arXiv:2305.03605v2 [math.OC] 6 May 2024

非单调性缺失情形下预条件近端点方法与 Douglas–Rachford 分裂的收敛性

Brecht Evens PieterPas PuyaLatafat Panagiotis Patrinos

摘要

近端点算法（PPA）是求解包含问题最广为人知的方法，也是许多数值算法的基础。尽管广受欢迎，其收敛结果在很大程度上仍局限于单调情形。本文研究满足斜弱 Minty 条件的一类非单调问题中（松弛）预条件 PPA 的收敛性。此外，我们通过建立 DRS 与采用半正定预条件子的预条件 PPA 之间的联系，研究非单调情形下的（松弛）Douglas-Rachford 分裂（DRS）方法。为了更好地刻画分析所涵盖的问题类别，我们引入半单调算子这一类算子，它自然地扩展了（弱）单调算子和余弱单调算子，并描述了其若干性质。我们给出了涉及两个半单调算子之和的 DRS 全局收敛的充分条件。值得注意的是，即使所涉及算子（或其逆）的和是非单调的，DRS 仍能收敛。我们还给出了多个示例问题，展示了收敛结果的紧确性，并凸显了该理论所能涵盖的广泛应用范围。

关键词。凸/非凸优化 ·单调/非单调变分不等式 ·包含问题 ·预条件近端点算法 · Douglas–Rachford 分裂 · 半单调算子**AMS 主题分类**。47H04 · 49J52 · 49J53 · 65K05 · 65K15 · 90C26.

1 引言

预条件近端点算法（PPPA）是求解优化和变分分析中产生的包含问题的最基本方法之一。其目标是寻找包含问题的一个解。

$$\text{find} \quad z \in \mathbb{R}^n \quad \text{such that} \quad 0 \in Tz, \tag{G-I}$$

其中 $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 是一个集合值算子，通常假设其具有单调性。该算子的迭代点的（松弛）预条件近端点算法（PPPA）可表示如下：

$$\begin{cases} z^k \in (P+T)^{-1} P z^k \\ z^{k+1} = z^k + \lambda_k (z^k - z^k). \end{cases} \tag{PPPA}$$

其中，$P \in \mathbb{R}^{n\times n}$ 是一个对称半正定矩阵，且 $(\lambda_k)_{k\in\mathbb{N}}$ 是由严格正松弛参数组成的序列。经典近端点算法（PPA） [56] 对应于预条件子P 取 $1/\gamma \mathrm{I}$ 的形式，其中 γ 是一个正步长参数，松弛参数λ_k 等于 1。

已知有多种方法属于PPPA，每种方法都采用特定的预条件子，以利用问题中存在的结构。其中包括众所周知的技术，例如 Douglas-Rachford 分裂（DRS） [31, 定理 6], 交替方向乘子法（ADMM） [28,§ 8], （近端）增广拉格朗日法（ALM） [73, §4 与 5],原始-对偶混合梯度（PDHG）方法 [42,

鲁汶大学电气工程系（ESAT-STADIUS），比利时鲁汶 Arenberg 城堡公园 10 号，邮编 3001电子邮箱：{brecht.evens,pieter.pas,puya.latafat,panos.patrinos}@kuleuven.be。本研究获得以下资助：弗兰德研究基金会（FWO）博士奖学金 No.1183822N、No. 11M9523N；博士后奖学金 No. 12Y7622N 及研究项目 G081222N、G033822N、G0A0920N；欧盟“地平线 2020”研究与创新计划玛丽·斯克沃多夫斯卡-居里资助协议 No. 953348。

引理 2。2], [32] （也称为 Chambolle-Pock（CP） [18]）、迭代细化 [64, §4.1。2], 部分逆方法 [31, §5], 渐进对冲 [76,定理。 1],以及渐进解耦 [75, 定理 4]。

上述方法的收敛性分析通常依赖于某种单调性/凸性假设 [73, 74, 75]。另请参阅 [82, 55, 30, 17, 34] ， 了解相关研究及其推广。不同于这一经典框架，我们旨在针对一类通过斜弱 Minty 条件（见定义 2.1）定义的非单调算子，建立松弛预条件近端点算法（PPPA）的收敛性。该定义涉及一个对称矩阵 V，但并不要求其为正定或负定矩阵，并且该矩阵控制着（非）单调性的程度。当 V 等于零矩阵时，该条件便退化为 Minty 变分不等式（MVI） [58, 35], ，有时也称为变分相干性。例如，所有伪凸函数和星凸函数 [87] 均满足这一条件。当 V 取为 ρI 的形式且 ρ 为负数时，便得到所谓的弱 MVI，该条件已被用于通过外梯度法和前向-后向-前向分裂求解非凸非凹极小极大问题 [24, 68, 69, 12, 1, 2]。弱 MVI 与余拟单调性的概念 [10, 50, 85]，密切相关，针对后者，已有研究 [67, 46, 19] 对经典近端点算法的收敛性进行了分析。弱 MVI 与余拟单调性之间的关键区别在于，其所涉及的不等式并非在 T 的图中的任意两点之间成立，而仅在 T 的图中的任意一点与 T 的零点集合（的某个子集）之间成立（另请参见例 2.6）。

与传统上依赖预解算子映射的坚定非扩张性的收敛性分析不同，我们的收敛性分析借鉴了 [81, 80, 47]中针对单调情形提出的射影解释。我们证明，只要考虑一个不同的半空间，同一思路也可应用于非单调情形下的预条件近端点算法。更具体地说，PPA的每次迭代PPPA 都可解释为投影到某个半空间上，该半空间将当前迭代点与满足斜弱 Minty 假设的一部分解分隔开来。我们收敛性分析的第二个关键方面是允许使用半正定预条件子（而非正定预条件子）；这对于建立与分裂技术（如（松弛）Douglas-Rachford 分裂）的联系至关重要（见第 3.1 节）。值得注意的是，已有研究 PPPA在单调情形下使用半正定预条件子时的收敛性，见 [49, 定理 3。4], [13, §2。1]。在非单调情形下，上述半空间将利用所谓的影序列构造，这些序列通过将序列 $(z^k)_{k\in\mathbb{N}}$ 和 $(\bar{z}^k)_{k\in\mathbb{N}}$ 投影到预条件子 P的值域上而得到。

在本文的第二部分中，我们考虑 Douglas/Peaceman-Rachford 分裂方法这一类，并将注意力转向以下形式的结构化包含问题

$$\text{find} \quad x \in \mathbb{R}^n \quad \text{such that} \quad 0 \in T_{\mathrm{P}} x := Ax + Bx, \tag{P-I}$$

其中 $A: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 和 $B: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 是两个（可能非单调的）算子。给定步长 $\gamma > 0$、严格为正的松弛参数 $\lambda_k > 0$ 以及某个初始猜测 $s^0 \in \mathbb{R}^n$， （松弛）Douglas-Rachford 分裂方法由以下迭代点构成：

$$\begin{cases} u^k \in J_{\gamma A}(s^k) \\ v^k \in J_{\gamma B}(2u^k - s^k) \\ s^{k+1} = s^k + \lambda_k(v^k - u^k). \end{cases} \tag{DRS}$$

当 $\lambda_k = 1$ 时，该方法对应于经典 Douglas-Rachford 分裂 [26], ；而当 $\lambda_k = 2$ 时，它则退化为 Peaceman-Rachford 分裂（PRS） [66]。这些方法最初用于求解热传导问题中出现的线性方程组，后来在 [54] 中得到扩展，用于求两个极大单调算子之和的零点，并证明序列 $(u^k)_{k\in\mathbb{N}}$ 收敛到一个解 [54, 83]。

在优化情形下，经典 Douglas–Rachford 分裂（对应于 $\lambda_k = 1$）在非凸情形下的全局收敛性已在 [53] 中得到证明，前提是其中一个函数具有 Lipschitz 连续的可微性。类似地，PRS（对应于 $\lambda_k = 2$）在 [52] 中进行了研究，并额外要求光滑函数是强凸的。随后，这些结果在 [84] 中得到了统一，并给出了（松弛）DRS 在非凸情形下的紧步长范围。当限制在其情形（其中一个函数具有 Lipschitz 连续的可微性）时，我们将证明，当非光滑项具有所谓的 半单调 次微分时，相较于 [84]（见注 5.3 (*iii*) ），我们的理论允许更大的步长范围。

关于 DRS 最小化一个函数与一个 α-凸函数之和的收敛性分析，已在 [41] 在 $\alpha + \beta > 0$ 的假设下给出，即在凸情形下。对于 DRS 应用于涉及一个 α-单调算子与一个 β-单调算子之和的包含问题 [21, 71]，以及一个 α-共单调算子与一个 β-共单调算子之和的包含问题 [8], ，在 $\alpha+\beta > 0$ 的假设下，也得到了相似结果。在前一项工作中，它

显然，所谓的原始包含（P-I）是单调的；而在后一项工作中，所谓的对偶包含（D-I）由于 [10, 引理 26] 而是单调的。因此，这一研究方向所取得的结果仅限于原始问题或对偶问题单调的情形。

1.1 主要贡献

我们的主要贡献可概括如下：

1. 我们证明了 PPPA 对于一类非单调算子的收敛性；这些算子在其一个（或多个）零点处具有斜弱 Minty 解（参见定义 2.1 和定理 2.4）。我们的收敛结果仅依赖于预条件矩阵半正定（参见 表 1），且步长范围的紧性已在示例 2.6中得到证明。此外，在更严格的斜共单调性假设下，我们在定理 2.11 中获得了次线性末迭代收敛速率。

2. 利用 DRS 与 PPPA 之间的原始-对偶联系，在相关原始-对偶算子具有斜弱 Minty 解时，以及在限制性更强的共单调情形下，均可推导出 DRS 的收敛结果（见 第 3 节）。此外，我们在示例 6.1（*ii*）中展示了这一结果的紧性。

3. 尽管存在底层的非单调性，PPPA 的（局部）线性收敛性定理 2.13在附加的度量次正则性假设下得以确立。进而，该结果给出了（可能）非单调分段多面体映射的充分条件，以保证DRS 的（局部）线性收敛性（见定理 3.6）。

4. 我们针对 (μ,ρ)-半单调算子这一类引入并发展了微积分法则。该类算子可视为（弱）单调算子和余弱单调算子的自然推广（见注 4.2）。此外，我们还给出了若干具有半单调次微分的常见（非凸）函数类示例。

5. 我们基于算子 A 和 B 的半单调性，给出了保证*DRS*收敛的充分条件（见定理 *5.2*）。通过我们的分析得到的步长区域涵盖并推广了现有的松弛 Douglas/Peaceman-Rachford 分裂方法结果，例如 [21, 8], 见注 5.3。据作者所知，这是首项能够涵盖原始包含、对偶包含以及原始-对偶包含均非单调情形下DRS 收敛性的工作。我们还提供了若干示例问题，展示了该理论在标准单调情形之外的适用性。

1.2 文章结构

本文结构如下。 第1.3节 介绍一些记号，并回顾若干标准定义。在第2节 中，我们正式引入斜弱 Minty 解的概念（参见定义 2.1），并研究至少存在一个此类解的算子类上的PPPA的收敛性。 第3节建立了PPPA 与DRS之间的原始-对偶联系，进而得到在相关原始-对偶算子满足斜弱 Minty 假设时 DRS 的收敛结果。第第4节 介绍半单调算子类，并为其提供若干微积分法则和例子。我们还在充第5节中给出基于底层算子半单调性的DRS收敛的充分条件。第。在 第6节 展示了本理论涵盖的若干DRS应用问题。最后，第7节 总结全文。为便于阅读，若干证明及辅助结果放在附录中。

1.3 记号

包含零的自然数集合记为 $\mathbb{N} := \{0, 1, 。\}$。实数和扩展实数集合分别记为 $\mathbb{R} := (-\infty, \infty)$ 和 $\overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$，而非负实数和严格正实数分别记为$\mathbb{R}_+ := [0, \infty)$ 和 $\mathbb{R}_{++} := (0, \infty)$。实数的正部记为 $[\cdot]_+ := \max\{0, \cdot\}$，负部记为 $[\cdot]_- := \min\{0, \cdot\}$。用 id 表示定义在适当空间上的恒等函数 $x \mapsto x$。用 $\mathbb{R}^n$ 表示带有内积 $\langle \cdot, \cdot \rangle$ 和诱导范数 $\|\cdot\|$的标准n维欧几里得空间。对于向量 $w = (w_1, 。_N, w_N) \in \mathbb{R}^n$，用$w_i \in \mathbb{R}^{n_i}$ 表示其第i个（块）坐标。单位矩阵

