arXiv:2301.00573v4 [math.OC] 6 Jul 2023

混合整数线性规划的拉格朗日松弛综述：重要性、挑战、历史回顾、最新进展与机遇

Mikhail A. Bragin 康涅狄格大学电气与计算机工程系，美国斯托斯mikhail.bragin@uconn.edu

摘要

社会重要领域中的运作常常被建模为混合整数线性规划（MILP）问题。尽管混合整数线性规划问题具有组合复杂性，拉格朗日松弛一直为通过分解解决相关难题带来了希望。由于拉格朗日对偶函数具有非光滑特性，该方法的协调环节面临着严峻挑战。本文回顾了若干重要的历史里程碑（始于 Polyak 于 1967 年开展的开创性工作），这些进展旨在通过改进非光滑泛函的优化，提升拉格朗日松弛的协调能力。最后，本文介绍了拉格朗日松弛在快速求解混合整数线性规划问题方面的最新进展，并简要探讨了拉格朗日松弛在当前所能带来的机遇。

关键词： 组合优化；分解与协调；离散优化；对偶性；拉格朗日松弛；混合整数线性规划；机器学习；Polyak 步长

1 引言

本文旨在综述基于拉格朗日松弛的方法在混合整数线性规划（MILP）问题中的应用。由于存在整数变量，拉格朗日松弛会导致对偶空间中的非光滑优化。因此，本文也将综述关键的非光滑优化方法。

本文重点研究如何高效求解可分离的混合整数线性规划（MILP），其形式化定义如下:

$$\min_{(x,y):=\{x_i,y_i\}_{i=1}^{I}} \left\{ \sum_{i=1}^{I} \left((c_i^x)^T \cdot x_i + (c_i^y)^T \cdot y_i \right) \right\}, \tag{1}$$

其中， I 个子系统通过以下约束相互耦合

$$s.t. \ \sum_{i=1}^{I} A_i^x \cdot x_i + \sum_{i=1}^{I} A_i^y \cdot y_i - b = 0, \ \{x_i, y_i\} \in \mathcal{F}_i, i = 1, \ldots, I. \tag{2}$$

假设原始问题 (1)-(2)是可行的，并且可行域 $\mathcal{F} \equiv \prod_{i=1}^{I} \mathcal{F}_i$with $\mathcal{F}_i \subset \mathbb{Z}^{n_i^x} \times \mathbb{R}^{n_i^y}$ 有界且有限。

1.1 混合整数线性规划问题的重要性与难点

混合整数线性规划在许多对社会具有重要意义的问题中有广泛应用：救护车重新配置（ Lee 等人，2022）、共享乘车的成本分摊（Hu 等人，2021）、投递箱选址（Schmidt 和 Albert，2023 ）、大规模分布式系统中的高效故障检测（ Er-Rahmadi 和 Ma，2022 ）、居家医疗服务路径规划（Dastgoshade 等人，2020）、上门服务路径规划与预约排程（Tsang 和 Shehadeh，2023）、作业车间调度（Liu 等人，2021）、设施选址（ Basciftci 等人，2021）、流水车间调度（Hong 等人，2019；¨Balogh 等人，2022；Oztop 等人，2022）、货运（Archetti 等人，2021）、救灾物资的选址与库存前置部署（Shehadeh 和 Tucker，2022）、具有与序列相关的准备时间的机器调度（Yalaoui 和 Nguyen，2021）、海上库存路径规划（Gasse 等人，2022）、基于冲突的搜索的多智能体路径规划（Huang 等人，2021）、多车场电动公交车调度（Gkiotsalitis 等人，2023）、多级多设施绿色逆向物流网络设计（Reddy 等人，2022）、最优医生人员配置（ Prabhu 等人，2021）、具有可见性的最优搜索路径（Morin 等人，2023）、口服霍乱疫苗配送（Smalley 等人，2015）、门诊结肠镜检查排程（Shehadeh 等人，2020）、药品配送（Zhu 和 Ursavas，2018）、植物工厂作物排程（Huang 等人，2020）、灾后血液供应（Hamdan 和 Diabat，2020；Kamyabniya 等人，2021）、人机协作下的实际装配线平衡（Nourmohammadi 等人，2022）、降低遭受人口贩运侵害的脆弱性（ Kaya 等人，2022）、恢复规划与作业人员路径规划（Morshedlou 等人，2021）、拼车（Gaul 等人，2022）、非常规油田开发调度（Soni 等人，2021）、安全约束最优潮流（Velloso 等人，2021）、半导体

制造业（Chang and Dong，2017）、手术排程（Kayvanfar 等人，2021）、机组组合（Kim 等人，2018；Chen 等人，2019；Li 和 Zhai，2019；Chen 等人，2020；Li 等人，2020；van Ackooij 等人，2021）、车辆共享与任务分配（Arias-Melia 等人，2022）、工作量分配（Gasse 等人，2022）等。然而，混合整数线性规划问题通常属于NP难问题，因为其中包含整数变量 x。实际规模的混合整数线性规划问题通常难以求解，这是由其组合复杂性造成的。随着问题规模增大，求得最优解所需的计算开销会超线性增长，例如呈指数增长；对于许多实际问题，即使只要求得一个可行解，计算开销也可能十分巨大。因此，由于这些问题具有NP难的性质，几乎所有优化算法的运行时间都是超线性的，这意味着除非 P=NP，否则目前尚无已知的多项式时间算法能够求得这些问题的最优解。此外，许多重要问题通常要求在较短时间内求解（根据应用场景不同，时间从20分钟到几秒不等），同时还需要高质量的解。

随着问题规模缩小，NP 难问题的复杂度会呈指数级降低。对偶分解与协调拉格朗日松弛法有望利用这种复杂度的降低；该方法本质上通过分解“逆转”组合复杂性，从而大幅减少求解子问题所需的工作量（每个子问题 i 对应一个子系统 i）。拉格朗日松弛也深深植根于经济学理论，其所得解遵循“供需”这一经济学原理。当“需求”超过“供给”时，拉格朗日乘子（可视为“影子价格”）便会上升，反之亦然，从而抑制各子系统作出不那么“经济可行”的决策。尽管分解这一方面具有优势，但该方法基于“价格”的协调机制（用于适当协调各子问题）数十年来一直是深入研究的课题，因为原始空间中存在整数变量，导致相关对偶函数所涉及的底层非光滑优化存在根本性困难。因此，本文也将回顾非光滑优化的关键方法。

本文旨在简要回顾拉格朗日松弛法在混合整数线性规划问题中的发展历程及其关键里程碑，以及

凸非光滑函数优化领域的发展。本文其余部分的结构安排如下：

1. 第2节开头介绍拉格朗日对偶问题，并说明通过拉格朗日松弛求解混合整数线性规划问题时面临的困难。随后各小节将逐一解决这些困难；

2. 在第2.1小节中，介绍非光滑优化的早期研究（Polyak，1967、1969），为后续进展奠定基础；具体而言，介绍依赖于最优对偶值 $q(\lambda^*)$的 Polyak 公式（Polyak，1969），以确保几何（也称为线性 收敛速率；

3. 在第2.2小节中，讨论了20世纪70年代兴起的拉格朗日对偶基础研究及其在求解离散优化问题中的应用；

4. 在第2.3小节中，介绍次梯度水平 方法（Goffin 和 Kiwiel，1999），以便无需知道 $q(\lambda^*)$即可确保收敛；

5. 在第2.4小节中，解释了与次梯度方法相关的基本困难（高计算开销和乘子之字形摆动）；

6. 在第2.5和第2.6小节中，回顾了两种旨在降低计算开销并缓解乘子之字形摆动的独立研究方向：代理（Kaskavelis 和 Caramanis，1998；Zhao 等人，1999）和增量（Nedic 和 Bertsekas，2001）；前一种研究方向仍需要 $q(\lambda^*)$ 才能实现收敛；后一种研究方向遵循第2.3小节提出的“次梯度水平”思想，避免了获知 $q(\lambda^*)$ 的必要性；

