高级算法/ 2026 年秋季

线性规划与优化专题

原始对偶算法

顶点覆盖与度量设施选址

第5 讲：100 分钟

本讲义仅覆盖栗师老师负责的线性规划专题。本专题共八讲，讲次编号与课件一致。第一、第二部分是本讲内部的阅读划分；选读附录供课后拓展。

授课教师：栗师
南京大学

以授课教师的课件为依据，借助AI 将讲解、证明和课后练习展开为文字。中英文版本的数学内容相互对应。

学习目标与阅读路线

LP 舍入先计算最优分数解；原始对偶算法则逐步增长一个可行的下界证书，并利用紧约束决定选择哪些整数对象。即使算法不调用LP 求解器，LP 仍可解释算法并提供保证。

1. 区分极大对偶可行向量与最优对偶向量。
2. 推导并分析带权顶点覆盖的原始对偶算法。
3. 解释设施选址基本表述与星形表述的等价性。
4. 理解对偶增长、设施筛选以及三段连接距离界。

第一部分建立证书视角并推导设施选址LP，第二部分给出完整算法与分析。分析中必须使用度量的对称性和三角不等式。

目录

1 第一部分：顶点覆盖与设施选址LP **1**
1.1 用边预算支付顶点成本 1
1.2 度量无容量设施选址 1
1.3 星形表述把LP 变成覆盖问题 2

2 第二部分：设施选址算法与分析 **2**
2.1 统一增长对偶预算 2
2.2 筛选临时开设设施 3
2.3 支付归属客户及开设成本 3
2.4 剩余客户的三段距离界 3
2.5 近似保证与实现 4

3 课后独立练习 **4**

1 第一部分：顶点覆盖与设施选址LP

1.1 用边预算支付顶点成本

对带权顶点覆盖，省略上一单元中冗余的变量上界，得到原始与对偶：

$$\begin{array}{ll}\min & \sum_v w_v x_v \\ \text{约束为} & x_u + x_v \ge 1 \quad (uv \in E) \\ & x \ge 0\end{array} \qquad \begin{array}{ll}\max & \sum_e y_e \\ \text{约束为} & \sum_{e\in\delta(v)} y_e \le w_v \quad (v \in V) \\ & y \ge 0.\end{array}$$

可把每条边看作不断积累预算。一个顶点的关联边预算之和达到其权重时，就有足够资金选择该顶点；对偶可行性则禁止在某顶点处积累超过其成本的预算。

带权顶点覆盖的原始对偶算法

1. 令$y = 0$、$S = \varnothing$。
2. 只要有未覆盖边uv，就把y_{uv} 增加两个端点剩余预算中的较小值。
3. 将一个达到紧约束的端点加入S，覆盖其全部关联边；若某端点已经紧，则允许增加量为零并立即选择它。
4. 返回S 及证书y。

每轮选择一个此前未选的顶点，因此至多有$|V|$ 轮。对偶始终可行，因为只有两个端点的约束发生变化，且增长在第一个约束变紧时停止。已选顶点保持紧：后续选择的未覆盖边不可能与其关联。终止时所有边都被覆盖。

定理1.1. 该算法是二近似算法。

证明. 利用已选顶点处的紧性并交换求和次序，有

$$\sum_{v\in S} w_v = \sum_{v\in S}\sum_{e\in\delta(v)} y_e = \sum_{uv\in E} y_{uv}|\{u,v\}\cap S| \le 2\sum_e y_e \le 2\,\mathrm{OPT}\,.$$

最后一个不等式来自弱对偶，因此构造出的对偶向量无需最优。 □

更一般地，取任意逐坐标极大的对偶可行y，选择所有紧顶点。若某条边的两个端点都不紧，则仍能增大该边预算，与极大性矛盾。因此紧顶点覆盖所有边，同一计费论证仍适用。“极大”是指固定其他坐标时不能再增加任何坐标，并不意味着目标值全局最大。

1.2 度量无容量设施选址

设F 为候选设施，C 为客户，$F \cup C$ 上给定对称度量d，开设成本为$f_i \ge 0$。选取$S \subseteq F$ 并把每个客户连接到某个已开设施，最小化

$$\sum_{i\in S} f_i + \sum_{j\in C} d(j,S), \qquad d(j,S) = \min_{i\in S} d(i,j).$$

设施没有容量限制。有客户时假设至少存在一个候选设施；没有客户时空解最优。核心权衡是开设较多的近距离设施，还是少开设施而支付更长的连接距离。

令y_i 表示开设程度，x_{ij} 表示分配程度，基本松弛为

$$\begin{aligned}\min \quad & \sum_i f_i y_i + \sum_{i,j} d(i,j)x_{ij} \\ \text{约束为} \quad & \sum_i x_{ij} \ge 1 \quad (j \in C), \\ & 0 \le x_{ij} \le y_i \quad (i \in F, j \in C), \qquad y_i \ge 0.\end{aligned} \tag{1.1}$$