记为 $I_n \in \mathbb{R}^{n\times n}$；无歧义时简记为 I。给定矩阵 $P \in \mathbb{R}^{n\times n}$，用 $\mathcal{R}(A)$ 和 $\mathcal{N}(A)$ 分别表示 A 的值域和核。对称、对称半正定和对称正定 n 阶矩阵集合分别记为 $\mathbb{S}^n$、 $\mathbb{S}^n_+$ 和 $\mathbb{S}^n_{++}$。对于 $P \in \mathbb{S}^n_+$ 和 $P \in \mathbb{S}^n_{++}$，我们也分别将 $P \succeq 0$ 和 $P \succ 0$记作。若存在非奇异矩阵 $X \in \mathbb{R}^{n\times n}$，使得 $P = X^{-1}QX$，则称矩阵 P、Q $\in \mathbb{R}^{n\times n}$ 相似。对于具有实特征值的任意矩阵 $P \in \mathbb{R}^{n\times n}$，其最小特征值记为 $\lambda_{\min}(P)$。给定对称矩阵 $P \in \mathbb{S}^n$，定义标量积 $\langle x,y\rangle_P = \langle x, Py\rangle$，以及二次函数 $q_P(x) := \langle x, x\rangle_P$。若 $P \in \mathbb{S}^n_{++}$，则其关联的诱导范数定义为 $\|x\|_P = \sqrt{q_P(x)}$。任意尺寸的两个矩阵之间的克罗内克积记为 $\otimes$。

我们用记号 $(z^k)_{k\in I}$ 表示指标属于集合 $I \subseteq \mathbb{N}$的序列。处理标量序列时，我们使用下标记号 $(\gamma_k)_{k\in I}$。若存在指标 $k_0 \in \mathbb{N}$ 和正标量 c，使得对所有 $k \geq k_0$，$|z^k| \leq cg(k))$，则称 $(z^k)_{k\in\mathbb{N}}$ 以大-$\mathcal{O}(g(k))$速率收敛（次线性收敛）；若 $z^k/_{g(k)} \to 0$，则称 $(z^k)_{k\in\mathbb{N}}$ 以小-o(g(k)) 速率收敛。若存在指标 $k_0 \in \mathbb{N}$，使得对所有 $k \geq k_0$，$\|z^{k+1} - z^\star\| \leq c\|z^k - z^\star\|$，则称 $(z^k)_{k\in\mathbb{N}}$ 以 Q-因子 $c \in (0, 1)$ Q-线性（至少）收敛到点 $z^\star$。若存在非负标量序列 $(v^k)_{k\in\mathbb{N}}$，使得 $\|z^k - z^\star\| \leq v^k$，且 $(v^k)_{k\in\mathbb{N}}$ Q-线性收敛到零，则称 $(z^k)_{k\in\mathbb{N}}$ 以 R-线性（至少）收敛到点 $z^\star$。

算子或集值映射 $A\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^d$ 将每个点 $x \in \mathbb{R}^n$ 映射到 $\mathbb{R}^d$ 的一个子集 A(x)。我们将交替使用记号 A(x) 和 Ax。我们用 $\operatorname{dom} A := \{x \in \mathbb{R}^n \mid Ax \neq \emptyset\}$ 表示 A 的定义域，用 $\operatorname{gph} A := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^d \mid y \in Ax\}$ 表示其图，并用 $\operatorname{zer} A := \{x \in \mathbb{R}^n \mid 0 \in Ax\}$ 表示其零点集。A 的逆通过其图定义，即 $\operatorname{gph} A^{-1} := \{(y, x) \mid (x, y) \in \operatorname{gph} A\}$，我们用 $\mathcal{R}(A) := \operatorname{dom} A^{-1}$ 表示 A 的值域。A 的预解算子定义为 $J_A := (\mathrm{id} + A)^{-1}$。两个算子 A 和 B 的复合记为 $A \circ B$。

定义 1.1（算子的并联和）。 算子 A，$B\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 之间的并联和定义为 $A \square B := (A^{-1} + B^{-1})^{-1}$。

定义 1.2（扩展实数的并联和）。 设 $a,b \in \mathbb{R}$。如果 $a = b = 0$ 或 $\overline{a} + b \neq 0$，则称 a 和 b 可并联求和，其并联和定义为

$$a \square b := \begin{cases} 0, & \text{if } a = b = 0, \\ \frac{ab}{a+b}, & \text{otherwise}, \end{cases}$$

其中我们采用约定：$\square\, \infty = a$。

定义 法（共）单调性。如果存在某个 μ，使得：$\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 满足以下条件，则称算子 A 为 μ-单调的，其中 $\mu \in \mathbb{R}$ 如果

$$\langle x - x', y - y'\rangle \geq \mu\|x - x'\|^2, \qquad \text{for all } (x, y), (x', y') \in \operatorname{gph} A,$$

并称其为某个 $\rho \in \mathbb{R}$ 下的 ρ-共单调，若

$$\langle x - x', y - y'\rangle \geq \rho\|y - y'\|^2, \qquad \text{for all } (x, y), (x', y') \in \operatorname{gph} A.$$

如果算子 A 的图未被包含于另一个（共）单调算子的图中，则称算子 A 为极大（共）单调。

我们称 A 在 $\bar{x} \in \operatorname{dom} A$ 处是外半连续（osc），如果

$$\limsup_{x\to\bar{x}} Ax := \{y \mid \exists x^k \to \bar{x}, \exists y^k \to y \text{ with } y^k \in Ax^k\} \subseteq A\bar{x}. \tag{1.1}$$

A 在所有地方的外半连续性等价于其图是 $\mathbb{R}^n \times \mathbb{R}^d$ 的闭子集。

扩展实值函数f的定义域是集合 $\operatorname{dom} f := \{x \in \mathbb{R}^n \mid f(x) < \infty\}$。如果 $\operatorname{dom} f \neq \emptyset$，则称$f$为真函数；如果上图 $\operatorname{epi} f := \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq \alpha\}$是中的闭子集，则称$f$为下半连续（lsc）。我们用 ∂f表示f的极限次微分。称f为 l-光滑，表示f连续可微，且 ∇f 是模数为 l 的 Lipschitz连续函数。

集合$E \subseteq \mathbb{R}^n$ 的示性函数记为 δ_E，即 $\delta_E(x) = 0$ 当 $x \in E$ 时，否则为 ∞。集合E 的法锥记为 N_E。关于 $\|\cdot\|_Q, Q \in \mathbb{S}^n_{++}$，到集合$E$ 的投影以及与集合E 的距离记为

$$\Pi^Q_E(x) := \arg\min_{z\in E}\{\|z - x\|_Q\}, \quad \operatorname{dist}_Q(x, E) := \inf_{z\in E}\{\|z - x\|_Q\},$$

分别地。上标/下标 Q 的缺失意味着相对于典范范数的定义相同。

2 预条件近端点法

在本节中，我们研究PPPA 在半正定预条件下对于一类非单调算子的收敛性，这类算子由斜弱 Minty 假设定义。

我们首先介绍这一类非单调算子，并详细说明基本假设。随后，借助近端点算法的射影解释，我们给出 PPPA 相应的收敛结果。接着，我们通过一个简单示例证明这些收敛结果是紧的。最后，在额外的度量次正则性假设下，我们建立（局部）线性收敛结果。

2.1 斜弱Minty条件下 PPPA 的收敛分析

考虑以下具有斜弱 Minty 解的算子类，它推广了在外梯度法和前向-后向-前向分裂背景下使用的具有弱 Minty 解的算子类 [24, 69, 12]。

定义 2.1 (V-斜弱 Minty 解)。 若算子 $T:\ \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 在（一个非空集合） $\mathcal{S}^\star \subseteq$ 零点集T 上，对于某个对称矩阵 $V \in \mathbb{S}^n$ 满足以下条件，则称其具有 V-斜弱 *Minty* 解

$$\langle v, z - z^\star \rangle \geq \mathrm{q}_V(v), \qquad \textit{for all } z^\star \in \mathcal{S}^\star, (z, v) \in \operatorname{gph} T, \tag{2.1}$$

其中，二次型 $\mathrm{q}_V(v) := \langle v, Vv \rangle$。当 $V = \rho \mathrm{I}$ 对于某个 $\rho \in \mathbb{R}$成立时，我们将其称为 ρ-弱 *Minty* 解。

将弱 MVI 推广为包含一般对称矩阵 V 而非标量值的形式ρ ，可以更细致地刻画问题类别，进而为PPPA和 DRS （参见第 3 节）的收敛性导出更紧致的结果。

上述定义还可以进一步放宽，要求 (2.1) 仅在 (z, v) $\in$gph T $\cap(\mathcal{R}((\mathrm{P} + \mathrm{T})^{-1}\mathrm{P}) \times \mathcal{R}(\mathrm{P}))$ 上成立，其中 P 是 PPPA 中的预条件子。在这一较弱条件下，定理 2.4 和 2.13 的结果仍然成立，因为 (2.1) 仅在该受限集合中的点处被调用（见 (2.9)）。鉴于这种松弛会将问题类别与预条件子耦合起来，并且在半单调性方面似乎不会导出更紧致的充分条件（参见第 5 节），我们不采用这一松弛。

如引言中所述，我们接下来对PPPA 的收敛性进行分析，其中一个关键方面是考虑半正定预条件子，这主要受以下收敛性分析的启发： [49,定理 3。4] 和 [20, 定理 3。3]。与上述研究不同，我们不假设T 是极大单调算子与线性斜对称算子之和。相反，我们要求它至少存在一个斜弱 Minty 解。具体而言，我们采用以下假设。

假设 I。式(G-I)中的算子 T 和式(PPPA) 中的预条件子 P 满足以下性质。

A1 $T:\ \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 是外半连续的。

A2 预条件化预解算子$(P + T)^{-1}P$ 的定义域为全定义域。

A3 存在V-斜非空集合 $\mathcal{S}^\star \subseteq$ 零点集T，以及一个对称但可能不定的矩阵 $V \in \mathbb{S}^n$，使得 T 具有 *ue* 弱*Minty*解 $\mathcal{S}^\star$ 对 V。

A4 $P \in \mathbb{S}^n_+$ 是一个对称半正定矩阵，满足

$$\eta_{\min} := 1 + \lambda_{\min}(U^\top V P U) > 0, \tag{2.2}$$

其中 U 是 P 值域的任意标准正交基。

请注意，预条件化预解算子（P+T）$^{-1}\circ P$ 可视为在/单调情形下所讨论的翘曲 [15, 36]的一个特例。

注 2.2. 假设 I。A4 对预条件子 P施加了限制，具体取决于斜弱 Minty 矩阵 V。为更好地理解这一条件，考虑矩阵 $X := U^\top P U$，该矩阵根据构造是正定的

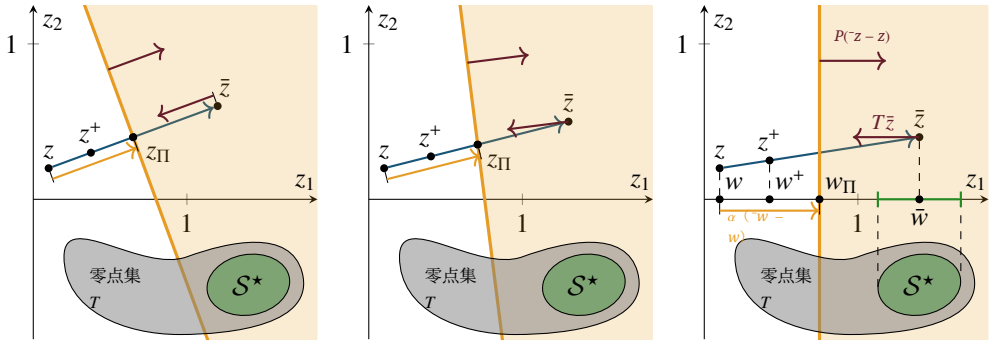

图 *1*：不同类型预条件器下引理 *2.3* 中涉及变量的可视化，其中 $V = -\frac{3}{2}I$，α 如 *(2.7)* 所示，且 $\lambda = 1/2$。这里，我们使用简写记号 $w_\Pi = \Pi^Q_{\mathscr{D}} w$，$\bar{w}(w)$ 和 $z^\Pi = \Pi^P_{\mathscr{D}} z$，$\bar{z}(z)$。在这些图中，向量 $\bar{z}$ 满足更新规则 *(PPPA)*（见 *(2.8)*）。注意，由于 α 对 V 的依赖，半空间的精确位置由斜弱 *Minty* 矩阵 V 决定。例如，若 $V = 0$，则 $\alpha = 1$，且相应的超平面将经过 $\bar{w}$。（左图和中图）正定预条件器，使得 $\Pi_{\mathcal{R}(P)} = I$，$P = Q$，且变量 z、$\bar{z}$、z^+、z_Π 分别等于变量 w、$\bar{w}$、w^+、w_Π。（左图）单位算子的倍数预条件器 $P = \frac{1}{3}I$。（中图）$P = diag(1/3, 1/6)$。（右图）半正定预条件器 $P = diag(1/3, 0)$。