7. 在第2.7小节中，回顾了代理拉格朗日松弛（SLR）（Bragin 等人，2015）：该方法通过利用“压缩映射”，在无需 $q(\lambda^*)$ 的情况下证明了收敛性，同时继承了第2.5小节代理方法的收敛性质；

8. 在第2.8小节中，介绍了 SLR 的进一步方法学进展；为加快约束违反的减少速度，同时支持使用混合整数线性规划求解器，引入了“绝对值”罚项（Bragin 等人，2018）；为在避免同步开销的同时高效协调分布式实体，开发了 SLR 的计算分布式版本，以异步方式高效协调分布式子系统（Bragin 等人，2020）；

9. 在第2.9小节中，回顾了代理“基于水平”的拉格朗日松弛（Bragin 和 Tucker，2022）；这一首创方法利用了第2.1小节所介绍的 Polyak 公式中固有的线性速率收敛潜力，但无需了解 $q(\lambda^*)$，也无需对第2.3小节介绍的其估计值进行启发式调整。相反， $q(\lambda^*)$ 的估计值（即“层级值”）通过一个简单的约束满足问题纯粹地确定；代理概念（Zhao 等人，1999；Bragin 等人，2015）确保了较低的计算需求，并减轻了之字形振荡；通过“绝对值”罚项（Bragin 等人，2018）实现了约束违反的加速减少，从而支持使用混合整数线性规划求解器；

10. 第3节将简要总结相关内容，并阐述未来的发展方向 ed.

2 离散规划与非光滑优化的拉格朗日对偶性

与原始混合整数线性规划问题（1）—（2）相对应的拉格朗日对偶问题如下，这是一个非光滑优化问题:

$$\max_{\lambda}\{q(\lambda) : \lambda \in \Omega \subset \mathbb{R}^m\}, \tag{3}$$

其中，凸对偶函数定义如下:

$$q(\lambda) = \min_{(x,y)} \big\{L(x, y, \lambda), \{x_i, y_i\} \in \mathcal{F}_i, i = 1, \ldots, I\big\}. \tag{4}$$

其中， $L(x,\ y,\ \lambda) \equiv \sum_{i=1}^{I} \big((c_i^x)^T \cdot x_i + (c_i^y)^T \cdot y_i\big) + \lambda^T \cdot \big(\sum_{i=1}^{I} A_i^x \cdot x_i + \sum_{i=1}^{I} A_i^y \cdot y_i - b\big)$ 是通过使用拉格朗日乘子 λ 松弛耦合约束（2）而得到的拉格朗日函数。在（4）中关于 $\{x, y\}$ 最小化 $L(x,\ y,\ \lambda)$ 被称为松弛问题，由于 $L(x,\ y,\ \lambda)$ 的可加性，该问题可分解为子问题。这一特性将从第2.5小节开始加以利用。

尽管原始原始问题 (1)是非凸的，但 $q(\lambda)$始终是连续且凹的，并且可行集 Ω 始终是凸集。由于原始空间中的整数变量 x ， $q(\lambda)$是不光滑的，具有相交于脊线的面（每个面都代表(4)的一个特定解）

在脊线处相交， $q(\lambda)$的导数表现出不连续性；特别是， $q(\lambda)$在 λ^*处不可微。

以下定义有助于进一步讨论：

定义 2.1 向量 $g(\lambda^k) \in \mathbb{R}^m$ 是 $q: \mathbb{R}^m \to \mathbb{R}$ 在 $\lambda^k \in \Omega$ 处的次梯度，若对于所有$\lambda \in \Omega$，均有： $q(\lambda) \leq q(\lambda^k) + \left(g(\lambda^k)\right)^T \cdot (\lambda - \lambda^k)$。

请注意，即使 q 在 λ^k处不可微，也可能存在次梯度。此外，函数 q 在点 λ^k处可能有多个次梯度。理解次梯度重要性的一种方式是， $q(\lambda^k) + \left(g(\lambda^k)\right)^T \cdot (\lambda - \lambda^k)$ 是 $q(\lambda)$ 的全局上估计函数。

定义 2.2 q 在 λ^k 处的次梯度集合称为 q在 λ^k处的次微分，记作 $\partial q(\lambda^k)$。

如上所述，梯度给出函数增长最快的方向，而次梯度并不唯一，也不一定能指示不可微函数的增长方向（参见图1中A点处的次梯度方向）。此外，次梯度可能几乎与指向最优乘子的方向垂直 λ^*，从而导致 λ 沿着对偶函数的脊线之字形振荡（图1中B点处相应的次梯度方向可作说明）。

虽然在式（4）中，拉格朗日乘子 λ 通常是固定参数， λ 但对于对偶问题（3）而言，它们是“对偶”决策变量。传统上，通过沿次梯度 g（x^k， y^k） s^k 进行一系列步长更新 λ，以迭代方式求解（3）的最大化问题：

$$\lambda^{k+1} = \lambda^k + s^k \cdot g(x^k, y^k), \tag{5}$$

其中， $\{x^k, y^k\}$是最优解 $\{x^*$（λ^k） y^*（λ^k）$\}$的简洁表示，表示乘子等于 λ^k. 时松弛问题（4）的最优解。在拉格朗日松弛中，次梯度定义为约束违反的程度 g（x^k， y^k） $\equiv \sum_{i=1}^{I} A_i^x \cdot x_i^k + \sum_{i=1}^{I} A_i^y \cdot y_i^k - b$。严格来说，由于 $\{x^k, y^k\}$ 是一个函数 λ^k，如上所述，适当的记法应为 g（x^k， y^k） $= g$（λ^k）。这两种记法可以互换使用；为简洁起见，视情况均简记为 g^k。

要了解为什么 g^k 确实是次梯度，请考虑以下内容：

$$q(\lambda) \leq L(x, y, \lambda), \forall (x, y) \in \mathcal{F}. \tag{6}$$

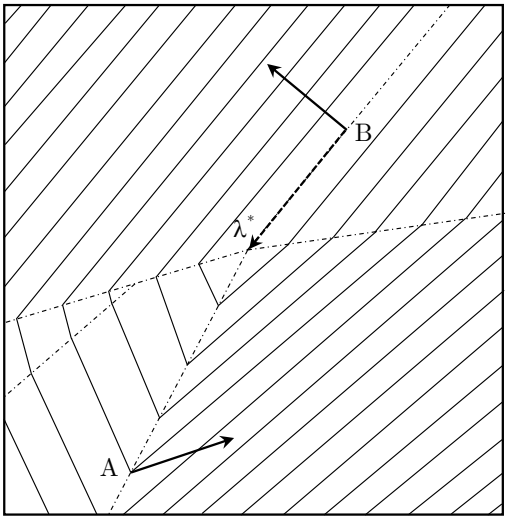


图1：一个对偶函数示例，展示了与次梯度方法相关的困难。实线表示等值曲线，点划线表示对偶函数的脊线，在这些位置梯度未定义（点 A 和 B 处可能的次梯度方向以实心箭头表示），虚线表示从点 B 指向最优乘子的次梯度方向。经许可，本图取自（Bragin 和 Tucker，2022）。

根据 (4)，该式成立。加上并减去 $\left(g^k\right)^T \cdot (\lambda^k)$，得到

$$q(\lambda) \leq L(x, y, \lambda^k) + \left(g^k\right)^T \cdot (\lambda - \lambda^k), \forall (x, y) \in \mathcal{F}. \tag{7}$$

由于上述不等式对所有可行的 x 和 y 均成立，因此对特定的可行值（例如 x^k 和 y^k. ）也成立。因此，

$$q(\lambda) \leq L(x^k, y^k, \lambda^k) + \left(g^k\right)^T \cdot (\lambda - \lambda^k). \tag{8}$$

