 如果玩家 B 始终使用单一策略，那一定是 T_2
- 如果任何一方偏离，对手都会获利
- 因此玩家 A 始终选择 S_2 以最大化她的最小收益
- 而玩家 B 始终选择 T_2 以最小化他的最大损失
- 博弈值 $= 0$

: 最优策略是始终使用单一策略。这种最优策略还是为："纯最优策略"（pure optimal strategy）

鞍点（Saddle Point）

Definition

对于一个 A 与 B 的零和博弈：

- 考虑每一行的最小值
- 考虑每一列的最大值

如果行最小值中的最大值和列最大值中的最小值在支付表中的同一个位置达到，则该位置称为鞍点（saddle point）

表中这个共同的位置给出了博弈值

玩家 A 的最优策略是该位置所在行对应的策略

玩家 B 的最优策略是该位置所在列对应的策略

【Remark】:

鞍点可能不存在

B (min profit by playing S_i)

A (max loss by playing T_i)

	T_1	T_2	min
S_1	1	−1	−1
S_2	−1	1	−1
max	1	1	

- 优势法则不能剔除任何策略
- 没有鞍点，由此没有纯最优策略，因此需要寻找混合最优策略（mixed optimum）

示例过程演绎-3

B

A		T_1	T_2	T_3
x	$\|S_1$	0	-2	2
$1-x$	$\|S_2$	5	4	-3

考虑玩家 A

- 如果玩家 B 始终选择 T_1，则玩家 A 的收益为 $0x + 5(1-x)$
- 如果玩家 B 始终选择 T_2，则玩家 A 的收益为 $-2x + 4(1-x)$
- 如果玩家 B 始终选择 T_3，则玩家 A 的收益为 $2x - 3(1-x)$
- 玩家 A：最大化最小值准则（Maximin criterion）（最大化最小收益）

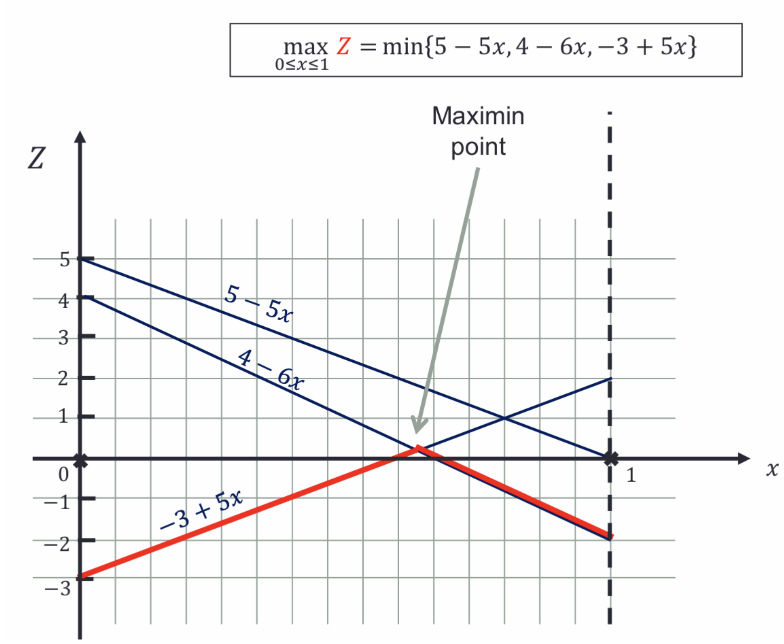
$$\boxed{\max_{0\le x\le 1} Z = \min\{5-5x,\ 4-6x,\ -3+5x\}}$$

- 玩家 A 的最优混合策略为 $\left(x = \frac{7}{11},\ 1-x = \frac{4}{11}\right)$
- 博弈值为 $-3 + 5\left(\frac{7}{11}\right) = \frac{2}{11}$

示例过程演绎-3

B

A	y_1 $\|T_1$	y_2 $\|T_2$	$1-y_1-y_2$ $\|T_3$
S_1	0	-2	2
S_2	5	4	-3



玩家 B 的最优策略是什么?

