# 线性规划

算法设计与分析小班课

2022 年 4 月

**注意:** 以下内容主要是给课本作一些注解并补充习题, 可能不能准确地代表课程内容和考核的方向, 也不能替代阅读课本、听课、完成作业题、做往年题等活动; 其中的内容没有经过课程主管老师的审核, 也可能存在错误. 不过出现的所有错误都由作者本人负责.

线性规划是 1939 年苏联数学家 Kantorovich 在《组织和计划生产的数学方法》中最早提出的, 它是一个非常广泛的数学模型, 可以说每一个运用优化工具的研究者都应该了解线性规划的基本技术.

## 1 建模问题

除了最直接的 "目标函数 + 建模" 的问题, 许多看上去比较复杂的事情都可以用线性规划来处理, 以下选取几个比较有意思的例子.

**问题 1** 有 $n$ 项活动同时申请使用同一个礼堂, 每个活动 $i$ 都有一个开始时间 $s_i$ 和结束时间 $f_i$, 任何两个活动不能同时举行 (端点除外). 举办活动 $i$ 的收益是 $v_i \in \mathbb{R}_{\geq 0}$. 设计一个算法选择一些活动, 使得收益最大. 写出问题的整数线性规划模型.

**解** 根据给出的信息构造冲突图 (冲突的活动连一条边), 设该冲突图的关联矩阵是 $M$, 并用 $\boldsymbol{x} \in \{0,1\}^n$ 代表最终的选择, 那么活动选择的限制就是取一个独立集 $M^\top \boldsymbol{x} \leq \mathbf{1}$, 目标函数是 $\boldsymbol{v}^\top \boldsymbol{x}$. (注: 用课上讲的方法也可以.) □

**注 2** 这个问题的有趣之处在于: 虽然以上问题是整数线性规划, 但我们可以证明, 若放松限制使得 $x_i \in [0,1]$, 那么最优解中 $x_i$ 只能是 0 或者 1, 所以它可以直接用普通的线性规划方法直接求解. ♤

**例 3** 图的最大匹配可以写为 (整数) 线性规划问题.

设图 $G$ 的关联矩阵 $M \in \{0,1\}^{n\times m}$, 令 $\boldsymbol{x} \in \{0,1\}^m$ 的含义为边集 $X \subseteq E(G)$ 的特征函数 (即是 $\mathbf{1}_{e\in X}$). 考虑约束 $M\boldsymbol{x} \leq \mathbf{1}$, 它表示 $X$ 中的边同 $G$ 中的每个点至多有一次邻接, 即代表一个匹配; 如果再设定最大化目标函数 $\max \mathbf{1}^\top \boldsymbol{x} = |X|$, 则这就对应于最大匹配.

即是:

$$\begin{aligned}\max_{\boldsymbol{x}} \quad & \mathbf{1}^\top \boldsymbol{x}, \\ \text{s.t.} \quad & M\boldsymbol{x} \le \mathbf{1}, \boldsymbol{x} \ge \mathbf{0}.\end{aligned}$$

类似地, 若给出带权图中每条边的权值 $\boldsymbol{w} \in \mathbb{R}^m$, 目标函数改为 $\boldsymbol{w}^\top \boldsymbol{x}$, 则就对应于带权图匹配的优化问题.

一般地求解整数线性规划是 NP-难的. 但通过放松对 $\boldsymbol{x}$ 的要求, 允许其取非整数值, 那么仍然是可以进行合理的估计的, 因为对于最大化问题 ILP $\le$ LP, 而对于最小化问题 LP $\le$ ILP. 而且幸运的是, 对于像上面这样的二部图中的问题, 可以论证最优解 $\boldsymbol{x}^*$ 一定代表一个合法的边集, 即每个分量只可能是 0 或者 1. 这样, 二部图的匹配问题就可以完全由线性规划的方法予以研究. ◇

**例 4** 让我们来写出例 3 中的问题的对偶, 它是

$$\begin{aligned}\min_{\boldsymbol{y}} \quad & \mathbf{1}^\top \boldsymbol{y}, \\ \text{s.t.} \quad & M^\top\boldsymbol{y} \ge \mathbf{1}, \boldsymbol{y} \ge \mathbf{0}.\end{aligned}$$

若将 $\boldsymbol{y} \in \{0,1\}^n$ 理解为点集 $X \subseteq V(G)$ 的特征函数 (即是 $\mathbf{1}_{v\in X}$), 约束此时表示 $G$ 中的每一条边都和 $X$ 中的某个点相邻, 最小化目标函数就是要求最小的满足这个条件的点集的大小.

在图 $G$ 中, 如果 $S \subseteq V(G)$ 使得每条边至少有一个端点在 $S$ 中, 则把 $S$ 称为 $G$ 的一个**点覆盖**. 因此最大匹配的对偶问题是最小点覆盖问题! ◇

**问题 5** 甲乙两人玩改版的石头剪子布游戏, 这个游戏中有一个新策略 X. 该策略会输给布, 但能够胜过石头和剪刀. 设胜利的收益是 1, 平手的收益是 0, 失败的收益是 −1. 游戏允许甲方使用四种策略中的任意一种, 乙方只允许出石头、剪刀或者布.

(1) 设甲分别以 $\boldsymbol{p} = (p_1, p_2, p_3, p_4)^\top$ 的概率出石头、剪刀、布和策略 X, 乙则以 $\boldsymbol{q} = (q_1, q_2, q_3)$ 的概率出石头、剪刀和布. 求此时甲和乙的期望收益.

(2) 假设甲、乙都是绝顶聪明的人, 他们的最佳策略分别是什么? 用线性规划建立该问题的数学模型, 尝试求解.

