面向生产级数学优化的凸对偶性契约

Juan Pablo Vielma，[a] Ross Anderson，[b] JoeyHuchette[c]Google Research，

[a] jvielma@google.com；b rander@google.com；c jhuchette@google.com；

arXiv:2602.04048v1 [math.OC] 3 Feb 2026

摘要。 在自主生产系统中部署数学优化，需要为优化求解器返回的对象制定精确的契约。然而，不同求解器和不同类别问题在对偶解及不可行性证书（射线）的约定上差异很大。本文介绍 MathOpt（由 Google 开发并使用的一种领域专用语言）用于统一这些概念的理论框架。我们基于简化的 Fenchel 对偶方案提出一个抽象的原始-对偶对，从而能够以机械化的方式推导对偶问题，以及 MathOpt 当前支持的所有类别问题所对应的契约（包括目标函数为线性或二次、约束为线性、锥、二次或双边线性的情形）。我们还展示了这些契约如何提升某些类别问题中基于互补松弛的最优性条件的清晰度。

1. 引言

数学优化领域专用语言（DSL）提供了一种接口，用户可借此对复杂的决策问题进行建模并求解。DSL 提供易于使用的建模接口，并充当问题高层描述与底层优化求解器之间的桥梁，后者用于生成答案。随着可建模问题的范围扩大，以及可用于生成答案的求解器数量增加，这座桥梁的价值也随之提升。然而，这种增长使抽象和统一化变得更加困难。

MathOpt（Google 2025）是 Google 开发并在其内部广泛使用的开源数学优化领域专用语言（DSL）。由于 MathOpt 广泛应用于自主生产系统，因此它需要提供契约，这些契约应当：(i) 精确且完整；(ii) 便于缺乏深厚理论优化背景的用户遵循；(iii) 适用于单元测试等标准软件工程实践（例如 Kanewala 和 Bieman（2014年）、Winters 等人（2020年）。

凸优化对偶性是一种用途极其广泛的工具，可用于证明优化问题的最优性、无界性和不可行性，也是大多数凸优化求解器的核心。然而，根据所考虑的优化问题类型（例如线性、二次、锥规划等），对偶性有许多可能的表达形式。事实上，即使是线性规划对偶性，也取决于所采用的特定“标准形式”（见第 1.1 节），这意味着即便针对同一类问题，不同的求解器也可能提供不同的对偶性契约。因此，对偶性是 MathOpt 契约中最棘手、最难妥善处理的部分。

本文介绍了 MathOpt 内部开发并使用的对偶框架。该框架采用抽象的原始—对偶优化对形式，其灵感来自近期的相关研究（Rockafellar2023、Roos 等人 2020年）。该框架的主要目标是：

G1 具备精确、机械且模块化的契约：该框架提供清晰、形式化的数学契约，可在代码中高效实现。G2 充分表达、通用且可扩展：该框架为对偶问题的所有性质与用途（例如证明最优性、无界性和不可行性）提供契约，并适用于 MathOpt 当前支持的所有目标函数和约束。G3 归约为现有对偶框架：该框架可通过简单方法归约为标准教科书及求解器特定的对偶性推导。特别是，MathOpt 所支持求解器提供的全部对偶性信息都可通过该框架传达给用户。据我们所知，我们的框架是首个同时提供上述全部能力的对偶框架。

MathOpt 对偶框架的原始问题由下式给出

$$\mathrm{opt}_P := \inf_{x\in\mathbb{R}^J} c\cdot x + \sum_{i\in I} g_i\left(F^i(x)\right) \tag{1}$$

其中 J 和 I 是有限集[1]，$c\in\mathbb{R}^J$ 且对于每个 $i\in I$，M_i 都是有限集，$F^i\colon \mathbb{R}^J\to\mathbb{R}^{M_i}$ 是连续可微函数，而 $g_i\colon \mathbb{R}^{M_i}\to\mathbb{R}\cup\{+\infty\}$ 是真闭凸函数。

MathOpt 对（1）的对偶是对 Fenchel 对偶的简化，它利用了将 F^i 和 g_i 分量拆分（参见第6.1节），其形式如下：

$$\mathrm{opt}_D := \sup_{y\in\mathbb{R}^M, w\in\mathbb{R}^{I\times J}} -\sum_{i\in I} F^i_\nabla\left(-y^i, w^i\right) + g_i^*\left(-y^i\right) \tag{2a}$$

$$s.t.$$

$$\sum_{i\in I} \nabla F^i\left(w^i\right)^T y^i = c, \tag{2b}$$

其中 $M := \bigtimes_{i\in I} M_i$，$\nabla F^i\colon \mathbb{R}^J\to\mathbb{R}^{M_i\times J}$ 是 F^i 的雅可比矩阵，由 $(\nabla F^i(x))_{k,j} = \frac{\partial F^i_k}{\partial x_j}(x)$ 给出，对所有 $k\in M_i$ $j\in J$，$g_i^*\colon \mathbb{R}^{M_i}\to\mathbb{R}\cup\{+\infty\}$ 都成立；g_i 的凸共轭函数由 $g_i^*(\bar{v}) := \sup_{z\in\mathbb{R}^{M_i}}\{\bar{v}\cdot z - g(z)\}$ 给出；$F^i_\nabla\colon \mathbb{R}^{M_i}\times\mathbb{R}^J\to\mathbb{R}\cup\{+\infty\}$ 则由 $F^i_\nabla\left(y^i, w^i\right) := \left(\nabla F^i\left(w^i\right) w^i - F^i\left(w^i\right)\right)\cdot y^i$ 给出。表 1 详细列出了为 MathOpt 支持的所有目标函数和约束构造（1）/（2）所需的函数。

	$g_i(z)$	$F_i(x)$	$\nabla F\left(w^i\right)$	$F^i_\nabla\left(-y^i, w^i\right)$	$g_i^*\left(-y^i\right)$
线性不等式	$\delta_{\mathbb{R}^{M_i}_+}(z)$	$Ax-b$	A	$-b\cdot y^i$	$\delta_{\mathbb{R}^{M_i}_-}\left(-y^i\right)$
锥约束	$\delta_K(z)$	$Ax-b$	A	$-b\cdot y^i$	$\delta_{K^\circ}\left(-y^i\right)$
双边不等式	$\delta_{[l,u]}(z)$	$a\cdot x$	a	0	最大值 $\{-ly^i, -uy^i\}$
二次目标函数	z	x^TQx	$2Qw^i$	$-y^i\left(w^i\right)^TQw^i$	$\delta_{\{1\}}(-y^i)$
二次不等式	$\delta_{\mathbb{R}_-}(z)$	$x^TQx+q\cdot x-u$	$2Qw^i+q$	$-y^i\left(\left(w^i\right)^TQw^i+u\right)$	$\delta_{\mathbb{R}_-}(-y^i)$

表 1 MathOpt 支持的目标函数和约束所对应的函数：线性不等式 $Ax\geq b$、锥约束 $Ax-b\in K$（目前仅支持线性规划和二阶锥），双边线性不等式 $l\leq a\cdot x\leq u$、凸二次目标函数 x^TQx 以及凸二次不等式 $x^TQx+q\cdot x\leq u$，其中 $A\in\mathbb{R}^{M_i\times J}$、$b\in\mathbb{R}^{M_i}$、$K\subseteq\mathbb{R}^{M_i}$ 是闭凸锥，$K^\circ := \{y\in\mathbb{R}^{M_i}\colon y\cdot x\leq 0\ \forall x\in K\}$ 是 K 的极锥，l，$u\in\mathbb{R}$ 满足 $l\leq u$，$Q\in\mathbb{R}^{J\times J}$ 是半正定的，$q\in\mathbb{R}^J$，而 $\delta_S\colon \mathbb{R}^L\to\mathbb{R}\cup\{+\infty\}$ 是闭凸集 $S\subseteq\mathbb{R}^L$ 的示性函数（即，当 $z\in S$ 时，$\delta_S(z) := 0$；否则，$\delta_S(z) := +\infty$）

[1]我们使用抽象集合，而不是诸如 $\{1,\ldots,n\}$ 这样的有序集，以使记号简洁，并与 MathOp 所支持的内容保持一致 t.

关于目标G1，对偶问题（2）显然定义明确且模块化。例如，对于原始问题（（y^i, w^i）中的每个模块 $i \in I$ （1），都有一组变量，且这些变量仅通过方程（2b）相连。只要 ∇F^i 和 g_i^* 均可机械地构造，对偶问题（ 2 ）也可以机械地构造。关于 ∇F^i 的这一假设相当标准；当 F^i为仿射函数或二次函数时，该假设成立。关于 g_i^*的假设则不那么标准，但限制性不强，并且适用于 MathOpt 当前支持的所有类别的目标函数和约束，详见表1。

关于目标 G2，我们显然可以看出，该框架具有通用性和可扩展性，并涵盖了 MathOpt 当前支持的所有目标和约束（见表 1）。假设（1）/（2）满足弱对偶性（我们将在定理 1 中将其形式化），便可推导出对偶解作为最优性证书的契约。我们将在第 1.1 节更正式地描述这些契约及其他契约。

关于目标 G3，我们可以使用表 1 中的函数来恢复标准对偶形式，详见第 5 节和附录 A （第 5 节还包含对表 1中推导的形式化命题）。特别要注意，在（ $-y^i$ 2a）中使用是为了确保符号约定符合锥规划和线性规划对偶性的标准约定（见第 5.1 节和 A.1节）。

最后，关于（1）/（2）的结构，有两点值得注意。首先，表 g_i1 中的所有函数要么是示性函数，要么是平凡的恒等函数。我们没有明确限制函数 g_i 只能属于这两种情况，因为更一般的情况可以稍微简化抽象描述，且不会造成实际负担（前提是 g_i^*具有如上所述的机械构造）。此外，更一般的情况使得表 1 能够扩展到求解器支持的一些非可微目标，例如凸分段线性函数（例如，双边线性不等式 $\{-ly^i, -uy^i\}$的目标中出现的最大值函数（2）[2]）。

第二个观察是，（2）中由于 w^i和 y^i 变量的乘积而显现出的非凸性。这只是将（2）背后更明显凸的问题写成机械描述所产生的假象，在 Wolfe 对偶中（见第 6.2 节）以及写出 Fenchel 对偶的具体形式时（例如，Roos 等人（Roos 等人（2020年））也会出现这种情况。如第 5 节所示，对于表 1 中的所有函数，都可以消除这种表面上的非凸性。

[2]特别是，这 使得 (2) 成为 (1) 的特例，但让原始-对偶对完全对称并非我们的目标 .

1.1. 线性规划中的对偶性契约

对于（1）/（2），我们希望获得的契约是线性规划（LP）在最一般情形下所具备的那些契约，下面将对此作详细说明。我们从基于不等式的“教科书式”线性规划开始，其形式为

$$\mathrm{opt}_P := \min_{x\in\mathbb{R}^J} \quad c\cdot x \tag{3a}$$

$$s.t.$$

$$A^i x \geq b^i, \quad \forall i \in I, \tag{3b}$$

其中 I与 J为有限集，$c \in \mathbb{R}^J$，且对每个 $i \in I$，M_i为一个有限集，$A^i \in \mathbb{R}^{M_i \times J}$，$b^i \in \mathbb{R}^{M_i}$，并且 $\mathbb{R}^{M_i}$ 中元素的不等式按分量取。

对偶（3）由下式给出

$$\mathrm{opt}_D := \max_{y\in\mathbb{R}^M} \quad \sum_{i\in I} b^i \cdot y^i \tag{4a}$$

$$s.t.$$

$$\sum_{i\in I} \left(A^i\right)^T y^i = c, \tag{4b}$$

$$y^i \geq 0 \quad \forall i \in I. \tag{4c}$$

第一类性质与契约与解有关，可规定如下。对于任意原始解 $\bar{x} \in \mathbb{R}^J$ 和对偶解 $\bar{y} \in \mathbb{R}^M$，有如下内容：

- **明确的可行性概念：**满足模约束（3b）对于 $\bar{x}$，并满足模约束（4c）加上简单方程（4b）对于 $\bar{y}$。
- **即使对于不可行解，目标值的概念也十分明确：** $c\cdot\ \bar{x}$ 与 $\sum_{i\in I} b^i \cdot\ \bar{y}^i$。
- **弱对偶性：**[3] $c\cdot\ \bar{x} \geq \sum_{i\in I} b^i \cdot\ \bar{y}^i$ 对于任意可行的 $\bar{x}$ 和 $\bar{y}$。
- **基于模互补松弛（互补松弛）的最优性条件：**如果 $\bar{x}$ 和 $\bar{y}$ 满足

$$\bar{y}^i \cdot (A^i\bar{x} - b^i) = 0 \quad \forall i \in I, \tag{5}$$

那么 $c \cdot \bar{x} = \sum_{i\in I} b^i \cdot \bar{y}^i$ 即使任一解或两个解都不可行；对于可行的 $\bar{x}$ 和 $\bar{y}$，条件（5）等价于两个解均为最优解。

[3]我们没有纳入强对偶性，因为在从求解器获取解时，它的相关性较低（例如，强对偶性使求解器能够返回最优解，但求解器仍可能因数值原因或用户设定的时间限制而失败）。

我们纳入了这样一个事实：即使在缺乏可行性的情况下，互补松弛仍意味着目标值相等；这是因为，当从基于单纯形法的线性规划求解器获得次优基可行解时，这一点具有重要意义。具体而言，为不可行解定义目标值，对于互补松弛这一性质至关重要。同样，低复杂度（在描述和计算成本方面）的可行性概念，对于利用弱对偶性获得原始和对偶界（例如，分别从原始单纯形法和对偶单纯形法的次优解中获得）至关重要。对于线性规划，可行解和不可行解的可行性与目标值概念都很直观，但对于（1）/（2），这些概念则不那么明确。在第 2节和第 3 节中，我们将说明如何把这些概念（以及相关的互补松弛性质）推广到（1）/（2）。在第6.1 节中，我们讨论对偶问题（2）如何降低检查对偶可行性的复杂度，相较于直接描述（1）的 Fenchel 对偶而言。

第二类性质和契约涉及射线，可定义如下：

- **原始射线：** 原始射线 $\hat{x} \in \mathbb{R}^J$ 满足（3b），其中 $b^i = 0$ 对所有 $i \in I$ 以及 $c \cdot \hat{x} < 0$。
- **对偶射线：** 对偶射线 $\hat{y} \in \mathbb{R}^M$ 满足（4b），其中 $c = 0$、（4c）以及 $\sum_{i \in I} b^i \cdot \hat{y}^i > 0$。

- **原问题无界：** （3）若存在原始可行解和原始射线，则称其无界。
- **对偶问题无界：** （4）若存在对偶问题可行解和对偶射线，则称其无界。
- **原问题不可行：** （3）若存在对偶射线，则称其不可行。
- **对偶问题不可行：** (4) 若存在原始射线，则该问题不可行。

射线的第一个作用是与相应问题的一个解一起证明该问题无界。第二个作用是独立证明另一个问题无界。具体而言，两个问题甚至可以同时通过射线证明其不可行性。

现在，线性规划形式（3）只考虑基于 $\geq$ 的不等式，但我们当然也可以考虑基于 $\leq$ 的不等式，或基于 $=$ 的不等式（即方程）。加入这些额外的不等式，只会改变对偶中的符号约束（4c）（对于基于 $\leq$ 的不等式，对偶变量为非正；对于方程，对偶变量的符号不受限制）。不那么显而易见的是，我们还可以考虑上下界不一定相同的不等式（即不一定是方程），我们称之为 双边线性不等式。例如，设 J

和 M 是有限集， $c \in \mathbb{R}^J$、 $A \in \mathbb{R}^{M \times J}$、 $b \in \mathbb{R}^M$，并考虑由下式给出的线性规划

$$\text{opt}_P := \min_{x \in \mathbb{R}^J} \quad c \cdot x \tag{6a}$$

$$s.t.$$

$$Ax = b, \tag{6b}$$

$$l_j \le x_j \le u_j \quad \forall j \in J. \tag{6c}$$

将 (6) 拆分为 $l_j \le x_j$和 $x_j \le u_j$，我们可以将该问题视为包含三个模块及其对应的对偶变量集合：一组符号不受限的变量，对应于 (6b)；另有两组符号受限的变量，对应于 (6c)。然而，针对 (6) 的基于单纯形法的求解器只返回两组对偶变量。第一组对偶变量 $\pi \in \mathbb{R}^M$ 是预期中对应于 (6b) 的无约束对偶变量集合。与 (6c) 关联的第二组对偶变量 $r \in \mathbb{R}^J$ 也不受符号约束，其元素r_j称为变量 x_j的约化成本。我们可以通过单纯形算法语境下自然的规则，推断基可行解中 r_j的符号（例如，若基变量对 $\bar{x}/(\bar{y}, \bar{r})$ 满足 $\bar{x}_j = l_j$，则 $\bar{r}_j \ge 0$）。然而，对于不熟悉单纯形算法的用户而言，这些规则可能不够直观。幸运的是，利用单调规划对偶性（Rockafellar 1981），我们可以构造出 (3) 的如下对偶问题（例如见 Fourer (1994)），由此可见，制定更清晰的契约是可行的。

$$\text{opt}_D := \max_{\pi \in \mathbb{R}^M, r \in \mathbb{R}^J} \quad b \cdot \pi + \sum_{j \in J} \min\{l_j r_j, u_j r_j\} \tag{7a}$$

$$s.t.$$

$$A^T \pi + r = c. \tag{7b}$$

如第 5.2 节所示，(7) 也是 MathOptdual 问题 (2) 的一个特例。在本文中，我们将展示所有这些概念及相关契约如何扩展到我们提出的对偶框架。

1.2. 现有对偶形式

1.2.1. 锥对偶及其扩展 满足 MathOpt 大多数要求的一个对偶框架是锥对偶（Ben-Tal 和 Nemirovski 2001）。其中缺少的一点是，纳入双边线性不等式、二次目标函数和二次约束需要进行重新表述