根据定义 (4)，并且由于 $\{x^k, y^k\} := \{x^*(\lambda^k)\ y^*(\lambda^k)\}$，

$$L(x^k, y^k, \lambda^k) \equiv q(\lambda^k). \tag{9}$$

因此，

$$q(\lambda) \leq q(\lambda^k) + \left(g^k\right)^T \cdot (\lambda - \lambda^k). \tag{10}$$

证明至此结束。

如果存在不等式约束 $\sum_{i=1}^{I} A_i^x \cdot x_i + \sum_{i=1}^{I} A_i^y \cdot y_i \leq b$，通常会通过引入非负实值松弛变量 z 将其转换为等式约束，使得 $\sum_{i=1}^{I} A_i^x \cdot x_i + \sum_{i=1}^{I} A_i^y \cdot y_i + z = b.$。随后按照 (5) 更新乘子，并将其投影到正正交象限 $\{\lambda : \lambda \geq 0\}$。

由于对偶问题源于对耦合约束的松弛，因此对偶值通常小于原始值 $q(\lambda^k) < f(x^k, y^k)$,[1] 即存在对偶间隙—— $q(\lambda^k)$ 与 $f(x^k, y^k)$ 之间的相对差异。[2] 由于原始问题 (1)-(2) 的离散性质，即使在最优时，对偶间隙通常也不为零，即$q(\lambda^*) < f(x^*, y^*)$。因此，最大化对偶函数并不能得到最优原始解 (x^*, y^*)，甚至无法得到可行解。为了获得关于原问题 (1)-(2) 可行的解，松弛问题 $\{x_i^k, y_i^k\}$ 的解通常会通过启发式方法进行扰动。[3] 通常，乘子越接近其最优值 λ^*，约束违反程度就越小（这是由于对偶函数的凹性），因此搜索可行解也就越容易。

简而言之，拉格朗日松弛法高效求解混合整数线性规划问题时面临的障碍如下：

1. 对偶函数不可微；

   (a) 次梯度方向并非下降方向；

   (b) 极值的必要充分条件不适用；

2. 当子系统数量较多时，计算次梯度方向需要高计算开销；

3. 乘子在对偶函数的脊线上之字形摆动，导致需要许多次迭代才能实现收敛；这一困难源于对偶函数的不可微性，但该困难值得单独解决；

[1] 在这种特定情况下，$f(x^k, y^k) = \sum_{i=1}^{I} \left((c_i^x)^T \cdot x_i^k + (c_i^y)^T \cdot y_i^k\right)$ 满足以下条件：$\{x^k, y^k\}$ 满足约束 (2)。[2] 对偶值可用于量化解 $\{x^k, y^k\}$ 的质量。[3] 要解决混合整数线性规划问题，拉格朗日松弛通常被视为一种启发式方法。然而，在对偶空间中，拉格朗日松弛法是精确的；与许多其他启发式方法不同，该方法还能够通过更新乘子来改进解。

4. 松弛问题的解 $\{x_i^k, y_i^k\}$ 组合在一起时，不满足约束条件 (2)。此外，

5. 即使在 λ^*，约束违反程度也可能很大，而且为“修复”松弛问题解 $\{x_i^k, y_i^k\}$ 所需的启发式工作仍可能相当繁重。

为解决难点 1(b)，步长需要趋近于零（但仅满足这一条件还不够），下文小节将对此进行讨论。这一要求限制了本文将回顾的方法。

范围。 通过考察上述难点（下文也将称为 D1(a)、D1(b)、D2、D3、D4 和 D5），本文将按时间顺序回顾拉格朗日松弛的发展历程及其在优化非光滑对偶函数和求解混合整数线性规划问题中的应用，并介绍应对上述难点的方法的具体特点。鉴于 D1(b) 和 D2 等上述难点，尽管某些研究方向各有价值，本文仍将不予讨论，例如：

1.**乘子法**。 交替方向乘子法（ADMM）源自增广拉格朗日松弛（ALR）（即“乘子法”），它引入二次惩罚来惩罚松弛约束的违反，从而提高拉格朗日松弛的收敛性。然而，这两种方法（ADMM 和 ALR）仅在求解连续原问题时才能收敛。若步长不趋近于零，在求解离散原问题时，两种方法在对偶空间中均不收敛 [，也无法解决 D1(b)。尽管如此，ALR 背后的惩罚思想促成了其他基于拉格朗日松弛的方法的发展，这些方法具有更好的收敛性，详见第 2.8 小节。

2. **束方法**。 束方法的理念是获得所谓的 $\varepsilon-$上升方向，以更新乘子（Zhao 和 Luh，2002）。鉴于对偶函数的不可微性通常会导致非上升的次梯度方向，束方法解决了 D1(a)。然而，由于需要多次求解松弛问题（Zhao 和 Luh，2002），获取乘子更新方向所需的工作量超过了次梯度方法，因此该方法无法解决 D2。

2.1 1960年代：最小化“非光滑泛函”

非光滑凸函数优化是一个比拉格朗日松弛中的 q（λ）优化更广泛的主题，其方法源自 Polyak（1967）的开创性工作。为了介绍支撑拉格朗日松弛高效求解混合整数线性规划问题的基本原理，接下来将讨论 Polyak（1967）的工作。

采用“非可求和”步长的次梯度法。尽管对于最小化（最大化）问题，次梯度通常并非下降（上升）方向（Polyak，1967，第33页），但在满足以下条件的步长公式下，已证明可收敛到优化非光滑函数的最优解（例如，λ^* 最大化 $q(\lambda)$）。该公式常被称为非可求和步长公式：

$$s^k > 0, \quad \lim_{k\to\infty} s^k = 0, \quad \sum_{k=1}^{\infty} s^k = \infty. \tag{11}$$

采用 Polyak 步长的次梯度法。正如 Polyak 在其后续研究中指出的（Polyak，1969，第15页），非可求和步长会导致收敛速度非常缓慢。为了实现线性[4] 收敛速度，使得 $\|\lambda^k - \lambda^*\|$ 单调递减，Polyak 提出了一种步长公式，可表示如下：

$$s^k = \gamma \cdot \frac{q(\lambda^*) - q(\lambda^k)}{\left\|g(x^k, y^k)\right\|^2}, 0 < \varepsilon_1 \le \gamma \le 2 - \varepsilon_2, \varepsilon_2 > 0. \tag{12}$$

假设函数 $q(\lambda)$ 是强凸的（$q(\lambda^*) - q(\lambda^k) \ge m \cdot \|\lambda^* - \lambda^k\|^2$），并且满足利普希茨条件（$g^k \le M \cdot \|\lambda^k - \lambda^*\|$）（这两个条件均见 Polyak，1969，第17页），则 Polyak 结果的一种表述如下。首先，将 $\|\lambda^* - \lambda^{k+1}\|^2$ 进行二项式展开：

$$\|\lambda^* - \lambda^{k+1}\|^2 = \|\lambda^* - \lambda^k\|^2 - 2 \cdot s^k \cdot (g^k)^T \cdot (\lambda^* - \lambda^k) + (s^k)^2 \cdot \|g^k\|^2. \tag{13}$$

由于对偶函数的凹性，

$$q(\lambda^*) - q(\lambda^k) \le (g^k) \cdot (\lambda^* - \lambda^k). \tag{14}$$

因此，(13) 变为：

$$\|\lambda^* - \lambda^{k+1}\|^2 \le \|\lambda^* - \lambda^k\|^2 - 2 \cdot s^k \cdot (q(\lambda^*) - q(\lambda^k)) + (s^k)^2 \cdot \|g^k\|^2. \tag{15}$$