（见 引理 A.1*(iv)*）。通过使用相似矩阵 $X^{1/2}$ [44, 推论 1.3。4], 进行相似变换，可得 $U^\top VPU$ 与对称矩阵 $X^{1/2}U^\top VUX^{1/2}$, 相似，即

$$\eta_{\min} = 1 + \lambda_{\min}(X^{1/2}U^\top VUX^{1/2}). \tag{2.3}$$

因此，假设 I。A4 当且仅当

$$(U^\top PU)^{-1} + U^\top VU \succ 0. \tag{2.4}$$

当$U^\top VU \succeq 0$时，此条件显然成立，这包括 WMI 设定。只有当P 不是正半定时，才需要适当选择预条件子$U^\top VU$ 时。例如，当$P = 1/\gamma \mathrm{I}$，其中 $\gamma > 0$时，假设 I。A4 可化简为 $\gamma > [-\lambda_{\min}(V)]_+$。此外，当 $V = \rho \mathrm{I}$ 时，它与余拟单调算子的步长条件相符（更多细节见例如 [19, 定理 3。1] 和表 1）。

□

我们的收敛性分析依赖于单调情形下预条件近端点算法的射影解释，这一解释可追溯至 [81, 80, 47]。特别是，已有研究表明，当预条件子正定时，PPPA 的每次迭代都可解释为向某个半空间的投影，该半空间利用序列 $(z^k)_{k\in\mathbb{N}}$ 和 $(\bar{z}^k)_{k\in\mathbb{N}}$构造。特别地，在假设$T$ 是极大单调的且$P = \gamma^{-1}\mathrm{I}$ 的条件下， [80] 考虑半空间

$$\mathscr{D}_{z,\bar{z}} := \left\{ r \in \mathbb{R}^n \mid \langle \gamma^{-1}(z - \bar{z}), \bar{z} - r \rangle \geq 0 \right\}. \tag{2.5}$$

我们将证明，只要考虑不同的半空间，就可以将这一思路推广到我们更一般的情形。具体而言，PPPA 的每次迭代都可解释为投影到一个半空间上，该半空间将当前迭代点与满足斜弱 Minty 假设的解的子集分隔开来。构造该半空间时，将使用所谓的影序列（$w^k)_{k\in\mathbb{N}} := (\Pi_{\mathcal{R}(P)} z^k)_{k\in\mathbb{N}}$ 和（$\bar{w}^k)_{k\in\mathbb{N}} := (\Pi_{\mathcal{R}(P)} \bar{z}^k)_{k\in\mathbb{N}}$，以处理这样一个事实：在我们的工作中，预条件子仅为正半定。

接下来，我们将在以下引理中介绍这个分离半空间及其主要性质，并在图 1中提供相应的示意图。请注意，矩阵$Q := P + \Pi_{\mathcal{N}(P)}$ 的引入，受到 [49,定理 34] 和 [20, 定理 33]中所进行的收敛性分析的启发。

引理 *2.3*（半空间解释）。假设假设 *I* 成立。考虑满足更新规则（*PPPA*）的点 $z \in \mathbb{R}^n$, $\bar{z} \in (P + T)^{-1}Pz$ 和 $z^+ = z + \lambda(\bar{z} - z)$，并令 $w := \Pi_{\mathcal{R}(P)} z$，$\bar{w} := \Pi_{\mathcal{R}(P)} \bar{z}$ 和 $w^+ := \Pi_{\mathcal{R}(P)} z^+$。定义矩阵 $Q := P + \Pi_{\mathcal{N}(P)}$ 和半空间

$$\mathscr{D}_{w,\bar{w}} := \{ r \in \mathbb{R}^n \mid \langle w - \bar{w}, \bar{w} - r \rangle_Q \geq \mathrm{q}_V(Q(w - \bar{w})) \}. \tag{2.6}$$

那么，以下结论成立。

(i) 集合 $\Pi_{\mathcal{R}(P)}\mathcal{S}^\star$ 是 $\mathscr{D}_{w,\ \bar{w}}$ 的子集。

(ii) 若 $w\in\mathscr{D}_{w,\ \bar{w}}$ 则 $P(z-\bar{z})=0$，特别地，$\bar{z}\in$零点集 T。

(iii) 若 $\bar{z}\notin$零点集 T，则 $q_P(z-\bar{z})\neq 0$ 且

$$\alpha:=1+\frac{q_V(P(z-\bar{z}))}{q_P(z-\bar{z})}\geq\eta_{\min}>0. \tag{2.7}$$

(iv) $w^+=(1-\lambda/\alpha)w+\lambda/\alpha\,\Pi^Q_{\mathscr{D}_{w,\ \bar{w}}}(w)$ 对于任意 $\lambda\in\mathbb{R}$。

证明。

♠ 2.3*(i)*：根据更新规则（PPPA），有

$$P(z-\bar{z})\in T\bar{z}. \tag{2.8}$$

因此，由于 V-斜弱 Minty 假设（假设 I。A3），对于所有 $z^\star\in\mathcal{S}^\star$

$$\langle P(z-\bar{z}),\,\bar{z}-z^\star\rangle\geq q_V(P(z-\bar{z})). \tag{2.9}$$

因此，对于所有 $w^\star:=\Pi_{\mathcal{R}(P)}z^\star\in\Pi_{\mathcal{R}(P)}\mathcal{S}^\star\subseteq\Pi_{\mathcal{R}(P)}$ 零点集T

$$\langle Q(w-\bar{w}),\,\bar{w}-w^\star\rangle=\langle P(z-\bar{z}),\,\bar{z}-z^\star\rangle\overset{(2.9)}{\geq}q_V(P(z-\bar{z}))=q_V(Q(w-\bar{w})),$$

其中，我们在第一次和最后一次等式中利用引理 A.1*(ii)* 来建立P 与Q 之间的关系。显然，根据构造，$\Pi_{\mathcal{R}(P)}\mathcal{S}^\star\subseteq\mathscr{D}_{w,\bar{w}}$。

♠ 2.3*(ii)*：假设 $w\in\mathscr{D}_{w,\ \bar{w}}$ 并定义简写记号 $\tilde{w}:=w-\bar{w}$，$\tilde{z}:=z-\bar{z}$。由 引理 A.1*(i)* 和 A.1*(ii)*可知

$$P=Q\,\Pi_{\mathcal{R}(P)}=\Pi_{\mathcal{R}(P)}\,Q\,\Pi_{\mathcal{R}(P)}=UU^\top QUU^\top=UXU^\top. \tag{2.10}$$

因此，根据 $\mathscr{D}_{w,\bar{w}}$ 的定义，有

$$\begin{aligned}0\geq\|\tilde{w}\|_Q^2+q_V(Q\tilde{w})&=\langle Q\,\Pi_{\mathcal{R}(P)}\,\tilde{w},\,(\mathrm{I}+VQ\,\Pi_{\mathcal{R}(P)})\tilde{w}\rangle\\&=\langle UXU^\top\tilde{w},\,(\mathrm{I}+VUXU^\top)\tilde{w}\rangle=\langle X^{1/2}U^\top\tilde{w},\,X^{1/2}U^\top(\mathrm{I}+VUXU^\top)\tilde{w}\rangle\\&=\langle X^{1/2}U^\top\tilde{w},\,(\mathrm{I}+X^{1/2}U^\top VUX^{1/2})X^{1/2}U^\top\tilde{w}\rangle.\end{aligned} \tag{2.11}$$

此外，由假设 I.a4 和注 2.2 可知，$\mathrm{I}+X^{1/2}U^\top VUX^{1/2}$ 是正定的；结合 (2.11) 和 $X\succ 0$，可得 $U^\top\tilde{w}=0$。进而，利用引理 A.1(i) 和 A.1(ii)，有 $P\tilde{z}=P\,\Pi_{\mathcal{R}(P)}\tilde{w}=PUU^\top\tilde{w}=0$；由 (2.8) 可知，$\bar{z}\in$ 属于 T 的零点集。

♠ 2.3(iii)：若 $\bar{z}\notin$T 的零点集，则包含关系 (2.8) 意味着 $P(z-\bar{z})\neq 0$。结合 $Q(w-\bar{w})=P(z-\bar{z})$（此处用到了引理 A.1(ii)）以及 Q 的正定性（见引理 A.1(iii)），可得 $q_P(z-\bar{z})=\|w-\bar{w}\|_Q^2>0$，从而确保 $\alpha=1+q_V(Q\bar{w})/\|\bar{w}\|_Q^2$ 有有限值。因此，利用 (2.11) 以及 $\|\tilde{w}\|_Q^2=\langle Q\,\Pi_{\mathcal{R}(P)}\tilde{w},\ \tilde{w}\rangle=\|X^{1/2}U^\top\tilde{w}\|^2$ 由 (2.10) 可写为 $\mathrm{I}+X^{1/2}U^\top VUX^{1/2}$ 和 $X^{1/2}U^\top\tilde{w}$ 的瑞利商 [44, 定理 4.2.2]，即，

$$\alpha=1+\frac{q_V(Q\tilde{w})}{\|\tilde{w}\|_Q^2}=\frac{\langle X^{1/2}U^\top\tilde{w},(\mathrm{I}+X^{1/2}U^\top VUX^{1/2})X^{1/2}U^\top\tilde{w}\rangle}{\|X^{1/2}U^\top\tilde{w}\|^2}\geq\lambda_{\min}(\mathrm{I}+X^{1/2}U^\top VUX^{1/2})\overset{(2.3)}{=}\eta_{\min}.$$

♠ 2.3(iv)：对于任意 $u\notin\mathscr{D}_{w,\bar{w}}$，成立 $\Pi^Q_{\mathscr{D}_{w,\bar{w}}}(u)=u+\frac{\langle w-\bar{w},u-\bar{w}\rangle_Q+q_V(Q(w-\bar{w}))}{\|w-\bar{w}\|_Q^2}(\bar{w}-w)$ [9, 例 29.20]。因此，$\Pi^Q_{\mathscr{D}_{w,\bar{w}}}(w)=w+\alpha(\bar{w}-w)$，且对于任意 $\lambda\in\mathbb{R}$，均成立

$$(1-\lambda/\alpha)w+\lambda/\alpha\,\Pi^Q_{\mathscr{D}_{w,\bar{w}}}(w)=w+\lambda(\bar{w}-w)=\Pi_{\mathcal{R}(P)}(z+\lambda(\bar{z}-z))=\Pi_{\mathcal{R}(P)}z^+=w^+.$$

□

最值得注意的是，引理 2.3*(iv)* 证明，由 PPPA生成的（影子）序列的更新规则可以解释为对由（2.6）定义的半空间进行松弛投影，而 引理 2.3*(i)* 和 2.3*(ii)*表明，如果在某次迭代中，当前迭代点属于包含投影斜弱 Minty 解的半空间 $\Pi_{\mathcal{R}(P)}\mathcal{S}^\star$，则可推出该点具有最优性。利用投影算子的非扩张性，可得到预条件近端点算法的如下收敛结果。

定理 2.4 (PPPA 的收敛）。假设 假设 *I* 成立，并考虑由 $(z^k,\ \bar{z}^k)_{k\in\mathbb{N}}$ 生成的序列 *PPPA* 从 $z^0\in\mathbb{R}^n$ 开始，松弛参数为 $\lambda_k\in(0,\ 2\alpha_k)$，满足 $\liminf_{k\to\infty}\lambda_k(2\alpha_k-\lambda_k)>0$，其中

$$\alpha_k=1+\frac{\mathrm{q}_V(P(z^k-\bar{z}^k))}{\mathrm{q}_P(z^k-\bar{z}^k)}. \tag{2.12}$$

那么，要么经过有限次迭代到达一点 $\bar{z}^k\in$ 零点集T，要么序列$(z^k,\ \bar{z}^k)_{k\in\mathbb{N}}$满足以下性质。

(i) $\bar{v}^k:=P(z^k-\bar{z}^k)\in T\bar{z}^k$ 对所有 k 均成立，且$(\bar{v}^k)_{k\in N}$ 收敛到零。

(ii) $(\bar{z}^k)_{k\in\mathbb{N}}$ 的每个极限点（如果存在）都属于零点集 T。

(iii) 影序列 $(\Pi_{\mathcal{R}(P)}z^k)_{k\in\mathbb{N}}$，$(\Pi_{\mathcal{R}(P)}\bar{z}^k)_{k\in\mathbb{N}}$ 有界，且其极限点属于 $\Pi_{\mathcal{R}(P)}\mathrm{T}$ 的零点集 T。