为锥形式。这些重新表述很简单，并且适用于更广泛的约束类别（例如 Mosek （2020年）），这也是 JuMP 等其他领域特定语言（Lubin 等人 2023）采用锥对偶的原因之一。然而，相关的重新表述至少存在两个潜在的复杂性，这些复杂性通过约束/解“桥接”系统在 JuMP 中得到了一定程度的缓解（Legat 等人 2022）。

第一个复杂之处涉及具有上界和下界的变量，其约化成本的符号不受约束。如第 1.1 节所述，这样的双侧变量界 可以拆分为两个 单侧变量界，但与这种定义相关的符号规则可能令人困惑。幸运的是，现有的线性规划和锥对偶扩展（例如 Fourer (1994)、Karimi 和 Tunçel(2020年、2024年）、Rockafellar (1981)）表明，非锥凸集（例如双侧变量界中的区间和线性不等式）可以直接纳入考量。这些方法符合我们提出的框架（见第 5.2 节和第 6.4节）。

第二个复杂之处涉及二次目标函数和二次约束。虽然相关的重新表述很简单，但由此得到的对偶解可能会让非专业用户感到困惑。例如，将二次目标函数重新表述为锥形式时，会增加一个新的锥约束，而该约束对应的对偶变量可能令人意外。幸运的是，锥对偶也可以扩展到二次目标函数（Vandenberghe 2010年），且这种扩展同样符合我们提出的框架（见第 5.3节）。相比之下，二次约束更为复杂；目前的方法更接近于将拉格朗日对偶扩展到锥约束，而不是将锥对偶扩展到二次约束，因此我们将在第 1.2.2节详细讨论。

另一个缺失之处是可扩展性，即考虑尚无已知锥重构的约束和目标。这包括非线性目标函数和非线性不等式（即定义为非线性函数水平集的约束）。针对这种情况的现有扩展更接近于将拉格朗日对偶扩展到锥约束，因此我们在第1.2.2节中详细讨论它们。

1.2.2. 拉格朗日对偶及其扩展一种极为通用的对偶形式是拉格朗日对偶（例如Boyd and Vandenberghe （2004）），尤其是其考虑非可微函数的扩展。此类扩展的例子包括：针对将基于可微函数的非线性不等式与二阶锥约束相结合的问题，Hours （2018）；以及针对将二次目标函数、二次不等式和锥约束相结合的问题（JuMP-dev 2025）。更一般地，Rockafellar（2023）的广义非线性规划（GNLP）框架详细阐述了一种涵盖所有凸优化问题的扩展。

拉格朗日对偶的一个问题在于，相应的对偶优化问题与原始优化问题的形式并不相同，因为它结合了两种优化方向（例如，如果

原始问题采用最小化方向，则拉格朗日对偶采用最大最小方向）。众所周知，可以消去拉格朗日对偶的其中一个方向，从而得到通常记为 Wolfe 对偶的单一方向版本（Dorn 1960, Wolfe 1961）。例如，这被用于 JuMP-dev （2025），以表明二次目标 + 锥约束的拉格朗日对偶可简化为锥对偶扩展，来自 Vandenberghe （2010年）。然而，正如我们在第6.2节中详述的那样，即便经过这一简化，该对偶仍不具有模块化结构，因为它包含原来的原始变量，而这些变量与所有约束所关联的对偶变量直接相互作用。尽管如此，正如我们同样在第 6.2节中详述的那样，我们提出的框架可被解释为 GNLP 拉格朗日对偶的一种变体，来自 Rockafellar（2023）。因此，我们的框架也涵盖所有凸优化问题。

最后，与之密切相关的是 Artelys (2025) 中讨论的双边线性和非线性约束的最优性条件。如第 6.3 节所述，这些条件可以解释为源自 GNLP 拉格朗日对偶，因此与我们关于互补松弛的最优性条件类似。不过，正如我们在第 6.3 节中也详细说明的那样，Artelys(2025) 中的条件在精确、简洁地描述对偶变量符号规则方面存在类似问题，正如第 1.1 节所指出的，对于具有下界和上界的变量的约化成本也存在这一问题。

1.2.3. Fenchel 对偶 Fenchel 对偶定理可用于推导适用于任意凸优化问题的抽象对偶形式。此外，正如 Roos 等人 (2020年) 所指出的，利用凸函数的凸共轭的微积分规则，通常可以将这些抽象对偶形式具体化。正如第 6.1 节所述，这种对偶形式唯一的问题在于，只有将其化简为适用于特定问题类别的具体形式后，才能机械地构造并详细描述其性质。正如我们在第 6.1 节中也详细说明的那样，我们提出的框架可以解释为对 Roos 等人 (2020年) 所述对偶的抽象版本与具体版本进行中间简化，从而能够以通用但机械的方式描述我们所关注的所有性质。

1.3. 贡献与概述

MathOpt 对偶框架的首项贡献，是简化标准 Fenchel 对偶框架，从而能够为更广泛的约束提供机械契约。如第6.1节所述，这种简化是通过利用原始问题（Rockafellar (2023）中的复合结构实现的。这里的关键观察是，（1）的 Fenchel 对偶包含函数$(g_i \circ F^i)^*$（其中$g_i \circ F^i(x) = g_i(F^i(x))$），很难将其机械地描述为$g_i \circ F^i$的函数。不过，我们证明，可以将$(g_i \circ F^i)^*$的一个恰当的紧逼近机械地描述为 ∇F^i 和 g_i^*的函数。

第二项贡献，是将对偶（2）的机械构造扩展为契约，以涵盖第1.1节所述对偶性的所有相关用途。尤其是，对于某些类别的约束，这些契约可以显著提升基于互补松弛的最优性条件的清晰度（例如，见第6.3节）。

本文其余部分安排如下。首先，我们在第 1.4 节介绍一些符号。在第 2 节中，我们给出原始解的定义和契约，并详细说明本文结果所依据的一些基准假设。随后，在第 3 节中，我们给出对偶解的定义和契约。具体而言，定理1为 (1)/(2) 确立了弱对偶性，推论1则为 (1)/(2) 确立了基于互补松弛的最优性条件。第 4 节介绍原始射线和对偶射线的契约，以及它们作为无界性和不可行性证书的作用。第 2、3 和 4 节给出通用或抽象的描述，因此我们在第 5 节针对 MathOpt 当前支持的目标和约束，将所有内容具体化。特别是，第 5 节说明，在适用时，所有相关契约都可机械地导出，并与标准对偶形式相符。最后，第 6 节将本文方法与其他对偶框架进行比较，尤其是 Rockafellar (2023) 和 Roos等人 (2020) 提出的框架。所有证明均放在第 7 节。为简洁起见，本文始终讨论原始最小化问题。原始最大化问题所需的相应修改简单且同样可机械完成，详见附录 A。

1.4. 符号约定

对于任意有限集 L 和 $S \subseteq \mathbb{R}^L$，我们令 $\delta_S \colon \mathbb{R}^L \to \mathbb{R} \cup \{+\infty\}$ 表示 S 的示性函数（即 $\delta_S(z) := 0$ 当 $z \in S$ 时成立，否则 $\delta_S(z) := +\infty$ ），并令 $\sigma_S \colon \mathbb{R}^L \to \mathbb{R} \cup \{+\infty\}$ 表示 S 的支撑函数（即 $\sigma_S(d) := \sup\{d \cdot z \colon z \in S\}$）。对于任意 $\mathrm{h} : \mathbb{R}^L \to \mathbb{R} \cup \{+\infty\}$，我们令 $\mathrm{dom(h)} := \{z \in \mathbb{R}^L \colon \mathrm{h}(z) < +\infty\}$ 表示 h 的定义域，$\mathrm{rge(h)} := \{\mathrm{h}(z) \in \mathbb{R} \cup \{+\infty\} : z \in \mathbb{R}^L\}$ 表示 h 的值域。最后，对于任意 $F^i \colon \mathbb{R}^J \to \mathbb{R}^{M_i}$ 和 $g_i \colon \mathbb{R}^{M_i} \to \mathbb{R} \cup \{+\infty\}$，我们令 $g_i \circ F^i \colon \mathbb{R}^J \to \mathbb{R} \cup \{+\infty\}$ 表示由 $g_i \circ F^i(x) = g_i\left(F^i(x)\right)$ 给出的函数。

2. 原始问题的定义与假设

由于函数 g_i 可以取值 $+\infty$，因此它们隐含地编码了（1）中的约束。

定义 1（原始 可行性）。若原始解 $\bar{x} \in \mathbb{R}^J$ 满足 可行条件 $g_i(F^i(\bar{x})) < +\infty \forall i \in I$，则称其为可行解。原始问题（1）可行，当且仅当它存在一个可行的原始解。

当唯一取无穷值的函数是示性函数（如表 1所示）时，可以如下定义一个在不可行解处也取有限值的目标函数。

D定义 2（P原始目标值）。对于任意 $h:\ \mathbb{R}^L \to \mathbb{R}\cup\{+\infty\}$，若 h 是示性函数，则令 $h^0:\ \mathbb{R}^L \to \mathbb{R}\cup\{+\infty\}$ 为恒等于零的函数；否则令其为 h。也就是说

$$h^0(x) := \begin{cases} 0 & \mathrm{rge}(h) \subseteq \{0, +\infty\} \\ h(x) & \text{o.w.} \end{cases}$$

任意（可行或不可行）原始解的目标值 $\bar{x} \in \mathbb{R}^J$ 为

$$\mathrm{obj_P}(\bar{x}) := c \cdot \bar{x} + \sum_{i \in I} g_i^0\left(F^i\left(\bar{x}\right)\right). \tag{8}$$

最后，本文始终对函数 g_i 和 F^i作出以下假设。

假设 1。对于每个 $i \in I$,

$$\exists \bar{x}^i \in \mathbb{R}^J \ s.t.\ g_i\left(F^i\left(\bar{x}^i\right)\right) < +\infty, \ and \tag{9a}$$

$$F^i \ is\ continuously\ differentiable, \tag{9b}$$

$$g_i \ is\ a\ proper,\ closed\ and\ convex\ function, \tag{9c}$$

$$g_i\left(F^i(x) + u\right) \ is\ jointly\ convex\ in\ (x, u). \tag{9d}$$

假设（9a）排除了会独立导致不可行性的函数。表 1 中的函数始终满足其余假设。然而，需要注意的是，尽管（9d）对于凸优化问题通常成立，但它比 $g_i \circ F^i$ [4] 的凸性略强。

3. 对偶问题的定义和性质

（2）中的唯一显式约束是方程（2b），通常称为驻点条件。然而，由于目标函数（2a）中的函数 g_i^* 也可以取值 $+\infty$，因此它们隐含地为（2）编码了额外的约束。遵循 KKT 条件中的惯例，我们将这些额外的模块级约束与联合驻点条件（2b）部分分离（即，在考虑单个 $i \in I$的对偶可行性时，不包括驻点条件）。

D定义 3（D对偶 F可行性）。对于每个 $i \in I$ a（逐模块）对偶解（$\bar{y}^i,\ \bar{w}^i$）$\in \mathbb{R}^{M_i} \times \mathbb{R}^J$，若且唯若 $g_i^*(-\bar{y}^i) < +\infty$，则其（逐模块）可行。

对偶解 （$\bar{y},\ \bar{w}$）$\in \mathbb{R}^M \times \mathbb{R}^{I\times J}$ 在且仅在 可行当 （$\bar{y}^i,\ \bar{w}^i$）对每个 $i \in I$ 且（$\bar{y},\ \bar{w}$）满足驻点条件（2b）时可行。

对偶问题（2）可行，当且仅当它存在一个可行的对偶解。

[4] 例如，若 $F^i(x) = x^2$ 和 $g_i(z) = \delta_{[0,1]}$，则 $g_i \circ F^i$ 是凸的，但（9d）不成立

我们还定义了一个对偶目标函数：当解仅因示性函数 g_i^* 而不可行时，该对偶目标函数 (2a)取有限值。

定义 4（对偶目标函数与值）。令 $\mathrm{O}_{F^i,g_i}\left(y^i,w^i\right):=F_\nabla^i\left(y^i,w^i\right)+g_i^*\left(y^i\right)$，则对偶目标函数（2a）等于 $-\sum_{i\in I}\mathrm{O}_{F^i,g_i}(-y^i,w^i)$。

对偶目标函数值（无论可行与否）对于任意对偶解（ $\bar{y},\bar{w}$） $\in\mathbb{R}^M\times\mathbb{R}^{I\times J}$ 为 s

$$\mathrm{obj_D}(\bar{y},\bar{w}):=-\sum_{i\in I}O^0_{F^i,g_i}(-\bar{y}^i,\bar{w}^i), \tag{10}$$

其中 $\overset{0}{\mathrm{O}}_{F^i,g_i}(\bar{y}^i,\bar{w}^i):=F_\nabla^i\left(\bar{y}^i,\bar{w}^i\right)+(g^*)^0\left(\bar{y}^i\right)$。

即使没有 除（9）之外的约束品性条件，我们仍有以下弱对偶性结果。

定理 1 **(弱对偶性)**。若 $\bar{x}\in\mathbb{R}^J$ 且 $(\bar{y},\bar{w})\in\mathbb{R}^M\times\mathbb{R}^{I\times J}$ 满足以下条件：$(\bar{y},\bar{w})$ 满足驻点条件 (2b)，则

$$c\cdot\bar{x}+\sum_{i\in I}g_i\left(F^i\left(\bar{x}\right)\right)\geq-\sum_{i\in I}O_{F^i,g_i}(-\bar{y}^i,\bar{w}^i). \tag{11}$$

特别地，若 $\bar{x}$ 和 $(\bar{y},\bar{w})$ 均可行，则

$$\mathrm{obj_P}(\bar{x})\geq\mathrm{obj_D}(\bar{y},\bar{w}). \tag{12}$$

⟶ 证明见第27 页。

要保证存在满足弱对偶不等式取等的原始-对偶对，确实需要额外的约束资格条件。然而，以下按模块的最优性条件不需要这类约束资格条件（见下文第一条观察）。

定理 2 **(最优性条件)**。设 $\bar{x}\in\mathbb{R}^J$ 且 $(\bar{y},\bar{w})\in\mathbb{R}^M\times\mathbb{R}^{I\times J}$ 满足以下条件：$(\bar{y},\bar{w})$满足驻点条件 (2b)。则以下条件等价

- 原始-对偶对$(\bar{x},(\bar{y},\bar{w}))$ 是最优的，其中 $c\cdot\bar{x}+\sum_{i\in I}g_i\left(F^i\left(\bar{x}\right)\right)=-\sum_{i\in I}O_{F^i,g_i}(-\bar{y}^i,\bar{w}^i)$。
- 对于所有 $i\in I$，

$$g_i\left(F^i\left(\bar{x}\right)\right)+O_{F^i,g_i}(-\bar{y}^i,\bar{w}^i)=-\left(\bar{y}^i\right)^T\nabla F^i\left(\bar{w}^i\right)\bar{x}. \tag{13}$$

此外，如果$(\bar{x},(\bar{y},\bar{w}))$ 是最优的，那么$(\bar{x},(\bar{y},(\bar{x})_{i\in I}))$ 也同样是最优的。