- 如果玩家 A 始终选择 S_1，则玩家 B 的损失为 $0y_1 - 2y_2 + 2(1 - y_1 - y_2)$
- 如果玩家 A 始终选择 S_2，则玩家 B 的损失为 $5y_1 + 4y_2 - 3(1 - y_1 - y_2)$

得到:

$$\min_{\substack{0 \le y_1 \le 1 \\ 0 \le y_2 \le 1 \\ y_1 + y_2 \le 1}} w = \max\{2 - 2y_1 - 4y_2,\ -3 + 2y_1 + y_2\}$$

两个变量 y_1, y_2，因此无法使用图解法。

示例过程演绎-4

			B	
			T_1	T_2
A	x	S_1	3	-2
	$1-x$	S_2	1	2

考虑玩家 A

- 如果玩家 B 始终选择 T_1，则玩家 A 的收益为 $3x + (1 - x)$
- 如果玩家 B 始终选择 T_2，则玩家 A 的收益为 $-2x + 2(1 - x)$

$$\max_{0 \le x \le 1} Z = \min\{2x + 1,\ -4x + 2\} \quad = \quad \boxed{\begin{array}{l} \max u \\ u \le 2x + 1 \\ u \le -4x + 2 \\ 0 \le x \le 1 \end{array}}$$ 可划归为线性规划

化为标准型后，求函数 $\max u + 0 \times x$ 的**最大化**：
求解得 $x = \frac{1}{6}$
博弈值 = 玩家 A 的收益 $= \frac{4}{3}$

示例过程演绎-4

			B	
			T_1	T_2
A	x	S_1	3	-2
	$1-x$	S_2	1	2

考虑玩家 B

- 如果玩家 A 始终选择 S_1，则玩家 B 的损失为 $3y - 2(1-y)$
- 如果玩家 A 始终选择 S_2，则玩家 B 的损失为 $y + 2(1-y)$
- 玩家 B：最小最大准则（Minimax criterion）（最小化最大损失）

$$\min_{0 \le y \le 1} v = \max\{5y-2,\ -y+2\} \quad = \quad \boxed{\begin{array}{l} \min v \\ v \ge 5y-2 \\ v \ge -y+2 \\ 0 \le y \le 1 \end{array}}$$

可以使用单纯形法求解

线性规划公式

			B		
			$y_1 \quad T_1$	$y_2 \quad T_2$	$y_3 \quad T_3$
A	x_1	S_1	0	-2	2
	x_2	S_2	5	4	-3

玩家 A

$$\max u$$
$$u \leq 5x_2 + 4x_2$$
$$u \leq -2x_1 + 4x_2$$
$$u \leq 2x_1 - 3x_2$$
$$x_1 + x_2 = 1$$
$$x_1, x_2 \geq 0$$
u 无约束

玩家 A：最大化最小值准则

最大化最小收益

玩家 B

$$\min v$$
$$v \geq -2y_2 + 2y_3$$
$$v \geq 5y_1 + 4y_2 - 3y_3$$
$$y_1 + y_2 + y_3 = 1$$
$$y_1, y_2, y_3 \geq 0$$
v 无约束

玩家 B：最小最大准则

最小化最大损失

|| 原问题

|| 对偶问题

$$\max u$$
$$u - 5x_2 \leq 0 \quad \xrightarrow{q_1}$$
$$u + 2x_1 - 4x_2 \leq 0 \quad \xrightarrow{q_2}$$
$$u - 2x_1 + 3x_2 \leq 0 \quad \xrightarrow{q_3}$$
$$x_1 + x_2 = 1 \quad \xrightarrow{w}$$
$$x_1, x_2 \geq 0$$
u 无约束

$$\min w$$
$$w + 2q_2 - 2q_3 \geq 0$$
$$w - 5q_1 - 4q_2 + 3q_3 \geq 0$$
$$q_1 + q_2 + q_3 = 1$$
$$q_1, q_2, q_3 \geq 0$$
w 无约束