[4]也有可能实现超线性收敛，但需要满足以下条件：1. 对偶问题必须重新表述（Charisopoulos 和 Davis，2022）；2. 如第 2.4 小节所述，在拉格朗日松弛框架下，通常无法得到对偶函数。

根据 Polyak 步长公式 (12)，方程 (15) 变为：

$$\|\lambda^* - \lambda^{k+1}\|^2 = \|\lambda^* - \lambda^k\|^2 - 2 \cdot \gamma \cdot \frac{(q(\lambda^*) - q(\lambda^k))^2}{\|g^k\|^2} + \gamma^2 \cdot \frac{(q(\lambda^*) - q(\lambda^k))^2}{\|g^k\|^2}. \tag{16}$$

强凸条件 $q(\lambda^*) - q(\lambda^k) \geq m \cdot \|\lambda^* - \lambda^k\|^2$ 意味着 $-\left(q(\lambda^*) - q(\lambda^k)\right)^2 \leq -m^2 \cdot \|\lambda^* - \lambda^k\|^4$，因此，鉴于 $(2 \cdot \gamma - \gamma^2) > 0$，对于 $0 < \varepsilon_1 \leq \gamma \leq 2 - \varepsilon_2$，方程（16）变为

$$\|\lambda^* - \lambda^{k+1}\|^2 = \|\lambda^* - \lambda^k\|^2 - (2 \cdot \gamma - \gamma^2) \cdot \frac{m^2 \cdot \|\lambda^* - \lambda^k\|^4}{\|g^k\|^2}. \tag{17}$$

G鉴于利普希茨条件为 $g^k \leq M \cdot \|\lambda^k - \lambda^*\|$，方程（17）变为

$$\|\lambda^* - \lambda^{k+1}\|^2 = \|\lambda^* - \lambda^k\|^2 - (2 \cdot \gamma - \gamma^2) \cdot \frac{m^2 \cdot \|\lambda^* - \lambda^k\|^4}{M^2 \cdot \|\lambda^* - \lambda^k\|^2}. \tag{18}$$

因此

$$\|\lambda^* - \lambda^{k+1}\|^2 = \|\lambda^* - \lambda^k\|^2 \cdot \left(1 - (2 \cdot \gamma - \gamma^2) \cdot \frac{m^2}{M^2}\right). \tag{19}$$

在（12）至（19）中及本文后续部分，均使用标准欧几里得范数（除非另有说明）。

Polyak 步长（12）可视为对 D1(a) 的一种创新性变通：它以一个可证明更易解决的问题——每一步降低 $\|\lambda^* - \lambda^k\|$，从而保证收敛——取代了一个计算难度更高的问题，即在每次迭代时获取上升方向（如 Bundle 方法中所做的），以确保收敛。

在 1969 年之后的数十年间，针对两种不同步长方向——“非可求和” (11) 和“Polyak” (12)——的研究持续开展，各研究团队分别侧重于其中一种方向。两种方向都保证收敛至使对偶函数 $q(\lambda)$ 达到最大值的 λ^* （从而解决 D1(b)），不过，截至目前的讨论，使用 Polyak 步长的收敛性纯属理论，因为 (12) 中所需的最优对偶值 $q(\lambda^*)$ 尚不知晓。

在深入探讨上述难题的解决方法之前，我们有必要回顾奠定我们对离散规划问题中的拉格朗日松弛理解基础的研究。若不了解那些深刻影响离散规划研究发展方向的开创性研究，我们对这一主题的历史考察便是不完整的。尽管当时尚不清楚最优对偶值

$q(\lambda^*)$尚未得到确定，但公式 (12) 所提供的几何收敛潜力，以及拉格朗日松弛中通过分解实现的复杂度指数级降低，均得到了广泛利用，并由此推动了混合整数线性规划问题的解决，后文将对此展开讨论。

2.2 20世纪70年代和20世纪80年代：拉格朗日松弛与非光滑优化在数学规划和运筹研究问题中的应用

20世纪70年代和20世纪80年代，拉格朗日松弛领域既涌现了大量基础研究，也迎来了广泛的实际应用。Shapiro（1971年）、Geoffrion（1974年）以及Fisher和Shapiro（1974年）的研究是基础研究的代表，为该领域未来的发展奠定了基础。与此同时，非光滑优化技术的研究与应用也显著增多。值得注意的是，Polyak方法（见公式（12））成为一种广泛使用的技术，用于求解各种离散优化问题，包括纯整数问题和混合整数问题。由于式（12）中相关对偶问题的最优对偶值 $q(\lambda^*)$ 通常未知，因此在计算步长时，需要估计 $q(\lambda^*)$，例如采用一个“目标”值q （如Held和Karp（1971年）），或使用原始问题的一个可行成本值（如Fisher（1976年））。该方法的重要应用包括旅行商问题（Held和Karp（1970年）、Held和Karp（1971年））、调度问题 （Fisher（1973年）、Fisher（1976年）、Muckstadt和Koenig（1977）、Shepardson和Marsten（1980年））、选址问题（Cornuejols 等（1977）、Erlenkotter （1978 年））等。Fisher （1981 年）和Fisher（1985年）对拉格朗日松弛在离散规划问题中的早期应用作了出色的总结。回顾上述研究，了解塑造离散规划中拉格朗日松弛方法发展的基本原理、见解和技术，能为我们深入探讨本综述后文所述的先进方法和创新应用提供宝贵视角。

进入20世纪90年代，一种变革性方法——次梯度水平法

——被提出，用以解决Polyak方法常见的实际收敛问题。

2.3 20世纪90年代：次梯度水平法

次梯度水平法（Goffin 和 Kiwiel，1999）通过根据对偶解的充分下降情况和振荡情况，自适应地调整水平估计值，以计算 Polyak 步长（12），从而解决了无法获得最优 [对偶] 值所带来的困难。

对于问题（3），该方法的步骤如下：用“水平”估计值 $q_{lev}^k = q_{rec}^{k_j} + \delta_j$ 代替最优对偶值 $q(\lambda^*)$，其中 q_{rec}^k是截至第 k 次迭代时获得的最佳（最高）对偶值（“记录目标值”），而 δ_j 是一个可调参数，j 表示 j^{th} 对 q_{lev}^k. 的更新。其基本前提是：当 δ_j “过大”时，乘子在移动了相当长的（预先设定的）距离 R 后仍未改善“记录”值，因而会出现振荡。在这种情况下，参数 δ_j 更新为 $\delta_{j+1} = \beta \cdot \delta_j$，其中 $\beta = \frac{1}{2}$. 。另一方面，如果 δ_j 使得对偶值充分增加：$q(\lambda^k) \geq q_{lev}^k + \tau \cdot \delta_j$，且 $\tau = \frac{1}{2}$，，则参数 δ_j 保持不变，并将乘子的移动距离重置为 0，以避免在未来迭代中因 β 而过早减小 δ_j 。

该方法开创了非光滑函数“基于水平”的优化新时代，后文将对此展开讨论，以最终提高混合整数线性规划问题的求解效率。这一进展奠定了坚实基础，并持续在后续方法中产生影响，也将在本综述的后续章节中持续体现。

在考察 D2 和 D3 的求解方法之后，关于实施 Polyak 步长以求解 D1 的进一步讨论将留待后续小节。

2.4 次梯度方法的根本困难

高计算开销（D2）。在迄今评述的方法中，假设非光滑函数是已知的。然而，对偶函数 $q(\lambda)$无法以较高的计算效率求得。事实上，即使给定乘子λ^k的值，通过式（4）中的最小化来求得对偶值 $q(\lambda^k)$及相应的次梯度，也十分耗时。即便如此