⟶ 证明见第 27 页。

关于定理 2，有三个重要的观察。首先，定理 2 并非强对偶定理，因为它给出的是原始-对偶对的最优性条件，而非原始问题的最优性条件

解。要确保原始最优解能够补全为原始-对偶对，确实需要额外的约束品性条件；这超出了本文的讨论范围[5]。

第二个观察涉及数值 $\bar{w}^i$。定理 2的最后一个结论$\bar{w}^i = \bar{x}$ 表明，对于任意 $i \in I$ 最优解，我们都可以取 w^i 。因此，我们可以放心地将对偶问题（2）中的所有变量替换为一组变量 $w \in \mathbb{R}^J$。如第 5.4 节和第 6.2节所示，对所有约束施加仅使用一组对偶变量 w 的限制，会得到一种更接近拉格朗日对偶的对偶形式。我们保留按模副本有两个原因。首先，这为对偶问题增加了灵活性（例如，如果我们碰巧从一个具有不相等 $\bar{w}^i$ 的解中获得了良好的对偶界，就无需寻找共同的 $\bar{w}$）。其次，这保留了对偶变量按模块分离的形式，这种形式常见于锥对偶和 Fenchel 对偶中，其中每个 $i \in I$ 的变量仅由一个平稳方程联系起来（例如，见第 5.1 节和第 6.1节）。

第三个观察是，条件（13）总是蕴含原始解 $\bar{x}$ 与按模块对偶解$(\bar{y}^i, \bar{w}^i)$的可行性。当这些可行性仅依赖于示性函数时，我们可以将它们从（13）中移除，从而得到如下互补松弛条件。

定义 5（互补松弛）。我们称 $\bar{x} \in \mathbb{R}^J$ 与（$\bar{y}$， $\bar{w}$） $\in \mathbb{R}^M \times \mathbb{R}^{I \times J}$（可行或不可行）满足互补松弛，对 $i \in I$ 当且仅当

$$g_i^0\left(F^i\left(\bar{x}\right)\right) + O^0_{F^i, g_i}\left(-\bar{y}^i, \bar{w}^i\right) = -\left(\bar{y}^i\right)^T \nabla F^i\left(\bar{w}^i\right) \bar{x}. \tag{14}$$

如第5 节与第6.2节所示，（14）对特定函数类可化为常见条件。此外，我们在一般情况下恢复如下常见最优性条件。

推论 1. 一对（$\bar{x}$，（$\bar{y}$， $\bar{w}$））$\in \mathbb{R}^J \times \left(\mathbb{R}^M \times \mathbb{R}^{I \times J}\right)$ 为最优，当且仅当

1. $\bar{x}$ 是原始可行的，2. $(\bar{y}^i$， $\bar{w}^i)$ 对于每个 $i \in I$，均按模块对偶可行3. $(\bar{x}$，$(\bar{y}^i$， $\bar{w}^i))$ 对所有 (14) 均满足互补松弛条件 $i \in I$，4. $(\bar{y}$， $\bar{w})$ 满足驻点条件 (2b)。

⟶ 证明见第27 页。

最后，即使对于不可行解，互补松弛条件（14）和驻点条件（2b）也意味着忽略示性函数的目标值（8）与（10）相等。特别是，通过这一结果，我们可以得出线性规划中互补解目标值相等这一熟悉性质，详见第1.1节。

[5]这里我们重点讨论复原 ering求解器返回的解，而不是使求解器能够返回此类解的条件 .

引理*1.* 设 $\bar{x} \in \mathbb{R}^J$ 与*(* $\bar{y}, \bar{w}$*)* $\in \mathbb{R}^M \times \mathbb{R}^{I\times J}$ （可行或不可行）使得*(* $\bar{y}, \bar{w}$*)*满足*(2b)*，且对所有 $i \in I$，*(* $\bar{x}$*, (* $\bar{y}, \bar{w}$*))*满足*(14)*。则obj_{P}*(* $\bar{x}$*)* $= obj_{\mathrm{D}}$*(* $\bar{y}, \bar{w}$*)*。

⟶ 证明见第 28 页。

4. 射线与不可行性证书

我们使用以下定义，以简洁地描述射线和不可行性证书。

定义 6. 对任意有限集 L 和真闭凸函数 $h: \mathbb{R}^L \to \mathbb{R} \cup \{+\infty\}$，渐近函数或地平线函数 h 记作 $h_\infty: \mathbb{R}^L \to \mathbb{R} \cup \{+\infty\}$，定义为

$$h_\infty\left(\bar{d}\right) := \sup_{t>0} \frac{h\left(\bar{z} + t\bar{d}\right) - h\left(\bar{z}\right)}{t},$$

其中 $\bar{z} \in \mathbb{R}^L$ 是任意一点，满足$h(\bar{z}) < +\infty$。

问题 (1)/(2) 的原始/对偶射线概念可定义如下。

定义 7. 对于 原始射线 (1) 是 $\hat{} x \in \mathbb{R}^J$，满足

$$c \cdot \hat{x} + \sum_{i\in I} (g_i \circ F^i)_\infty(\hat{x}) < 0. \tag{15}$$

对于 对偶射线 (2) 是 $(\hat{} y, \hat{} w) \in \mathbb{R}^M \times \mathbb{R}^{I\times J}$，满足

$$\sum_{i\in I} \nabla F^i(\hat{w}^i)^T \hat{y}^i = 0 \tag{16a}$$

$$-\sum_{i\in I} F^i_\nabla(-\hat{y}^i, \hat{w}^i) + (g_i^*)_\infty\left(-\hat{y}^i\right) > 0. \tag{16b}$$

如预期所示，只要存在可行解，射线就意味着相应问题无界（即第 1.1 节中针对线性规划描述的射线的第一个性质）。

定理 3 （**无界性条件**）。 如果原始问题有一个可行解 $\bar{x} \in \mathbb{R}^J$和一条原始射线$\hat{} x \in \mathbb{R}^J$，那么原始问题无界。

如果对偶问题有一个可行解$(\bar{y}, \bar{w}) \in \mathbb{R}^M \times \mathbb{R}^{I\times J}$和一条对偶射线$(\hat{} y, \hat{} w) \in \mathbb{R}^M \times \mathbb{R}^{I\times J}$，满足

$$\nabla F^i\left(\bar{w}^i\right)^T \hat{y}^i = \nabla F^i\left(\hat{w}^i\right)^T \hat{y}^i \quad \forall i \in I \tag{17}$$

那么对偶问题无界。

⟶ 请参阅第 28页了解证明。

根据定理 1，原始可行解和原始射线的存在意味着对偶问题不可行，反之亦然。不过，即使不存在可行解（即第 1.1 节中线性规划射线的第二种作用），某个问题存在射线也意味着另一个问题不可行。

定理 4 （**不可行性证书**）。如果存在原始射线，则对偶问题不可行；如果存在对偶射线，则原始问题不可行。

⟶ 证明见第 28 页。

5. 特定约束

5.1. 锥约束（包括单边线性不等式）

目前，MathOpt 仅支持非负正交象限的锥约束（即单边线性不等式），以及二阶锥（SOC）或锥二次锥（Ben-Tal 和 Nemirovski 2001，Mosek 2020年）。不过，我们的结果适用于任意闭凸锥。

命题 *1.* 设 $K_i \subseteq \mathbb{R}^{M_i}$ 是一个闭凸锥，$K_i^\circ := \{y \in \mathbb{R}^{M_i} : \ y \cdot x \leq 0 \quad \forall x \in K_i\}$ 是 K_i的极锥，$F^i : \mathbb{R}^J \to \mathbb{R}^{M_i}$ 满足：对于 $A^i \in \mathbb{R}^{M_i \times J}$ 和 $b^i \in \mathbb{R}^{M_i}$，$F^i(x) = A^i x - b^i$，且$g_i : \mathbb{R}^{M_i} \to \mathbb{R} \cup \{+\infty\}$ 满足 $g_i = \delta_{K_i}$。若 $g_i \circ F^i$ 满足 *(9a)*，则 $g_i^0 \equiv 0$, $g_i^* = (g_i^*)^\infty = \delta_{K_i^\circ}$、$(g_i^*)^0 \equiv 0$、$\nabla F^i\left(w^i\right) = A^i$、$F_\nabla^i\left(y^i, \ w^i\right) = b^i \cdot y^i$，且 $\left(g_i \circ F^i\right)_\infty(\hat{x}) = \delta_{K_i}\left(A^i \hat{x}\right)$。

⟶ 请参见第 29 页的证明。

为了说明命题 1的用法，注意，原始问题（1）中所有的 F^i 和 g_i 均如命题1 所示，因此可化简为

$$\begin{aligned} \text{opt}_P := \inf_{x \in \mathbb{R}^J} \quad & c \cdot x \\ s.t. \quad & \\ & A^i x - b^i \in K_i \quad \forall i \in I \end{aligned}$$

相应的对偶问题（2）可化简为

$$\operatorname{opt}_D := \sup_{y\in\mathbb{R}^M} \sum_{i\in I} b^i \cdot y^i \tag{18a}$$

$$s.t.$$

$$\sum_{i\in I} (A^i)^T y^i = c, \quad -y^i \in K_i^\circ \quad \forall i \in I. \tag{18b}$$

如果令 $K_i^* := -K_i^\circ = \{y \in \mathbb{R}^{M_i} : \quad y \cdot x \geq 0 \quad \forall x \in K_i\}$ 为 K_i 的对偶锥，我们可以在（18）中识别出标准锥对偶（即，将（18）中的 $-y^i \in K_i^\circ$替换为 $y^i \in K_i^*$）。

正如预期，由于 F^i 是仿射函数，变量 w^i 不会出现在（18 ）.

原始可行性变为标准锥原始可行性 $A^i x - b^i \in K_i$ 对所有 $i \in I$均成立；对偶可行性变为标准锥对偶可行性 $\sum_{i\in I} (A^i)^T y^i = c$ 以及 $y^i \in K_i^*$对所有 $i \in I$均成立。互补松弛是标准锥互补松弛 $\bar{y}^i \cdot (A^i \bar{x} - b^i) = 0$ 对所有$i \in I$均成立。此外，原始射线定义（15）变为标准锥原始射线条件

$$A^i \hat{x} \in K_i \quad \forall i \in I \quad \text{and} \quad c \cdot \hat{x} < 0,$$

以及对偶射线定义（16）变为标准锥对偶射线条件

$$\sum_{i\in I} (A^i)^T \hat{y}^i = 0, \quad \hat{y}^i \in K_i^* \quad \forall i \in I$$
$$\sum_{i\in I} b^i \cdot \hat{y}^i > 0.$$

同样，我们可以看到，变量 w^i 也不是描述对偶射线所必需的。

5.2. 双边（或区间）线性不等式

命题 *2.* 设 $a^i \in \mathbb{R}^J$, $l_i \in \mathbb{R}$ 和 $u_i \in \mathbb{R}$ 满足 $l_i \leq u_i$, $F^i : \mathbb{R}^J \to \mathbb{R}$ 满足$F^i(x) = a^i \cdot x$，且 $g_i : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ 满足 $g_i = \delta_{[l_i,u_i]}$。则 $g_i^0 \equiv 0$, $g_i^*(y_i) = (g_i^*)^\infty(y_i) = (g_i^*)^0(y_i) =$最大值$\{l_i y_i, u_i y_i\}$, $\nabla F^i(w^i) = a^i$, $F_\nabla^i \equiv 0$，且 $(g_i \circ F^i)_\infty(\hat{x}) = \delta_{\{0\}}(a^i \cdot \hat{x})$。

⟶ 证明见第30页。

为说明命题2的用法，请注意，原始问题（1）在所有 F^i 与 g_i 均如命题2 中所述时可化为

$$\mathrm{opt}_P := \inf_{x\in\mathbb{R}^J} \quad c\cdot x \tag{19a}$$

$$s.t.$$

$$l_i \le a^i\cdot x \le u_i \quad \forall i\in I \tag{19b}$$

以及相应的对偶问题（2）化简为

$$\mathrm{opt}_D := \sup_{y\in\mathbb{R}^M} \quad \sum_{i\in I}\min\{l_i y_i, u_i y_i\} \tag{20a}$$

$$s.t.$$

$$\sum_{i\in I} a^i y_i = c. \tag{20b}$$

正如预期，由于 F^i 是线性函数，变量 w^i 不会出现在（20）中。我们还可以看到，具有变量下界和上界的线性规划的对偶问题（7）是（2）的一个特例。具体而言，当（20）中不存在（7）的等式约束时，（7）是（20）的特例。

原始可行性变为 $l_i \le a^i\cdot x \le u_i$ 对所有 $i\in I$ 均成立，所有逐模块对偶解 $\bar{y}_i$ 均可行，而完整对偶解 $\bar{y}$ 的可行性则简化为满足驻点条件 $\sum_{i\in I} a^i y_i = c$。互补松弛变为

$$\min\{l_i\bar{y}_i, u_i\bar{y}_i\} = \bar{y}_i a^i\cdot\bar{x} \quad \forall i\in I. \tag{21}$$

最后，原始射线定义（15）变为

$$a^i\cdot\hat{x} = 0 \quad \forall i\in I \quad \text{and} \quad c\cdot\hat{x} < 0, \tag{22}$$

而对偶射线定义（16）变为

$$\sum_{i\in I} a^i\hat{y}_i = 0 \quad \text{and} \quad \sum_{i\in I}\min\{l_i\hat{y}_i, u_i\hat{y}_i\} > 0. \tag{23}$$

再次可以看出，描述对偶射线也不需要变量 w^i .

5.3. 二次目标

在表 1 中，我们用矩阵记法描述了（二次齐次）二次函数。不过，这种矩阵记法尚未完全标准化（例如，只有当函数乘以 /2 时，与之关联的矩阵才对应于该二次函数的海森矩阵）。同样，我们也希望契约在采用非矩阵记法约定时仍然适用。幸运的是，利用二次函数以下简单的性质即可解决这些问题。

定义 8. 若 $\mathrm{h}:\mathbb{R}^J\to\mathbb{R}$ 是凸二次型，则称 h 为凸二次型，即它是一个二次的凸齐次多项式（也就是说，它是一个非零项次数全为 2 的凸多项式）。

引理 2. 设 $h:\mathbb{R}^J\to\mathbb{R}$ 是一个凸二次型。那么，对于任意 $w,w'\in\mathbb{R}^J$，有 $h(w)=\frac{1}{2}w^T\nabla^2h(w')w$, $\nabla h(w)=\nabla^2h(w')w$，并且 $\nabla h(w)\cdot w=2h_i(w)$。

此外，$\nabla h(w)$ 对 h 的矩阵或非矩阵描述有一个简单的表述：

- 若 $h(w)=w^TQw$ 对于 $Q\in\mathbb{R}^{J\times J}$，则 $\nabla h(w)=2Qw$。
- 若 $h(w)=(1/2)w^TQw$ 对 $Q\in\mathbb{R}^{J\times J}$ 成立，则 $\nabla h(w)=Qw$。
- 若 $h(w)=\sum_{(j,j')\in K}q_{j,j'}w_jw_{j'}$ 对于 $K:=\{(i,j')\in J\times J:\ i\le j'\}$ 和 $q_{j,j'}\in\mathbb{R}$ 对所有 $(j,j')\in K$ 均成立，则 $(\nabla h(w))_j=q_{j,j'}+\sum_{j'\in J}q_{\min\{j,j'\},\max\{j,j'\}}w_{j'}$ 对所有 $j\in J$ 均成立。

感谢引理 2，我们可以用 h 和 ∇h 抽象地描述二次目标函数（及约束）的契约，之后再针对特定的符号约定将其具体化。

命题 3. 设 $h_i:\mathbb{R}^J\to\mathbb{R}$ 是一个凸二次型，$F^i=h_i$ 和 $g_i:\mathbb{R}\to\mathbb{R}\cup\{+\infty\}$ 满足 $g_i(z)=z$。则 $g_i^0(z)=g_i(z)$, $g_i^*=\delta_{\{1\}}$, $(g_i^*)^0\equiv 0$, $(g_i^*)^\infty=\delta_{\{0\}}$, $\nabla F^i=\nabla h_i$, $F^i_\nabla(y_i,w^i)=y_ih_i(w^i)$，且 $(g_i\circ F^i)_\infty(\hat{x})=\delta_{\{0\}}(\nabla h_i(\hat{x}))$。

⟶ 参见第30 页的证明。

为说明命题3的用法，注意，1）中的原始问题，对于 $I=\{i_0\}$ 以及 F^{i_0} 和 g_{i_0}，如命题 3 所示，可化简为

$$\mathrm{opt}_P:=\inf_{x\in\mathbb{R}^J}c\cdot x+h_{i_0}(x)=\inf_{x\in\mathbb{R}^J}c\cdot x+g_{i_0}\left(F^{i_0}(x)\right) \tag{24}$$

且相应的对偶问题（2）可化简为

$$\mathrm{opt}_D:=\sup_{\substack{y_{i_0}\in\mathbb{R},\\ w^{i_0}\in\mathbb{R}^J}} y_{i_0}h_{i_0}\left(w^{i_0}\right) \tag{25a}$$

$$s.t.$$

$$y_{i_0}\nabla h\left(w^{i_0}\right)=c,\quad -y_{i_0}=1. \tag{25b}$$

由于 F^i 不是仿射函数，变量 w^i 确实出现在（25）中。

所有原始解都是可行的（因为原始问题是无约束的线性加二次最小化），对偶可行性为 $-\bar{y}_{i_0} = 1$，互补松弛为

$$h_{i_0}(\bar{x}) - \bar{y}_{i_0} h_{i_0}\left(\bar{w}^{i_0}\right) = -\bar{y}_{i_0} \nabla h_{i_0}\left(\bar{w}^{i_0}\right) \cdot \bar{x}. \tag{26}$$