求解玩家 B 的最优混合策略

原问题	对偶问题
$\max u$	$\min v$
$u-5x_2\le 0 \quad \xrightarrow{y_1}$	$v+2y_2-2y_3\ge 0 \quad \xrightarrow{x_1}$
$u+2x_1-4x_2\le 0 \quad \xrightarrow{y_2}$	$v-5y_1-4y_2+3y_3\ge 0 \quad \xrightarrow{x_2}$
$u-2x_1+3x_2\le 0 \quad \xrightarrow{y_3}$	$y_1+y_2+y_3=1 \quad \xrightarrow{u}$
$x_1+x_2=1 \quad \xrightarrow{v}$	$y_1,y_2,y_3\ge 0$
$x_1,x_2\ge 0$	v 无约束
u 无约束	
$x_1^*=\frac{7}{11},\ x_2^*=\frac{4}{11},\ u^*=\frac{2}{11}$	$y_1^*=0,\ y_2^*=\frac{5}{11},\ y_3^*=\frac{6}{11},\ v^*=\frac{2}{11}$

- 玩家 B 的最优混合策略为 $(y_1^*=0,\ y_2^*=\frac{5}{11},\ y_3^*=\frac{6}{11})$
- 强对偶性意味着 $v^*=\frac{2}{11}$
- 解出原问题后，再用互补松弛条件求对偶问题的解：应用互补松弛条件：

$$
\begin{aligned}
&y_1^*(u^*-5x_2^*)=0 &&\Rightarrow\quad y_1^*\left(-\frac{18}{11}\right)=0 \quad\Rightarrow\quad y_1^*=0\\
&y_2^*(u^*+2x_1^*-4x_2^*)=0 &&\Rightarrow\quad 0y_2^*=0\\
&y_3^*(u^*-2x_1^*+3x_2^*)=0 &&\Rightarrow\quad 0y_3^*=0\\
&x_1^*(v^*+2y_2^*-2y_3^*)=0 &&\Rightarrow\quad 2y_2^*-2y_3^*=-\frac{2}{11}\\
&x_2^*(v^*-5y_1^*-4y_2^*+3y_3^*)=0 &&\Rightarrow\quad -4y_2^*+3y_3^*=-\frac{2}{11}\\
&y_1^*+y_2^*+y_3^*=1
\end{aligned}
$$

$$\Rightarrow\quad y_2^*=\frac{5}{11},\ y_3^*=\frac{6}{11}$$

线性规划公式：一般情况

$$
\begin{array}{c|cc|cc|c|cc|c|cc}
 & & & & & & B & & & & \\
 & & & y_1 & T_1 & \dots & y_j & T_j & \dots & y_n & T_n \\
\hline
 & x_1 & S_1 & & p_{11} & \dots & & p_{1j} & \dots & & p_{1n} \\
 & & \vdots & & \vdots & \ddots & & \vdots & \ddots & & \vdots \\
A & x_i & S_i & & p_{i1} & \dots & & p_{ij} & \dots & & p_{in} \\
 & & \vdots & & \vdots & \ddots & & \vdots & \ddots & & \vdots \\
 & x_m & S_m & & p_{m1} & \dots & & p_{mj} & \dots & & p_{mn} \\
\end{array}
$$

已知：玩家 A 的收益表

玩家 A	玩家 B
$\max u$ $u \leq \sum_{i=1}^{m} p_{ij} x_i \quad \forall j = 1, \ldots, n$ $\sum_{i=1}^{m} x_i = 1$ $x_i \geq 0 \quad \forall i = 1, \ldots, m$ u 无约束 玩家 A：最大化最小值准则 最大化最小收益	$\min v$ $v \geq \sum_{j=1}^{n} p_{ij} y_j \quad \forall i = 1, \ldots, m$ $\sum_{j=1}^{n} y_j = 1$ $y_j \geq 0 \quad \forall j = 1, \ldots, n$ v 无约束 玩家 B：最小最大准则 最小化最大损失

由此我们挖掘出对偶性的一个重要结论：

Theorem

最小最大定理（Minimax Theorem）：对于一对根据最大化最小值和最小最大值准则为最优的混合策略 (x^*, y^*):

1. 值 u^* 和 v^* 将相等;

2. 任何一方都不能通过单方面改变自己的策略而获得更好的结果

即，当玩家 B 继续采用其最优策略 y^* 时，玩家 A 不能通过改变策略而获得超过 u^* 的收益

当玩家 A 继续采用其最优策略 x^* 时，玩家 B 不能通过改变策略而使损失小于 v^*.