(iv) 若对所有 k，$\lambda_k(2\alpha_k-\lambda_k)\geq\kappa>0$ 均一致成立，则

$$\min_{k=0,1,\ldots,N}\mathrm{q}_P(z^k-\bar{z}^k)\leq\frac{1}{(N+1)\kappa}\mathrm{q}_P(z^0-z^\star),\qquad \textit{for all } z^\star\in\mathcal{S}^\star.$$

(v) 若在假设 $I.a_3$ $\Pi_{\mathcal{R}(P)}\mathcal{S}^\star=\Pi_{\mathcal{R}(P)}T$ 的零点集 中，则 $(\Pi_{\mathcal{R}(P)}z^k)_{k\in\mathbb{N}}$ 收敛到 $\Pi_{\mathcal{R}(P)}T$ 的零点集 中的某个元素。若此外 $(P+T)^{-1}\circ P$ 是（单值）连续的，则 $(\bar{z}^k)_{k\in\mathbb{N}}$ 收敛到某个 $z^\star\in T$ 的零点集。最后，若 λ_k 在区间 $(0, 2)$ 内还一致有界，则 $(z^k)_{k\in\mathbb{N}}$ 也收敛到 $z^\star\in T$ 的零点集。

证明。 注意，只要 $\bar{z}^k\notin$ 零点集T 根据 引理 2.3*(iii)* 即可得出 $\alpha_k\geq\eta_{\min}>0$。考虑影序列 $(w^k)_{k\in\mathbb{N}}:=(\Pi_{\mathcal{R}(P)}z^k)_{k\in\mathbb{N}}$ 和 $(\bar{w}^k)_{k\in\mathbb{N}}:=(\Pi_{\mathcal{R}(P)}\bar{z}^k)_{k\in\mathbb{N}}$ 及其对应的更新规则

$$w^{k+1}=w^k+\lambda_k(\bar{w}^k-w^k). \tag{2.13}$$

若 $w^k\in\mathcal{D}_{w^k,\bar{w}^k}$ 对某个 $k\in\mathbb{N}$ 成立，则根据引理 2.3*(ii)*，算法已在有限次迭代内到达点 $\bar{z}^k\in$ 零点集 T。因此，在证明余下部分中，我们考虑$w^k\notin\mathcal{D}_{w^k,\bar{w}^k}$ 的情形。

♠ 2.4(i)：由更新规则（PPPA）可直接得出，对所有 k，$\bar{v}^k:=P(z^k-\bar{z}^k)\in T\bar{z}^k$。尚需证明 $(\bar{v}^k)_{k\in\mathbb{N}}$ 收敛到零。根据引理 2.3(iv)，更新步骤 (2.13) 等价于松弛投影 $\tilde{\Pi}_k:=(1-\zeta_k)\mathrm{id}+\zeta_k\Pi^Q_{\mathcal{D}_{w^k,\bar{w}^k}}$，其中 $\zeta_k:=\lambda_k/\alpha_k$。由于 $\Pi^Q_{\mathcal{D}_{w^k}}$ 的坚定非扩张性，根据$_{\bar{w}^k}$ [9, 命题 4.16], ，在内积空间 $\langle\cdot,\cdot\rangle_Q$ [9，中，松弛投影 Π_k 是 $\zeta_k/2$-平均映射。根据引理 2.3(i)，注意到对任意 $\mathrm{w}^\star\in\Pi_{\mathcal{R}(P)}\mathcal{S}^\star$，均有 $\mathrm{w}^\star\in\mathcal{D}_{w^k,\bar{w}^k}$。因此，由 [9, 命题 4.35(iii)] 可知，对任意 $\mathrm{w}^\star\in\Pi_{\mathcal{R}(P)}\mathcal{S}^\star$，有

$$\begin{aligned}\|w^{k+1}-w^\star\|_Q^2=\|\tilde{\Pi}_k(w^k)-\tilde{\Pi}_k(w^\star)\|_Q^2&\leq\|w^k-w^\star\|_Q^2-\tfrac{1-\zeta_k/2}{\zeta_k/2}\|w^{k+1}-w^k\|_Q^2\\&=\|w^k-w^\star\|_Q^2-\tfrac{2\alpha_k-\lambda_k}{\lambda_k}\|\lambda_k(w^k-\bar{w}^k)\|_Q^2\\&=\|w^k-w^\star\|_Q^2-\lambda_k(2\alpha_k-\lambda_k)\|w^k-\bar{w}^k\|_Q^2,\end{aligned} \tag{2.14}$$

由此可得 $(w^k)_{k\in\mathbb{N}}$ 相对于 $\Pi_{\mathcal{R}(P)}\mathcal{S}^\star$ [9, 定义 51]是 Fejér 单调的。通过对（2.14）进行望远镜求和，并且由于 $\liminf_{k\to\infty}\lambda_k(2\alpha_k-\lambda_k)>0$，可得 $\sum_{k=0}^\infty\|w^k-\bar{w}^k\|_Q^2<+\infty$，这意味着

$$\|w^k-\bar{w}^k\|_Q\to0. \tag{2.15}$$

通过注意到 $(\bar{v}^k)_{k\in N}$ 收敛到零，可以证明

$$\|\bar{v}^k\|^2=\|P(w^k-\bar{w}^k)\|^2\leq\|P^{1/2}\|^2\|P^{1/2}(w^k-\bar{w}^k)\|^2=\|P\|\|w^k-\bar{w}^k\|_Q^2. \tag{2.16}$$

表 1： 定理 *2.4* 与 *PPPA*的现有收敛结果之间的联系。

	$P = \gamma^{-1}\mathrm{I}$		$P \succeq 0$	
	条件	定理	条件	定理
$V = 0$	$\gamma > 0, \lambda_k = 1$	[56][1], [74][1,2], [3, 定理 3.2]	$\lambda_k \in (0, 2)$	[49, 定理 3.4][1], [13, 定理 2.9][1]
	$\gamma > 0$, $\lambda_k \in (0, 2)$	[31, 定理 3][1,2]		
$V = \rho\mathrm{I}$	$\gamma > [-2\rho]_+$, $\lambda_k = 1$	[67, 定理 9][1,2], [46, 定理 1][1,2], [39, 定理 3.1]	$1 + \lambda_{\min}(\rho U^\top P U) > 0$, $\lambda_k \in (0, 2\alpha_k)$	定理 2.4
	$\gamma > [-\rho]_+$, $\lambda_k \in \left(0, 2(1 + \frac{\rho}{\gamma})\right)$	[19, 定理 3.1][1,2]		
$V \in \mathbb{S}$	$\gamma > [-\lambda_{\min}(V)]_+$, $\lambda_k \in (0, 2\alpha_k)$	定理 2.4	$1 + \lambda_{\min}(U^\top V P U) > 0$, $\lambda_k \in (0, 2\alpha_k)$	定理 2.4

[1]考虑最大单调性，而非 MVI；或考虑最大 ρ-共单调性，而非 ρ-弱 Minty。[2]考虑步长 γ_k 可变，且预解算子的计算为非精确计算的情形。

♠ 2.4*(ii)*：由于 $(\bar{v}^k)_{k\in\mathbb{N}}$ 收敛到零，且 $\bar{v}^k \in T\bar{z}^k$，因此根据 T (由于假设 I。A1) 的外半连续性，$(\bar{z}^k)_{k\in\mathbb{N}}$ 的任意极限点都属于 T 的零点集。

♠ 2.4(iii)：由 (2.14) 可知，$(\|w^k - w^\star\|_Q)_{k\in\mathbb{N}}$ 收敛，特别地，$(w^k)_{k\in\mathbb{N}}$ 有界。进而，利用 (2.15) 和三角不等式 $\|\bar{w}^k\|_Q \le \|\bar{w}^k - w^k\|_Q + \|w^k\|_Q$，可知 $(\bar{w}^k)_{k\in\mathbb{N}}$ 有界，因而至少有一个极限点。取一个子序列 $(\bar{w}^k)_{k\in K}$，使其收敛到某个极限点 w^∞。由于 $(\bar{w}^k - w^k)_{k\in\mathbb{N}}$ 根据 (2.15) 收敛到零，因此 $(w^k)_{k\in K}$ 也收敛到同一个极限点 w^∞。注意到根据定理 2.4(ii)，$w^\infty \in \Pi_{\mathcal{R}(P)}$ 属于 T 的零点集，即可得证。

♠ 2.4*(iv)*：通过望远镜求和（2.14），并利用一致有界性 $\forall k \in \mathbb{N}$： $\lambda_k(2\alpha_k - \lambda_k) \ge \kappa > 0$，可得

$$\min_{0\le k\le N} \|w^k - \bar{w}^k\|_Q^2 \le \tfrac{1}{(N+1)\kappa}\|w^0 - w^\star\|_Q^2. \tag{2.17}$$

所声 称的上界由此可得，只需注意到 $\|w^k - \bar{w}^k\|_Q^2 = \mathrm{q}_P(z^k - \bar{z}^k)$ 和 $\|w^0 - w^\star\|_Q^2 = \mathrm{q}_P(z^0 - z^\star)$.

♠ 2.4(v)：若 $\Pi_{\mathcal{R}(P)}\mathcal{S}^\star = \Pi_{\mathcal{R}(P)}\operatorname{zer} \mathrm{T}$，则由 [9, 定理 5.5] 可知，序列 $(\Pi_{\mathcal{R}(P)}\mathrm{z}^k)_{k\in\mathbb{N}}$ 收敛到 $\Pi_{\mathcal{R}(P)}\operatorname{zer}\mathrm{T}$ 中的某个元素。设 $\mathrm{G} = (\mathrm{P} + \mathrm{T})^{-1} \circ \mathrm{P}$ 为（单值）连续映射，因此 $\bar{\mathrm{z}}^k = \mathrm{G}\mathrm{z}^k = \mathrm{G}\mathrm{w}^k$。取一个子序列 $(\mathrm{w}^k)_{k\in K}$，它收敛到某个极限点 w^∞。于是，由 G 的连续性，序列 $(\bar{\mathrm{z}}^k)_{k\in\mathbb{N}}$ 收敛到 $\bar{\mathrm{z}}^\infty = \mathrm{G}\mathrm{w}^\infty$，并且根据定理 2.4(ii)，有 $\bar{\mathrm{z}}^\infty \in \operatorname{zer}\mathrm{T}$。

还需证明，如果 λ_k 在区间 $(0, 2)$ 内一致有界，那么 $(z^k)_{k\in\mathbb{N}}$ 也收敛到 $\bar{z}^\infty$。注意，

$$\|z^{k+1} - \bar{z}^\infty\| \le \lambda_k\|z^k - \bar{z}^\infty\| + |1 - \lambda_k|\|z^k - \bar{z}^\infty\|.$$

因此，由 λ_k的一致有界性，即 $(\lambda_k)_{k\in\mathbb{N}} \subseteq [\epsilon, 2 - \epsilon]$ 对某个 $\epsilon > 0$成立，并由 [70, §2 的引理 32] 可得 $\lim_{k\to\infty}\|z^{k+1} - \bar{z}^\infty\| = 0$，也就是说，$(z^k)_{k\in\mathbb{N}}$ 也收敛到 $\bar{z}^\infty \in$ 零点集 T。 □

<escape>注意，只要尚未达到解，即只要 $\bar{z}^k \notin$ 零点集 T，那么 $\alpha_k \ge \eta_{\min} > 0$ 根据引理 2.3*(iii)*。因此，定理 2.4 的所有结论对于（限制性更强的）固定步长范围$\lambda_k \in (0, 2\eta_{\min})$ 也同样成立。</escape>

由于我们的基本假设具有较高的普遍性，定理 2.4 统一并推广了文献中已有的关于PPPA 的收敛结果，如表 1所总结的那样。定理 2.4 的一个值得注意的特例是$V = \rho\mathrm{I}$ 且 $P = \gamma^{-1}\mathrm{I}$，此时 $\alpha_k = \eta_{\min} = 1 + \rho/\gamma$。为完整起见，我们在下文给出这一情形下的简化定理表述。

推论 *2.5.* 假设 T 是外半连续的，存在非空集合 $\mathcal{S}^\star \subseteq T$ 的零点集，使得 T 在 $\mathcal{S}^\star$ 处具有 ρ-弱 *Minty* 解，并且给定步长 $\gamma > [-\rho]_+$ 时，预解算子 $J_{\gamma T}$ 的定义域为全空间。考虑由 *PPPA* 生成的序列 $(z^k, \bar{z}^k)_{k\in\mathbb{N}}$，其初始点为 $z^0 \in \mathbb{R}^n$，预条件子 P 为 $= 1/\gamma I$，松弛参数 $\lambda_k \in (0, 2(1 + \rho/\gamma))$ 满足 $\liminf_{k\to\infty} \lambda_k(2(1 + \rho/\gamma) - \lambda_k) > 0$。那么，定理 *2.4* 中的所有结论均成立。