此外，原始射线定义（15）变为 $\nabla h_{i_0}(\hat{x}) = 0$ 以及 $c \cdot \hat{x} < 0$。类似地，对偶射线定义（16）变为

$$\hat{y}_{i_0} \nabla h\left(\hat{w}^{i_0}\right) = 0, \quad \hat{y}_{i_0} = 0, \quad \hat{y}_{i_0} h_{i_0}\left(\hat{w}^{i_0}\right) > 0.$$

特别地，我们可以验证，对偶射线并不存在，这与问题无约束时所预期的一致。

正如引理1所述，我们可以验证，任何对 $(\bar{x}, (\bar{y}, \bar{w}))$ 只要满足（26）和驻点条件 $\bar{y}_{i_0} \nabla h_{i_0}(\bar{w}^{i_0}) = c$，就有 $\mathrm{obj}_P(\bar{x}) = c \cdot \bar{x} + h(\bar{x}) = \bar{y}_{i_0} h_{i_0}\left(\bar{w}^{i_0}\right) = \mathrm{obj}_D(\bar{y}, \bar{w})$，即使 $\bar{y}$ 不满足对偶可行性约束 $-\bar{y}_{i_0} = 1$。然而，如果允许 y 为（先验）未固定变量，则（25b）的目标函数和约束中包含非凸函数。

为了得到一个明显凸的对偶问题重新表述，我们可以利用对偶可行性方程 y_{i_0} 消去变量 $-y_{i_0} = 1$，从而得到更常见的对偶问题（例如参见（Vandenberghe2010，(3)）以及（Roos 等人 2020，第6.3节）），如下所示：

$$\mathrm{opt}_D := \sup_{w^{i_0} \in \mathbb{R}^J} \quad -h_{i_0}\left(w^{i_0}\right) \tag{27a}$$

$$s.t.$$

$$0 = c + \nabla h_{i_0}\left(w^{i_0}\right) \tag{27b}$$

这种重新表述不需要变量 y_{i_0}，且其所有对偶解都逐模块可行（即对偶可行性可归结为平稳方程）。此外，互补松弛条件（26）变为 $h_{i_0}(\bar{x}) + h_{i_0}\left(\bar{w}^{i_0}\right) = \nabla h_{i_0}\left(\bar{w}^{i_0}\right) \cdot \bar{x}$，根据引理2，这等价于

$$\left(\nabla h_{i_0}(\bar{x}) - \nabla h_{i_0}\left(\bar{w}^{i_0}\right)\right) \cdot \left(\bar{x} - \bar{w}^{i_0}\right) = 0. \tag{28}$$

特别是，当满足 28 时，互补松弛条件 $\bar{x} = \bar{w}^{i_0}$；此时，驻点条件就变为线性加二次无约束最小化的标准最优性条件，其形式为 $c = -\nabla h_{i_0}(\bar{x})$。然而，$\bar{x}$ 与 $\bar{w}^{i_0}$ 之间的互补松弛关系也可以通过 $\nabla h_{i_0}(\bar{x}) = \nabla h_{i_0}\left(\bar{w}^{i_0}\right)$ 来实现；当 $\bar{x} = \bar{w}^{i_0}$ 为 $\nabla^2 h$ 奇异时，这种方式可能较弱。

5.4. 二次约束

命题 4. 设 $h_i\colon \mathbb{R}^J \to \mathbb{R}$ 是一个凸二次型，$q^i \in \mathbb{R}^J$，$u_i \in \mathbb{R}$，$F^i(x) = h_i(x) + q^i \cdot x - u_i$ 且 $g_i = \delta_{\mathbb{R}_-}$。那么，$g_i^0 \equiv 0$，$g_i^* = (g_i^*)^\infty = \delta_{\mathbb{R}_+}$，$(g_i^*)^0 \equiv 0$，$\nabla F^i(w^i) = \nabla h_i(w^i) + q^i, F_\nabla^i(y^i, w^i) = (h_i(w^i) + u) y_i$，且 $(g_i \circ F^i)_\infty = \delta_{\mathcal{D}}$ 对于 $D := \{d \in \mathbb{R}^J\colon \nabla h_i(d) = 0, \quad q^i \cdot d \le 0\}$。

⟶ 证明见第 30 页。

为说明命题 4 的用法，注意到命题 1 中关于 $I = \{i_0\}$、F^{i_0} 和 g_{i_0} 的原始问题（4）可化简为

$$\begin{aligned} \operatorname{opt}_P := \inf_{x \in \mathbb{R}^J} \quad & c \cdot x \\ s.t. \quad & \\ & h_{i_0}(x) + q^{i_0} \cdot x \le u_{i_0} \end{aligned}$$

相应的对偶问题（2）可化简为

$$\operatorname{opt}_D := \sup_{y_{i_0} \in \mathbb{R}, w^{i_0} \in \mathbb{R}^J} \left(h_{i_0}\left(w^{i_0}\right) + u_{i_0}\right) y_{i_0} \tag{29a}$$

$$s.t. \quad \left(\nabla h_{i_0}\left(w^{i_0}\right) + q^{i_0}\right) y_{i_0} = c, \quad y_{i_0} \le 0. \tag{29b}$$

由于 F^i 不是仿射函数，因此变量 w^i 确实出现在（29）中。

原始可行性变为 $h_{i_0}(\bar{x}) + q^{i_0} \cdot \bar{x} \le u_{i_0}$，对偶可行性为 $\bar{y}_{i_0} \le 0$ 加上 $y_{i_0}\left(\nabla h_{i_0}\left(w^{i_0}\right) + q^{i_0}\right) = c$，利用引理 2 可将互补松弛写为

$$\begin{aligned} & \bar{y}_{i_0}\left(h\left(\bar{w}^{i_0}\right) + q^{i_0} \cdot \bar{w}^{i_0} - u_{i_0}\right) \\ & \qquad = \bar{y}_{i_0}\left(\nabla h\left(\bar{w}^{i_0}\right) + q^{i_0}\right) \cdot \left(\bar{w}^{i_0} - \bar{x}\right). \end{aligned} \tag{30}$$

此外，原始射线定义（15）变为

$$\nabla h_{i_0}(\hat{x}) = 0, \quad q^{i_0} \cdot \hat{x} \le 0 \quad \text{and} \quad c \cdot \hat{x} < 0, \tag{31}$$

而对偶射线定义（16）变为 $\hat{}y_{i_0}\left(\nabla h_{i_0}\left(\hat{}y^{i_0}\right)+q^{i_0}\right)=0,\ \hat{}y_{i_0}\le 0$ 和 $\left(h_{i_0}\left(\hat{}y^{i_0}\right)+u_{i_0}\right)\hat{}y_{i_0}>0$。

注意，互补松弛条件（30）在 $\bar{w}^{i_0}=\bar{x}$ 时成为标准拉格朗日对偶的互补松弛条件。然而，对偶可行性 $\bar{y}_{i_0}\le 0$ 是标准非负拉格朗日乘子的否定。这是我们与标准锥规划及线性规划对偶性的符号约定保持一致所带来的副作用（见第 5.1 节和 A.1），也源于这些约定与拉格朗日乘子约定并不完全一致（例如，见（Boyd and Vandenberghe 2004，第5.2.1节））。

现在，与第 5.3 节中的二次目标类似，对偶问题（29）看似非凸，因而需要凸重构。如 Roos 等人（2020年）所指出的，这可以通过变量替换 $z^{i_0}=y_{i_0}w^{i_0}$ 以及与 h_{i_0} 相关联的如下函数来完成。

命题 5. 设 $h_i:\mathbb{R}^J\to\mathbb{R}$ 是一个凸二次型，且 $\mathcal{G}_{h^i}:\mathbb{R}^J\times\mathbb{R}\to\mathbb{R}\cup\{+\infty\}$ 是 h_i 的透视函数的闭包，定义如下：

$$\mathcal{G}_{h^i}\left(z^i,\mu_i\right):=\begin{cases}\frac{h_i\left(z^i\right)}{\mu_i} & \mu_i>0\\ 0 & z^i=0,\quad \mu_i=0\\ \infty & o.w.\end{cases}.$$

那么 $\mathcal{G}_{h^i}$ 是一个真闭凸函数。此外，对于

$$z^i\left(w^i,y_i\right):=y_iw^i,\quad w^i\left(z^i,y_i\right):=\begin{cases}z^i/y_i & y_i<0\\ 0 & o.w.\end{cases}\tag{32}$$

我们有 $-\mathcal{G}_{h^i}\left(z^i\left(w^i,y_i\right),-y_i\right)=y_ih_i\left(w^i\right)$ 和 $-\mathcal{G}_{h^i}\left(z^i,-y_i\right)=y_ih_i\left(w^i\left(z^i,y_i\right)\right)$ 对于任意 $y_i\le 0$.

⟶ 证明见第 31页。

利用命题 5 以及 $y_i\nabla h_i\left(w^i\right)=\nabla h_i\left(y_iw^i\right)$ 对于任意二次型 h_i，我们可以将（29）重写为如下凸优化问题

$$\mathrm{opt}_D:=\sup_{\substack{y_{i_0}\in\mathbb{R},\\ z^{i_0}\in\mathbb{R}^J}} -\mathcal{G}_{h^{i_0}}\left(z^{i_0},-y_{i_0}\right)+u_{i_0}y_{i_0}\tag{33a}$$

$$s.t.$$

$$\nabla h_{i_0}\left(z^{i_0}\right)+q^{i_0}y_{i_0}=c,\quad y_{i_0}\le 0,\tag{33b}$$

并且我们可以通过（32）在（29）和（33）的解之间进行转换。

6. 与其他对偶框架的联系

6.1. 与 Fenchel 对偶的联系

（1）的 Fenchel 对偶可以写作

$$\operatorname{opt}_{D_{\text{Fenchel}}} := \sup_{v \in \mathbb{R}^{I \times J}} \quad -\sum_{i \in I} \left(g_i \circ F^i\right)^* \left(-v^i\right) \tag{34a}$$

$$s.t.$$

$$\sum_{i \in I} v^i = c. \tag{34b}$$

正如 Roos等人 (2020) 所指出的，要使（34）具有实用性，就需要得到共轭函数 $\left(g_i \circ F^i\right)^*$ 的实用形式。Roos等人 (2020) 展示了如何以特设方式为各种函数类获得此类实用形式。相比之下，通过函数 O_{F^i,g_i}（见定义4），（2）提供了一种更机械的构造（我们只需要 $F^i, \nabla F^i$ 和 g_i^* 的形式，而不是 $\left(g_i \circ F^i\right)^*$ 的形式；如下文所述，前者应该更容易）。在本节中，我们探讨 O_{F^i,g_i} 如何从 $\left(g_i \circ F^i\right)^*$ 的一般近似中产生，并讨论该近似何时是精确的。为此，我们需要以下来自 Rockafellar和Wets（2009）的约束资格条件定义。

定义 9. h： $\mathbb{R}^J \to \mathbb{R} \cup \{+\infty\}$ 是分段线性二次（PWLQ）当且仅当

$$h(x) = \begin{cases} h_l(x) + q^l \cdot x - u_l & x \in P_l, l \in L \\ +\infty & \text{o.w.} \end{cases},$$

其中 L 是有限的，且对每个 $l \in L$， P_l 是多面体，h_l 是二次型， $q^l \in \mathbb{R}^J$ 且 $u_l \in \mathbb{R}$。对于 F^i： $\mathbb{R}^J \to \mathbb{R}^{M_i}$ 与 g_i： $\mathbb{R}^{M_i} \to \mathbb{R} \cup \{+\infty\}$，令 U_i： $= \left\{u \in \mathbb{R}^{M_i} : \inf_{x \in \mathbb{R}^J} g_i\left(F^i(x) - u\right) < +\infty\right\}$。我们称 $g_i \circ F^i$ 满足约束资格条件，当且仅当

$$0 \in \operatorname{int}\left(U_i\right), \text{or} \tag{35a}$$

$$0 \in U_i, \text{ and } g_i \circ F^i \text{ is PWLQ}. \tag{35b}$$

请注意，对于第5节/表1中的函数，（35a）可化为标准的逐模块Slater条件，且（35b）包含逐模块多面体Slater条件。

以下推论确立了 O_{F^i,g_i} 与 $\left(g_i \circ F^i\right)$ 之间的联系 *.

推论 2. 对于 $F^i:\ \mathbb{R}^J \to \mathbb{R}^{M_i}$ 和 $g_i:\ \mathbb{R}^{M_i} \to \mathbb{R} \cup \{+\infty\}$ 满足条件 (9)，对于所有 $\bar{v}^i \in \mathbb{R}^J$

$$-\left(g_i \circ F^i\right)^*\left(-\bar{v}^i\right) \geq \sup_{\substack{y^i \in \mathbb{R}^{M_i}, \\ w^i \in \mathbb{R}^J}} -O_{F^i,g_i}\left(-y^i, w^i\right) \tag{36}$$

$$s.t.$$

$$\nabla F^i\left(w^i\right)^T y^i = \bar{v}^i.$$

此外，对于所有 $\bar{v}^i \in \mathbb{R}^J$ 和 $\bar{y}^i \in$定义域(g_i^*)，

$$\left(\bar{y}^i \cdot F^i\right)^*\left(\bar{v}^i\right) = \sup_{w^i \in \mathbb{R}^J} \bar{v}^i \cdot w^i - \bar{y}^i \cdot F^i\left(w^i\right) \tag{37}$$

是一个凸优化问题。

最后，令 $S\left(\bar{v}^i, \bar{y}^i\right) =$使$_{w^i \in \mathbb{R}^J} \bar{v}^i \cdot w^i - \bar{y}^i \cdot F^i\left(w^i\right)$达到最大值，并考虑以下条件

$$\left(\bar{y}^i \cdot F^i\right)^*\left(\bar{v}^i\right) < +\infty \quad \Rightarrow \quad S\left(\bar{v}^i, \bar{y}^i\right) \neq \emptyset, \tag{38}$$

该条件表明，(37)只要最优值有限，就能取到该最优值。

如果(38) 对所有 $\bar{v}^i \in \mathbb{R}^J$ 和 $\bar{y}^i \in$ 定义域(g_i^*)都成立，且 $g_i \circ F^i$ 满足 (35)，那么 (36) 以等式成立，其中将 上确界替换为最大值。

⟶ 证明见第 32页。

不等式（36）表 明对偶问题（2）是 Fenchel 对偶（34）的保守近似，并且

$$\text{opt}_{D_{\text{Fenchel}}} \geq \text{opt}_D, \tag{39}$$

在推论 2 的附加条件下等式成立。下面的引理表明，对于表 1 中的所有函数，更为技术性的条件（38）始终成立，因此要保证（39）中的等式，我们只需要 Slater 条件（35）。

引理 3. 若 （F^i， g^i）属于表 *1*，则 （38）对所有 $\bar{v}^i \in \mathbb{R}^J$ 和 $\bar{y}^i \in$ 定义域（g_i^*） 成立。

⟶ 证明见第 35 页。

现在，考虑一个满足驻点条件（ $\bar{v} \in \mathbb{R}^{I \times J}$ 34）的对偶解(34b)。计算 $\bar{v}$ 的目标值以得到对偶界，需要求解以下凸优化问题

$$\left(g_i \circ F^i\right)^*\left(-\bar{v}^i\right) := \sup_{w^i \in \mathbb{R}^J} -\bar{v}^i \cdot w^i - g\left(F^i\left(w^i\right)\right). \tag{40}$$

反过来，对于满足驻点条件（2b）的式（2）的对偶解 $\bar{y},\bar{w}$，我们只需计算 $\mathrm{O}_{F^i,g_i}\left(-\bar{y}^i,\right.$ $\left.\bar{w}^i\right)=-\left(\nabla F^i\left(\bar{w}^i\right)\bar{w}^i-F^i\left(\bar{w}^i\right)\right)\cdot\ y^i+g_i^*\left(-\bar{y}^i\right)$，这只需解决更简单的凸优化问题

$$g_i^*\left(-\bar{y}^i\right):=\sup_{z^i\in\mathbb{R}^{M_i}}-\bar{y}^i\cdot z^i-g\left(z^i\right). \tag{41}$$