11.2 Nash Equilibrium

Definition

纳什均衡是非合作博弈中的一种策略组合：若策略组合 (x^*, y^*) 对参与者 A 和 B 满足以下" 任何参与者都无法通过单方面改变策略来提高自身收益。"，则称为（纳什）均衡：

对参与者 A：当 B 采用 y^* 时，没有任何策略 x 的收益优于 x^*。

对参与者 B：当 A 采用 x^* 时，没有任何策略 y 的收益优于 y^*。

1. 纯策略纳什均衡："百分百选某一策略" 参与者选择确定的行动（而非概率性选择），且该行动是对他人策略的 "最优反应"。

两个囚犯被隔离审讯，规则为：

- 都坦白：各判 5 年；

- 都抵赖：因证据不足各判 1 年；
- 一人坦白、一人抵赖：坦白者释放，抵赖者判 8 年。

对每个囚犯而言：

- 若对方“坦白”，自己“坦白”（判 5 年）比“抵赖”（判 8 年）更优；
- 若对方“抵赖”，自己“坦白”（释放）比“抵赖”（判 1 年）更优。

因此，(,) 是唯一的纯策略纳什均衡——双方都没有动力单方面改变策略。

此外，纳什均衡还揭示了“个体理性”与“集体理性”的冲突（如囚徒困境中，个体选“坦白”是理性的，但集体结果“都坦白”比“都抵赖”更差，坦白从严牢底坐穿，而抗拒能回家过年），为理解“非合作互动下的局限与可能”提供了关键视角。

2. 混合策略纳什均衡：“概率性选择策略”参与者以概率分布随机选择不同行动（而非百分百选某一个），且这种概率选择是对他人策略的“最优反应”。

案例：政府救济流浪汉：

政府策略：“救济”或“不救济”；流浪汉策略：“找工作”或“游荡”。

若政府选“救济”，流浪汉最优策略是“游荡”；若政府选“不救济”，流浪汉最优策略是“找工作”。此时无纯策略纳什均衡，但存在混合策略均衡：

假设政府以 p 概率选“救济”、$1-p$ 概率选“不救济”；流浪汉以 q 概率选“找工作”、$1-q$ 概率选“游荡”。

通过“最大化期望收益”的推导，最终会得到一组概率（如流浪汉以 0.2 概率找工作、0.8 概率游荡），此时双方都无法通过单方面改变概率分布来提高收益，这就是混合策略纳什均衡。

用上面的例子，考虑收益矩阵（参与者 A：行；参与者 B：列）：

	T_1	T_2	T_3	最小值（参与者 A 的收益）
S_1	-3	-2	4	-3
S_2	2	0	2	0
S_3	5	-2	-4	-4
最大值（参与者 B 的损失）	5	0	4	

关键要点：

参与者 A 选择 S_2（最大最小准则：最大化最小收益）。

参与者 B 选择 T_2（最小最大准则：最小化最大损失）。

任何一方偏离策略都会让对手获利。

博弈值：0；$(x=(0,1,0), y=(0,1,0))$（纯策略 S_2, T_2）是纳什均衡。

存在性：双人零和博弈：均衡一定存在（通过线性规划对偶性）。

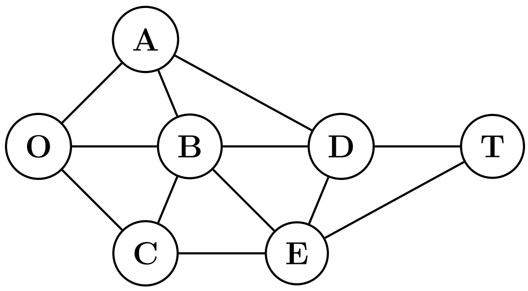
纳什（作为学生）证明了大量博弈类型（包括双人零和博弈）中均衡的存在性。

我们观察到：

- 所有纯最优策略都是纳什均衡。
- 所有鞍点都是纳什均衡。
- 纳什均衡可包含混合策略（概率性策略组合）。
- 验证方法：检查当对手策略固定时，双方均无法提高收益。

11.3 网络优化 Network Opyimization Models

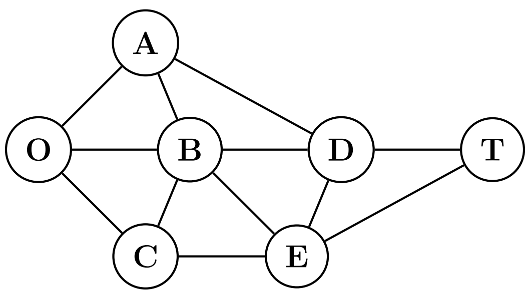
网络（图）$\mathcal{G}$ 是**节点**（顶点）和**弧**（边）的集合，记为 $\mathcal{G} = (\mathcal{V}, \mathcal{A})$。其中，**弧**表示节点之间的连接关系。