通常也只能得到次梯度的一个可能取值；一般无法完整描述次梯度（Goffin，1977）。

乘子之字形摆动（D3）。正如（Goffin，1977）所假设的，次梯度方法收敛缓慢是由病态条件造成的。条件数 μ 的形式定义如下（Goffin，1977）： $\mu = \inf\{\mu(\lambda) : \lambda \in \mathbb{R}^m / P\}$，其中 $P = \{\lambda \in \mathbb{R}^m : q(\lambda) = q(\lambda^*)\}$（最优解集合），而 $\mu(\lambda) = \min_u \frac{u^T \cdot (\lambda^* - \lambda)}{\|u^T\| \cdot \|\lambda^* - \lambda\|}$（次梯度与指向最优乘子的方向所成角度的余弦）。[5] 随后，在求解调度问题（Czerwinski 和 Luh，1994，图 3(b)，第 104 页）以及电力系统问题（Guan 等人，1995，图 4，第 774 页）时，通过实验确认，病态条件会导致乘子沿着对偶函数的脊线之字形振荡。

为了解决这两个难题，20世纪90年代末至21世纪初出现了“代理”“交错”和“增量”次梯度的概念。正如下一节将回顾的那样，这些方法无需将松弛问题完全优化即可加快收敛。

2.5 20世纪90年代末：交错次梯度法与代理次梯度法

在 Kaskavelis 和 Caramanis（1998）提出的交错次梯度方法中，每次求解一个子问题后，便更新乘子

$$\min_{(x_i, y_i)} \left\{ (c_i^x)^T \cdot x_i + (c_i^y)^T \cdot y_i + \lambda^T \cdot \left(A_i^x \cdot x_i + A_i^y \cdot y_i \right), \{x_i, y_i\} \in \mathcal{F}_i \right\}, \tag{20}$$

而不是像次梯度方法那样求解所有子问题。这显著降低了计算开销，尤其是对于子系统数量众多的问题。

随后，Zhao 等人（1999）提出了更通用且已证明具有收敛性的代理次梯度方法，该方法不要求松弛问题（甚至子问题）达到精确最优。只要满足以下“替代最优性条件”

$$L(\tilde{x}^k, \tilde{y}^k, \lambda^k) < L(\tilde{x}^{k-1}, \tilde{y}^{k-1}, \lambda^k) \tag{21}$$

[5]通过目视检查图1所示的对偶函数，条件数很可能为 0，因为从点 B 出发的次梯度似乎与指向最优乘子的方向成直角。

乘子便按如下方式更新

$$\lambda^{k+1} = \lambda^k + s^k \cdot g(\tilde{x}^k, \tilde{y}^k), \tag{22}$$

通过使用以下公式

$$0 < s^k < \gamma \cdot \frac{q(\lambda^*) - L(\tilde{x}^k, \tilde{y}^k, \lambda^k)}{\|g(\tilde{x}^k, \tilde{y}^k)\|^2}, \ \gamma < 1. \tag{23}$$

保证收敛至 λ^*（Zhao 等人，1999年）。与原始 Polyak 公式（12）不同，参数 γ 小于 1，以确保 q（λ^*）$> L$（$\tilde{x}^k$，$\tilde{y}^k$，λ^k），从而使步长公式（23）首先有定义；这一点已由（Zhao 等人，1999年，命题 3.1，第 703 页）证明。此处的“波浪号”表示相应的解不一定是子问题最优解。解 $\{\tilde{x}^k, \tilde{y}^k\}$ 构成一个集合 $\mathcal{S}$（$\tilde{x}^{k-1}$，$\tilde{y}^{k-1}$，λ^k）$\equiv \{$（x，y）：L（x，y，λ^k）$< L$（$\tilde{x}^{k-1}$，$\tilde{y}^{k-1}$，λ^k）$\}$。一旦找到一个成员 $\{\tilde{x}^k, \tilde{y}^k\} \in \mathcal{S}$（$\tilde{x}^{k-1}$，$\tilde{y}^{k-1}$，$\lambda^k$），即满足替代最优性条件（21），松弛问题的优化便停止，并按式（22）更新乘子。当 $\mathcal{S}$（$\tilde{x}^{k-1}$，$\tilde{y}^{k-1}$，λ^k）$= \emptyset$时，这表示对于给定的 λ^k，无法找到优于 $\{\tilde{x}^{k-1}, \tilde{y}^{k-1}\}$ 的解，这说明（$\tilde{x}^{k-1}$，$\tilde{y}^{k-1}$）$=$（x^*（λ^k）y^*（λ^k））对于给定值 λ^k是子问题最优的，并且乘子通过次梯度方向更新。

收敛性证明与式 (13)–(19) 中给出的证明相当类似。唯一需要注意的是，严格来说，L($\tilde{x}^k$，$\tilde{y}^k$，λ^k) 并不是一个函数；与对偶函数 q(λ^k) 不同，L($\tilde{x}^k$，$\tilde{y}^k$，λ^k) 对给定的 λ^k 可以取多个值。因此，式 (14) 的对应结论不能由 L($\tilde{x}^k$，$\tilde{y}^k$，λ^k) 的凹性推出。该结论源于替代对偶值无需通过求解所有子问题即可获得，因此

$$q(\lambda^*) \leq L(\tilde{x}^k, \tilde{y}^k, \lambda^*). \tag{24}$$

在前一个不等式中加上并减去 g($\tilde{x}^k$，$\tilde{y}^k$)$^T \cdot \lambda^k$，可得

$$q(\lambda^*) - L(\tilde{x}^k, \tilde{y}^k, \lambda^k) \leq g(\tilde{x}^k, \tilde{y}^k)^T \cdot (\lambda^* - \lambda^k). \tag{25}$$

式 (15)–(19) 中所述的过程可类似地进行。

除了计算开销降低之外，还观察到乘子之字形摆动也随之减少。事实上，除上述情况 $\mathcal{S}$($\tilde{x}^{k-1}$，$\tilde{y}^{k-1}$，λ^k) $= \emptyset$ 外，一个子问题 (20) 的解就足以

满足 (21)。在这种情况下，每个代理次梯度求和中只有一项会更新，因此可避免代理次梯度发生剧烈变化，也避免乘子因此出现之字形振荡。

2.6 21 世纪初：增量次梯度法

在增量次梯度法中，在更新乘子之前会先求解一个子问题 i （类似于交错法）。然而，与一次性更新所有乘子不同，增量次梯度法会逐步更新乘子。在计算出i^{th} 次梯度分量后，乘子更新为

$$\psi_i^k = \psi_{i-1}^k + s^k \cdot \left(A_i^x \cdot x_i^k + A_i^y \cdot y_i^k - \beta_i \right). \tag{26}$$

这里， β_i 是满足 $\sum_{i=1}^{I} \beta_i = b$的向量，例如$\beta_i = \frac{b}{I}$.。只有在求解完所有 i子问题后，乘子才会按如下方式进行“完全”更新

$$\lambda^{k+1} = \psi_I^k. \tag{27}$$

次梯度水平法的收敛结果（Goffin 和 Kiwiel，1999）已扩展至次梯度法。研究者提出了该方法的变体，其中 β 和τ属于区间 $[0,1]$，而不是等于1 2。 此外，为了提高收敛性，不再使用常数 R， 而是提出了一个 R_l 序列，使其满足 $\sum_{l=1}^{\infty} \underline{R}_l = \infty$.