特别是，如果我们只考虑一个固定的函数类 g_i，其中 g_i^*具有闭式（例如表1中的那些函数），那么计算$(\bar{y},\bar{w})$ 的目标值无需求解任何优化问题。相比之下，即使我们将函数 g_i 限制为具有闭式 g_i^*且将仿射函数 F^i限制为，(40) 仍可能是一个非平凡的优化问题，其最优值可能无法取到（例如，若 g_i 是闭凸锥的示性函数（40) 是一个锥优化问题）。因此，(2) 可能优于 (34)，尤其是在 (39) 以等式成立时。

6.2. 与拉格朗日对偶和广义拉格朗日对偶的联系

通过令 (Rockafellar 2023, (1.1)) 中的 $f_0\colon\mathbb{R}^J\to\mathbb{R}, g\colon\mathbb{R}^M\to\mathbb{R}\cup\{+\infty\}$ 和 $F\colon\mathbb{R}^J\to\mathbb{R}^M$ 分别取值为 $f_0(x)=c\cdot\ x, g\left(\left(z^i\right)_{i\in I}\right)=\sum_{i\in I}g_i\left(z^i\right)$ 和 $F(x)=\left(F^i(x)\right)_{i\in I}$，我们可以将 (1) 纳入 Rockafellar (2023) 的广义非线性规划（GNLP）框架。于是，(Rockafellar 2023, (1.1)) 中的拉格朗日函数 $l\colon\mathbb{R}^J\times\mathbb{R}^M\colon\to\mathbb{R}\cup\{-\infty\}$ 变为 $l(x,y)=c\cdot\ x+\sum_{i\in I}y^i\cdot\ F^i(x)-g_i^*(y^i)$；当对所有 $i\in I$ 都有 $|M_i|=1$ 和 $g_i=\delta_{\mathbb{R}_-}$ 时，它就化为标准拉格朗日函数。

最优性条件 $0\in\partial_x l(\bar{x},\ \bar{y})=0$对于上确界$_{x\in\mathbb{R}^J}\ l(x,\ y)$（参见（Rockafellar2023，（1.18））变为

$$c+\sum_{i\in I}\nabla F^i\ (x)^T\ y^i=0, \tag{42}$$

由此我们可以使用基于 $\sup^{y\in\mathbb{R}^M}\inf_{x\in\mathbb{R}^J}\ l(x,\ \ y)$ 的 Wolfe 对偶版本，表示为

$$\mathrm{opt}_D:=\sup_{y\in\mathbb{R}^M,x\in\mathbb{R}^J}\{l(x,y):(42)\}. \tag{43}$$

然后，将（42）乘以 x 以替换 $c\cdot\ x$ 在 $l(x,y)$ 的定义中的项，我们可以看到，（43）中的 $l(x,y)$ 可以替换为 $l(x,y)=-\sum_{i\in I}F_\nabla^i\left(y^i、\ \ x\right)+g_i^*(y^i)$，因此（43）是对偶（2）的一种形式，其中$-y$ 被 y 替换，并且对所有 $i\in I$都有 $w^i=x$。也就是说，除了对每个$i\in I$都有 x 的副本这一点之外，对偶（2）就是 Rockafellar (2023) 的 GNLP 拉格朗日对偶的 Wolfe 对偶类比（更准确地说，是 Dorn (1960) 提出的、如 Wolfe (1961) 所述的 Wolfe 对偶版本）。

特别是，对于标准拉格朗日设定（即 $|M_i|=1$且 $g_i=\delta_{\mathbb{R}_-}$对所有 $i\in I$），不含按模副本 14的互补松弛条件（ w^i （即对 $w^i=x$ 所有 $i\in I$）可简化为 y_iF^i（x）$=0$，这就是我们所熟悉的标准拉格朗日互补松弛条件。

6.3. 与双边最优性条件的联系

Artelys (2025）针对含有双边非线性不等式的优化问题提出了最优性条件。在凸情形下，真正的双边不等式（即上下界均有限的不等式）只能是双边线性不等式，因此我们将比较范围限定在第 5.2 节的情境中。通过 $\lambda_i = -y_i$，Artelys (2025）针对（19b）的最优性条件为

$$l_i \le a^i \cdot \bar{x} \le u_i \quad \forall i \in I \tag{44a}$$

$$\sum_{i \in I} a^i \bar{y}_i = c \tag{44b}$$

$$\bar{y}_i \min\{a^i \cdot \bar{x} - l_i, u_i - a^i \cdot \bar{x}\} = 0 \quad \forall i \in I \tag{44c}$$

再加上“乘子的适当符号 $[y_i]$ 取决于在解处哪个界（如果有的话）是紧的（活跃的）。”特别要注意，当某个约束不活跃时，(44c) 会达到预期效果，强制 $y_i = 0$。然而，如果某个约束活跃，(44c) 并不会对 y_i施加预期的符号约束，这就是为什么还需要用文字补充说明。遗憾的是，对于不熟悉优化理论的人来说，推断 y_i 的正确符号可能并不容易。

我们可以验证，(44a) 和 (44b) 分别是第 5.2 节中的原始可行性条件和对偶驻点条件。此外，第 5.2 节中的对偶变量没有符号限制，且对偶可行性始终成立。因此，我们的最优性条件与此处唯一的区别在于 (44c) 和互补松弛条件 (21) 之间；后者为 $\min\{l_i \bar{y}_i, u_i \bar{y}_i\} = \bar{y}_i a^i \cdot \bar{x}$对所有 $i \in I$均成立。我们还可以验证，(21) 不仅在相关约束不活跃时施加 $y_i = 0$，而且在约束活跃时也会施加正确的符号（例如，如果 $a_i \cdot \bar{x} = l_i$，那么 (21) 会通过假设 $\bar{y}_i \ge 0$施加 $l_i \le u_i$）。

6.4. 域驱动对偶

Karimi 和 Tuncel (2020年, 2024) 的原始问题，在令 $I = \{1\}$, $A \in \mathbb{R}^{M_1 \times J}$, $F^1 \colon \mathbb{R}^J \to \mathbb{R}^{M_1}$由 $F^1(x) := Ax$, $D \subseteq \mathbb{R}^{M_1}$为闭凸

集，并且 $g_1 : \mathbb{R}^{M_1} \to \mathbb{R} \cup \{+\infty\}$ 满足 $g_1 = \delta_D$。该问题的域驱动对偶为

$$\text{opt}_{DD} := \inf_{y \in \mathbb{R}^{M_1}} \sigma_D(y) \tag{45a}$$

$$s.t.$$

$$A^T y = -c,\ y \in D_*, \tag{45b}$$

其中 D_∞ 是 D 和 $D_* := \{y \in \mathbb{R}^{M_1} : \quad y \cdot w \le 0 \quad \forall w \in D_\infty\}$。根据（Hiriart-Urruty and Lemar´echal 2004，命题 C.2.2.4），我们可以得出结论：$y \in D_*$ 在（45）中是冗余的。接着利用 $\delta_D^* = \sigma_D$ 这一事实，我们可以得出结论：（45）等价于（2）的相应实现，其中 $-y$ 替换为 y，且 $\text{opt}_{DD} = -\text{opt}_D$

7. 证明

7.1. 第 3 节3和 4、4 节中的证明

为了证明第 3 和 4 节中的结果，我们使用下面这个推论，该推论源自 $g_i \circ F^i$ 与 O_{F^i,g_i} 之间的联系，相关讨论见第 6.1节。

推论 推论 3. 令 $F^i : \mathbb{R}^J \to \mathbb{R}^{M_i}$ 且 $g_i : \mathbb{R}^{M_i} \to \mathbb{R} \cup \{+\infty\}$ 满足 (9)。那么，对于任意 $\bar{w}^i$, $\tilde{w}^i \in \mathbb{R}^J$和 $y^i \in \mathbb{R}^{M_i}$满足 $\nabla F^i \left(\bar{w}^i\right)^T y^i = \nabla F^i \left(\tilde{w}^i\right)^T y^i$,

$$F_\nabla^i(y^i, \bar{w}^i) = F_\nabla^i(y^i, \tilde{w}^i). \tag{46}$$

此外，对于任意$(\bar{x}, \bar{w}^i, y^i) \in \mathbb{R}^J \times \mathbb{R}^J \times \mathbb{R}^{M_i}$，我们有

$$g_i \circ F^i(\bar{x}) + O_{F^i,g_i}\left(y^i, \bar{w}^i\right) \ge \left(\nabla F^i \left(\bar{w}^i\right)^T y^i\right) \cdot \bar{x}, \tag{47}$$

并且对于任意$(\bar{x},\ y^i) \in \mathbb{R}^J \times \mathbb{R}^{M_i}$存在 $\bar{w}^i \in \mathbb{R}^J$使得(47) 当且仅当(47) 对 $\bar{w}^i = \bar{x}$成立且取等号。

最后，对于任意$\hat{x} \in \mathbb{R}^J$

$$\left(g_i \circ F^i\right)_\infty (\hat{x}) \ge \sup_{w \in \mathbb{R}^J} \sigma_{\text{dom}(g_i^*)} \left(\nabla F^i(w) \hat{x}\right) \tag{48}$$

并且对任意$(\hat{w}^i, y^i) \in \mathbb{R}^J \times \mathbb{R}^{M_i}$且 $S_i :=$定义域 $\left(g_i \circ F^i\right)$

$$\sigma_{S_i}\left(\nabla F^i \left(\hat{w}^i\right)^T y^i\right) \le F_\nabla^i\left(y^i, \hat{w}^i\right) + \left(g_i^*\right)_\infty \left(y^i\right) \tag{49}$$

⟶ 参见第34页的证明。

定理 1 （**弱对偶性**）。 若 $\bar{x} \in \mathbb{R}^J$ 且 $(\bar{y}, \bar{w}) \in \mathbb{R}^M \times \mathbb{R}^{I \times J}$ 满足以下条件： $(\bar{y}, \bar{w})$ 满足驻点条件 (2b)， 则

$$c \cdot \bar{x} + \sum_{i \in I} g_i \left(F^i \left(\bar{x} \right) \right) \geq - \sum_{i \in I} O_{F^i, g_i} (-\bar{y}^i, \bar{w}^i). \tag{11}$$

特别地， 若 $\bar{x}$ 且 $(\bar{y}, \bar{w})$ 可行， 则

$$\mathrm{obj_P}(\bar{x}) \geq \mathrm{obj_D}(\bar{y}, \bar{w}). \tag{12}$$

定理 1 将推论 47）中的（3 对所有 $i \in I$ （其中 $y^i = -\bar{y}^i$）求和， 得到$\sum_{i \in I} g_i \left(F^i (\bar{x}) \right) + \left(\nabla F^i \left(\bar{w}^i \right)^T \bar{y}^i \right) \cdot x \geq - \sum_{i \in I} O_{F^i, g_i} (-\bar{y}^i, \bar{w})$。 由 （2b） 即可得出结论。 □

定理 2 **(最优性条件)**。 设 $\bar{x} \in \mathbb{R}^J$ 和 $(\bar{y}, \bar{w}) \in \mathbb{R}^M \times \mathbb{R}^{I \times J}$ 满足 $(\bar{y}, \bar{w})$驻点条件 (2b)。那么， 以下条件等价

- 原始-对偶对 $(\bar{x}, (\bar{y}, \bar{w}))$ 是最优的， 其中 $c \cdot \bar{x} + \sum_{i \in I} g_i \left(F^i (\bar{x}) \right) = - \sum_{i \in I} O_{F^i, g_i} (-\bar{y}^i, \bar{w}^i)$。
- 对于所有 $i \in I$,

$$g_i \left(F^i \left(\bar{x} \right) \right) + O_{F^i, g_i} (-\bar{y}^i, \bar{w}^i) = - \left(\bar{y}^i \right)^T \nabla F^i \left(\bar{w}^i \right) \bar{x}. \tag{13}$$

此外， 如果 $(\bar{x}, (\bar{y}, \bar{w}))$ 是最优的， 那么 $(\bar{x}, (\bar{y}, (\bar{x})_{i \in I}))$ 也同样最优。

定理2的证明 若 $(\bar{y}, \bar{w})$ 满足 (2b)， 则 $c \cdot \bar{x} + \sum_{i \in I} g_i \left(F^i (\bar{x}) \right) + \sum_{i \in I} \mathrm{O}_{F^i, g_i} (-\bar{y}^i, w^i) = \sum_{i \in I} \alpha_i$对 $\alpha_i := g_i \left(F^i (\bar{x}) \right) + \mathrm{O}_{F^i, g_i} (-\bar{y}^i, \bar{w}^i) - \left(\nabla F^i \left(\bar{w}^i \right)^T (-\bar{y}^i) \right) \cdot \bar{x}$成立。由推论 3（取 $y^i = -\bar{y}^i$）， $\alpha_i \geq 0$ 对所有 $i \in I$ 均成立； 再由定理 1 $\sum_{i \in I} \alpha_i \geq 0$， 可得 $\sum_{i \in I} \alpha_i = 0$ 当且仅当 $\alpha_i = 0$ 对所有 $i \in I$ 均成立， 结论随之得证。最后， 由推论 3（取 $y^i = -\bar{y}^i$）， 若 (13) 对某个 $(\bar{x}, \bar{y}^i, \bar{w}^i) \in \mathbb{R}^J \times \mathbb{R}^{M_i} \times \mathbb{R}^J$成立， 那么它对 $(\bar{x}, \bar{y}^i, \bar{x})$ 也成立， 这就说明了最后一个结论。□

推论 1. 一对 $(\bar{x}, (\bar{y}, \bar{w})) \in \mathbb{R}^J \times \left(\mathbb{R}^M \times \mathbb{R}^{I \times J} \right)$ 是最优的当且仅当

1. $\bar{x}$ 是原始可行的， 2. $(\bar{y}^i, \bar{w}^i)$ 对每个 $i \in I$是按模块对偶可行的， 3. $(\bar{x}, (\bar{y}^i, \bar{w}^i))$ 满足互补松弛条件 (14) 对所有 $i \in I$， 4. $(\bar{y}, \bar{w})$ 满足驻点条件 (2b)。

推论1的证明 若 $\bar{x}$原始可行， 则 $\mathrm{obj_P}(\bar{x}) = c \cdot \bar{x} + \sum_{i \in I} g_i \left(F^i (\bar{x}) \right)$。类似地， 若对每个 $i \in I$， $(\bar{y}^i, \bar{w}^i)$均对偶可行， 则 $\overset{0}{\mathrm{O}}_{F^i, g_i} (-\bar{y}^i, \bar{w}^i) = \mathrm{O}_{F^i, g_i} (-\bar{y}^i, \bar{w}^i)$， 且 $\mathrm{obj_D}(\bar{y}, \bar{w}) = - \sum_{i \in I} \mathrm{O}_{F^i, g_i} (-\bar{y}^i, \bar{w}^i)$。于是， 对所有 $i \in I$， 基于互补松弛的最优性条件等价于(2b)和(13)， 结论由定理2得出。□

引理*1.* 设 $\bar{x} \in \mathbb{R}^J$ 和*(* $\bar{y}, \bar{w}$*)* $\in \mathbb{R}^M \times \mathbb{R}^{I \times J}$*(*可行或不可行*)*满足以下条件：*(* $\bar{y}, \bar{w}$*)*满足*(2b)*，且对所有 $i \in I$，*(* $\bar{x}$*, (* $\bar{y}, \bar{w}$*))*满足*(14)*。则 $obj_{\mathrm{P}}(\bar{x}) = obj_{\mathrm{D}}(\bar{y}, \bar{w})$。

引理的证明 1 该结果通过将（14）对所有 $i \in I$ 求和并应用（2b）得到。 □

定理 3 （**无界性条件**）。 若原始问题有可行解 $\bar{x} \in \mathbb{R}^J$ 以及原始射线 $\hat{x} \in \mathbb{R}^J$，则原始问题无界。

若对偶问题有可行解 $(\bar{y}, \bar{w}) \in \mathbb{R}^M \times \mathbb{R}^{I \times J}$ 以及对偶射线 $(\hat{y}, \hat{w}) \in \mathbb{R}^M \times \mathbb{R}^{I \times J}$ 使得

$$\nabla F^i\left(\bar{w}^i\right)^T \hat{y}^i = \nabla F^i\left(\hat{w}^i\right)^T \hat{y}^i \quad \forall i \in I \tag{17}$$

则对偶问题无界。

定理 3 的证明 对于原始情形，令 $\mathrm{h}(x) = c \cdot x + \sum_{i \in I} g_i\left(F^i(x)\right)$。若 $\bar{x} \in \mathbb{R}^J$ 是原始可行解且满足 $\mathrm{h}(\bar{x}) < +\infty$，则由（Hiriart-Urruty 与 Lemaréchal 2004，命题 B.3.2.8）可知 $\mathrm{h}_\infty(d) = c \cdot d + \sum_{i \in I}(g_i \circ F^i)_\infty(d)$。此外，对所有 $t > 0$，有 $\mathrm{h}(\bar{x} + t\hat{x}) \leq \mathrm{h}(\bar{x}) + t\mathrm{h}_\infty(\hat{x})$。于是，若 $\hat{x}$ 是原始射线，则令 t 趋于无穷即可得到该结果。