**解** 令矩阵

$$A = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & -1 \end{bmatrix},$$

则甲的期望收益应该是 $\boldsymbol{p}^\top A\boldsymbol{q}$ (甲出策略 $i$ 且乙出策略 $j$ 的概率是 $p_i q_j$, 故贡献 $A_{ij} p_i q_j$). 因为甲、乙都是绝顶聪明的人, 所以甲可以考虑对于任何的 $\boldsymbol{p}$, 乙都出合适的最差策略使

得自己的期望收益最小, 也就是说甲希望最大化最坏情况下的收益:

$$\begin{aligned} \max_{\boldsymbol{p}} \min_{1\le j\le 3} \quad & \boldsymbol{p}^\top A_{\cdot j}, \\ \text{s.t.} \quad & \mathbf{1}^\top \boldsymbol{p} = 1, \boldsymbol{p} \ge \mathbf{0}. \end{aligned}$$

这里 $A_{\cdot j}$ 表示矩阵的第 $j$ 列. 为了将它改写成比较标准的线性规划问题, 我们可以引入变量 $z$ 来表示内层优化, 即:

$$\begin{aligned} \max_{\boldsymbol{p}, z} \quad & z \\ \text{s.t.} \quad & z \le \boldsymbol{p}^\top A_{\cdot j}\ (\forall 1 \le j \le 3), \\ & \mathbf{1}^\top \boldsymbol{p} = 1, \boldsymbol{p} \ge \mathbf{0}. \end{aligned}$$

同理可处理乙的情形. 通过求解上述线性规划可知道最优策略是 $\boldsymbol{p}^* = \left(0, \frac{2}{9}, \frac{4}{9}, \frac{1}{3}\right), \boldsymbol{q}^* = \left(\frac{1}{3}, \frac{4}{9}, \frac{2}{9}\right)$. (这类博弈问题和对偶性有着紧密的联系.) □

网络流、最短路问题也可以写成线性规划, 这些问题的建模都是类似的, 详参 [Cor+09, §29.2 节], 另外在网络流的讨论中我们还会提及.

# 2 对偶性

**2.1 怎样取对偶?** 以下我们用 Lagrange 乘子的方法引出对偶性. 考虑一个简单的、一般的优化问题:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f_0(\boldsymbol{x}), \\ \text{s.t.} \quad & f_1(\boldsymbol{x}) \le 0. \end{aligned}$$

如果没有约束, 直接优化 $f_0(\cdot)$ 是比较容易的. 回忆我们或多或少听过的 “正则化” 方法, 我们可以暴力地决定直接优化

$$J(\boldsymbol{x}) = f_0(\boldsymbol{x}) + \mathbf{1}_{f_1(\boldsymbol{x})>0} \cdot \infty.$$

这样如果存在可行解, 则上式就不会出现 $\infty$ 项. 当然, 这是一个很糟糕的函数, 因为它的性质很差. 作为一个替代, 考虑

$$L(\boldsymbol{x}, \lambda) \stackrel{\text{def}}{=} f_0(\boldsymbol{x}) + \lambda f_1(\boldsymbol{x}) \quad (\lambda \ge 0).$$

(这是大家熟悉的东西, 即 Lagrange 乘子.) 一个很重要的观察是 $\max_{\lambda\ge 0} L(\boldsymbol{x},\lambda) = J(\boldsymbol{x})$. 因此我们的问题就是

$$\min_{\boldsymbol{x}} \max_{\lambda\ge 0} f_0(\boldsymbol{x}) + \lambda f_1(\boldsymbol{x}).$$

其实这种 max–min 问题和我们前面提到的剪刀石头布的问题很像, 就是两个人对抗地进行优化. 对于优化 $\boldsymbol{x}$ 的玩家来说, 她不会希望约束被违反, 否则优化 $\lambda$ 的玩家可以将 $\lambda$ 选得很大 (记住这个理解).

现在这些变换都还没有对我们的问题带来什么帮助, 尝试交换顺序得到以下辅助问题:

$$\max_{\lambda\ge 0} \underbrace{\min_{\boldsymbol{x}} f_0(\boldsymbol{x}) + \lambda f_1(\boldsymbol{x})}_{\overset{\text{def}}{=} g(\lambda)}.$$

我们把 $g(\lambda)$ 称为**对偶函数**, 最大化它的问题称为**对偶问题**. 对于凸函数和凸约束, 我们可以一般地证明: 原问题最优解的值大于等于对偶问题最优解的值 (即**弱对偶性**). 对偶问题的意义在于, 它可能比原问题容易很多 (例如在一些优化问题中, 原问题非凸而对偶是凸的), 或者有意义/容易理解/容易导出最优性条件 (线性规划是这样). 不过在一般的优化问题中, 强对偶性不一定成立.

到这里, 对偶究竟长什么样还是很抽象的, **我们接下来希望把 $g(\lambda)$ 的优化问题 "倒回去"**, 就能重新得到一个带约束的优化问题, 这样就能看到最后对偶的样子了. 我们用一个例子来演示接下来的工作.

**例 6** 考虑问题 (用这种方法的时候, $\boldsymbol{x} \ge \boldsymbol{0}$ 的约束最好放在最优化目标下面):

$$\begin{aligned}
\min_{x_1\ge 0, x_2\le 0, x_3} \quad & -v_1x_1 - v_2x_2 - v_3x_3, \\
\text{s.t.} \quad & a_1x_1 + x_2 + x_3 - b_1 \le 0, \\
& x_1 + a_2x_2 - b_2 = 0, \\
& -a_3x_3 + b_3 \le 0.
\end{aligned}$$

写出 Lagrange 乘子形式的优化问题并交换 max–min:

$$\begin{aligned}
\max_{\lambda_1\ge 0, \lambda_2, \lambda_3\ge 0} \min_{x_1\ge 0, x_2\le 0, x_3} & -v_1x_1 - v_2x_2 - v_3x_3 \\
& + \lambda_1(a_1x_1 + x_2 + x_3 - b_1) \\
& + \lambda_2(x_1 + a_2x_2 - b_2) \\
& + \lambda_3(-a_3x_3 + b_3).
\end{aligned}$$

(思考为什么 $\lambda_2$ 无约束?) 下面我们来恢复出对偶问题! 上面的式子是 "目标函数 + 对偶变量 × 关于原问题变量的式子", 为了交换原问题变量和对偶变量, 将问题改写成