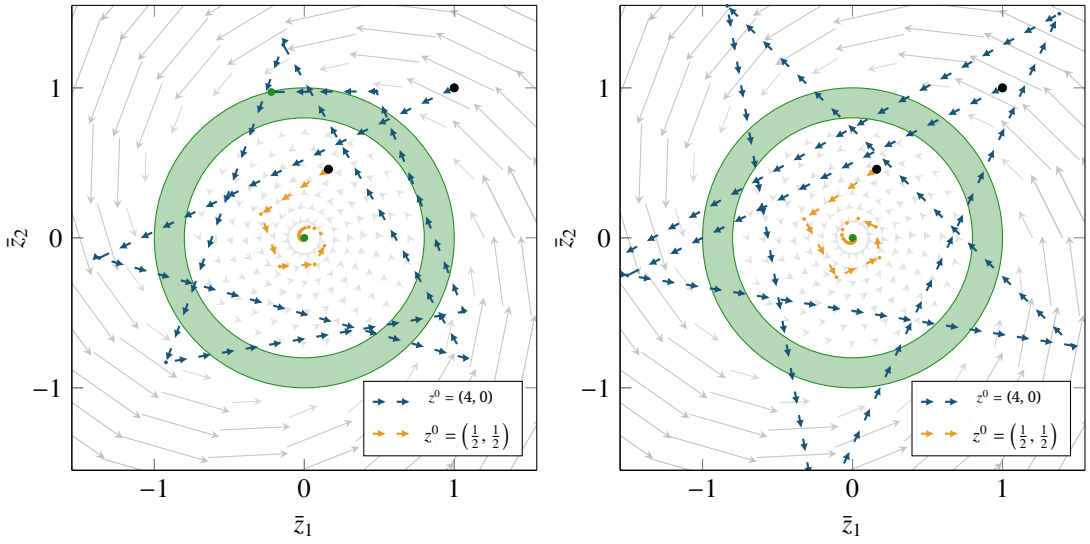

图2：松弛近端点算法在步长为 $\gamma = 1$、固定松弛参数为 λ 时应用于例 *2.6* 所生成的序列 $(\bar{z}^k)_{k\in\mathbb{N}}$ 的可视化，其中 $a = 2$ 且 $b = 1$。T 的零点以绿色标出，其梯度流 $-T$ 用灰色箭头表示。（左）使用 $\lambda = 2.3 < \bar{\lambda}$ 生成的序列，其中 $\bar{\lambda}$ 是 *(2.18)* 中 λ_k 的上界。橙色序列收敛到唯一的 (b/a^2+b^2)-弱 *Minty* 解 *(0, 0)*，而蓝色序列收敛到 T 的一个零点，但该点不是弱 *Minty* 解（参见定理 *2.4(ii)*）。（右）使用 $\lambda = 2.5 > \bar{\lambda}$ 生成的序列。在这种情况下，定理 *2.4* 不再保证序列 $(\bar{z}^k)_{k\in\mathbb{N}}$ 收敛，蓝色序列（发散）便说明了这一点。

值得再次强调的是，定理 2.4 中的所有结果（最后一项除外）即使只有零点集的一个子集T是斜弱 Minty 解时也适用。尤其值得注意的是，定理 2.4*(ii)* 指出，序列 $(\bar{z}^k)_{k\in\mathbb{N}}$ 的极限点是T的零点，但它们不一定是斜弱 Minty 解。以下示例的第二部分凸显了这一结果的重要性，其中展示了这种收敛行为在实践中的表现（另见图2）。该示例的第一部分通过考察一个与鞍点问题相关的简单线性包含，展示了我们结果的紧性。详细内容见附录C。本文数值示例的补充代码可在GitHub上找到。

例 2.6 （玩具示例）。 考虑算子

$$T(z) := f\big(\|z\|\big)\begin{bmatrix} b & a \\ -a & b \end{bmatrix} z,$$

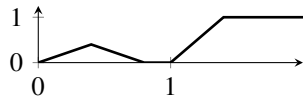

图 3： f 的定义： $\mathbb{R}_+ \to [0,1]$ 在 例 *2.6*。

其中 f： $\mathbb{R}_+ \to [0,1]$，且 a、$b \in \mathbb{R}$ 不同时等于零。

(i) 若$f \equiv 1$，则以下结论成立。

1. T是 (b/a^2+b^2)-共单调的，因此 T在 $\mathcal{S}^\star = \operatorname{zer} T = (0,0)$ 处具有唯一的 (b/a^2+b^2)-弱 Minty 解。
2. 根据 定理 2.4*(v)*，由固定松弛参数 λ 下的近端点算法生成的序列 $(\bar{z}^k)_{k\in\mathbb{N}}$ 收敛到 $\operatorname{zer}T$，若 $\gamma > {}^{[-}b]_+\, a^2+b^2$ 且 λ 位于该区间内。 ——

$$0 < \lambda < \bar{\lambda} := 2\Big(1 + \frac{b}{\gamma(a^2+b^2)}\Big). \tag{2.18}$$

通过考察算法算子的谱半径，可以看出这一结果是紧的。

(ii) 若 f 如 图 3 所定义，且 b 为非负数，则以下结论成立。

1. T 有一个唯一的 (b/a^2+b^2)-弱 Minty 解，位于点 $\mathcal{S}^\star = (0,0) \subset$ T 的零点集，而 T 的其他零点对于任何 $\rho \in \mathbb{R}$ 都不是 ρ-弱 Minty 解。2. 根据定理 2.4(ii)，由采用固定松弛参数 λ 的松弛近端点算法生成的序列 $(\bar{z}^k)_{k\in\mathbb{N}}$ 的极限点属于 T 的零点集（不一定属于 $\mathcal{S}^\star$），条件是 $\gamma > 0$ 和 λ 按照 (2.18) 选取。该结果已在图2中通过数值验证。

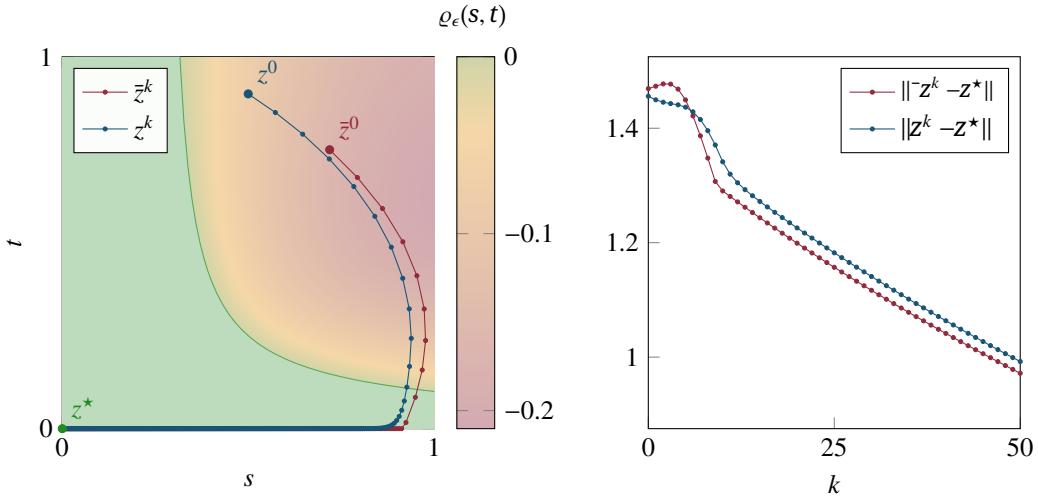

图 4：（左）松弛近端点算法应用于例2.7中的极小极大问题时，序列 $(z^k)_{k\in\mathbb{N}}$ 和 $(\bar{z}^k)_{k\in\mathbb{N}}$ 的可视化，其中问题参数为 $\epsilon = 1/10$，，弱 *Minty* 常数为 $\rho = \frac{\epsilon^{-1}}{4} = -9/40$，，步长为 $\gamma = 1/4 > [-\rho]_+$，松弛参数为 $\alpha = 1/3 \in (0, 2(1+\rho/\gamma))$。这些选择符合推论 2.5。定义域 $z \in \Delta_2 \times \Delta_2$ 由 $z = (s, 1-s, t, 1-t)$ 参数化。热图显示了 $\varrho_\epsilon(z) = \inf_{v\in N_C(z)} \langle F(z)+v, z-z^\star\rangle / \|F(z)+v\|^2$ 的取值，可将其视为算子的局部非单调性度量（另见附录C）。$\varrho_\epsilon(z)$ 为正的区域以绿色标出。尽管迭代点经过非单调区域，两个序列都收敛到纳什均衡。（右）序列 $(z^k)_{k\in\mathbb{N}}$ 和 $(\bar{z}^k)_{k\in\mathbb{N}}$ 与纳什均衡 $z^\star$ 之间距离的可视化。请注意，在前几次迭代中，经过非单调区域时，序列 $\|\bar{z}^k - z^\star\|$ 增大。另一方面，序列 $\|z^k - z^\star\|$ 单调递减（另见 *(2.14)*）。

□

作为第二个示例，我们将考虑把我们的理论应用于冯·诺依曼的经济均衡模型 [61]。对于该示例，最近在 [22, 命题 2] 中观察到，对于某些选择R 和 S，算子 F 并非单调，但满足弱 MVI。在此，我们通过证明（2.19）的纳什均衡是完整算子 $T := F + N_C$的弱 Minty 解，拓展这一结果。详情见附录C。

例2.7（冯·诺伊曼经济均衡模型）。采用单纯形 $\Delta_d := \left\{z \in \mathbb{R}^d \mid z \geq 0, \sum^d i{=}1\, z_i = 1\right\}$ 的标准定义，考虑一个有 n 种商品的经济体，其相对价格为 $y \in \Delta_n$，这些商品可由 m 种生产过程生产，各生产过程的相对强度为 $x \in \Delta_m$。令 $R_{ij} \geq 0$和 $S\, ij \geq 0$分别表示第 i 个生产过程生产和消耗的第 j 种商品的单位数。根据 [61,§5]，使该经济体得以扩张的一对 $(x^\star, y^\star)$ 对应于以下问题的一个极小极大解

$$\underset{x\in\Delta_m}{\text{minimize}}\ \underset{y\in\Delta_n}{\text{maximize}}\ f(x,y) := \frac{\langle x, Ry\rangle}{\langle x, Sy\rangle}. \tag{2.19}$$

定义$z := (x, y)$，求（2.19）的驻点对应于求以下映射的零点：$T(z) := F(z) + N_C(z)$，其中$F(z) = \begin{pmatrix} \nabla_x \\ -\nabla_y \end{pmatrix} f(x,y) f(x,y)$ 且 $C = \Delta_2 \times \Delta_2$。令 $\epsilon \in (0, 1)$，并令

$$R = \begin{bmatrix} 0 & 1+\epsilon/2 \\ 2-\epsilon/2 & 2 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} 1/2 & 1/2 \\ 1 & 1 \end{bmatrix}. \tag{2.20}$$

然后，唯一的纳什均衡 $z^\star := (x^\star, y^\star) = [0\ 1\ 0\ 1]^\top$（2.19）的$\left(\frac{\epsilon^{-1}}{4}\right)$-弱 Minty 解属于$T$。利用这一结果的数值实验见 图 4。

□

2.2 共单调性下 PPPA 的最后迭代点收敛性

定理 定理 2.4*(iv)* 中的收敛结果以最佳已知的 $\mathcal{O}(1/(N+1))$ 速率来表示，其中 $(\mathrm{q}_P(z^k - \bar{z}^k))_{k\in\mathbb{N}}$。若要获得更强的最后迭代收敛速率，控制不等式 (2.1) 必须对任意两点之间都成立

在 T 的图中，而不仅仅是在 T 的图中任意一点与（某个子集的）零点集 $\operatorname{zer} T$ 之间成立。这一观察促使我们引入V-共单调算子类。

定义 2.8 (V-余单调算子)。 若某个对称矩阵 $V \in \mathbb{S}^n$ 满足以下条件，则称算子 $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 为 V-余单调算子：

$$\langle x - x', y - y' \rangle \geq \mathrm{q}_V(y - y'), \qquad \textit{for all } (x, y), (x', y') \in \operatorname{gph} T. \tag{2.21}$$

若其图未严格包含于另一个 V-余单调算子的图中，则称其为极大 V-余单调算子。

当$V = \rho \mathrm{I}$ 时，此定义自然退化为（极大）ρ-共单调算子的概念。具有V-斜弱 Minty 解的算子类，远大于（极大）V-共单调算子类。例如，以下命题表明，极大V-共单调算子只包含没有孤立零点的算子。