对于对偶情形，首先注意到，根据推论 3 中的 (46)（其中 $y^i = -\bar{y}^i$）和 (17)，$(\hat{y}, \bar{w})$ 也是一条对偶射线，因此我们可以假设 $\bar{w} = \hat{w}$。令 $\mathrm{h} = \sum_{i \in I} \mathrm{O}_{F^i, g_i}(-y^i, \bar{w}^i) = \sum_{i \in I} \mathrm{O}_{F^i, g_i}(-y^i, \hat{w}^i)$。由于 $\mathrm{h}(y)$ 是真、闭且凸的，且 $\mathrm{h}(\bar{y}) < +\infty$，根据（Hiriart-Urruty 和 Lemaréchal，2004，命题 B.3.2.8），可得 $\mathrm{h}_\infty(\hat{y}) = \sum_{i \in I}\left(\nabla F^i\left(\hat{w}^i\right)\hat{w}^i - F^i\left(\hat{w}^i\right)\right) \cdot \hat{y}^i + \left(g_i^*\right)_\infty\left(\hat{y}^i\right)$。此外，对于所有 $t > 0$，$\mathrm{h}(\bar{y} + t\hat{y}) \leq \mathrm{h}(\bar{y}) + t\mathrm{h}_\infty(\hat{y})$。最后，对于所有 t，$(\bar{y} + t\hat{y}, \bar{w}) = (\bar{y} + t\hat{y}, \hat{w})$ 满足 (2b)。令 t 趋于无穷即可得到所需结果。□

定理 4 （**不可行性证书**）。 若存在原始射线，则对偶问题不可行；若存在对偶射线，则原始问题不可行。

定理 4 的证明 设 $\hat{x} \in \mathbb{R}^J$ 是一条原始射线，并且对于每个 $w \in \mathbb{R}^{I \times J}$，令 $v(w) = \left(\nabla F^i\left(w^i\right)\hat{x}\right)_{i \in I} \in \mathbb{R}^M$，$S(w) = \{y \in \mathbb{R}^M : \sum_{i \in I} \nabla F^i\left(w^i\right)^T y^i = c\}$，且 $C = \left\{y \in \mathbb{R}^M : -y^i \in \text{定义域}\left(g_i^*\right)\ \forall i \in I\right\}$。当 $S(w) \neq \emptyset$ 时，我们有 $\sigma_{S(w)}(v(w)) = c \cdot \hat{x}$；当 $S(w) = \emptyset$ 时，$\sigma_{S(w)}(v(w)) = -\infty$。此外，

$$\begin{aligned} c \cdot \hat{x} &< -\sum_{i \in I}(g \circ F)_\infty(\hat{x}) \\ &\leq -\sum_{i \in I} \sup_{w^i \in \mathbb{R}^J} \sigma_{\mathrm{dom}\left(g_i^*\right)}\left(\nabla F^i\left(w^i\right)\hat{x}\right) \\ &= -\sup_{w \in \mathbb{R}^{I \times J}} \sigma_C\left(-v(w)\right) = \inf_{w \in \mathbb{R}^{I \times J}, y \in C} v(w) \cdot y \end{aligned}$$

其中，第一个不等式由 (15) 得出，第二个不等式由推论 3 中的 (48) 得出。因此，对于任意满足 $S(w) \neq \emptyset$ 的 $w \in \mathbb{R}^{I\times J}$，我们有 $\sup\{v(w)\cdot y\colon\ y \in S(w)\} < \inf\{v(w)\cdot y : y \in C\}$，所以对于所有 $w \in \mathbb{R}^{I\times J}$，$S(w) \cap C = \emptyset$，且对偶问题不可行。

设 $(\hat{y}, \hat{w}) \in \mathbb{R}^M \times \mathbb{R}^{I\times J}$ 为一条对偶射线，并对每个 $i \in I$ 令 $S_i :=$ 属于定义域 $(g_i \circ F^i)$，且 $v^i := \nabla F^i(\hat{w}^i)^T \hat{y}^i$。根据推论 3 中的 (49)（取 $y^i = -\hat{y}^i$），对于所有 $i \in I$，我们有

$$\sigma_{S_i}\left(-v^i\right) \leq F_\nabla^i\left(-\hat{y}^i, \hat{w}^i\right) + \left(g_i^*\right)_\infty\left(-\hat{y}^i\right). \tag{50}$$

反设 $\bigcap_{i\in I} S_i \neq \emptyset$，并令 $k \in I$ 使得 $\bigcap_{i\in I\setminus\{k\}} S_i \neq \emptyset$。注意，对偶射线的条件（16a）等价于 $v^k = \sum_{i\in I\setminus\{k\}} -v^i$，因此（Rockafellar 和 Wets 2009，定理 11.24(e)）意味着

$$\sigma_{\bigcap_{i\in I\setminus\{k\}} S_i}\left(v^k\right) \leq \sum_{i\in I\setminus\{k\}} \sigma_{S_i}(-v^i). \tag{51}$$

将（51）与（50）结合，对于 $i \in I\setminus\{k\}$ 以及对偶射线条件（16b），可得

$$\sigma_{\bigcap_{i\in I\setminus\{k\}} S_i}\left(v^k\right) < -F_\nabla^k\left(-\hat{y}^k, \hat{w}^k\right) - \left(g_k^*\right)_\infty\left(-\hat{y}^k\right). \tag{52}$$

将（52）与（50）结合起来，对于 $i = k$ 我们得到 $\sigma_{\bigcap_{i\in I\setminus\{k\}} S_i}\left(v^k\right) < -\sigma_{S_k}\left(-v^k\right)$，或者等价地，上确界 $\{v^k\cdot x : x \in \bigcap_{i\in I\setminus\{k\}} S_i\} <$ 下确界 $\{v^k\cdot x\colon\ x \in S_k\}$，这与 $\bigcap_{i\in I} S_i \neq \emptyset$ 矛盾。□

第 8 节的证明[5]

命题 *1*. 设 $K_i \subseteq \mathbb{R}^{M_i}$ 为闭凸锥，$K_i^\circ := \{y \in \mathbb{R}^{M_i}\colon\ y\cdot x \leq 0 \quad \forall x \in K_i\}$ 为 K_i 的极锥，$F^i\colon \mathbb{R}^J \to \mathbb{R}^{M_i}$ 满足，对于 $A^i \in \mathbb{R}^{M_i\times J}$ 和 $b^i \in \mathbb{R}^{M_i}$，$F^i(x) = A^i x - b^i$，且 $g_i\colon \mathbb{R}^{M_i} \to \mathbb{R}\cup\{+\infty\}$ 满足 $g_i = \delta_{K_i}$。若 $g_i \circ F^i$ 满足（*9a*），则 $g_i^0 \equiv 0$, $g_i^* = (g_i^*)^\infty = \delta_{K_i^\circ}$, $(g_i^*)^0 \equiv 0$, $\nabla F^i\left(w^i\right) = A^i$, $F_\nabla^i\left(y^i, w^i\right) = b^i\cdot y^i$，且 $\left(g_i \circ F^i\right)_\infty(\hat{x}) = \delta_{K_i}\left(A^i\hat{x}\right)$。

命题的证明 1 的刻画 g_i^* 由示性函数共轭函数的已知性质得出（例如（Hiriart-Urruty 和 Lemar´echal2004，例 E.1.1.5）。渐近函数的刻画由（Hiriart-Urruty 和 Lemar´echal2004，命题 B.3.2.1、命题 B.3.2.8）以及以下事实得出：闭凸锥的示性函数的上图也是闭凸锥，而闭凸锥的衰退锥就是其本身。其他刻画则可直接得出。□

命题 *2*。设 $a^i \in \mathbb{R}^J$、$l_i \in \mathbb{R}$ 和 $u_i \in \mathbb{R}$ 满足 $l_i \leq u_i$、$F^i\colon\ \mathbb{R}^J \to \mathbb{R}$ 满足 $F^i(x) = a^i\cdot x$，且 $g_i\colon\ \mathbb{R} \to \mathbb{R}\cup\{+\infty\}$ 满足 $g_i = \delta_{[l_i,u_i]}$。则 $g_i^0 \equiv 0$、$g_i^*(y_i) = (g_i^*)^\infty(y_i) = (g_i^*)^0(y_i) =$ 最大值 $\{l_i y_i, u_i y_i\}$，$\nabla F^i\left(w^i\right) = a^i$、$F_\nabla^i \equiv 0$，且 $\left(g_i \circ F^i\right)_\infty(\hat{x}) = \delta_{\{0\}}\left(a^i\cdot\hat{x}\right)$。

命题 2 的证明 g_i^* 的刻画源于示性函数的共轭函数的已知性质（例如（Hiriart-Urruty and Lemar´echal 2004，例 E.1.1.5））。渐近函数的刻画源于（Hiriart-Urruty and Lemar´echal 2004，命题 B.3.2.1、命题 B.3.2.8）以及以下事实：任一支撑函数的上图都是闭凸锥，闭凸锥的衰退锥就是其自身，并且$\delta^{\infty}_{[l_i,u_i]} = \delta_{\{0\}}$。其他刻画可直接得到。□

命题 *3.* 设 $h_i: \mathbb{R}^J \to \mathbb{R}$ 是凸二次型，$F^i = h_i$ 且 $g_i: \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ 满足 $g_i(z) = z$。则 $g_i^0(z) = g_i(z)$，$g_i^* = \delta_{\{1\}}$，$(g_i^*)^0 \equiv 0$，$(g_i^*)^{\infty} = \delta_{\{0\}}$，$\nabla F^i = \nabla h_i$，$F^i_{\nabla}(y_i, w^i) = y_i h_i(w^i)$，且 $(g_i \circ F^i)_{\infty}(\hat{x}) = \delta_{\{0\}}(\nabla h_i(\hat{x}))$。

命题 3 的证明 关于 g^* 的刻画可直接由凸共轭函数的定义得到。渐近函数的刻画则由（Hiriart-Urruty 和 Lemar´echal 2004，命题 B.3.2.1）以及凸二次型上图的简单性质得到。关于 ∇F^i 的刻画是直接的，而 F^i_{Δ} 的刻画 $\nabla h_i(w^i) \cdot w^i = 2h_i(w^i)$ 由引理 2中的事实可得。□

命题 *4.* 设 $h_i: \mathbb{R}^J \to \mathbb{R}$ 为凸二次型，$q^i \in \mathbb{R}^J$，$u_i \in \mathbb{R}$，$F^i(x) = h_i(x) + q^i \cdot x - u_i$ 且 $g_i = \delta_{\mathbb{R}_-}$. 则，$g_i^0 \equiv 0$, $g_i^* = (g_i^*)^{\infty} = \delta_{\mathbb{R}_+}$, $(g_i^*)^0 \equiv 0$, $\nabla F^i(w^i) = \nabla h_i(w^i) + q^i$, $F^i_{\nabla}(y^i, w^i) = (h_i(w^i) + u) y_i$，且 $(g_i \circ F^i)_{\infty} = \delta_{\mathcal{D}}$ 对于 $D := \{d \in \mathbb{R}^J: \nabla h_i(d) = 0, \quad q^i \cdot d \le 0\}$。

命题 4 的证明 4 对 g_i^* 的刻画可由共轭函数的示性函数的已知性质推出（例如（Hiriart-Urruty 和 Lemar´echal2004，例 E.1.1.5）。渐近函数的刻画可由（Hiriart-Urruty 和 Lemar´echal2004，命题 B.3.2.1）、闭凸锥的示性函数的上图也是闭凸锥、闭凸锥的衰退锥就是其自身，以及二次函数水平集的简单性质推出。对 ∇F^i 的刻画是直接的，而 F^i_{Δ} 的刻画可由以下事实推出：对任意二次型 $h_i(w^i) \cdot w^i = 2h_i(w^i)$ 均成立。h_i.□

命题 5. 设 h_i： $\mathbb{R}^J \to \mathbb{R}$ 为凸二次型，且 $\mathcal{G}_{h^i}$： $\mathbb{R}^J \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ 为 h_i 的透视函数的闭包，由下式给出

$$\mathcal{G}_{h^i}(z^i, \mu_i) := \begin{cases} \frac{h_i(z^i)}{\mu_i} & \mu_i > 0 \\ 0 & z^i = 0, \quad \mu_i = 0 \cdot \\ \infty & o.w. \end{cases}$$

则 $\mathcal{G}_{h^i}$ 是真闭凸函数。此外，对于

$$z^i\left(w^i, y_i\right) := y_i w^i, \quad w^i\left(z^i, y_i\right) := \begin{cases} z^i / y_i & y_i < 0 \\ 0 & o.w. \end{cases} \tag{32}$$

我们有 $-\mathcal{G}_{h^i}\left(z^i\left(w^i, y_i\right), -y_i\right) = y_i h_i\left(w^i\right)$ 以及 $-\mathcal{G}_{h^i}\left(z^i, -y_i\right) = y_i h_i\left(w^i\left(z^i, y_i\right)\right)$ 对任意 $y_i \leq 0$.

命题 5 的证明 由简单的分类讨论以及以下事实可得：$h_i\left(y_i w^i\right) = y_i^2 h_i\left(w^i\right)$ 对任意二次型 h_i。 □

9. 第 6 节中的证明

我们使用 Vielma (2026 的定理 3.1 和引理 3.2；将其专门化到我们的情形后，可写成如下形式（另见 Burke 等人 (2021), Bot.等人 (2007、2008、2009), Combari 等人(1994), Gissler和 Hoheisel (2023), Pennanen (1999)).

定理 5 (**Vielma (2026)**)：若 $F^i: \mathbb{R}^J \to \mathbb{R}^{M_i}$ 且 $g_i: \mathbb{R}^{M_i} \to \mathbb{R} \cup \{+\infty\}$ 满足 (9)，则对任意 $\bar{v}^i \in \mathbb{R}^J$,

$$\left(g_i \circ F^i\right)^*\left(\bar{v}^i\right) \leq \inf_{y^i \in Y^i}\left(y^i \cdot F^i\right)^*\left(\bar{v}^i\right) + g_i^*\left(y^i\right), \tag{53}$$

其中，Y^i 等价地可以取为等于 $\mathbb{R}^{M_i}$ 或 定义域(g_i^*)。

若 $g_i \circ F^i$ 还满足约束资格条件 (35)，则 (53) 在等式处成立，且 (53) 中的下确界是最小值（可能存在 $\left(g_i \circ F^i\right)^*\left(\bar{v}^i\right) = \left(y^i \cdot F^i\right)^*\left(\bar{v}^i\right) + g_i^*\left(y^i\right) = +\infty$ 对所有 $y^i \in \mathbb{R}^{M_i}$）。

最后，$\bar{y}^i \cdot F^i(.)$ 对所有情况都是凸的 $\bar{y}^i \in$ 定义域(g_i^*).