“目标函数 + 原问题变量 × 关于对偶变量的式子”:

$$\max_{\lambda_1\geq 0,\lambda_2,\lambda_3\geq 0}\min_{x_1\geq 0,x_2\leq 0,x_3} -b_1\lambda_1-b_2\lambda_2+b_3\lambda_3$$

$$+x_1(a_1\lambda_1+\lambda_2-v_1) \tag{2.1}$$

$$+x_2(\lambda_1+a_2\lambda_2-v_2) \tag{2.2}$$

$$+x_3(\lambda_1-a_3\lambda_3-v_3). \tag{2.3}$$

现在想一下怎么把 $g(\lambda)$ 求出来. 首先我们观察一下式 (2.1)~(2.3) 带来的影响. 因为 $x_1\geq 0$, 所以优化 $\lambda$ 的玩家肯定不想让 (2.1) 小于零 (否则目标可以任意小), 因此它化成约束 $a_1\lambda_1+\lambda_2-v_1\geq 0$. 这时内层玩家的最佳选择就是取 $x_1=0$. 同理式 (2.2) 要小于等于零, 而式 (2.3) 应该等于零. 这些约束都成立的时候, $g(\lambda)$ 就等于 $-b_1\lambda_1-b_2\lambda_2+b_3\lambda_3$! 于是, 对偶问题就出来了:

$$\begin{aligned}
\max_{\lambda_1\geq 0,\lambda_2,\lambda_3\geq 0}\quad & -b_1\lambda_1-b_2\lambda_2+b_3\lambda_3,\\
\text{s.t.}\quad & a_1\lambda_1+\lambda_2-v_1\geq 0,\\
& \lambda_1+a_2\lambda_2-v_2\leq 0,\\
& \lambda_1-a_3\lambda_3-v_3=0.
\end{aligned}$$

◇

总结一下, 这种取对偶的方法由以下几步完成:

(i) 确保目标函数是最小化, 所有约束都是等式或者 $\leq$;

(ii) 对于每一个约束引入一个对偶变量, 其中不等式型约束引入的对偶变量为非负实数, 等式型约束引入的则为任意实数. 然后把约束嵌入到最优化问题中, 将问题改写为一个无约束的 max–min 问题;

(iii) 整理, 将此时的目标函数中的原变量和对偶变量的角色互换;

(iv) 恢复约束. 对于要求非负的原变量, 括号中的变量对应的约束为 $\geq 0$; 对于要求非正的原变量, 相应约束为 $\leq 0$; 对于无约束的原变量, 相应约束则为 $=0$.

这种方法用起来通常会比记忆矩阵形式要快很多, 更重要的是, 有实际意义的对偶通常都需要用这种方法. 可以尝试一下给下面的问题取对偶:

**问题 7** 设有限集 $U$, 函数 $u,v:\mathscr{P}(U)\longrightarrow\mathbb{R}_{\geq 0}$ 和正实常数 $\gamma$ 皆已给定, 写出问题

$$\begin{aligned}
\max_{\substack{\ell_+,\ell_-\\ p_1,\ldots,p_{|U|}}}\quad & \ell_+-\ell_-,\\
\text{s.t.}\quad & \sum_{S\subseteq U}u(S)\sum_{j\in S}p_j-\sum_{j\in T}p_j+(\ell_+-\ell_-)\leq\sum_{S\subseteq U}u(S)v(S\setminus T)-\gamma v(U)\quad(\forall T\subseteq U),\\
& p_j\geq 0\quad(\forall j\in U),\ell_+,\ell_-\geq 0.
\end{aligned}$$

的对偶. (答案参看 [DKL20, 引理 3.1 的证明].)

**2.2 附: 强对偶性 (另证) 和互补松弛性** 强对偶性就是说在线性规划的一对对偶都有解时, 二者的值相等. 课文中的证明是基于单纯形法的, 以下我们给出一个传统的证明, 它基于最优化中的 Farkas 引理. 证明之后我们还会讨论一下强对偶性的一些重大意义.

**定理 8 (弱对偶性)** *$P$ 的值不超过 $D$ 的值. 特别地, 若 $\boldsymbol{x}, \boldsymbol{y}$ 分别为 $P, D$ 的可行解, 满足 $\boldsymbol{c}^\top \boldsymbol{x} = \boldsymbol{b}^\top \boldsymbol{y}$, 则 $\boldsymbol{x}, \boldsymbol{y}$ 都是相应问题的最优解.*

**证明** 因为 $\boldsymbol{x}, \boldsymbol{y}$ 分别为 $P, D$ 的可行解, 所以

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

"特别地" 一句是上面不等式显见的推论. □

接续例 3, 4 中的符号, 由弱对偶性我们导出了: 最大匹配大小不大于最小顶点覆盖. 当然, 可以证明, 在二部图中二者实际上是相等的 (这是大家在集合论和图论中学到的), 而这可以用强对偶性导出.

为了讨论强对偶性, 先来考察原问题和对偶问题是否有可行解这一问题, 在这里的论证中, 我们会慢慢从松弛形过渡到标准形[1].

**引理 9** *设 $A \in \mathbb{R}^{m\times n}, \boldsymbol{b} \in \mathbb{R}^m$, 则下面两个叙述有且仅有一个成立:*

(i) $A\boldsymbol{x} = \boldsymbol{b}$ *有解;*

(ii) $A^\top \boldsymbol{y} = \boldsymbol{0}, \boldsymbol{b}^\top \boldsymbol{y} = -1$ *有解.*

这一引理的几何意义是明确的: 如果 $\boldsymbol{b}$ 不在 $A$ 的列空间 $V$ 中, 那么一定存在一个正交于 $V$ 的向量, 满足它和 $\boldsymbol{b}$ 成钝角.

**证明** 如果 (i)、(ii) 同时成立, 则 $-1 = \boldsymbol{b}^\top \boldsymbol{y} = \boldsymbol{x}^\top A^\top \boldsymbol{y} = 0$, 矛盾.

假设 (i) 不成立, 则 $\boldsymbol{b}$ 不在 $A$ 的列空间中, 即 $\operatorname{rank}\,[A \quad \boldsymbol{b}] = \operatorname{rank} A + 1$, 从而以下矩阵

$$\begin{bmatrix} A & \boldsymbol{b} \\ \boldsymbol{0}^\top & -1 \end{bmatrix}$$

的秩也是 $\operatorname{rank} A + 1$. 所以 $\begin{bmatrix} \boldsymbol{0} \\ -1 \end{bmatrix}$ 可被 $[A \quad \boldsymbol{b}]^\top$ 线性表出, 故存在所说的 $\boldsymbol{y}$. □

**引理 10 (Farkas)** *设 $A \in \mathbb{R}^{m\times n}, \boldsymbol{b} \in \mathbb{R}^m$, 则下面两个叙述有且仅有一个成立:*

(i) $A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \ge \boldsymbol{0}$ *有解;*

(ii) $A^\top \boldsymbol{y} \ge \boldsymbol{0}, \boldsymbol{b}^\top \boldsymbol{y} < 0$ *有解.*

Farkas 引理是最优化理论的基石之一. 其几何意义和前一个引理差不多: $A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \ge \boldsymbol{0}$ 有解实质上是指 $\boldsymbol{b}$ 落在 $A$ 的列向量张成的凸锥中; 如果不是这样, 因为锥是凸

[1]这里我们认为标准形是均为不等式约束, 松弛形是都为等式约束, 这和课本上的定义不同, 请注意.