固定可行的开设向量y 后，每个客户可按距离从近到远购买总量为一的分配，每个设施至多提供y_i。不同客户之间不竞争容量；距离相同时，最优分配未必唯一。

1.3 星形表述把LP 变成覆盖问题

星形对象(i, J) 由设施i 与客户子集$J \subseteq C$ 组成，成本为$f_i + \sum_{j\in J} d(i, j)$，对应非负变量$z_{iJ}$。星形松弛及其对偶为

$$\min \quad \sum_{i,J} \left(f_i + \sum_{j\in J} d(i,j) \right) z_{iJ} \qquad \text{约束为} \quad \sum_{i,J:j\in J} z_{iJ} \geq 1, \quad z \geq 0; \tag{1.2}$$

$$\max \quad \sum_{j} \alpha_j \qquad \text{约束为} \quad \sum_{j\in J} \alpha_j \leq f_i + \sum_{j\in J} d(i,j), \quad \alpha \geq 0. \tag{1.3}$$

每个整数设施选址解都是一组覆盖客户的星形对象。如果同一设施出现多个星形对象，只开一次设施并服务其客户并集，成本不会增加。因此整数解释正确，分数层面的等价性也成立。

命题1.2. 基本LP 与星形LP 具有相同最优值。

证明. 从星形表述到基本表述，令$y_i = \sum_J z_{iJ}$、$x_{ij} = \sum_{J:j\in J} z_{iJ}$，通过求和即可验证可行性与成本相等。反之，对基本解中的每个设施i，令阈值t 取遍$[0, y_i]$，定义$J_i(t) = \{j : x_{ij} \geq t\}$，把产生同一星形对象的阈值区间长度作为其质量。设施i 的总星形质量为y_i，包含客户j 的质量为x_{ij}，所以客户覆盖与两项目标成本都保持。只有断点$0, y_i, x_{ij}$ 重要，每个设施至多需要$|C| + 1$ 个星形对象即可完成转换。 □

一个设施对应的指数多个对偶约束，等价于一个可计算的不等式：

$$\sum_{j\in C} (\alpha_j - d(i,j))_+ \leq f_i, \qquad (a)_+ = \max\{a, 0\}. \tag{1.4}$$

因为违反程度最大的星形对象恰好选择剩余预算$\alpha_j - d(i, j)$ 为正的客户。将α_j 理解为客户预算：先支付连接距离，余款贡献给设施开设。弱对偶保证每个可行预算向量都满足$D := \sum_j \alpha_j \leq$ OPT。

2 第二部分：设施选址算法与分析

2.1 统一增长对偶预算

初始预算全为零，所有客户处于活跃状态。活跃预算随全局时钟t 以单位速率增长，冻结客户的预算保持不变。设施在 (1.4) 首次取等号时临时开设，记该时刻为t_i。任何能在当前预算内到达临时开设设施的活跃客户都冻结；后续客户的预算一旦达到到已临时开设设施的距离，也立即冻结。推进时钟前一致地处理同时发生的事件。客户j 的最终预算α_j 就是其冻结时刻。为每个客户记录一个见证设施i，满足

$$t_i \leq \alpha_j, \qquad d(i,j) \leq \alpha_j. \tag{2.1}$$

对偶可行性始终保持。在设施变紧前，增长会在首次紧事件处停止；变紧后，已向它作正贡献的客户全部冻结。新的活跃客户在达到距离等号时就冻结，不会开始对该设施作正贡献。因此该设施的总贡献保持恰为f_i。

当开设成本为零时，即使没有贡献者，设施也可能在时刻零变紧，应将其纳入临时开设设施。这样距离等号冻结规则仍有定义，也不会等待本来不需要的正开设付款。所有客户最终都会冻结，否则某个有限距离设施会不断收到该客户增长的贡献，最终必然变紧。

2.2 筛选临时开设设施

令$N_i = \{j : \alpha_j > d(i,j)\}$ 为设施i 的正贡献客户集合，预算恰等于距离的客户不属于N_i。不能简单地开设每个紧设施，因为同一客户可能同时为多个设施付款。

筛选与最终连接

1. 按t_i 非减顺序处理临时设施，平局采用固定规则；初始$S = \varnothing$，所有客户未被归属。
2. 若N_i 中所有客户都未被归属，则开设i，加入S，并令其拥有N_i 中全部客户；否则舍弃i。
3. 把有归属的客户连到其所属设施，其余客户连到S 中最近的设施。