2.7 2010年代：代理拉格朗日松弛法

代理拉格朗日松弛（SLR）（Bragin 等人，2015）基于压缩映射，克服了因缺乏最优对偶值的相关知识而带来的困难。在连续迭代中，乘子之间的距离必须减小，即:

$$\left\| \lambda^{k+1} - \lambda^k \right\| \leq \alpha_k \cdot \left\| \lambda^k - \lambda^{k-1} \right\|, \quad 0 \leq \alpha_k \leq 1. \tag{28}$$

根据（22）和（28），推导出步长公式:

$$s^k = \alpha_k \cdot \frac{s^{k-1} \left\| g(\tilde{x}^{k-1}, \tilde{y}^{k-1}) \right\|}{\| g(\tilde{x}^k, \tilde{y}^k) \|}. \tag{29}$$

Mor此外，已有一种用于设定 α_k 的特定公式，可保证收敛 :

$$\alpha_k = 1 - \frac{1}{M \cdot k^{1-\frac{1}{k^r}}},\ M \geq 1,\ 0 \leq r \leq 1. \tag{30}$$

由于 $\alpha_k \to 1$，SLR 中的步长是“非可求和”的。只有在 λ^* 的某个邻域之外，才能保证线性收敛（Bragin 等人，2015，命题 2.5，第 187 页）。

当乘子接近其最优值时[6]，，代理次梯度会逐渐接近实际次梯度，从而由于对偶函数的凹性，约束违反会减少。这表明各子问题协调良好，松弛问题解接近可行解。因此，只有少数子问题会导致不可行。这极大地有助于解决难点 D4（“松弛问题的解组合在一起后，可能无法构成原问题的可行解”）：Bragin 等人在 Bragin 等人（2018）中提出了一种“自动”流程，用于识别并“修复”导致原问题不可行的少数子问题解。

2.8 2010年代末至2020年代初：方法论的进一步进展

替代绝对值拉格朗日松弛（Bragin 等人，2018）。替代绝对值拉格朗日松弛（SAVLR）方法旨在保证收敛，并加快约束违反的减少，同时避免使用传统二次项时可能出现的非线性和非凸性。在 SAVLR 方法中，考虑以下对偶问题：

$$\max_{\lambda}\{q_\rho(\lambda) : \lambda \in \Omega \subset \mathbb{R}^m\}, \tag{31}$$

其中

$$q_\rho(\lambda) = \min_{(x,y)} \left\{ \sum_{i=1}^{I} \left((c_i^x)^T \cdot x_i + (c_i^y)^T \cdot y_i\right) + \lambda^T \cdot \left(\sum_{i=1}^{I} A_i^x \cdot x_i + \sum_{i=1}^{I} A_i^y \cdot y_i - b\right) + \right.$$
$$\left. \rho \cdot \left\| \sum_{i=1}^{I} A_i^x \cdot x_i + \sum_{i=1}^{I} A_i^y \cdot y_i - b \right\|_1, \{x_i, y_i\} \in \mathcal{F}_i, i = 1, \ldots, I \right\}. \tag{32}$$

上述最小化问题涉及可精确线性化的分段线性惩罚，这些惩罚用于惩罚约束违反，从而最终减少子问题的数量

[6]将在第 2.9 小节讨论一种用于量化乘子质量的指标（即衡量乘子与其最优值的接近程度）。

导致第 2.7 节所述的不可行性，从而减少启发式算法寻找原始解所需的工作量。这解决了困难 D5。[7]

分布式异步代理拉格朗日松弛（DA-SLR）（Bragin 等人，2020）。随着增强分布式计算能力的技术不断涌现，再加上网络基础设施的支持，与使用单台计算机相比，利用分布式计算资源能够更高效地完成计算任务。在假设存在单一协调器的前提下，DA-SLR 方法被开发出来，以异步方式高效协调分布式子系统，同时避免同步开销。与顺序式代理拉格朗日松弛（Bragin 等人，2015）相比，数值测试显示其收敛更快（在广义指派问题的一个实例中，达到 0.03% 的间隙所需时间缩短至原来的十二分之一）。

这里有必要做一个简短总结。从理论上讲，Polyak 公式具有线性收敛率；然而，无论是在 Goffin 和 Kiwiel（1999）的原始论文中，还是在后续对次梯度“基于水平”的思想的应用中（例如 Nedic 和 Bertsekas（2001）），都没有讨论次梯度水平法的收敛率。可能是因为乘子需要显式或隐式地经过无限距离，因而无法实现线性收敛。基于 SLR 的方法虽然无需估计最优对偶值，但为避免过早终止而设定的要求（30）会导致 [步长] 不可求和性。理想情况下，目标是避免乘子之字形振荡，减少计算乘子更新方向所需的计算开销，并实现线性收敛。朝这一方向迈出的第一步将在下一小节中说明。

2.9 E2020 年代初：代理水平法拉格朗日松弛 n

为发挥 Polyak 步长公式固有的线性收敛潜力，近年来开发了代理“基于水平”的拉格朗日松弛（SLBLR）方法

[7]参数 ρ 会不断增大，直到替代最优性条件 (21) 不再成立，这意味着无法再获得适当的“替代”次梯度方向。此外，过度惩罚约束违反会过分强调可行性，子问题解可能会陷入次优解。在这些情况下，ρ 不再增大；相反，替代最优性条件不成立时，会促使惩罚系数 $\rho^{k+1} = \rho^k/\beta$降低。

(Bragin 和 Tucker，2022)。研究证明，一旦以下“乘子发散检测”可行性问题

$$\begin{aligned}
&\|\lambda - \lambda^{k_j+1}\| \le \|\lambda - \lambda^{k_j}\|,\\
&\|\lambda - \lambda^{k_j+2}\| \le \|\lambda - \lambda^{k_j+1}\|,\\
&\dots\\
&\|\lambda - \lambda^{k_j+n_j}\| \le \|\lambda - \lambda^{k_j+n_j-1}\|,
\end{aligned} \tag{33}$$

对于某些 k_j 和 n_j，不存在关于 λ（即上述问题中的决策变量）的可行解，此时“层级值”等于

$$\overline{q}_j = \max_{\kappa\in[k_j,k_j+n_j]} \overline{q}_{\kappa,j} > q(\lambda^*), \tag{34}$$

其中

$$\overline{q}_{\kappa,j} = \frac{1}{\gamma}\cdot s^{\kappa}\cdot \|g(\tilde{x}^{\kappa},\tilde{y}^{\kappa})\|^2 + L(\tilde{x}^{\kappa},\tilde{y}^{\kappa},\lambda^{\kappa}). \tag{35}$$

在后续迭代中，使用 Polyak 步长公式

$$s^k = \zeta\cdot\gamma\cdot\frac{\overline{q}_j - L(\tilde{x}^k,\tilde{y}^k,\lambda^k)}{\|g(\tilde{x}^k,\tilde{y}^k)\|^2}, \zeta < 1, k = k_{j+1},\dots,k_{j+1}+n_{j+1}-1. \tag{36}$$

本质上，上述公式会一直使用，直到上述可行性问题再次无解；此时，层级值将从 q_j 重置为 q_{j+1}，并继续乘子更新过程。值得注意的是，优化问题 (3) 是最大化问题，其解（拉格朗日乘子）的质量可以通过 $\{q_j\}$ 提供的上界进行量化。

–

这里的假设是，上述可行性问题 (33) 会“周期性地”且“无限多次地”变得不可行，从而触发对“层级”值 q_j 的重新计算；这些值将逐渐减小，并从上方趋近于 q（λ^*）。这一假设是合理的，因为确保 (33) 始终可行的唯一可靠方法，就是知道最优对偶值。

鉴于上述内容，并加入“绝对值”罚项以促进可行解搜索，SLBLR 方法解决了全部 D1-D5 难题。该方法继承了 Polyak 公式（12）的特点、替代方法降低计算开销并缓解之字形振荡的优势（Zhao 等人，1999；Bragin 等人，2015），以及加速减少约束违反的优势（Bragin 等人，2018）。

SLBLR 的决定性优势在于，它通过上述高效的、基于决策的流程来确定“水平”值，从而能够实际应用 Polyak 公式（12），而无需估算或通过启发式方法调整最优对偶值的估计。Bragin 和 Tucker（2022）报告的结果表明，与其他方法相比，SLBLR 方法能够将更广泛的广义指派问题（GAPs）求解至最优。在其他条件相同的情况下，SLBLR 方法的“基于水平”的步长策略（Bragin 和 Tucker，2022）比 SAVLR 方法的“非可求和”步长策略（Bragin 等人，2018）更具优势。此外，SLBLR 还能成功求解随机作业车间调度和制药调度等其他问题，其表现至少比商业求解器快两个数量级。