的, 所以直观地存在一个超平面分离 $\boldsymbol{b}$ 和这个锥, 这个超平面对应的方程就是 $\boldsymbol{x}^\top \boldsymbol{y} = 0$, 对应于 (ii). 这里我们还是另外给出一个纯代数的证明.

**证明** (i)、(ii) 同时成立则导出 $0 \le \boldsymbol{x}^\top A^\top \boldsymbol{y} = (A\boldsymbol{x})^\top \boldsymbol{y} = \boldsymbol{b}^\top \boldsymbol{y} < 0$, 矛盾. 若 (i) 不成立, 则我们可不妨设 $A\boldsymbol{x} = \boldsymbol{b}$ 有解, 但是不满足 $\boldsymbol{x} \ge \boldsymbol{0}$, 否则可直接用引理 9 得到结论.

设 $A = (\boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$, 现在对 $n$ 归纳来证明此时 (ii) 成立. 当 $n = 1$ 时, 条件为 $x_1 \boldsymbol{a}_1 = \boldsymbol{b}$ 有解 $x_1 < 0$, 令 $\boldsymbol{y} = -\boldsymbol{b}$, 代入得

$$\boldsymbol{a}_1^\top \boldsymbol{y} = -\frac{\boldsymbol{b}^\top \boldsymbol{b}}{x_1} > 0, \quad \boldsymbol{b}^\top \boldsymbol{y} = -\boldsymbol{b}^\top \boldsymbol{b} < 0,$$

满足 (ii).

若命题对一切 $k \le n-1$ 都成立, 因为 (i) 没有满足 $\boldsymbol{x} \ge 0$ 的解, 所以 $\sum_{i=1}^{n-1} x_i \boldsymbol{a}_i = \boldsymbol{b}$ 一定没有满足 $x_1, \dots, x_{n-1} \ge 0$ 的解 (否则可令 $x_n = 0$). 由归纳假设知, 存在 $\boldsymbol{v}$ 满足 $\boldsymbol{a}_i^\top \boldsymbol{v} \ge 0\,(1 \le i \le n-1)$ 及 $\boldsymbol{b}^\top \boldsymbol{v} < 0$. 假如 $\boldsymbol{a}_n^\top \boldsymbol{v} \ge 0$ 恰好也成立, 那么 (ii) 已经成立. 故下面可设 $\boldsymbol{a}_n^\top \boldsymbol{v} < 0$. 取

$$\begin{aligned} \tilde{\boldsymbol{a}}_i &= (\boldsymbol{a}_i^\top \boldsymbol{v})\boldsymbol{a}_n - (\boldsymbol{a}_n^\top \boldsymbol{v})\boldsymbol{a}_i \quad (1 \le i \le n-1), \\ \tilde{\boldsymbol{b}} &= (\boldsymbol{b}^\top \boldsymbol{v})\boldsymbol{a}_n - (\boldsymbol{a}_n^\top \boldsymbol{v})\boldsymbol{b}. \end{aligned} \tag{2.4}$$

我们指出 $\sum_{i=1}^{n-1} x_i \tilde{\boldsymbol{a}}_i = \tilde{\boldsymbol{b}}$ 亦没有满足 $x_1, \dots, x_{n-1} \ge 0$ 的解, 否则整理可得

$$-\frac{1}{\boldsymbol{a}_n^\top \boldsymbol{v}} \left( \sum_{i=1}^{n-1} x_i (\boldsymbol{a}_i^\top \boldsymbol{v}) - \boldsymbol{b}^\top \boldsymbol{v} \right) \boldsymbol{a}_n + \sum_{i=1}^{n-1} x_i \boldsymbol{a}_i = \boldsymbol{b},$$

与 (i) 不成立矛盾. 故再次使用归纳假设得存在 $\boldsymbol{w}$, 使得 $\tilde{\boldsymbol{a}}_i{}^\top \boldsymbol{w} \ge 0\,(1 \le i \le n-1), \tilde{\boldsymbol{b}}^\top \boldsymbol{w} < 0$. 这样, 可以验证

$$\boldsymbol{y} = (\boldsymbol{a}_n^\top \boldsymbol{w})\boldsymbol{v} - (\boldsymbol{a}_n^\top \boldsymbol{v})\boldsymbol{w}$$

就是 (ii) 的解. 事实上, 由式 (2.4) 得

$$\begin{aligned} &\boldsymbol{a}_i^\top \boldsymbol{y} = (\boldsymbol{a}_n^\top \boldsymbol{w})(\boldsymbol{a}_i^\top \boldsymbol{v}) - (\boldsymbol{a}_n^\top \boldsymbol{v})(\boldsymbol{a}_i^\top \boldsymbol{w}) = \tilde{\boldsymbol{a}}_i{}^\top \boldsymbol{w} \ge 0\,(1 \le i \le n-1), \quad \boldsymbol{a}_n^\top \boldsymbol{y} = 0, \\ &\boldsymbol{b}^\top \boldsymbol{y} = (\boldsymbol{a}_n^\top \boldsymbol{w})(\boldsymbol{b}^\top \boldsymbol{v}) - (\boldsymbol{a}_n^\top \boldsymbol{v})(\boldsymbol{b}^\top \boldsymbol{w}) = \tilde{\boldsymbol{b}}^\top \boldsymbol{w} < 0. \end{aligned}$$

□

**推论 11** 设 $A \in \mathbb{R}^{m \times n}, \boldsymbol{b} \in \mathbb{R}^m$, 则下面两个叙述有且仅有一个成立:

(i) $A\boldsymbol{x} \le \boldsymbol{b}, \boldsymbol{x} \ge \boldsymbol{0}$ 有解;

(ii) $A^\top \boldsymbol{y} \ge \boldsymbol{0}, \boldsymbol{y} \ge \boldsymbol{0}, \boldsymbol{b}^\top \boldsymbol{y} < 0$ 有解.