命题 2.9。 若算子 $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 是极大 V-余单调的，则零点集T 是凸集。

证明。 首先，注意到 $\operatorname{zer}(T^{-1} + V)^{-1} = (T^{-1} + V)(0) = \mathrm{T}$ 的零点集 T。此外，注意到根据定义，$T^{-1} + V$是极大单调的，因此 $(T^{-1} + V)^{-1}$也具有这一性质。极大单调算子的零点集是凸的 [9,命题 23。39], 故结论得证。 □

对于极大 V-共单调算子，PPPA的收敛仅取决于预条件子的选择 P 以及松弛参数 λ_k。这是下述命题的推论，该命题允许直接验证 定理 2.4 的基本假设。

命题 2.10。 设 T 是极大 V-共单调的，且 T 的零点集T 非空，并且 P 按照 假设 I。A4 选取。那么，假设 I 成立。

证明。设 $(\mathrm{x}^k, \mathrm{y}^k) \in \mathrm{T}$ 的图，其中 $\mathrm{x}^k \to \bar{\mathrm{x}}$ 且 $\mathrm{y}^k \to \bar{\mathrm{y}}$。注意到 $\mathrm{x}^k \in (\mathrm{T}^{-1} - \mathrm{V})^{-1}(\mathrm{y}^k - \mathrm{V}\,\mathrm{x}^k)$。由于 T 是极大 V-共单调的，$(\mathrm{T}^{-1} - \mathrm{V})^{-1}$ 是极大单调的，因此也是外半连续的 [77,例 12.8]。由此可得 $\bar{\mathrm{x}} \in (\mathrm{T}^{-1} - \mathrm{V})^{-1}(\bar{\mathrm{y}} - \mathrm{V}\,\bar{\mathrm{x}})$。这意味着 $\bar{\mathrm{x}} \in \mathrm{T}\,\bar{\mathrm{y}}$，表明 T 是外半连续的。假设 I.a2 由命题 A.2 直接推出，而假设 I.a3 根据 V-共单调性的定义以及 T 的零点集非空这一条件成立。 □

利用 命题 2.10，通过证明序列 $(\mathrm{q}_P(z^k - \bar{z}^k))_{k\in\mathbb{N}}$在极大 V-余单调情形下单调非增，并使用 定理 2.4*(iv)*中的最佳迭代点结果，即可得到次线性收敛速率，详情如下。

定理 2.11 (最后迭代收敛)。 假设 T 是极大 V-余单调的，且 零点集T 非空，并且 P 按照 假设 I.A4 选取。考虑序列 $(z^k, \bar{z}^k)_{k\in\mathbb{N}}$ 由 *PPPA* 从 起始点为 $z^0 \in \mathbb{R}^n$，松弛参数为 $\lambda_k \in (0, 2\eta_{\min})$ 且满足 $\lambda_k(2\eta_{\min} - \lambda_k) \geq \kappa > 0$ 对所有 k 均一致地生成。则对所有 $z^\star \in$ 属于零点集 T，以下收敛估计成立：

$$\mathrm{q}_P(z^N - \bar{z}^N) \leq \frac{\mathrm{q}_P(z^0 - z^\star)}{(N+1)\kappa} \quad \textit{and} \quad \mathrm{q}_P(z^N - \bar{z}^N) = o(1/(N+1)). \tag{2.22}$$

证明。 根据更新规则（PPPA），可得

$$P(z^k - \bar{z}^k) \in T\bar{z}^k \quad \text{and} \quad P(z^{k+1} - \bar{z}^{k+1}) \in T\bar{z}^{k+1}.$$

定义 $\tilde{z}^k := z^k - \bar{z}^k$ 并利用V下T 在点 $\bar{z}^k$ 与 $\bar{z}^{k+1}$之间的共单调性，这意味着

$$\langle P(\tilde{z}^k - \tilde{z}^{k+1}), z^k - z^{k+1} \rangle \geq \mathrm{q}_{P+PVP}(\tilde{z}^k - \tilde{z}^{k+1}),$$

根据更新规则 (2.13) 可化简为

$$\langle 2P(\tilde{z}^k - \tilde{z}^{k+1}), \tilde{z}^k \rangle \geq \tfrac{2}{\lambda_k}\, \mathrm{q}_{P+PVP}(\tilde{z}^k - \tilde{z}^{k+1}).$$

由于 $\mathrm{q}_P(u) - \mathrm{q}_P(v) = \langle 2Pu, u - v\rangle - \mathrm{q}_P(u - v)$ 对任意 $u,v \in \mathbb{R}^n$，并利用 $P = UXU^\top$，这意味着 th at

$$\begin{aligned}
\mathrm{q}_P(\tilde{z}^k) - \mathrm{q}_P(\tilde{z}^{k+1}) &= \langle 2P\tilde{z}^k, \tilde{z}^k - \tilde{z}^{k+1}\rangle - \mathrm{q}_P(\tilde{z}^k - \tilde{z}^{k+1}) \\
&\geq \tfrac{2}{\lambda_k}\,\mathrm{q}_{P+PVP}(\tilde{z}^k - \tilde{z}^{k+1}) - \mathrm{q}_P(\tilde{z}^k - \tilde{z}^{k+1}) \\
&= \tfrac{2}{\lambda_k}\,\mathrm{q}_{UX^{1/2}(\mathrm{I}+X^{1/2}U^\top VUX^{1/2})X^{1/2}U^\top}(\tilde{z}^k - \tilde{z}^{k+1}) - \mathrm{q}_P(\tilde{z}^k - \tilde{z}^{k+1}) \\
&\overset{(2.3)}{\geq} \left(\tfrac{2\eta_{\min}}{\lambda_k} - 1\right)\mathrm{q}_P(\tilde{z}^k - \tilde{z}^{k+1}),
\end{aligned}$$

证明 $\mathrm{q}_P(z^k - \bar{z}^k)$ 是非递增的。结合 定理 2.4*(iv)* 即可得到所声称的大$\mathcal{O}$收敛速率。最后，小o 收敛性由 [23, 引理3-(1a)]得出。 □

定理 2.11 为PPPA的次线性收敛速率提供了一个通用框架，统一了文献中的许多现有结果。除其他结果外，我们还得到了标准 PPA（其中 $\mathcal{O}$ 速率（其中 $P = \gamma^{-1}\mathrm{I}$ 且$\lambda_k = 1$）在单调情形下（其中 $V = 0$） [14,命题 8] 以及弱 Minty 情形（其中 $V = \rho\mathrm{I}$） [39,定理3。1], 中得到的速率，以及 PPPA 在单调情形下得到的速率 [49, 定理 3。4]。此外，值得强调的是，当 $P = \gamma^{-1}\mathrm{I}$ 时，定理 2.11 指出，最优速率在 $\lambda_k = \eta_{\min}$时取得，而非通常采用的 $\lambda_k = 1$。

度量次正则性下 PPPA 的线性收敛

度量次正则性是描述集值映射在给定点附近行为的一个众所周知的概念，也可以看作度量正则性的单点版本。事实表明，许多用于优化的算法（例如分裂方法）在应用于具有度量次正则性的算子时具有良好的收敛性质，参见例如 [86, 27, 49, 48, 36]。继而， [25, 例 3H。4], 我们将度量次正则性定义如下。

定义 2.12 （度量次正则性）。若对于 $\bar{x}$ 和$\bar{y}$，有： $\mathbb{R}^n \rightrightarrows \mathbb{R}^d$ 集值映射 T 在 $\bar{x}$ 处关于$\bar{y}$具有度量次正则性，则（$\bar{x}$， $\bar{y}$）$\in$属于图 T，并且存在一个正常数 ζ 以及一个次正则性邻域 $\mathscr{U}$，它是 $\bar{x}$ 的邻域，使得

$$\operatorname{dist}(x, T^{-1}\bar{y}) \leq \zeta \operatorname{dist}(\bar{y}, Tx), \qquad \textit{for all } x \in \mathscr{U}. \tag{2.23}$$

存在许多重要的度量次正则算子类实例。例如，所有分段多面体映射在其图中的所有点处都是度量次正则的 [25, §3]。另一个重要实例是满足二次增长条件的真下半连续凸函数的次微分算子 ∂f（参见 [4]）。感兴趣的读者可参阅 [25,§3] 和 [77, §9]， 了解更多讨论。

本小节的主要结果是定理 2.13，其中证明了，在附加度量次正则性假设下，PPPA可以实现局部线性收敛。若度量次正则性在全局成立，即 $\mathscr{U} = \mathbb{R}^n$ 满足定义 2.12，则线性收敛也在全局成立。该结果进而将给出（局部）线性收敛的充分条件，适用于（可能）非单调分段多面体映射的DRS（参见定理 3.6）。

有趣的是，当 $P = \gamma^{-1}\mathrm{I}$ 且 $V = 0$时，定理 2.13 证明，序列 $(\operatorname{dist}(z^k, \text{零点集 } T))_{k\in\mathbb{N}}$以 Q-线性方式收敛于零，其收敛因子与极大单调情形下得到的因子相同 [79,定理 3.2(b)]。关于单调情形下 PPA 线性收敛的更详细讨论，感兴趣的读者可参阅 [5, 51] 以及其中的参考文献。

定理 *2.13*（度量次正则性下 *PPPA* 的线性收敛）。除假设 *I* 外，假设 *T* 的零点集非空、闭且凸，假设 *I.a3* 对 $\mathcal{S}^\star = zer\ T$ 成立，$(P + T)^{-1} \circ P$ 连续，并且 *T* 在所有 $z^\star \in zer\ T$（满足 *0*）处均为度量次正则的，模数为 $\zeta \in \mathbb{R}_+$。考虑一个由 *PPPA* 生成、以 $z^0 \in \mathbb{R}^n$ 为初始点的序列 $(z^k, \bar{z}^k)_{k\in\mathbb{N}}$，其松弛参数为 $\lambda_k \in (0, 2\alpha_k)$，且当 k 足够大时 $\lambda_k(2\alpha_k - \lambda_k) \geq \kappa > 0$，其中 α_k 按式 *(2.12)* 定义。则 $(dist_Q(\Pi_{\mathcal{R}(P)} z^k, \Pi_{\mathcal{R}(P)} T$ 的零点集$))_{k\in\mathbb{N}}$ 以 Q-线性方式收敛至零，Q-因子为 $\sqrt{1 - \kappa/(1+\zeta\|P\|)^2}$；$(\Pi_{\mathcal{R}(P)} z^k)_{k\in\mathbb{N}}$以 R-线性方式收敛至 T 的零点集中的某个元素，且 $q_P(z^k - \bar{z}^k)$以 R-线性方式收敛至零。

证明。由定理 2.4(v)，序列 $(\bar{z}^k)_{k\in\mathbb{N}}$ 收敛到点 $z^\star \in$，其中 z 属于 T 的零点集。至多舍弃有限个初始迭代点后，有 $\bar{z}^k \in \mathscr{U}_{z^\star}$，其中 $\mathscr{U}_{z^\star}$表示与 $z^\star$相关联的次正则性邻域。T 在 $z^\star$ 处关于 0 的度量次正则性，以及 $P(z^k - \bar{z}^k) \in T\bar{z}^k$ 这一事实（参见定理 2.4(i)），意味着

$$\operatorname{dist}(z^k, zer\ T) \leq \zeta\|P(z^k - \bar{z}^k)\|, \tag{2.24}$$

对于某个 $\zeta \in (0,\infty)$。因此，影序列到集合 $\Pi_{\mathcal{R}(P)}\mathrm{T}$ 的零点集 T 的距离有界如下：

$$\begin{aligned}\operatorname{dist}(\bar{w}^k, \Pi_{\mathcal{R}(P)} \operatorname{zer} T) = \inf_{v\in\Pi_{\mathcal{R}(P)}\operatorname{zer} T}\{\|v-\bar{w}^k\|\} = \inf_{v\in\operatorname{zer} T}\{\|\Pi_{\mathcal{R}(P)}(v-\bar{z}^k)\|\} \le \|\Pi_{\mathcal{R}(P)}\|\operatorname{dist}(\bar{z}^k, \operatorname{zer} T)\\ \overset{(2.24)}{\le} \zeta\|P(z^k-\bar{z}^k)\| \qquad (2.25)\end{aligned}$$