SLBLR 方法（Bragin 和 Tucker，2022）并不局限于混合整数线性规划问题，因为上述层级值的确定并不要求线性。该方法具有模块化特性，并有望支持即插即用能力。例如，虽然制药调度应用已通过使用固定数据进行了测试（Bragin 和 Tucker，2022），但该方法也适用于处理制造新药品的紧急请求，因为此类产品可以“即时”引入松弛问题，而相应的子问题则可以继续通过拉格朗日乘子进行协调，无需调度员进行重大干预，也不会扰乱整体生产计划。

3 结论与未来方向

本文旨在总结高效求解混合整数线性规划问题所面临的关键难题，并简要回顾半个多世纪以来采用拉格朗日松弛应对这些难题的重要研究里程碑。此外，最新的 SLBLR 方法具有以下特点：1. 通用性强，有望处理一般的混合整数规划问题，因为即使不具备线性结构，也能利用线性收敛速率；2. 灵活且模块化，有望支持即插即用能力，并能实时响应自然灾害、运行故障和信息物理事件等不可预测和／或造成干扰的事件，包括电力系统中的发电机停运，

制造业中接到紧急订单，或运输过程中遭遇交通拥堵。借助通信和分布式计算能力，这些事件可通过持续更新拉格朗日乘子来应对，从而增强系统韧性；因此，该方法也适用于快速重新优化；3. 该方法兼容量子计算和机器学习方法等其他优化方法，这些方法有望进一步改进子问题求解，进而凭借该方法快速协调的特点，大幅缩短 CPU 时间。

致谢

本研究部分经费由美国国家科学基金会通过项目编号 ECCS-1810108 提供支持。

利益冲突声明

作者声明不存在利益冲突。

参考文献

Archetti, C., L. Peirano 和 M. G. Speranza（2021）。多式联运货运问题中的优化：综述。《欧洲运筹学杂志》*299*(1), 1–20。

Arias-Melia, P., J. Liu, 和 R. Mandania（2022）。车辆共享与任务分配问题：混合整数线性规划公式及一种启发式求解方法。《计算机与运筹学》*147*，105929。

Balogh, A., M. Garraffa, B. O’Sullivan, 和 F. Salassa（2022）。用于最小化无空闲置换流水车间问题总延迟时间的基于混合整数线性规划的局部搜索程序。《计算机与运筹学》*146*，105862。

Basciftci, B., S. Ahmed 和 S. Shen（2021）。决策相关随机需求下的分布鲁棒设施选址问题。欧洲运筹学杂志 *292*(2), 548–561。

Bragin, M. A., P. B. Luh, B. Yan 和 X. Sun（2018）。一种用于求解自动化领域中混合整数线性规划问题的可扩展方法。*IEEE* 自动化科学与工程汇刊 *16*(2), 531–541。

Bragin, M. A., P. B. Luh, J. H. Yan, N. Yu, 和 G. A. Stern（2015）。替代拉格朗日松弛法的收敛性。《优化理论与应用杂志》*164*(1), 173–201。

Bragin, M. A. 和 E. L. Tucker（2022）。用于混合整数线性规划的代理“基于水平”的拉格朗日松弛。《科学报告》*22*(1), 1–12。

Bragin, M. A., B. Yan, 和 P. B. Luh（2020）。通过代理拉格朗日松弛对混合整数线性系统进行分布式异步协调。*IEEE* 自动化科学与工程汇刊 *18*(3), 1191–1205。

Chang, X. 和 M. Dong（2017）。半导体制造中并行机器资质管理的随机规划。《计算机与运筹学》*79*, 49–59。

Charisopoulos, V. 和 D. Davis (2022). 一种用于尖锐半光滑问题的超线性收敛次梯度法。*arXiv* 预印本 *arXiv:2201.04611*。

Chen, Y.、F. Pan、J. Holzer、E. Rothberg、Y. Ma 和 A. Veeramany (2020). 一种基于高性能计算、由市场经济学驱动的邻域搜索与改进算法，用于安全约束机组组合。*IEEE Transactions on PowerSystems 36*(1), 292–302。

Chen, Y.、F. Wang、Y. Ma 和 Y. Yao (2019). 一种用于求解和基准测试安全约束机组组合的分布式框架，并支持热启动。*IEEE Transac-tions on Power Systems 35*(1), 711–720。

Cornuejols, G., M. L. Fisher 和 G. L. Nemhauser（1977）。杰出论文——通过银行账户位置优化浮存金：对精确算法和近似算法的分析研究。《管理科学》*23*(8), 789–810。

Czerwinski, C. S. 和 P. B. Luh（1994）。利用改进的拉格朗日松弛技术进行物料清单产品调度。*IEEE*机器人与自动化汇刊 *10*(2), 99–111。

Dastgoshade, S., A. Abraham 和 N. Fozooni（2020）。居家医疗保健问题的拉格朗日松弛方法。载于国际软计算与模式识别会议，第333–344页。Springer。

Er-Rahmadi, B. 和 T. Ma（2022）。用于大规模分布式系统高效故障检测的数据驱动的基于混合整数线性规划的优化。《欧洲运筹学杂志》*303*(1)，337–353。

Erlenkotter, D.（1978）。无容量限制设施选址的基于对偶的程序。《运筹学》*26*(6), 992–1009。

Fisher, M. L.（1973）。利用拉格朗日乘子求解调度问题的最优解：第一部分。《运筹学》*21*(5)，1114–1127。

Fisher, M. L. (1976). 单机调度问题的一个对偶算法。数学规划 *11*(1), 229–251. Fisher, M. L. (1981). 求解整数规划问题的拉格朗日松弛法。《管理科学》*27*(1), 1–18. Fisher, M. L. (1985). 拉格朗日松弛的面向应用指南。界面 *15*(2), 10–21. Fisher, M. L. and J. F. Shapiro (1974). 整数规划中的构造性对偶。《*SIAM*应用数学杂志》*27*(1), 31–52.
Gasse, M., S. Bowly, Q. Cappart, J. Charfreitag, L. Charlin, D. Ch´etelat, A. Chmiela, J. Dumouchelle, A. Gleixner, A. M. Kazachkov, et al. (2022). 组合优化机器学习竞赛（ML4CO）：结果与见解。载于*NeurIPS 2021*竞赛与演示专题，第220–231页。PMLR。
Gaul, D., K. Klamroth, and M. Stiglmayr (2022). 拼车应用的基于事件的混合整数线性规划模型。《欧洲运筹学杂志》*301*(3), 1048–1063。

Geoffrion, A. (1974). 整数规划的拉格朗日松弛。数学规划研究 *2*，82–114。

Gkiotsalitis, K., C. Iliopoulou 和 K. Kepaptsoglou（2023）。一种求解带时间窗的多车场电动公交车调度问题的精确方法。《欧洲运筹学杂志》*306*(1), 189–206。

Goffin, J.-L.（1977）。关于次梯度优化方法的收敛速度。*Mathe-matical programming 13*(1), 329–347。

Goffin, J.-L. 和 K. C. Kiwiel（1999）。一种简单的次梯度水平法的收敛性。*Mathematical Programming 85*(1), 207–211。

Guan, X., P. B. Luh 和 L. Zhang（1995）。用于基于拉格朗日松弛的水火电调度算法的非线性近似方法。*IEEE Transactions on PowerSystems 10*(2), 772–778。

Hamdan, B. 和 A. Diabat（2020）。利用拉格朗日松弛法在中断风险下对血液供应链进行稳健设计。《交通研究*E*辑：物流与交通评论》*134*, 101764。