**证明** 若 (i)、(ii) 同时成立, 则 $0 \le \boldsymbol{x}^\top A^\top \boldsymbol{y} = (A\boldsymbol{x})^\top \boldsymbol{y} \le \boldsymbol{b}^\top \boldsymbol{y} < 0$, 矛盾. 如果 (i) 不成立, 按

转换为松弛形 (课本上的标准形) 的方法, 这等价于

$$[A \quad I_m]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{x}'\end{bmatrix}=\boldsymbol{b},\quad \boldsymbol{x},\boldsymbol{x}'\geq \boldsymbol{0}$$

无解, 由 Farkas 引理, 存在 $\boldsymbol{y}$ 使得

$$\begin{bmatrix}A^\top\\ I_m\end{bmatrix}\boldsymbol{y}\leq \boldsymbol{0},\quad \boldsymbol{b}^\top\boldsymbol{y}<0,$$

就得到 (ii). □

我们可以证明强对偶性了.

**定理 12 (强对偶性)** 设线性规划的原问题及其对偶分别为 $P, D$, 则下面叙述有且仅有一个成立:

(i) $P, D$ 都有可行解, 且最优解的值相等;
(ii) $D$ 有可行解, $P$ 无可行解, 且 $D$ 的目标函数在约束下无界;
(iii) $P$ 有可行解, $D$ 无可行解, 且 $P$ 的目标函数在约束下无界;
(iv) 两个问题均无可行解.

**证明** 设 $P, D$ 都有可行解, 根据弱对偶性, 我们只要证明不等式组

$$\begin{aligned}
&A\boldsymbol{x}\leq \boldsymbol{b},\quad \boldsymbol{x}\geq\boldsymbol{0},\\
&A^\top\boldsymbol{y}\geq \boldsymbol{c},\quad \boldsymbol{y}\geq\boldsymbol{0},\\
&\boldsymbol{c}^\top\boldsymbol{x}-\boldsymbol{b}^\top\boldsymbol{y}\geq 0,
\end{aligned}$$

是有解的即可. 把它写成分块矩阵的形式, 就是

$$\begin{bmatrix}A & 0\\ 0 & -A^\top\\ -\boldsymbol{c}^\top & \boldsymbol{b}^\top\end{bmatrix}\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix}\leq\begin{bmatrix}\boldsymbol{b}\\ -\boldsymbol{c}\\ 0\end{bmatrix},\quad \begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix}\geq\boldsymbol{0}. \tag{2.5}$$

如果此系统无解, 则根据推论 11, 存在 $\boldsymbol{z},\boldsymbol{w}\geq\boldsymbol{0}$ 和实数 $\alpha\geq 0$ 使得

$$A^\top\boldsymbol{z}\geq\alpha\boldsymbol{c},\quad A\boldsymbol{w}\leq\alpha\boldsymbol{b},\quad \boldsymbol{b}^\top\boldsymbol{z}<\boldsymbol{c}^\top\boldsymbol{w}.$$

首先说明 $\alpha=0$ 是不可能的. 不然, 设 $\tilde{\boldsymbol{x}},\tilde{\boldsymbol{y}}$ 是 $P, D$ 的可行解, 有

$$0\leq\tilde{\boldsymbol{x}}^\top A^\top\boldsymbol{z}=(A\tilde{\boldsymbol{x}})^\top\boldsymbol{z}\leq\boldsymbol{b}^\top\boldsymbol{z}<\boldsymbol{c}^\top\boldsymbol{w}\leq(A^\top\tilde{\boldsymbol{y}})^\top\boldsymbol{w}=\tilde{\boldsymbol{y}}^\top A\boldsymbol{w}\leq 0,$$

矛盾. 此时再令 $\boldsymbol{x}=\boldsymbol{w}/\alpha, \boldsymbol{y}=\boldsymbol{z}/\alpha$, 则可验证 $\boldsymbol{x},\boldsymbol{y}$ 分别是 $P, D$ 的可行解, 故再次由弱对

偶性得

$$\boldsymbol{c}^\top \boldsymbol{w} = \alpha(\boldsymbol{c}^\top \boldsymbol{x}) = \alpha(\boldsymbol{b}^\top \boldsymbol{y}) = \boldsymbol{b}^\top \boldsymbol{z},$$

与 $\boldsymbol{b}^\top \boldsymbol{z} < \boldsymbol{c}^\top \boldsymbol{w}$ 矛盾. 所以式 (2.5) 一定有解, 即最优解的值相等.

再设 $P$ 无可行解, 而 $D$ 有可行解. 结合 $P$ 无可行解和推论 11 得 $A^\top \boldsymbol{w} \geq \boldsymbol{0}, \boldsymbol{w} \geq \boldsymbol{0}, \boldsymbol{b}^\top \boldsymbol{w} < 0$ 有解. 注意到, 如果 $\boldsymbol{y}$ 是 $D$ 的可行解, 则对每个非负实数 $\lambda$, $\boldsymbol{y} + \lambda \boldsymbol{w}$ 都是可行解, 这样 $D$ 的目标函数可写为 $\boldsymbol{b}^\top(\boldsymbol{y} + \lambda \boldsymbol{w}) = \boldsymbol{b}^\top \boldsymbol{y} + \lambda(\boldsymbol{b}^\top \boldsymbol{w})$. 因为 $\boldsymbol{b}^\top \boldsymbol{w}$, 所以取充分大的 $\lambda$ 可使目标函数充分小, 即其无界. 同理可证 (iii).

(iv) 则是剩余的可能情形, 对此没有可说的结论. □

至此, 我们用推论 11 刻画了可行解的存在性, 用强对偶性描写了最优解的性质. 强对偶性还将导出一个很重要的性质, 就是**互补松弛性**.

**推论 13** 设线性规划的原问题及其对偶如上所说, $\boldsymbol{x}, \boldsymbol{y}$ 分别是 $P, D$ 的最优解, 当且仅当下面的命题成立:

(i) $y_i > 0$ 推出 $\sum_{j=1}^n a_{ij} x_j = b_i \, (1 \leq i \leq m)$;

(ii) $x_j > 0$ 推出 $\sum_{i=1}^m b_{ij} y_i = c_j \, (1 \leq j \leq n)$.