其中我们利用了 $\|\Pi_{\mathcal{R}(P)}\| = 1$。注意，由于集合 $\mathcal{S}^\star = \operatorname{zer} T$ 是凸的，因此 $\Pi_{\mathcal{R}(P)}\operatorname{zer} T$ 也是凸的，这意味着 $\operatorname{dist}(\bar{}w^k, \Pi_{\mathcal{R}(P)}\operatorname{zer} T) = \|\bar{}w^k - \Pi_{\Pi_{\mathcal{R}(P)}\text{零点集}\,T}(\bar{}w^k)\|$ 可以取到。由三角不等式

$$\begin{aligned}\|w^k - \Pi_{\Pi_{\mathcal{R}(P)}\operatorname{zer} T}(\bar{w}^k)\| \le \|\bar{w}^k - \Pi_{\Pi_{\mathcal{R}(P)}\operatorname{zer} T}(\bar{w}^k)\| + \|w^k-\bar{w}^k\| = \operatorname{dist}(\bar{w}^k, \Pi_{\mathcal{R}(P)}\operatorname{zer} T) + \|w^k-\bar{w}^k\|\\ \overset{(2.25)}{\le} \zeta\|P(z^k-\bar{z}^k)\| + \|w^k-\bar{w}^k\|\\ \overset{(2.16)}{\le} \xi\|w^k-\bar{w}^k\|_Q,\end{aligned}$$

其中 $\xi = \zeta\|P\|^{1}\!/\!_{2} + \|P\|^{-1}\!/\!_{2}$。由此可得

$$\operatorname{dist}_Q^2(w^k, \Pi_{\mathcal{R}(P)}\operatorname{zer} T) = \inf_{v\in\Pi_{\mathcal{R}(P)}\operatorname{zer} T}\{\|w^k-v\|_Q^2\} \le \|w^k - \Pi_{\Pi_{\mathcal{R}(P)}\operatorname{zer} T}(\bar{w}^k)\|_Q^2 \le \xi^2\|P\|\|w^k-\bar{w}^k\|_Q^2. \qquad (2.26)$$

因此，由于(2.14)和(2.26)，我们对所有足够大的 k 得到

$$\begin{aligned}\operatorname{dist}_Q^2(w^{k+1}, \Pi_{\mathcal{R}(P)}\operatorname{zer} T) &\le \|w^{k+1} - \Pi^Q_{\Pi_{\mathcal{R}(P)}\operatorname{zer} T}(w^k)\|_Q^2\\ &\overset{(2.14)}{\le} \|w^k - \Pi^Q_{\Pi_{\mathcal{R}(P)}\operatorname{zer} T}(w^k)\|_Q^2 - \lambda_k(2\alpha_k-\lambda_k)\|w^k-\bar{w}^k\|_Q^2\\ &= \operatorname{dist}_Q^2(w^k, \Pi_{\mathcal{R}(P)}\operatorname{zer} T) - \lambda_k(2\alpha_k-\lambda_k)\|w^k-\bar{w}^k\|_Q^2 \qquad (2.27)\\ &\overset{(2.26)}{\le} \left(1-\tfrac{\kappa}{\xi^2\|P\|}\right)\operatorname{dist}_Q^2(w^k, \Pi_{\mathcal{R}(P)}\operatorname{zer} T),\end{aligned}$$

<escape>证明序列 $(\operatorname{dist}_Q^2(w^k, \Pi_{\mathcal{R}(P)}\text{零点集}\ T))_{k\in\mathbb{N}}$ Q-线性收敛于零。因此，由 (2.27) 可知 $(\|w^k - \bar{}w^k\|_Q)_{k\in\mathbb{N}}$ R-线性收敛于零，因为</escape>

$$\operatorname{dist}_Q^2(w^k, \Pi_{\mathcal{R}(P)}\operatorname{zer} T) \ge \lambda_k(2\alpha_k-\lambda_k)\|w^k-\bar{w}^k\|_Q^2 \ge \kappa\|w^k-\bar{w}^k\|_Q^2.$$

因此，由（2.13）可得

$$\|w^{k+1}-w^k\|_Q^2 = \lambda_k^2\|\bar{w}^k-w^k\|_Q^2 \le \tfrac{\lambda_k^2}{\kappa}\operatorname{dist}_Q^2(w^k, \Pi_{\mathcal{R}(P)}\operatorname{zer} T).$$

结合 $\lambda_k < 2\alpha_k < +\infty$ 有界这一事实（后一个不等式直接由 (2.12) 得出），并且由于右侧的距离 Q-线性收敛到零，$(\|\mathrm{w}^{k+1}-\mathrm{w}^k\|_Q)_{k\in\mathbb{N}}$ R-线性收敛到零。这意味着 $(\|\mathrm{w}^k-\mathrm{w}^\star\|_Q)_{k\in\mathbb{N}}$（其中 $\mathrm{w}^\star \in \Pi_{\mathcal{R}(P)}\mathrm{T}$ 的零点集是序列根据定理 2.4(v) 收敛到的点）也 R-线性收敛到零。注意到 $\xi^2\|\mathrm{P}\| = (1+\zeta\|\mathrm{P}\|)^2$，且 Q 正定，证毕。 □

3 Douglas-Rachford 分裂

在上一节中，针对一类非单调算子给出了预条件近端点方法的收敛结果。本节利用 DRS 与 DRS 和 PPPA之间的等价性，在非单调情形下得到了 的收敛结果。具体而言，我们给出了原始-对偶算子存在斜弱 Minty 解时的收敛结果，以及在限制更严格的共单调情形下的收敛结果，并证明了分段多面体映射的（局部）线性收敛。

3.1 DRS 与 PPPA 的等价性

<escape>在单调情形下，DRS 的收敛性已通过建立其与采用半正定预条件子的预条件近端点方法之间的联系，在 [20] 中得到了研究。在非单调情形下，这种等价性在很大程度上依赖于 [6], [7, §6.9]中引入的</escape>

此框架中，包含问题（P-I）称为原始包含。与原始包含相关联的是对偶包含，其定义为

$$\text{find} \quad y \in \mathbb{R}^n \quad \text{such that} \quad 0 \in T_{\mathrm{D}} y := -A^{-1}(-y) + B^{-1}(y). \tag{D-I}$$

对偶包含的一个关键性质是：原始包含可解，当且仅当对偶包含可解 [7, 命题 6.9。1]。最后，原始-对偶包含表示为

$$\text{find} \quad z = (x, y) \in \mathbb{R}^{2n} \quad \text{such that} \quad \begin{bmatrix} 0 \\ 0 \end{bmatrix} \in T_{\mathrm{PD}} z := \begin{bmatrix} Ax \\ B^{-1}y \end{bmatrix} + \begin{bmatrix} y \\ -x \end{bmatrix}. \tag{PD-I}$$

对于原始-对偶包含，由 [7, 命题 6.92]可知，

$$(x^\star, y^\star) \in \operatorname{zer} T_{\mathrm{PD}} \iff (x^\star, -y^\star) \in \operatorname{gph} A, \ (x^\star, y^\star) \in \operatorname{gph} B,$$

其中，T 的零点集$_{\mathrm{P}}$ = {x* | ∃y* : (x*, y*) ∈ T 的零点集$_{\mathrm{PD}}$}，且 T 的零点集$_{\mathrm{D}}$ = {y* | ∃x* : (x*, y*) ∈ T 的零点集$_{\mathrm{PD}}$}。因此，原始-对偶包含涵盖了原始包含和对偶包含的解，也就是说，(x*, y*) 是 (PD-I) 的解，当且仅当 x* 是 (P-I) 的解且 y* 是 (D-I) 的解。因此，要获得 T$_{\mathrm{P}}$的一个零点，可以从初始点 z^0 = (x^0, y^0) 出发，将 PPPA 应用于原始-对偶算子 T$_{\mathrm{PD}}$，其迭代点为 z^k = (x^k, y^k) 和 z̄^{k} = (x̄^{k}, ȳ^k)。尤其值得注意的是，当预条件子 P 选为等于.....时，将 PPPA 应用于原始-对偶算子 T$_{\mathrm{PD}}$ 会化为 DRS。

$$P = \begin{bmatrix} \gamma^{-1}\mathrm{I}_n & -\mathrm{I}_n \\ -\mathrm{I}_n & \gamma \mathrm{I}_n \end{bmatrix}, \tag{3.1}$$

其中 $\gamma \in \mathbb{R}_{++}$ 是一个严格为正的步长参数。在下一个引理中，PPPA 与 DRS 之间的等价关系将被正式总结如下。

引理 *3.1*（*DRS* 与 *PPPA* 的等价性）。设 $z^0 = (x^0, y^0) \in \mathbb{R}^{2n}$ 并令 $s^0 = x^0 - \gamma y^0$。那么，对于由 *DRS* 生成的任意序列 $(u^k, v^k, s^k)_{k\in\mathbb{N}}$（以 s^0 初始化），用于求解 *(P-I)*，都存在由 *PPPA* 生成的对应序列 $(z^k)_{k\in\mathbb{N}} = (x^k, y^k)_{k\in\mathbb{N}}$、$(\bar{z}^k)_{k\in\mathbb{N}} = (\bar{x}^k, \bar{y}^k)_{k\in\mathbb{N}}$（以 z^0 初始化），用于通过预条件子 *(3.1)* 求解 *(PD-I)*（反之亦然）；对应关系如下。

(i) $s^k = x^k - \gamma y^k$.

(ii) $\bar{x}^k = u^k$.

(iii) $\bar{y}^k = {}^1\!/\!_\gamma(2u^k - s^k - v^k)$.

证明。 参见附录B.1。 □

3.2 斜弱Minty条件下DRS的收敛性分析

尽管 DRS是一种原始算法，引理 3.1表明，它可以被解释为对相关的PPPA原始-对偶算子T_{PD}应用该方法。基于这一等价关系，我们利用第2节中得到的结果，进一步推导DRS的收敛结果。首先，将 PPPA 中的假设 I 所给出的基本假设转化为关于单个算子A 和B的如下条件。

假设 II。 在问题(P-I)中，以下条件成立。

A1 算子 A 和 B 是外半连续的。

A2 对于所选步长，预解算子具有全定义域，即 定义域$J_{\gamma A}$ = 定义域$J_{\gamma B} = \mathbb{R}^n$。

A3 存在非空集合 $\mathcal{S}^\star \subseteq$ 零点集T_{PD}，使得原始-对偶算子 (PD-I) 在 $\mathcal{S}^\star$处具有 V-斜弱 *Minty* 解，其中

$$V := \operatorname{blkdiag}(\beta_{\mathrm{P}}\mathrm{I}_n, \beta_{\mathrm{D}}\mathrm{I}_n) \tag{3.2}$$

以及 $\beta_{\mathrm{P}}, \beta_{\mathrm{D}}, \in \mathbb{R}$ 满足 $[\beta_{\mathrm{D}}]_-[\beta_{\mathrm{P}}]_- < {}^1\!/\!_4$.

表 2： 总结抽象对偶框架中的哪些算子满足明蒂变分不等式（*MVI*）的表，具体取决于 β_P 和 β_D的符号。

	$\beta_D \geq 0$	$\beta_D < 0$
$\beta_P \geq 0$	T_P, T_D, T_{PD}	T_P
$\beta_P < 0$	T_D	-

请注意，这些假设并不十分严格。例如，只需原始-对偶包含存在一个斜弱 Minty 解。此外，可以看出，β_P和 β_D 在假设 II。A3 中的值域涵盖了原始包含（P-I）和对偶包含（D-I）均不满足 Minty 变分不等式（MVI）的情形；这可由以下引理直接推出（另见表 2）。

引理 3.2（原始算子和对偶算子的弱明蒂条件）。 假设 假设 *II*。A3 成立，并设

$$\mathcal{S}_P^\star := \{x^\star \mid \exists y^\star : (x^\star, y^\star) \in \mathcal{S}^\star\} \quad and \quad \mathcal{S}_D^\star := \{y^\star \mid \exists x^\star : (x^\star, y^\star) \in \mathcal{S}^\star\}.$$

那么，原始算子 T_P 在 $\mathcal{S}_P^\star$ 处具有 β_P-弱*Minty*解，而对偶算子 T_D 在 $\mathcal{S}_D^\star$ 处具有 β_D-弱*Minty*解。

证明。 参见附录B.1。 □

引理 3.2的逆命题不成立。例如，考虑线性算子 $A = -\mathrm{I}_n$ 和 $B = \mathrm{I}_n$：在这种情况下，原始算子T_P 和对偶算子 T_D 都化为零映射，而该映射处处都有 0-弱 Minty 解；然而，原始-对偶算子化为 $T_{PD} = \begin{bmatrix} -\mathrm{I}_n & \mathrm{I}_n \\ \mathrm{I}_n & \mathrm{I}_n \end{bmatrix}$，它没有任何 0-弱 Minty 解。