证明 关于 $\left(g_i \circ F^i\right)^*$ 的陈述由 （Vielma 2026，定理 3.1）得出。关于 $\bar{y}^i \cdot F^i(\cdot)$ 的陈述由（Vielma 2026，引理 3.2）得出。 □

我们还将用到经典勒让德变换的以下性质（例如，参见（Rockafellar2015，第 26 节）和（Rockafellar andWets2009，第 11.C 节）。

命题 6. 设 $h: \mathbb{R}^J \to \mathbb{R}$ 是一个凸可微函数，并且给定 $\mathcal{L}_h: \mathbb{R}^J \to \mathbb{R} \cup \{+\infty\}$，定义为

$$\mathcal{L}_h(v) = \begin{cases} \nabla h(w) \cdot w - h(w) & \nabla h(w) = v \\ +\infty & o.w. \end{cases}.$$

那么，$\mathcal{L}_h$ 是良定义的（即对于任意满足 $\nabla h(w) = v$ 的 $w \in \mathbb{R}^J$，我们得到的值都相同），并且

$$h^*(v) \leq \mathcal{L}_h(v) \tag{54}$$

对所有 $v \in \mathbb{R}^J$ 均成立。此外，对于任意 x、w、$v \in \mathbb{R}^J$，我们有

$$h^*(\nabla h(w)) = \mathcal{L}_h(\nabla h(w)), \tag{55a}$$

$$h(x) + h^*(v) = v \cdot x \quad \Leftrightarrow \quad v = \nabla h(x). \tag{55b}$$

特别是，(54) 当存在 $\bar{w} \in \mathbb{R}^J$ 使得 $\nabla h(\bar{w}) = v$，或者当 $h^*(v) = +\infty$ 时，等式成立。

证明 条件 $\nabla h(\bar{w}) = v$ 等价于凸优化问题 $h^*(v) =$ 上确界$_w \{v \cdot w - h(w)\}$ 的最优值有限且由 $\bar{w}$ 取得。当这样的 $\bar{w}$ 存在时，最优值为 $\nabla h(\bar{w}) \cdot \bar{w} - h(\bar{w})$，这表明 L_h 定义良好，并且不等式 (54) 成立（当 $h^*(w)$ 的最优值有限但未取得时，不等式严格成立）。因此，等式 (55) 直接得出。□

推论 2. 对于 F^i： $\mathbb{R}^J \to \mathbb{R}^{M_i}$ 以及 g_i： $\mathbb{R}^{M_i} \to \mathbb{R} \cup \{+\infty\}$ 满足(9)，且对于所有 $\bar{v}^i \in \mathbb{R}^J$

$$-\left(g_i \circ F^i\right)^*\left(-\bar{v}^i\right) \geq \sup_{\substack{y^i \in \mathbb{R}^{M_i}, \\ w^i \in \mathbb{R}^J}} -O_{F^i, g_i}\left(-y^i, w^i\right) \tag{36}$$

$$s.t.$$

$$\nabla F^i\left(w^i\right)^T y^i = \bar{v}^i.$$

此外，对于所有 $\bar{v}^i \in \mathbb{R}^J$ 和 $\bar{y}^i \in$定义域(g_i^*)，

$$\left(\bar{y}^i \cdot F^i\right)^*\left(\bar{v}^i\right) = \sup_{w^i \in \mathbb{R}^J} \bar{v}^i \cdot w^i - \bar{y}^i \cdot F^i\left(w^i\right) \tag{37}$$

是一个凸优化问题。

最后，令 $S\left(\bar{v}^i, \bar{y}^i\right) =$使$_{w^i \in \mathbb{R}^J} \bar{v}^i \cdot w^i - \bar{y}^i \cdot F^i\left(w^i\right)$最大值，并考虑条件

$$\left(\bar{y}^i \cdot F^i\right)^*\left(\bar{v}^i\right) < +\infty \quad \Rightarrow \quad S\left(\bar{v}^i, \bar{y}^i\right) \neq \emptyset, \tag{38}$$

这说明 (37)在其最优值有限时可达到该最优值。

若 (38) 对所有 $\bar{v}^i \in \mathbb{R}^J$ 和 $\bar{y}^i \in$ 定义域(g_i^*)均成立，且 $g_i \circ F^i$ 满足 (35)，则 (36) 以等式成立，其中将 上确界替换为最大值。

推论 2 的证明 令 $\bar{y}^i \in \mathrm{dom}(g_i^*)$，且 $h = \bar{y}^i \cdot F^i$。则 h 可微，并且根据定理 5，h 也是凸函数。因此，(37) 是一个凸优化问题。此外，$\nabla h\left(w^i\right) = \nabla F^i\left(w^i\right)^T \bar{y}^i$，所以 $\nabla h\left(w^i\right) \cdot w^i - h\left(w^i\right) = F_\nabla^i\left(\bar{y}^i, w^i\right)$，因此 $L_h(-\bar{v}^i) =$inf$_{w^i \in \mathbb{R}^J}\left\{F_\nabla^i\left(\bar{y}^i, w^i\right): \nabla F^i\left(w^i\right)^T \bar{y}^i = -\bar{v}^i\right\}$。于是，根据命题 6，我们有 $-\left(\bar{y}^i \cdot F^i\right)^*\left(-\bar{v}^i\right) \geq$

$\sup_{w^i\in\mathbb{R}^J}\left\{-F_\nabla^i\left(\bar{y}^i,\ w^i\right):\ \nabla F^i\left(w^i\right)^T\bar{y}^i=-\bar{v}^i\right\}$，且等式成立；如果存在 $\bar{w}^i\in\mathbb{R}^J$ 使得 $\nabla F^i\left(\bar{w}^i\right)^T\bar{y}^i=-\bar{v}^i$，则将 sup 替换为最大值。

根据定理 5，$-\left(g_i\circ\ F^i\right)^*\left(-\bar{v}^i\right)\geq$定义域$(g_i^*)$ 的$y^i\in$上确界 $-\left(y^i\cdot\ F^i\right)^*\left(-\bar{v}^i\right)-g_i^*\left(y^i\right)$，其中为等式；如果 $g_i\circ F^i$ 满足 (35)，则可将上确界替换为最大值。结合上一段的结果，我们得到 $-\left(g_i\circ\ F^i\right)^*\left(-\bar{v}^i\right)\geq$定义域$(g_i^*)$ 的$y^i\in$上确界，$w^i\in\mathbb{R}^J\left\{-F_\nabla^i\left(y^i,\ w^i\right)-g_i^*\left(y^i\right):\right.$ $\left.\nabla F^i\left(w^i\right)^T y^i=-\bar{v}^i\right\}$或者等价地，$-\left(g_i\circ\ F^i\right)^*\left(-\bar{v}^i\right)\geq$的$y^i\in\mathbb{R}^{M_i}$上确界，$w^i\in\mathbb{R}^J\left\{-F_\nabla^i\left(-y^i,\right.\right.$ $\left.\left.w^i\right)-g_i^*\left(-y^i\right):\ \nabla F^i\left(w^i\right)^T y^i=\bar{v}^i\right\}$，其中为等式；如果 $g_i\circ F^i$ 满足 (35)，并且对于所有 $\bar{y}^i\in$定义域(g_i^*)，都存在 $\bar{w}^i\in\mathbb{R}^J$，使得只要 $-\left(\bar{y}^i\cdot\ F^i\right)^*\left(-\bar{v}^i\right)>-\infty$，就有 $\nabla F^i\left(\bar{w}^i\right)^T\bar{y}^i=-\bar{v}^i$。注意到，等式成立的这一最后条件，等价于 (37) 中的最优值在其有限时能够取得（即条件 (38)），即可得到该结果。□

为证明推论3，我们还需要以下引理。

引理 4. 令 $f^i:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ 对于 $i\in\{1,2\}$ 是满足$f_1\leq f_2$的真闭凸函数。那么 $(f_1)_\infty\leq(f_2)_\infty$。

证明 若 $f_1\leq f_2$，则 $f_2^*\leq f_1^*$，定义域$(f_1^*)\subseteq$定义域(f_2^*)且 $\sigma_{\mathrm{dom}(f_1^*)}\subseteq\sigma_{\mathrm{dom}(f_2^*)}$。结论由(Rockafellar 和 Wets 2009，命题 11.5)得出。该命题指出，对于任意真闭凸函数 f，均有$\sigma_{\mathrm{dom}(f^*)}=f_\infty$。□

推论 3. 设 $F^i:\mathbb{R}^J\to\mathbb{R}^{M_i}$和 $g_i:\mathbb{R}^{M_i}\to\mathbb{R}\cup\{+\infty\}$ 满足 (9)。那么，对于任意 $\bar{w}^i,\tilde{w}^i\in\mathbb{R}^J$和 $y^i\in\mathbb{R}^{M_i}$使得 $\nabla F^i\left(\bar{w}^i\right)^T y^i=\nabla F^i\left(\tilde{w}^i\right)^T y^i$,

$$F_\nabla^i(y^i,\bar{w}^i)=F_\nabla^i(y^i,\tilde{w}^i).\tag{46}$$

此外，对于任意 $(\bar{x},\bar{w}^i,y^i)\in\mathbb{R}^J\times\mathbb{R}^J\times\mathbb{R}^{M_i}$，我们有

$$g_i\circ F^i(\bar{x})+O_{F^i,g_i}\left(y^i,\bar{w}^i\right)\geq\left(\nabla F^i\left(\bar{w}^i\right)^T y^i\right)\cdot\bar{x},\tag{47}$$

并且对于任意 $(\bar{x},y^i)\in\mathbb{R}^J\times\mathbb{R}^{M_i}$ 存在 $\bar{w}^i\in\mathbb{R}^J$ 使得 (47) 成立且为等式，当且仅当(47) 对于 $\bar{w}^i=\bar{x}$成立且为等式。

最后，对于任意 $\hat{x}\in\mathbb{R}^J$

$$\left(g_i\circ F^i\right)_\infty(\hat{x})\geq\sup_{w\in\mathbb{R}^J}\sigma_{\mathrm{dom}(g_i^*)}\left(\nabla F^i(w)\hat{x}\right)\tag{48}$$

且对于任意 $(\hat{w}^i,y^i)\in\mathbb{R}^J\times\mathbb{R}^{M_i}$ 以及 $S_i:=$定义域 $\left(g_i\circ F^i\right)$

$$\sigma_{S_i}\left(\nabla F^i\left(\hat{w}^i\right)^T y^i\right)\leq F_\nabla^i\left(y^i,\hat{w}^i\right)+\left(g_i^*\right)_\infty\left(y^i\right)\tag{49}$$

推论 3的证明。对于 $y^i \in \mathbb{R}^{M_i}$，设 $h := y^i \cdot F^i$ 使得 $\nabla h\left(w^i\right) = \nabla F^i\left(w^i\right)^T y^i$ 且 $\mathcal{L}_h\left(\nabla h\left(w^i\right)\right) = F^i_\nabla\left(y^i,\ w^i\right)$ 对于任意 $w^i \in \mathbb{R}^{M_i}$。则（46）成立，因为 $\mathcal{L}_h$ 由命题 6 良定义。

不等式 (47) 对于 $(\bar{x}, \bar{w}^i, y^i) \in \mathbb{R}^J \times \mathbb{R}^J \times \mathbb{R}^{M_i}$ 由此得出

$$\begin{aligned}
\nabla h\left(\bar{w}^i\right) \cdot \bar{x} - g_i \circ F^i(\bar{x}) &\leq \left(g_i \circ F^i\right)^*\left(\nabla h\left(\bar{w}^i\right)\right) && \text{(56a)}\\
&\leq h^*\left(\nabla h\left(\bar{w}^i\right)\right) + g_i^*\left(y^i\right) && \text{(56b)}\\
&= \mathcal{L}_h\left(\nabla h\left(\bar{w}^i\right)\right) + g_i^*\left(y^i\right) && \text{(56c)}\\
&= O_{F^i,g_i}\left(y^i, \bar{w}^i\right), && \text{(56d)}
\end{aligned}$$

其中 (56a) 根据 Fenchel 不等式（例如，(Rockafellar and Wets 2009，命题 11.3）适用于 $g_i \circ F^i$，(56b) 根据定理 53中的 (5，而 (56c) 根据命题 55a中的 (6) 得出。

若 (47) 取等号成立，则 (56) 中的所有不等式也都取等号成立，因此特别有

$$\begin{aligned}
\nabla h\left(\bar{w}^i\right) \cdot \bar{x} - h^*\left(\nabla h\left(\bar{w}^i\right)\right) &= g_i\left(F^i(\bar{x})\right) + g_i^*\left(y^i\right) && \text{(57a)}\\
&\geq y^i \cdot F^i(\bar{x}) = h(\bar{x}), && \text{(57b)}
\end{aligned}$$

其中 (57a) 源于 (56a) 左侧与 (56b) 右侧相等，而 (57b) 源于对 g_i 应用 Fenchel 不等式。于是 $\mathrm{h}(\bar{x}) + \mathrm{h}^*\left(\nabla \mathrm{h}\left(\bar{w}^i\right)\right) = \nabla \mathrm{h}\left(\bar{w}^i\right) \cdot \bar{x}$，根据命题 6 中的 (55b)（对于 $v = \nabla \mathrm{h}\left(\bar{w}^i\right)$），可得 $\nabla F^i\left(\bar{w}^i\right)^T y^i = \nabla \mathrm{h}\left(\bar{w}^i\right) = \nabla \mathrm{h}\left(\bar{x}\right) = \nabla F^i(\bar{x})^T y^i$。然后，根据 (46) 还有 $F^i_\nabla\left(y^i, \bar{w}^i\right) = F^i_\nabla\left(y^i, \bar{x}\right)$, $\mathrm{O}_{F^i,g_i}\left(y^i, \bar{w}^i\right) = \mathrm{O}_{F^i,g_i}\left(y^i, \bar{x}\right)$，因此 (47) 对于 $\bar{w}^i = \bar{x}$ 也取等号成立。

对于 (48)，注意到根据推论 2 中的 (36)，定义域 $\left(\left(g_i \circ F^i\right)^*\right) \supseteq \bigcup_{y^i \in \mathrm{dom}\left(g_i^*\right)} \bigcup_{w^i \in \mathbb{R}^J} \left\{\nabla F^i\left(w^i\right)^T y^i\right\}$。于是 $\sigma_{\mathrm{dom}\left(\left(g_i \circ F^i\right)^*\right)}(\hat{x}) \geq \sup_{w^i \in \mathbb{R}^J} \sigma_{\mathrm{dom}\left(g_i^*\right)}\left(\nabla F^i\left(w^i\right)\hat{x}\right)$，结论由此得出，因为 $\sigma_{\mathrm{dom}\left(\left(g_i \circ F^i\right)^*\right)} = \left(g_i \circ F^i\right)_\infty$，根据（Rockafellar 和 Wets 2009，定理 11.5）。

对于 (49)，令 $\mathrm{h}^1_{w^i}\left(y^i\right) = \left(g_i \circ F^i\right)^*\left(\nabla F^i(w^i)^T y^i\right)$，且 $\mathrm{h}^2_{w^i}\left(y^i\right) = \mathrm{O}_{F^i,g_i}\left(y^i, w^i\right)$。于是，根据推论 2 中的 (36)，对于所有 $w^i \in \mathbb{R}^J$ 和 $y^i \in \mathbb{R}^{M_i}$，有 $\mathrm{h}^1_{w^i}\left(y^i\right) \leq \mathrm{h}^2_{w^i}\left(y^i\right)$。此外，$\mathrm{h}^1_{w^i}$ 和 $\mathrm{h}^1_{w^i}$ 是真闭凸函数，因此根据引理4，我们还有 $\left(\mathrm{h}^1_{w^i}\right)_\infty\left(y^i\right) \leq \left(\mathrm{h}^2_{w^i}\right)_\infty\left(y^i\right)$。注意到，根据 (Hiriart-Urruty and Lemaréchal 2004, 命题 B.3.2.8)，有 $\left(\mathrm{h}^2_{w^i}\right)_\infty\left(y^i\right) =$，结论即得。

$F_{\nabla}^{i}\left(y^{i},\ w^{i}\right)+\left(g_{i}^{*}\right)_{\infty}\left(y^{i}\right)$且 $\left(h_{w^{i}}^{1}\right)_{\infty}\left(y^{i}\right)=\left(\left(g_{i}\circ F^{i}\right)^{*}\right)_{\infty}\left(\nabla F(w^{i})^{T}y^{i}\right)$，并且根据 (Rockafellar 和 Wets 2009, 定理 11.5) $\left(\left(g_{i}\circ\ F^{i}\right)^{*}\right)_{\infty}\left(\nabla F(w^{i})^{T}y^{i}\right)=\sigma_{\mathrm{dom}\left(g_{i}\circ F^{i}\right)}\left(\nabla F^{i}\left(w^{i}\right)^{T}y^{i}\right)$。□

引理EMMA 3. 若 $(F^{i},\ g^{i})$ 位于表 *1*中，则 (38) 对所有 $\bar{v}^{i}\in\mathbb{R}^{J}$ 和 $\bar{y}^{i}\in$ 定义域(g_{i}^{*})均成立。

引理 3 的证明 对于表 1中的函数，我们有 F^{i} 是仿射函数或凸二次函数。因此，$\bar{v}^{i}\cdot w^{i}-\bar{y}^{i}\cdot F^{i}\left(w^{i}\right)$ 在（37）中的目标函数是仿射函数或凹二次函数，其无约束上确界在有限时总能取得。□

致谢

作者谨此感谢 Robert Fourer 指出 Rockafellar 单调规划对偶性 Rockafellar (1981）与具有下界和上界变量界的线性规划的分段线性对偶（7）（例如 Fourer (1994）之间的联系。

参考文献

Artelys（2025）Artelys Knitro 用户手册：终止准则。 https://www.artelys.com/app/docs/knitro/2_userGuide/termination.html，访问日期：2025-12-15。

Ben-Tal A, Nemirovski A(2001)《现代凸优化讲义：分析、算法与工程应用》（SIAM）。

Bot¸RI, Grad SM, Wanka G (2007) 复合凸优化问题的新约束资格条件与共轭对偶。《优化理论与应用杂志》135(2):241–255。

Bot¸RI, Grad SM, Wanka G (2008) 无限维空间中复合凸函数次微分公式的新约束资格条件。《数学新闻》281(8):1088–1107。

Bot¸RI, Grad SM, Wanka G (2009) 复合凸函数的广义 Moreau–Rockafellar 结果。《优化》58(7):917–933。

Boyd S, Vandenberghe L (2004)凸优化（剑桥大学出版社）。

Burke JV, Tim H, Nguyen QV (2021) 通过下确界卷积研究凸凸复合函数及其应用。《运筹学数学》46(4):1324–1348。

Combari C、Laghdir M、Thibault L（1994）凸复合函数的次微分。*Ann. Sci.Math.Qu´ebec*18(2)：119–148。

Dorn W S（1960）二次规划中的对偶性。《Quarterly of Applied Mathematics》18(2)：155–1 62.