用自然语言表述, 就是第 $i$ 个分量非负的约束和对偶问题的第 $i$ 个约束一定有一个取等.

**证明** 必要性. 由强对偶性, 我们有

$$\sum_{j=1}^n c_j x_j = \sum_{i,j} y_i a_{ij} x_j = \sum_{i=1}^m b_i y_i \implies \sum_{i=1}^m \underbrace{\left( b_i - \sum_{j=1}^n a_{ij} x_j \right)}_{\geq 0} \underbrace{y_i}_{\geq 0} = 0.$$

注意求和中每一项都非负, 故每一项都必须是 0, 因而第 $i$ 个分量 $\geq 0$ 的约束和对偶问题的第 $i$ 个约束一定有一个取等. (ii) 同理可证.

充分性. (i)、(ii) 成立时即有

$$\boldsymbol{c}^\top \boldsymbol{x} = \sum_{j=1}^n c_j x_j = \sum_{j=1}^n \left( \sum_{i=1}^m a_{ij} y_i \right) x_j = \sum_{i=1}^m \left( \sum_{j=1}^n a_{ij} x_j \right) y_i = \sum_{i=1}^m b_i y_i = \boldsymbol{b}^\top \boldsymbol{y},$$

由对偶性知道 $\boldsymbol{x}, \boldsymbol{y}$ 均为最优解. □

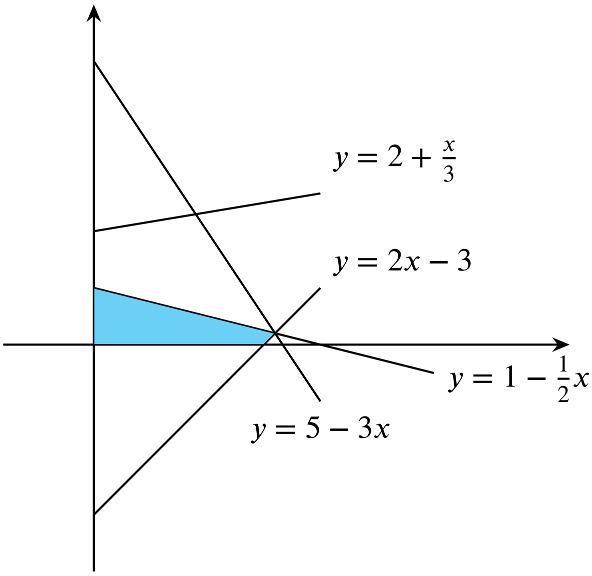
**例 14** 我们考虑用二维图解法确定下面线性规划问题的最优解:

$$\begin{aligned}
\min \quad & 2x_1 + 3x_2 + 5x_3 + 6x_4, \\
\text{s.t.} \quad & x_1 + 2x_2 + 3x_3 + x_4 \geq 2, \\
& -2x_1 + x_2 - x_3 + 3x_4 \leq 3, \\
& x_1, x_2, x_3, x_4 \geq 0.
\end{aligned}$$

首先取对偶得到

$$\begin{aligned} \max \quad & 2y_1 - 3y_2, \\ \text{s.t.} \quad & y_1 + 2y_2 \le 2, 2y_1 - y_3 \le 3, \\ & 3y_1 + y_2 \le 5, y_1 - 3y_2 \le 6, \\ & y_1, y_2 \ge 0. \end{aligned}$$

图解得到最优解 $(1.5, 0)$, 值为 3. 图解法过程如下图所示.




由互补松弛性, 蓝色约束均不是紧的, 所以原问题最优解在 $(0, x_2, 0, 0)$ 处取到, 故 $x_2 = 1$, 值也为 3. ◇

最后我们给出一个和对偶性的几何意义有关的比较深入的例题 (某年作业题).

**问题 15** 用线性规划的模型为下列**凸多面体分离**问题建模: 设 $A, C \in \mathbb{R}^{m \times n}$, $\boldsymbol{b}, \boldsymbol{d} \in \mathbb{R}^n$, 令

$$P_1 = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} \le \boldsymbol{b}\}, \quad P_2 = \{\boldsymbol{x} \in \mathbb{R}^n : C\boldsymbol{x} \le \boldsymbol{d}\}.$$

求 "最佳" 的 $\boldsymbol{w} \in \mathbb{R}^n$ 和实数 $\gamma$, 使得对于任意的 $\boldsymbol{x} \in P_1$ 有 $\boldsymbol{w}^\top \boldsymbol{x} < \gamma$, 对任意的 $\boldsymbol{x} \in P_2$ 有 $\boldsymbol{w}^\top \boldsymbol{x} > \gamma$. 对于 "最佳", 请根据几何意义自己选择一个较好的优化目标.

**解** 我们先不对 $\|\boldsymbol{w}\|$ 作优化, 那么我们只需要找到一个可行解. 现在计算约束

$$\begin{aligned} & A^\top \boldsymbol{x} + C^\top \boldsymbol{y} = \boldsymbol{0}, \\ & \boldsymbol{b}^\top \boldsymbol{x} + \boldsymbol{d}^\top \boldsymbol{y} < 0, \\ & \boldsymbol{x}, \boldsymbol{y} \ge \boldsymbol{0}. \end{aligned} \tag{2.6}$$

的一个可行解 $(\boldsymbol{x}_0, \boldsymbol{y}_0)$, 方法是将课本上寻找基本初始可行解的方法应用到该问题上. 令 $\boldsymbol{w} = C^\top \boldsymbol{y}_0, \gamma = \frac{1}{2}\left(\boldsymbol{d}^\top \boldsymbol{y}_0 - \boldsymbol{b}^\top \boldsymbol{x}_0\right)$ 即得.

下面说明该构造的正确性. 第一步: 证明式 (2.6) 一定有可行解. 考虑一组互为对偶的问题

$$\begin{array}{llll} (\mathsf{P}) & \max & \mathbf{0}^\top \boldsymbol{v}, \\ & \text{s.t.} & A\boldsymbol{v} \le \boldsymbol{b}, \\ & & C\boldsymbol{v} \le \boldsymbol{d}. \end{array} \qquad \begin{array}{llll} (\mathsf{D}) & \min & \boldsymbol{b}^\top \boldsymbol{x} + \boldsymbol{d}^\top \boldsymbol{y}, \\ & \text{s.t.} & A^\top \boldsymbol{x} + C^\top \boldsymbol{y} = \mathbf{0}, \\ & & \boldsymbol{x}, \boldsymbol{y} \ge \mathbf{0}. \end{array}$$

由于多面体不交, 所以 P 无解, 而 $\mathbf{0}, \mathbf{0}$ 是对偶问题的一个可行解, 由强对偶性质 D 必然无界. 因此必存在一个充分小的 $\alpha < 0$ 使得 $\boldsymbol{b}^\top \boldsymbol{x} + \boldsymbol{d}^\top \boldsymbol{y} = \alpha$, 即式 (2.6) 必然有解.