已选设施的正贡献客户集互不相交。每个被舍弃的设施都与一个更早选中的设施共享贡献者。没有贡献者的设施自动通过测试，而它的开设成本为零，因此不会造成多付款。

2.3 支付归属客户及开设成本

令$C_1 = \bigcup_{i\in S} N_i$、$C_2 = C \setminus C_1$，记总开设成本$f = \sum_{i\in S} f_i$，客户$j$ 的最终连接成本为c_j。对每个已选设施，由紧性有

$$f_i = \sum_{j\in N_i} (\alpha_j - d(i,j)).$$

其归属客户按距离$d(i,j)$ 连接，因此

$$f + \sum_{j\in C_1} c_j \le \sum_{j\in C_1} \alpha_j. \tag{2.2}$$

对规定的所属设施连接实际上取等号；改连到更近的已选设施只会改善成本。贡献者集合的不相交性，保证一份预算之和能够支付全部开设费用。

2.4 剩余客户的三段距离界

引理2.1. 每个$j \in C_2$ 都满足$c_j \le 3\alpha_j$。

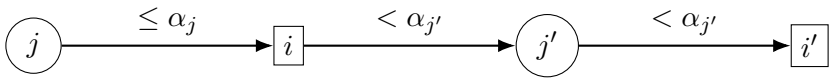
证明. 取 (2.1) 中客户j 的见证设施i。若$i \in S$，则连到最近已选设施的成本至多为$d(j,i) \le \alpha_j$。否则存在$j' \in N_i$ 已归属于某个已选设施i'，且$t_{i'} \le t_i$。向i' 作正贡献的客户不晚于i' 开设时冻结，因此

$$\alpha_{j'} \le t_{i'} \le t_i \le \alpha_j, \qquad d(i,j') < \alpha_{j'}, \quad d(i',j') < \alpha_{j'}.$$

沿$j \to i \to j' \to i'$ 使用度量对称性与三角不等式：

$$c_j \le d(j,i') \le d(j,i) + d(i,j') + d(j',i') \le \alpha_j + 2\alpha_{j'} \le 3\alpha_j.$$

□



按开设时间处理设施并非无关紧要的选择，它给出$\alpha_{j'} \le \alpha_j$，从而能把三段距离全部计入j 的预算。区分正贡献客户与预算等于距离的客户，对于付款与筛选论证也同样重要。

2.5 近似保证与实现

定理2.2. 算法返回可行的设施选址解，满足$f + \sum_j c_j \le 3\,\mathrm{OPT}$。实际上还有更强的不等式$3f + \sum_j c_j \le 3D \le 3\,\mathrm{OPT}$。

证明. 每个客户都有见证设施。见证设施被选中时可直接连接，被舍弃时可通过共享贡献者找到更早选中的设施，因此所有客户都能连接。把 (2.2) 乘以三，并利用连接成本非负，有

$$\begin{aligned} 3f + \sum_j c_j &\le 3\left(f + \sum_{j\in C_1} c_j\right) + \sum_{j\in C_2} 3\alpha_j \\ &\le 3\sum_{j\in C} \alpha_j = 3D \le 3\,\mathrm{OPT}\,. \end{aligned}$$

由于$f \ge 0$，普通的三倍近似保证随之成立。 □

不必枚举星形对象。对尚未变紧的设施，其总贡献是全局时间的连续分段线性函数，断点为到活跃客户的距离，可通过排序或扫描断点计算下一次变紧时间；同时跟踪活跃客户到达已紧设施的事件。每个主要事件都会新增一个临时设施或冻结一个客户，因此事件驱动实现为多项式时间。使用极小固定时间步长的朴素模拟则不能证明多项式界。

注2.3 (课件中的背景结果)**.** 课件提到栗师的1.488 近似算法，以及接近1.463 的困难性阈值，即$\rho = 1 + 2e^{-\rho}$ 的正解。这些是独立于此处三近似证明的深入结果。栗师2011 年论文中陈述的历史困难性假设为NP $\not\subseteq$ DTIME$(n^{O(\log\log n)})$。准确背景可参见原论文。

3 课后独立练习

1. 在一个带权路径上按两种不同边顺序运行顶点覆盖算法，比较返回覆盖及对偶值。
2. 为某个顶点覆盖实例构造逐坐标极大但非最优的对偶向量。
3. 固定设施i，把$y_i = 0.8$、分配值为$0.2, 0.5, 0.8$ 的基本LP 向量显式转换为至多四个星形对象。
4. 证明设施i 的全部星形对偶约束与 (1.4) 等价。
5. 解释开设所有紧设施为何会多次计入同一客户的剩余预算，以及筛选如何避免该问题。
6. 准确指出证明中分别在哪里使用了度量对称性、三角不等式、开设时间顺序以及开设成本非负性。
7. 说明如何处理零开设成本和同时变紧事件，而不破坏对偶可行性。

来源与阅读说明

主要依据：授课教师的课件文件夹

`5.PrimalDual`

`Topics.tex`.

选读部分对应课件中的选讲内容（如有）。补充的例子和练习服务于课后阅读，不占用新的授课讲次。

课程网站. 网站上前段课程的讲义用于参考组织方式与展开程度，本专题的内容仍以线性规划课件为准。