Fourer R（1994）关于对偶单纯形法的笔记。https://www.researchgate.net/publication/243776146_Notes_on_the_Dual_Simplex_Method，访问日期：2025-12-15。

Gissler A、Hoheisel T（2023）关于 K-上图的注记。《优化》72(9)：2251–2285。

Google (2025) OR-Tools MathOpt。https://developers.google.com/optimization/math_opt，访问日期：2025-12-15。

Hiriart-Urruty JB, Lemar´echal C (2004)《凸分析基础》（Springer Science & Business Media）。Hours JH (2018) Artelys Knitro 11.0：一款新的锥规划求解器及其他新进展，第二届 JuMP-dev 研讨会，https://jump.dev/meetings/bordeaux2018/，访问日期：2025-12-15。JuMP-dev (2025) 对偶性。https://jump.dev/JuMP.jl/stable/moi/background/duality/，访问日期：2025-12-15。Kanewala U, Bieman JM (2014) 科学软件测试：系统性文献综述。Information and Software Technology 56(10):1219–1232。Karimi M, Tunc¸el L (2020) 域驱动形式的原始—对偶内点法。Mathematics of Operations Research 45(2):591–621。Karimi M, Tunc¸el L (2024) 域驱动求解器 (DDS) Version2.1：一种用于域驱动形式凸优化问题的基于 MATLAB 的软件包。Mathematical Programming Computation 16(1):37–92。Legat B, Dowson O, Garcia JD, Lubin M (2022) MathOptInterface：数学优化问题的数据结构。INFORMS Journal on Computing 34(2):672–689。Lubin M, Dowson O, Garcia JD, Huchette J, Legat B, Vielma JP (2023) JuMP 1.0：数学优化建模语言的最新改进。Mathematical Programming Computation 15(3):581–589。Mosek (2020) Mosek 建模手册。Pennanen T (1999) 图凸映射与 K-凸函数。Journal of Convex Analysis 6(2):235–266。Rockafellar R (1981) 单调异性规划：下降算法与对偶性。Nonlinear programming 4, 327–366（Elsevier）。Rockafellar R (2015)《凸分析》。Princeton Landmarks in Mathematics and Physics（Princeton University Press）。Rockafellar RT (2023) 局部最优性的充分条件中的增广拉格朗日函数与隐藏凸性。Mathematical Programming 198(1):159–194。Rockafellar RT, Wets RJB (2009)《变分分析》，第317卷（Springer Science & Business Media）。Roos K, Balvert M, Gorissen BL, Den Hertog D (2020) 推导对偶优化问题形式的一种通用且结构化的方法。INFORMS Journal on Optimization 2(4):229–255。Vandenberghe L (2010) CVXOPT 线性与二次锥规划求解器。在线资源：https://www.seas. ucla.edu/ ~vandenbe/publications/coneprog.pdf 。Vielma JP (2026) 无需 K-凸性的复合函数凸分析。arXiv:2601.10843，https:// arxiv.org/abs/2601.10843，访问日期：2026-01-22。Winters T, Manshreck T, Wright H (2020)《Google 软件工程：长期编程的经验教训》（O’Reilly Media, Inc.）。Wolfe P (1961) 非线性规划的对偶定理。Quarterly of Applied Mathematics 19(3):239–244。

附录 A：最大化问题的变更

原始问题（1）的最大化版本为

$$\operatorname{opt}_P := \sup_{x\in\mathbb{R}^J} \; c\cdot x - \sum_{i\in I} g_i\left(F^i(x)\right) \tag{58}$$

相应的对偶问题（2）则变为

$$\operatorname{opt}_D := \inf_{y\in\mathbb{R}^M, w\in\mathbb{R}^{I\times J}} \quad \sum_{i\in I} F^i_\nabla\left(y^i, w^i\right) + g_i^*\left(y^i\right) \tag{59a}$$

$$s.t.$$

$$\sum_{i\in I} \nabla F(w^i)^T y^i = c. \tag{59b}$$

原问题目标函数（8）变为

$$\operatorname{obj_P}(\bar{x}) := c\cdot\bar{x} - \sum_{i\in I} g_i^0\left(F^i\left(\bar{x}\right)\right)$$

对偶目标函数（10）变为

$$\operatorname{obj_D}(\bar{y},\bar{w}) := \sum_{i\in I} O^0_{F^i,g_i}(\bar{y}^i,\bar{w}^i).$$

弱对偶不等式（11）/（12）变为

$$c\cdot\bar{x} - \sum_{i\in I} g_i\left(F^i\left(\bar{x}\right)\right) \le \sum_{i\in I} O_{F^i,g_i}(\bar{y}^i,\bar{w}^i),$$

$$\operatorname{obj_P}(\bar{x}) \le \operatorname{obj_D}(\bar{y},\bar{w}),$$

最优性条件（13）变为

$$g_i\left(F^i\left(\bar{x}\right)\right) + O_{F^i,g_i}(\bar{y}^i,\bar{w}^i) = \left(\nabla F^i\left(\bar{w}^i\right)^T \bar{y}^i\right)\cdot\bar{x}, \tag{60}$$

和互补松弛条件（14）变为

$$g_i^0\left(F^i\left(\bar{x}\right)\right) + O^0_{F^i,g_i}(\bar{y}^i,\bar{w}^i) = \left(\nabla F^i\left(\bar{w}^i\right)^T \bar{y}^i\right)\cdot\bar{x}. \tag{61}$$

类似地，射线的定义变为

$$c \cdot \hat{x} - \sum_{i \in I} (g_i \circ F^i)_\infty (\hat{x}) > 0, \tag{62}$$

以及对偶射线定义（16）变为

$$\sum_{i \in I} \nabla F^i (\hat{w}^i)^T \hat{y}^i = 0 \tag{63a}$$

$$\sum_{i \in I} F^i_\nabla (\hat{y}^i, \hat{w}^i) + (g_i^*)_\infty \left(\hat{y}^i\right) < 0. \tag{63b}$$

A.1. 锥约束（包括单边线性不等式）

带有所有 F^i 和 g_i （如命题 1所示）的最大化原始问题（58）归约为

$$\begin{aligned} \text{opt}_P := \sup_{x \in \mathbb{R}^J} \quad & c \cdot x \\ s.t. \quad & \\ & A^i x - b^i \in K_i \quad \forall i \in I \end{aligned}$$

以及相关的对偶问题（59）归约为

$$\begin{aligned} \text{opt}_D := \inf_{y \in \mathbb{R}^M} \quad & \sum_{i \in I} b^i \cdot y^i \\ s.t. \quad & \\ & \sum_{i \in I} (A^i)^T y^i = c, \quad y^i \in K_i^\circ \quad \forall i \in I. \end{aligned}$$

正如预期的那样，由于 F^i 是仿射函数，变量 w^i 不会出现在该对偶问题中。

原始可行性变为标准锥原始可行性 $A^i x - b^i \in K_i$ 对所有 $i \in I$，对偶可行性则变为标准锥对偶可行性 $\sum_{i \in I} (A^i)^T y^i = c$ 以及 $y^i \in K_i^\circ$对所有 $i \in I$。互补松弛即标准锥互补松弛 $\bar{y}^i \cdot (A^i \bar{x} - b^i) = 0$ 对所有 $i \in I$。此外，原始射线定义（ 62 ）变为标准锥原始射线条件

$$A^i \hat{x} \in K_i \quad \forall i \in I \quad \text{and} \quad c \cdot \hat{x} > 0,$$

而对偶射线定义（63）变为标准锥对偶射线条件

$$\sum_{i\in I}(A^i)^T\hat{y}^i=0,\quad \hat{y}^i\in K_i^\circ\quad \forall i\in I$$
$$\sum_{i\in I}b^i\cdot\hat{y}^i<0.$$

再次可以看出，变量 w^i 对于描述对偶射线也并非必需 .

A.2. 双边（或区间）线性不等式

原始问题（58）中所有 F^i 和 g_i 均如命题 2所述，化简为

$$\begin{aligned}\mathrm{opt}_P:=\sup_{x\in\mathbb{R}^J}\quad & c\cdot x\\ s.t.\quad & \\ & l_i\le a^i\cdot x\le u_i\quad \forall i\in I\end{aligned}$$

相应的对偶问题（59）化简为

$$\begin{aligned}\mathrm{opt}_D:=\inf_{y\in\mathbb{R}^M}\quad & \sum_{i\in I}\max\{l_iy_i,u_iy_i\}\\ s.t.\quad & \\ & \sum_{i\in I}a^iy_i=c.\end{aligned}$$

正如预期所料，由于 F^i 是一个线性函数，变量 w^i 不会出现在该对偶问题中 .

原始可行性变为 $l_i\le a^i\cdot x\le u_i$ 对于所有 $i\in I$，所有逐模块对偶解 $\bar{y}_i$ 均可行，而完整对偶解 $\bar{y}$ 的可行性则归结为满足驻点条件 $\sum_{i\in I}a^iy_i=c$。互补松弛变为

$$\max\{l_i\bar{y}_i,u_i\bar{y}_i\}=\bar{y}_ia^i\cdot\bar{x}\quad \forall i\in I.$$

最后，原始射线定义（62）变为

$$a^i\cdot\hat{x}=0\quad \forall i\in I\quad \text{and}\quad c\cdot\hat{x}>0,$$

而对偶射线定义（63）变为

$$\sum_{i\in I} a^i \hat{y}_i = 0 \quad \text{and} \quad \sum_{i\in I} \max\{l_i \hat{y}_i, u_i \hat{y}_i\} < 0.$$

再次可以看出，描述对偶射线也不需要变量 w^i。

A.3. 二次目标

原始问题 (58) 对于 $I = \{i_0\}$ 以及 F^{i_0} 和 g_{i_0} 如命题 3 所述，化简为

$$\text{opt}_P := \sup_{x\in\mathbb{R}^J} c\cdot x - h_{i_0}(x) = \sup_{x\in\mathbb{R}^J} c\cdot x - g_{i_0}\left(F^{i_0}(x)\right)$$

相应的对偶问题 (59) 化简为

$$\text{opt}_D := \inf_{\substack{y_{i_0}\in\mathbb{R},\\ w^{i_0}\in\mathbb{R}^J}} \quad y_{i_0} h_{i_0}\left(w^{i_0}\right) \tag{64a}$$

$$s.t.$$

$$y_{i_0}\nabla h\left(w^{i_0}\right) = c, \quad y_{i_0} = 1. \tag{64b}$$

由于 F^i 不是仿射函数，变量 w^i 确实会出现在 (64) 中。

所有原始解均可行（因为原始问题是不受约束的线性加二次最小化问题），对偶可行性为 $\bar{y}_{i_0} = 1$，而互补松弛为

$$h_{i_0}(\bar{x}) + \bar{y}_{i_0} h_{i_0}\left(\bar{w}^{i_0}\right) = \bar{y}_{i_0}\nabla h_{i_0}\left(\bar{w}^{i_0}\right)\cdot\bar{x}. \tag{65}$$

此外，原始射线定义（62）变为 $\nabla h_{i_0}(\hat{x}) = 0$ 且 $c\cdot\hat{x} > 0$。类似地，对偶射线定义（63）变为

$$\hat{y}_{i_0}\nabla h\left(\hat{w}^{i_0}\right) = 0, \quad \hat{y}_{i_0} = 0, \quad \hat{y}_{i_0} h_{i_0}\left(\hat{w}^{i_0}\right) < 0.$$

特别是，我们可以验证，正如问题无约束时所预期的那样，不存在对偶射线。

与第 5.3 节类似，为了得到一个显然凸的对偶问题重构，我们可以使用对偶可行性方程消去变量 y_{i_0} $y_{i_0}=1$，从而得到更为熟悉的对偶问题（例如，参见（Vandenberghe 2010年，(3)）和（Roos 等人 2020年，第6.3节）），其形式如下：

$$\operatorname{opt}_D := \inf_{w^{i_0}\in\mathbb{R}^J} \quad h_{i_0}\left(w^{i_0}\right) \tag{66a}$$

$$s.t.$$

$$\nabla h_{i_0}\left(w^{i_0}\right) = c \tag{66b}$$

这种重新表述不需要变量 y_{i_0}，且其所有对偶解均逐模块可行（即对偶可行性可归结为平稳方程）。此外，互补松弛条件（65）变为 $h_{i_0}(\bar{x}) + h_{i_0}\left(\bar{w}^{i_0}\right) = \nabla h_{i_0}\left(\bar{w}^{i_0}\right)\cdot \bar{x}$，根据引理 2，这等价于

$$\left(\nabla h_{i_0}(\bar{x}) - \nabla h_{i_0}\left(\bar{w}^{i_0}\right)\right)\cdot\left(\bar{x}-\bar{w}^{i_0}\right) = 0. \tag{67}$$

特别地，当 67）满足 $\bar{x}=\bar{w}^{i_0}$ 时，互补松弛条件成立；此时，驻点条件变为线性减二次无约束最大化的标准最优性条件，即 $c=\nabla h_{i_0}(\bar{x})$。然而，$\bar{x}$ 与 $\bar{w}^{i_0}$ 之间的互补松弛关系也可以通过 $\nabla h_{i_0}(\bar{x}) = \nabla h_{i_0}\left(\bar{w}^{i_0}\right)$来实现；当 $\bar{x}=\bar{w}^{i_0}$为 $\nabla^2 h$ 奇异时，这种关系可能弱于

A.4. 二次约束

原始问题（58）对于 $I=\{i_0\}$ 以及 F^{i_0} 和 g_{i_0} 如命题4 所示，化简为

$$\operatorname{opt}_P := \sup_{x\in\mathbb{R}^J} \quad c\cdot x$$

$$s.t.$$

$$h_{i_0}(x) + q^{i_0}\cdot x \le u_{i_0}$$

相应的对偶问题（59）化简为

$$\operatorname{opt}_D := \inf_{y_{i_0}\in\mathbb{R}, w^{i_0}\in\mathbb{R}^J} \left(h_{i_0}\left(w^{i_0}\right) + u_{i_0}\right) y_{i_0} \tag{68a}$$

$$s.t.$$

$$y_{i_0}\left(\nabla h_{i_0}\left(w^{i_0}\right) + q^{i_0}\right) = c, \quad y_{i_0} \geq 0. \tag{68b}$$

由于 F^i 不是仿射函数，变量 w^i 确实会出现在（68）中。

原始可行性变为 $h_{i_0}(\bar{x}) + q^{i_0}\cdot \bar{x} \leq u_{i_0}$，对偶可行性为 $\bar{y}_{i_0} \geq 0$ 加上 $y_{i_0}\left(\nabla h_{i_0}\left(w^{i_0}\right) + q^{i_0}\right) = c$，并且利用引理 2 我们可以将互补松弛写为

$$\begin{aligned}&\bar{y}_{i_0}\left(h\left(\bar{w}^{i_0}\right) + q^{i_0}\cdot\bar{w}^{i_0} - u_{i_0}\right)\\ &\qquad = \bar{y}_{i_0}\left(\nabla h\left(\bar{w}^{i_0}\right) + q^{i_0}\right)\cdot\left(\bar{w}^{i_0} - \bar{x}\right).\end{aligned} \tag{69}$$

此外，原始射线定义（62）变为

$$\nabla h_{i_0}(\hat{x}) = 0, \quad q^{i_0}\cdot\hat{x} \leq 0 \quad \text{and} \quad c\cdot\hat{x} > 0, \tag{70}$$

而对偶射线定义（63）变为 $\hat{y}_{i_0}\left(\nabla h_{i_0}\left(\hat{y}^{i_0}\right) + q^{i_0}\right) = 0$，$\hat{y}_{i_0} \geq 0$ 以及 $\left(h_{i_0}\left(\hat{y}^{i_0}\right) + u_{i_0}\right)\hat{}$。

最后，与第 5.4 节类似，我们可以使用命题 5 将（68）重写为凸优化问题

$$\operatorname{opt}_D := \inf_{\substack{y_{i_0}\in\mathbb{R},\\ z^{i_0}\in\mathbb{R}^J}} \mathcal{G}_{h^{i_0}}\left(z^{i_0}, y_{i_0}\right) + u_{i_0}y_{i_0} \tag{71a}$$

$$s.t.$$

$$\nabla h_{i_0}\left(z^{i_0}\right) + q^{i_0}y_{i_0} = c, \quad y_{i_0} \geq 0, \tag{71b}$$

并且可以通过以下方式在（68）和（71）的解之间进行转换

$$z^i\left(w^i, y_i\right) := y_i w^i, \quad w^i\left(z^i, y_i\right) := \begin{cases} z^i/y_i & y_i > 0 \\ 0 & \text{o.w.} \end{cases}.$$