第二步: 证明构造的超平面一定分离了两个多面体. 考虑计算 $A\boldsymbol{x} \le \boldsymbol{b}$ 到超平面距离的优化问题及其对偶

$$\begin{array}{llll} (\mathsf{P}) & \min & \boldsymbol{w}^\top \boldsymbol{v}, \\ & \text{s.t.} & A\boldsymbol{v} \le \boldsymbol{b}. \end{array} \qquad \begin{array}{llll} (\mathsf{D}) & \max & -\boldsymbol{b}^\top \boldsymbol{x}, \\ & \text{s.t.} & A^\top \boldsymbol{x} = -\boldsymbol{w}, \boldsymbol{x} \ge \mathbf{0}. \end{array} \tag{2.7}$$

注意到 $-\boldsymbol{x}_0$ 是 D 的可行解, 所以由弱对偶性质 P 的最优解 $m_1 \ge -\boldsymbol{b}^\top \boldsymbol{x}_0$.

同理, $C\boldsymbol{x} \le \boldsymbol{d}$ 到超平面距离的优化问题及其对偶是

$$\begin{array}{llll} (\mathsf{P}) & \max & \boldsymbol{w}^\top \boldsymbol{v}, \\ & \text{s.t.} & C\boldsymbol{v} \le \boldsymbol{d}. \end{array} \qquad \begin{array}{llll} (\mathsf{D}) & \min & \boldsymbol{d}^\top \boldsymbol{x}, \\ & \text{s.t.} & C^\top \boldsymbol{x} = \boldsymbol{w}, \boldsymbol{x} \ge \mathbf{0}. \end{array} \tag{2.8}$$

$\boldsymbol{y}_0$ 是 D 的可行解, 所以由弱对偶性质 P 的最优解 $m_2 \le \boldsymbol{d}^\top \boldsymbol{y}_0 = -\boldsymbol{b}^\top x_0 + \alpha$.

因此 $m_2 \le \boldsymbol{d}^\top \boldsymbol{y}_0 = -\boldsymbol{b}^\top \boldsymbol{x}_0 + \alpha < -\boldsymbol{b}^\top \boldsymbol{x}_0 \le m_1$. 于是可以取 $\gamma = -\boldsymbol{b}^\top \boldsymbol{x}_0 + \frac{\alpha}{2} = \frac{1}{2}\left(\boldsymbol{d}^\top \boldsymbol{y}_0 - \boldsymbol{b}^\top \boldsymbol{x}_0\right)$.

再考虑选择更好的解. 因为上面没有对超平面向量的范数作出什么约束, 所以距离的差可以任意大 ($\alpha$ 可以充分小), 无法定义最好的超平面. 如果给定一个限界, 比如 $-\mathbf{1} \le \boldsymbol{w} \le \mathbf{1}$, 就可最优, 直接求解 (2.7)、(2.8) 即可, 即

$$\begin{array}{ll} \max & -\boldsymbol{b}^\top \boldsymbol{x} - \boldsymbol{d}^\top \boldsymbol{y}, \\ \text{s.t.} & A^\top \boldsymbol{x} = -\boldsymbol{w}, C^\top \boldsymbol{y} = \boldsymbol{w}, \\ & \boldsymbol{x}, \boldsymbol{y} \ge \mathbf{0}, -\mathbf{1} \le \boldsymbol{w} \le \mathbf{1}. \end{array}$$

解取 $\left(\boldsymbol{w}^*, \frac{1}{2}\left(\boldsymbol{d}^\top \boldsymbol{y}^* - \boldsymbol{b}^\top \boldsymbol{x}^*\right)\right)$. □

### 2.3 附: 全幺模矩阵

回到例 3 最后提到的问题. 对于一些特别的 $A$ (比如二部图最大匹配), 我们可以保证一般线性规划的最优解恒为整数向量.

**定义 16** 设 $A \in \mathbb{Z}^{n\times n}$, 如果 $A$ 的行列式为 $+1$ 或 $-1$, 就称它是**幺模矩阵**. 如果一个整数矩阵的每个非奇异子方阵都是幺模矩阵 (换言之, 每个子方阵的行列式都是 $0, +1, -1$ 之

一), 则说它是**全幺模矩阵**.

**引理 17** 设 $A \in \mathbb{Z}^{m\times n}$, 则 $A$ 是全幺模矩阵当且仅当 $[A \quad I_m]$ 是全幺模矩阵.

**证明** 充分性显然. 对于必要性, 我们对选取的子方阵的大小 $k$ 作归纳法. 当 $k = 1$ 时, $I_m$ 的每个元素 $\in \{0, \pm 1\}$, 命题成立. 假设阶严格小于 $k$ 的全部非奇异子方阵都是幺模矩阵, 考虑一 $k \times k$ 的子方阵. 如果它的列不包括 $I_m$ 中的列, 那命题自然成立; 否则, 它选入了 $I_m$ 中的某一列的一部分, 这一列可能全为 0, 此时行列式为 0; 也可能恰有一个 1, 那按这一行展开其行列式, 由归纳假设知其结果仍属 $\{0, \pm 1\}$ 之一. □

**推论 18** 若线性规划标准形中的矩阵 $A$ 为全幺模矩阵, 则其最优解必为整数向量.

**证明** 首先将标准形化为松弛形, 这时 $[A \quad I_m]$ 也是全幺模矩阵.

依据课本中的定理, 最优解一定是基本解. 而当求基本解时, 线性方程组 $\sum_{i\in I} x_i \boldsymbol{a}_i = \boldsymbol{b}$ 的系数矩阵是幺模矩阵, 由 Cramer 法则知 $x_i \in \mathbb{Z}\,(i \in I)$, 于是基本解一定是整数向量, 从而最优解亦然. □

于是, 如果 $A$ 是全幺模矩阵, 则整数线性规划可以用普通的线性规划算法照常计算. 我们不加证明地叙述一个有用的判定全幺模矩阵的充分条件:

**定理 19 (Hoffman–Gale)** 设 $A \in \mathbb{Z}^{m\times n}$, 如果 $A$ 的每个元素只是 $\{0, \pm 1\}$ 之一, 每列至多有两个非零元素, 且其行可划分为两个集合 $B, C$, 使得

(i) 若某列中有两个非零元素 $+1, -1$, 则它们所在的行必同属 $B$ 或者 $C$;

(ii) 若某列中两个非零元素相同, 则它们所在的行必分属 $B, C$.