节点集合 $\mathcal{V} = \{A, B, C, D, E, T, O\}$,弧集合 $\mathcal{A} = \{AB, AO, AD, BC, BO, BD, BE, CE, CO, DE,$

子图：
图 $\mathcal{G}$ 的**子图** $\mathcal{H}$ 满足：

- $\mathcal{H}$ 包含 $\mathcal{G}$ 节点的一个子集 S；
- $\mathcal{H}$ 包含 $\mathcal{G}$ 中 S 内节点之间的弧的一个子集。

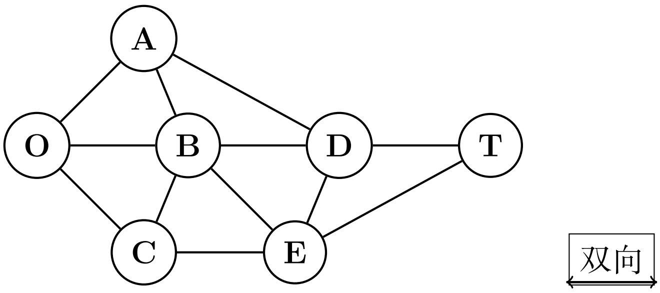


（注：子图通过选择节点和弧的子集得到，上图与“网络（图)”结构一致，仅需限制节点/弧范围即可构造子图。)

无向与有向网络

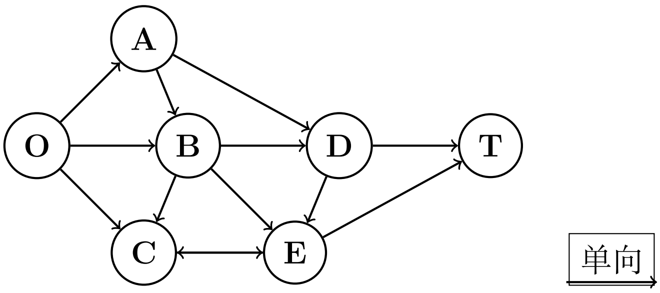
无向弧与无向网络

无向弧：弧没有方向，可双向遍历（此类弧也称为“链接”或“边”)。由无向弧组成的网络称为**无向网络**。



有向弧与有向网络

有向弧：弧有方向，只能沿指定方向遍历。由有向弧组成的网络称为**有向网络**。

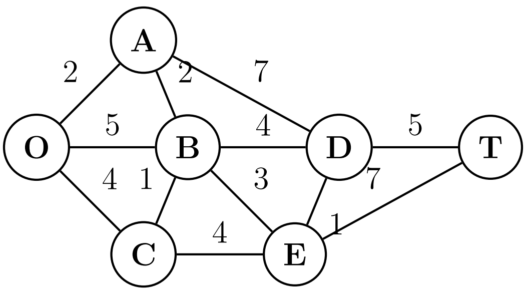


网络要么是有向网络，要么是无向网络（即所有弧同时为有向或无向)。

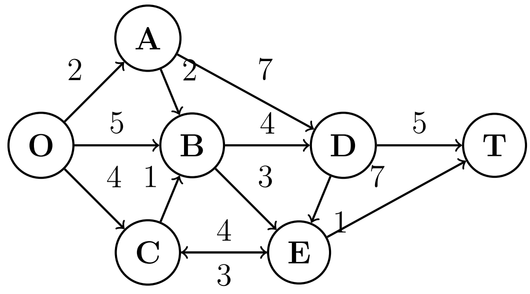
加权网络:

加权网络：通常在弧上赋予**权重**。例如，权重可表示“使用该弧的成本（时间/距离)”或“弧的容量”。若未指定权重，则默认权重为 1。

无向加权网络:



有向加权网络:

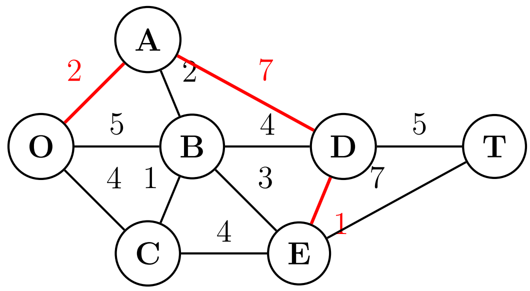


路径:

节点 i 到节点 j 的**路径**：连接 i 到 j 的一系列**不同弧**的序列，且每个节点最多被访问一次。

路径长度：路径中所有边的权重之和。

例如，下图中**红色路径**的长度为 $2+7+1=10$。



路径 $O \to A \to D \to E$ 的长度为 $2+7+1=10$。

最短路径问题：给定带权无向网络 $\mathcal{H}=(\mathcal{V},\mathcal{A})$、起点 O 和终点 T，其中弧的权重对应旅行时间（或距离、旅行成本），且所有弧的权重非负，目标是找到从起点到终点的最短路径。

对于有向图中的每条边 $(i,j)\in E$，定义：

$$x_{ij}=\begin{cases}1, & \text{若使用边}(i,j),\\ 0, & \text{否则}.\end{cases}$$

由于是有向图，x_{ij} 不一定等于 x_{ji}。此时可建立如下优化模型：

$$\begin{aligned}
\underset{\{x_{ij}\}}{\text{minimize}} \quad & \sum_{(i,j)\in E} w_{ij}x_{ij}\\
\text{subject to} \quad & x_{ij}\in\{0,1\}, \quad \forall (i,j)\in E,\\
& \sum_{\substack{j:\\(s,j)\in E}} x_{sj} - \sum_{\substack{j:\\(j,s)\in E}} x_{js} = 1,\\
& \sum_{\substack{j:\\(j,t)\in E}} x_{jt} - \sum_{\substack{j:\\(t,j)\in E}} x_{tj} = 1,\\
& \sum_{\substack{j:\\(i,j)\in E}} x_{ij} = \sum_{\substack{j:\\(j,i)\in E}} x_{ji}, \quad \forall i\neq s,t.
\end{aligned}$$

该优化问题属于带约束的整数优化问题，通常称为整数（线性）优化。

【Remark】:

起点不一定就没有流入的边；同理，终点不一定没有流出的边。

约束的核心是模拟“路径”的流量特性：从起点出发、到终点结束、中间节点“进多少出多少”。

（1）变量的二进制约束

$$x_{ij} \in \{0, 1\},\ \forall (i, j) \in E$$

保证每条边“选或不选”是二值决策，因此模型属于整数线性规划（变量取整数 0/1，目标和约束均为线性）。

（2）起点的流量约束

$$\sum_{\substack{j:\\(s,j)\in E}} x_{sj} - \sum_{\substack{j:\\(j,s)\in E}} x_{js} = 1$$

s 是起点（origin node）；左边第一项：从起点 s 出发的所有边（(s, j)）中，被选中的数量之和（“离开起点的次数”）；左边第二项：进入起点 s 的所有边（(j, s)）中，被选中的数量之和（“进入起点的次数”）；差值为 1：表示路径从起点出发（离开次数比进入次数多 1）。

（3）终点的流量约束

$$\sum_{\substack{j:\\(j,t)\in E}} x_{jt} - \sum_{\substack{j:\\(t,j)\in E}} x_{tj} = 1$$

t 是终点（destination node）；左边第一项：进入终点 t 的所有边（(j, t)）中，被选中的数量之和（“进入终点的次数”）；左边第二项：从终点 t 出发的所有边（(t, j)）中，被选中的数量之和（“离开终点的次数”）；差值为 1：表示路径到终点结束（进入次数比离开次数多 1）。

（4）中间节点的流量守恒约束

$$\sum_{\substack{j:\\(i,j)\in E}} x_{ij} = \sum_{\substack{j:\\(j,i)\in E}} x_{ji},\ \forall i \neq s, t$$

i 是中间节点（既不是起点也不是终点）；左边：从节点 i 出发的所有边中，被选中的数量之和（“离开中间节点的次数”）；右边：进入节点 i 的所有边中，被选中的数量之和（“进入中间节点的次数”）；等式要求：中间节点的“进入次数”等于离开次数，保证路径经过中间节点时“流量守恒”（不会在中间节点累积或中断）。