Held, M. 和 R. M. Karp（1970）。旅行商问题与最小生成树。《运筹学》*18*(6)，1138–1162。

Held, M. 和 R. M. Karp（1971年）。旅行商问题与最小生成树：第二部分。*Mathematical Programming 1*(1)，6–25。

Hong, I.-H.、C.-C. Chou 和 P.-K. Lee（2019）。采用拉格朗日松弛（MIPLAR）的排队时间循环生产混合整数规划中的准入控制。*Computers& Industrial Engineering 129*, 417–425。

Hu, S.、M. M. Dessouky、N. A. Uhan 和 P. Vayanos（2021）。共享乘车的成本分摊机制设计。*Transportation Research Part B: Methodological 150*，410–434。

Huang, K.-L.、C.-L. Yang 和 C.-M. Kuo（2020）。运用拉格朗日松弛，考虑体积、产量变化和多期收获的植物工厂作物调度。*BiosystemsEngineering 200*，328–337。

Huang, T.、S. Koenig 和 B. Dilkina（2021）。利用基于冲突的搜索学习解决多智能体路径规划中的冲突。*AAA*人工智能大会论文集 *35*(13)，11246–11253。

Kamyabniya, A., Z. Noormohammadzadeh, A. Saur´e, and J. Patrick（2021）。一种用于灾害救援行动中基于年龄的多组血小板的稳健集成物流模型。《交通研究*E*辑：物流与交通评论》*152*, 102371。

Kaskavelis, C. A. and M. C. Caramanis（1998）。适用于大型工业作业车间调度问题的高效拉格朗日松弛算法。《*IIE*汇刊》*30*(11), 1085–1097。

Kaya, Y. B., K. L. Maass, G. L. Dimas, R. Konrad, A. C. Trapp, and M. Dank（2022）。改善离家出走及无家可归青年获得住房和支持性服务的机会：降低纽约市人口贩运风险。《*IISE*汇刊》,1–15。

Kayvanfar, V., M. R. Akbari Jokar, M. Rafiee, S. Sheikh, and R. Iranzad（2021）。一种采用择期患者策略的手术室排程新模型。《*INFOR*：信息系统与运筹研究》*59*(2), 309–332。

Kim, K.、A. Botterud 和 F. Qiu（2018）。电力系统生产成本建模中用于改进机组组合的时间分解。*IEEE Transactions on PowerSystems 33*(5)，5276–5287。

Lee, Y.-C.、Y.-S. Chen 和 A. Y. Chen（2022）。考虑随机需求与截断泊松分布的救护车重新定位与路径规划的拉格朗日对偶分解。*Transportation research part B: methodological 157*，1–23。

Li, X. 和 Q. Zhai（2019）。多阶段鲁棒输电约束机组组合：一种具有隐式决策规则的分解框架。《国际电力与能源系统杂志》*108*，372–381。

Li, X., Q. Zhai, 和 X. Guan (2020). 鲁棒输电约束机组组合：列合并方法。《IET发电、输电与配电》*14*(15), 2968– 2975.

Liu, A., P. B. Luh, B. Yan, 和 M. A. Bragin (2021). 一种用于作业车间调度问题的新型整数线性规划表述。*IEEE*机器人与自动化快报 *6*(3), 5937–5944.

Morin, M., I. Abi-Zeid 和 C.-G. Quimper（2023）。用于搜救行动路径规划的蚁群优化。《欧洲运筹学杂志》*305*(1), 53–63。

Morshedlou, N., K.Barker, A.D.Gonz´alez, andA.Ermagun（2021）。一种针对相互依赖的恢复规划与人员路径规划问题的启发式方法。*Computers &Industrial Engineering 161*, 107626。

Muckstadt, J. A. 和 S. A. Koenig（1977）。拉格朗日松弛在发电系统调度中的一种应用。《运筹学》*25*(3), 387–403。

Nedic, A. 和 D. P. Bertsekas（2001）。用于不可微优化的增量次梯度方法。《*SIAM*优化杂志》*12*(1), 109–138。

Nourmohammadi, A., M. Fathi 和 A. H. Ng（2022）。具有人机协作任务的装配线平衡与调度。《计算机与运筹学》*140*, 105674。

¨Oztop, H.、M. F. Tasgetiren、L. Kandiller 和 Q.-K. Pan（2022）。采用重启与学习机制求解以最大完工时间准则为目标的无等待流水车间调度问题的元启发式算法。《计算机与运筹学》*138*，105616。

Polyak, B. T.（1967）。求解极值问题的一般方法。《科学院院报》*174*（1），33–36。

Polyak, B. T.（1969）。非光滑泛函的最小化。苏联计算数学与数学物理 *9*（3），14–29。

Prabhu, V. G., K. Taaffe, R. Pirrallo, W. Jackson 和 M. Ramsay（2021）。通过医生轮班排班改善急诊科的患者安全和患者流动。载于*2021*年冬季仿真会议（*WSC*），第 1–12 页。IEEE。

Reddy, K.N., A. Kumar, A. Choudhary 和 T.E. Cheng（2022）。多期绿色逆向物流网络设计：一种改进的基于 Benders 分解的启发式方法。《欧洲运筹学杂志》*303*(2)，735–752。

Schmidt, A. 和 L. A. Albert（2023）。投递箱选址问题。《*IISE* 汇刊》（已录用，即将发表），1–24。

Shapiro, J. F.（1971年）。整数规划中的广义拉格朗日乘子。《运筹学》*19*(1)，68–76。

Shehadeh, K. S.、A. E. Cohn 和 R. Jiang（2020）。门诊结肠镜检查排程的分布鲁棒优化方法。《欧洲运筹学杂志》*283*(2)，549–561。

Shehadeh, K. S. 和 E. L. Tucker（2022）。救灾物资选址与库存前置部署的随机优化模型。《交通运输研究*C*辑：新兴技术》*144*，103871。

Shepardson, F. 和 R. E. Marsten（1980年）。双值班时段排程问题的拉格朗日松弛算法。《管理科学》*26*(3)，274–281。

Smalley, H. K., P. Keskinocak, J. Swann, and A. Hinman (2015). 优化口服霍乱疫苗分配策略以最大限度降低疾病发病率：混合整数规划模型及孟加拉国情景分析。《疫苗》*Vaccine 33*(46), 6218–6223.

Soni, A., J. Linderoth, J. Luedtke, and F. Rigterink (2021). 用于非常规油田开发调度的混合整数线性规划。*Optimization and Engineer-ing 22*(3), 1459–1489.

Tsang, M. Y. and K. S. Shehadeh (2023). 针对具有随机出行与服务时间的上门服务路径规划和预约排程问题的随机优化模型。《欧洲运筹学杂志》*307*(1), 48–63.

van Ackooij, W., C. d’Ambrosio, D. Thomopulos 和 R. S. Trindade（2021）。用于求解水电流域水电机组组合与排程的分解与最短路径问题建模。《欧洲运筹学杂志》*291*(3), 935–943。

Velloso, A., P. Van Hentenryck 和 E. S. Johnson（2021）。用于求解安全约束最优潮流的精确且可扩展的问题分解方法。《电力系统研究》*195*, 106677。

Yalaoui, F. 和 N. Q. Nguyen (2021). 带有序列相关准备时间的相同机器调度问题：混合整数线性规划表述与计算研究。《美国运筹学杂志》*11*(1), 15–34.

Zhao, X. 和 P. Luh (2002). 求解拉格朗日松弛对偶问题的新型束方法。《优化理论与应用杂志》*113*(2), 373–397.

Zhao, X.、P. B. Luh 和 J. Wang (1999). 用于拉格朗日松弛的代理梯度算法。《优化理论与应用杂志》*100*(3), 699–712.

Zhu, S. X. 和 E. Ursavas (2018). 制药行业中采用直接配送的卫星网络设计与分析。《交通研究*E*辑：物流与交通评论》*116*, 190–207.