那么 $A$ 是全幺模矩阵.

所以, 二部图的关联矩阵是全幺模矩阵!

在上面结论的基础上, 我们可以很轻松地证明一个图论中的结论. 我们把图中最大独立集、最大匹配、最小点覆盖和最小边覆盖的大小分别记为 $\alpha(G), \nu(G), \tau(G), \rho(G)$.

**推论 20 (König–Gallai)** 对任意图 $G$, 有 $\nu(G) \le \tau(G)$ 以及 $\alpha(G) \le \rho(G)$. 特别地, 在二部图 $G$ 中, 还有等号成立:

(i) $\nu(G) = \tau(G)$;

(ii) $\alpha(G) = \rho(G)$;

(iii) $\nu(G) + \rho(G) = \alpha(G) + \tau(G) = |G|$.

**证明** 首先记 LP 为最大匹配问题的线性规划形式, 则根据例 3, 4 的论证, DLP 为最小点覆盖问题的线性规划形式. 由弱对偶性和整数线性规划的性质得

$$\mathrm{ILP} \le \mathrm{LP} \le \mathrm{DLP} \le \mathrm{IDLP},$$

所以命题成立. 特别地, 在二部图中最优解为整数, 因此由强对偶性等好全部成立. 这就证明了有关 $\nu(G), \tau(G)$ 的结论.

对于最大独立集问题, 设图 $G$ 关联矩阵的转置 $M^\top \in \{0,1\}^{m\times n}$, 令 $\boldsymbol{x} \in \{0,1\}^n$ 的含义为点集 $X \subseteq V(G)$ 的特征函数 (即是 $\mathbf{1}_{v\in X}$). 考虑约束 $A^\top \boldsymbol{x} \le \mathbf{1}$, 它表示 $X$ 中的点两两不邻接, 即代表一个独立集; 如果再设定最大化目标函数 $\max \mathbf{1}^\top \boldsymbol{x} = |X|$, 则这就对应于最大独立集. 即是:

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & \mathbf{1}^\top \boldsymbol{x}, \\ \text{s.t.} \quad & M^\top \boldsymbol{x} \le \mathbf{1}, \boldsymbol{x} \ge \mathbf{0}. \end{aligned}$$

其对偶问题 DLP 为

$$\begin{aligned} \min_{\boldsymbol{y}} \quad & \mathbf{1}^\top \boldsymbol{y}, \\ \text{s.t.} \quad & M \boldsymbol{y} \ge \mathbf{1}, \boldsymbol{y} \ge \mathbf{0}. \end{aligned}$$

若将 $\boldsymbol{y} \in \{0,1\}^m$ 理解为边集 $Y \subseteq V(G)$ 的特征函数 (即是 $\mathbf{1}_{e\in Y}$), 约束此时表示 $G$ 中的每一个点都和 $Y$ 中的某条边相邻, 最小化目标函数就是要求最小的满足这个条件的边集的大小. 再次利用弱对偶性和整数线性规划的性质便得到 $\alpha(G) \le \rho(G)$. 特别地, 在二部图中最优解为整数, 因此由强对偶性等好全部成立.

最后证明 (iii). 注意到一个简单的观察: 一个点集 $S$ 是独立集, 当且仅当 $V \setminus S$ 是点覆盖. 事实上, 若 $S$ 为独立集, 则每条边都和 $V \setminus S$ 上的点相邻; 若 $V \setminus S$ 为点覆盖, 则 $S$ 中点之间必然无边. 这导出 $\alpha(G) + \tau(G) = |G|$, 结合 (i)、(ii) 即知 (iii) 成立. □

**注 21** 对偶性在图论中也有很高的地位, 由此也可以看到线性规划的概括性. 同学们在集合论与图论中学到的 Menger 定理 (**最小**的 $XY$ 割大小等于**最大**的相互独立的 $XY$ 路径的集合的大小)、Dilworth 定理 (有限偏序集**最大**反链的大小等于其**最小**链覆盖的大小)、Hall 定理, 之后的**最大**流–**最小**割定理其实都是对偶性的例子. ♤

# 参考文献

线性规划是由苏联数学家 Kantorovich、荷兰数学家 Koopmans、美国数学家 Dantzig 和 von Neumann 等人共同发展的. Dantzig 曾写过一本介绍线性规划的著作 [DT97].

对偶理论是由 Dantzig 发现的, 而 von Neumann 则指出对偶理论和他对博弈论的研究有着等价性——就是我们在前面讲的取对偶的方法, 那和 von Neumann 的极大–极小定理有着紧密的联系. 这种方法是作者从 [Lah15] 中学到的.

我们没有提到单纯形法, 这是因为 [Cor+09, 第 29 章] 已经有了阐述, 而作者没有什么更有趣的理解, 所以就不重复造轮子了. 不过推荐大家理解证明之前先用 (不是课本中打表的通用方式) 单纯形法自己解几个简单的例子.

[Cor+09] T. H. Cormen et al. *Introduction to Algorithms*. MIT press, 2009. ISBN: 978-0-26-203384-8.

[DKL20] P. Dütting, T. Kesselheim, and B. Lucier. "An $O(\log\log m)$ prophet inequality for subadditive combinatorial auctions". 刊于: *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE. 2020, pp. 306–317.

[DT97] G. B. Dantzig and M. N. Thapa. *Linear Programming 1: Introduction*. Springer Series in Operations Research and Financial Engineering. Springer New York, 1997. ISBN: 978-0-38-794833-1.

[Lah15] S. Lahaie. *How to take the dual of a linear program*. 2015. URL: http://slahaie.net/docs/lpdual.pdf.

编写: WC
*E-mail*: wchang@pku.edu.cn