以下定理由定理 2.4 推得，并刻画了DRS 在非单调情形下的收敛性，前提是相应的原始-对偶算子具有 *V*-斜弱 Minty 解。为证明这一结果，我们依据引理 A.5，该引理表明，当假设 II 对DRS成立时，相应预条件近端点算法满足假设 I.A1 至I.A3。请注意，该引理未涵盖假设 I.A4 。为确保满足这一最后的假设，需要适当选择 γ，即DRS 的步长（参见条件（3.3））。

Theore命题 3.3（DRS的收敛性）。假设假设 *II*成立，且 γ 位于区间

$$\left(\frac{2[-\beta_P]_+}{1+\sqrt{1-4\beta_D\beta_P}}, \frac{1+\sqrt{1-4\beta_D\beta_P}}{2[-\beta_D]_+}\right) \tag{3.3}$$

（该区间非空，因为 $[\beta_D]_-[\beta_P]_- < 1/4$）。考虑由 *DRS* 生成的序列 $(u^k, v^k, s^k)_{k\in\mathbb{N}}$，其初始值为 $s^0 \in \mathbb{R}^n$，步长为 γ，松弛参数为 $\lambda_k \in (0, 2\alpha)$，并满足 $\liminf_{k\to\infty} \lambda_k(2\alpha - \lambda_k) > 0$，其中 $\alpha := 1 + \frac{1}{\gamma}\beta_P + \gamma\beta_D > 0$。那么，要么在有限次迭代后达到满足 $u^k = v^k$ 的点，这意味着 $(u^k, 1/\gamma(u^k - s^k)) \in$ 属于 T 的零点集$_{PD}$；要么序列 $(u^k, v^k, s^k)_{k\in\mathbb{N}}$满足以下性质。

(i) $\gamma^{-1}(u^k - v^k) \in Au^k + Bv^k$ *for all* k, *and* $(u^k - v^k)_{k\in\mathbb{N}}$ *converges to zero.*

(ii) $(u^k, 1/\gamma(u^k - s^k))_{k\in\mathbb{N}}$ 的每个极限点（如果存在）都属于零点集 T_{PD}。

(iii) 序列 $(s^k)_{k\in\mathbb{N}}$ 有界，且其极限点属于 $\{(x^\star - \gamma y^\star) \mid (x^\star, y^\star) \in$ 零点集 T_{PD} $\}$.

(iv) 若对所有 k，$\lambda_k(2\alpha - \lambda_k) \geq \kappa > 0$ 一致成立，则对所有 $s^\star \in \{(x^\star - \gamma y^\star) \mid (x^\star, y^\star) \in \mathcal{S}^\star\}$，均有

$$\min_{k=0,1,\ldots,N} \|u^k - v^k\|^2 \leq \frac{\|s^0 - s^\star\|^2}{(N+1)\kappa}.$$

(v) 如果在假设 *II.a3* $\Pi_{\mathcal{R}(P)}\mathcal{S}^\star = \Pi_{\mathcal{R}(P)}T$ 的零点集$_{PD}$中，则 $(s^k)_{k\in\mathbb{N}}$ 收敛到 $\{(x^\star - \gamma y^\star) \mid (x^\star, y^\star) \in T$ 的零点集$_{PD}\}$ 中的某个元素。此外，如果 $J_{\gamma A}$ 和 $J_{\gamma B}$ 是（单值）连续的，则 $(u^k, 1/\gamma(u^k - s^k))_{k\in\mathbb{N}}$收敛到 T 的零点集$_{PD}$中的某个元素。

证明。 回顾一下，DRS 可被视为PPPA的一个实例，正如引理 3.1中所总结的那样。因此，所有断言均可由定理 2.4 应用于原始-对偶包含（PD-I）并取P 为（3.1）中定义的形式而得出。此外，根据引理 A.5 可立即得出，假设 I.A1 至I.A3 均成立，因此只需证明假设 I.A4 在 γ 满足（3.3）时成立。

在整个证明过程中，T_{PD} 如（PD-I）中所述。预条件矩阵 P 有两个不同的特征值，即零和 $\gamma+\gamma^{-1}$，每个特征值的重数均为 n。因此，$P = (\gamma+\gamma^{-1})UU^\top$，其中 U 是 P 的值域的标准正交基，如假设 I.A4中所定义。特别地，我们有

$$U = \frac{1}{\sqrt{\gamma^2+1}}\begin{bmatrix} \mathrm{I} \\ -\gamma\mathrm{I} \end{bmatrix} \quad \text{and} \quad \Pi_{\mathcal{R}(P)} = UU^\top = \frac{1}{\gamma^2+1}\begin{bmatrix} \mathrm{I} & -\gamma\mathrm{I} \\ -\gamma\mathrm{I} & \gamma^2\mathrm{I} \end{bmatrix}.$$

因此，将 $V := \mathrm{blkdiag}(\beta_{\mathrm{P}}\mathrm{I}_n, \beta_{\mathrm{D}}\mathrm{I}_n)$ 代入 (2.2)，可得 $\eta_{\min} = 1 + \frac{1}{\gamma}\beta_{\mathrm{P}} + \gamma\beta_{\mathrm{D}}$。 由 事实 A.3*(ii)*， $\eta_{\min} > 0$i当且仅当 γ 满足 (3.3) 时成立，因此 假设 I。A4 成立。

在已确立定理 2.4 的基本假设得到满足后，接下来需要根据DRS的变量来表示其中的序列。利用引理 3.1中确立的等价关系，可得

$$P(z^k - \bar{z}^k) = (-\tfrac{1}{\gamma}(u^k - v^k), u^k - v^k) \tag{3.4}$$

$$\mathrm{q}_P(z^k - \bar{z}^k) = \tfrac{1}{\gamma}\|u^k - v^k\|^2 \tag{3.5}$$

$$\mathrm{q}_V(P(z^k - \bar{z}^k)) = (\tfrac{1}{\gamma^2}\beta_{\mathrm{P}} + \beta_{\mathrm{D}})\|u^k - v^k\|^2. \tag{3.6}$$

与定理 2.4 的证明类似，考虑采用更新规则 (2.13) 的影序列 $(\mathrm{w}^k)_{k\in\mathbb{N}} := (\Pi_{\mathcal{R}(P)}\,\mathrm{x}^k)_{k\in\mathbb{N}}$ 和 $(\bar{\ }\mathrm{w}^k)_{k\in\mathbb{N}} := (\Pi_{\mathcal{R}(P)}\,\bar{\ }\mathrm{x}^k)_{k\in\mathbb{N}}$。由引理 2.3(ii) 和 (3.4)，有 $\mathrm{w}^k \in \mathscr{D}_{w^k,\bar{}\mathrm{w}^k}$ 当且仅当 $\mathrm{u}^k = \mathrm{v}^k$。因此，若对某个 $\mathrm{k} \in \mathbb{N}$ 有 $\mathrm{w}^k \in \mathscr{D}_{w^k,\bar{}\mathrm{w}^k}$，则算法在有限次迭代后到达满足 $\mathrm{u}^k = \mathrm{v}^k$ 的点；根据 DRS 更新规则，这立即意味着 $(\mathrm{u}^k, 1/\gamma(\mathrm{u}^k - \mathrm{s}^k)) \in \mathrm{T}$ 的零点集$_{\mathrm{PD}}$。在证明的余下部分，我们将考虑 $\mathrm{w}^k \notin \mathscr{D}_{w^k,\bar{}\mathrm{w}^k}$ 的情形；此时，按 (2.12) 定义的 α_k 与本定理中定义的 $\alpha > 0$ 一致。

♠ 3.3*(i)*：根据定理 2.4*(i)* 和（3.4），序列 $(u^k - v^k)_{k\in\mathbb{N}}$ 收敛于零，且

$$\tfrac{1}{\gamma}(u^k - v^k) \in A\bar{x}^k + \bar{y}^k, \quad -(u^k - v^k) \in B^{-1}\bar{y}^k - \bar{x}^k,$$

从而得到所声称的包含关系。

♠ 3.3(ii)：根据引理 3.1，有 $(\bar{\ }\mathrm{x}^k, \bar{\ }\mathrm{y}^k)_{k\in\mathbb{N}} = (\mathrm{u}^k, 1/\gamma(2\mathrm{u}^k - \mathrm{s}^k - \mathrm{v}^k))_{k\in\mathbb{N}}$。因此，由定理 2.4(ii) 可知，$(\mathrm{u}^k, 1/\gamma(2\mathrm{u}^k - \mathrm{s}^k - \mathrm{v}^k))_{k\in\mathbb{N}}$ 的每一个极限点（若存在）都属于 T 的零点集$_{\mathrm{PD}}$。所声称的结果随即由 $(\mathrm{u}^k - \mathrm{v}^k)_{k\in\mathbb{N}}$ 收敛于零这一事实得到。

♠ 3.3*(iii)*：由定理 2.4*(iii)* 直接可得，只需注意到 $\Pi_{\mathcal{R}(P)}\,z^k = \frac{1}{1+\gamma^2}(s^k, \quad -\gamma s \qquad\qquad -\!\!- \qquad\qquad ^k)$.

♠ 3.3(iv)：所述收敛速率由定理 2.4(iv) 得出，因为 $\|\mathrm{P}(\mathrm{z}^k - \bar{\ }\mathrm{z}^k)\|^2 = (1 + \frac{1}{\gamma^2})\|\mathrm{u}^k - \mathrm{v}^k\|^2$，$\mathrm{q}_P(\mathrm{z}^0 - \bar{\ }\mathrm{z}^\star) = \frac{1}{\gamma}\|\mathrm{s}^0 - \mathrm{s}^\star\|^2$ 且 $\|\mathrm{P}\| = \gamma + \frac{1}{\gamma}\sigma$

♠ 3.3(v)：关于 $(\mathrm{s}^k)_{k\in\mathbb{N}}$ 的断言可根据定理 2.4(v) 中关于 $(\Pi_{\mathcal{R}(P)}\mathrm{z}^k)_{k\in\mathbb{N}}$ 的断言立即验证。此外，根据引理 3.1，我们知道 $\bar{\ }\mathrm{z}^k \in (\mathrm{P} + \mathrm{T}_{\mathrm{PD}})^{-1}\mathrm{P}z^k$ 当且仅当 $\bar{\ }\mathrm{x}^k \in \mathrm{J}_{\gamma A}(\mathrm{x}^k - \gamma\mathrm{y}^k)$ 且 $\bar{\ }\mathrm{y}^k \in 1/\gamma(2\mathrm{J}_{\gamma A}(\mathrm{x}^k - \gamma\mathrm{y}^k) - \mathrm{x}^k + \gamma\mathrm{y}^k - \mathrm{J}_{\gamma B}(2\mathrm{J}_{\gamma A}(\mathrm{x}^k - \gamma\mathrm{y}^k) - \mathrm{x}^k + \gamma\mathrm{y}^k))$。因此，若 $\mathrm{J}_{\gamma A}$ 和 $\mathrm{J}_{\gamma B}$ 均为（单值）连续的，则 $(\mathrm{P} + \mathrm{T}_{\mathrm{PD}})^{-1}\mathrm{P}$ 也为（单值）连续的。因此，关于 $(\mathrm{u}^k, 1/\gamma(\mathrm{u}^k - \mathrm{s}^k))_{k\in\mathbb{N}}$ 的结论可由定理 2.4(v) 中关于 $(\bar{\ }\mathrm{x}^k, \bar{\ }\mathrm{y}^k)_{k\in\mathbb{N}} = (\mathrm{u}^k, 1/\gamma(2\mathrm{u}^k - \mathrm{s}^k - \mathrm{v}^k))_{k\in\mathbb{N}}$ 的断言以及 $(\mathrm{u}^k - \mathrm{v}^k)_{k\in\mathbb{N}}$ 收敛到零这一事实推出。 □

请注意，在定理 3.3(v) 中，我们证明了在适当条件下，$(\mathrm{u}^k, 1/\gamma(\mathrm{u}^k - \mathrm{s}^k))_{k\in\mathbb{N}}$ 收敛到 T 的零点集中的某个元素$_{\mathrm{PD}}$。根据 [7, 命题 6.9.2],，这立即意味着 $(\mathrm{u}^k)_{k\in\mathbb{N}}$ 收敛到 T 的零点集中的某个元素$_{\mathrm{P}}$，且 $(1/\gamma(\mathrm{u}^k - \mathrm{s}^k))_{k\in\mathbb{N}}$ 收敛到 T 的零点集中的某个元素$_{\mathrm{D}}$。对于定理 3.3(ii) 中的极限点结果，也有类似的论证。

在附加的V-共单调性假设下，原始-对偶算子的最后迭代收敛速率可针对序列 $(\|u^k - v^k\|^2)_{k\in\mathbb{N}}$得到。该结果直接由定理 2.11 得出，方法是将DRS 视为PPPA的一个实例，这与定理 3.3的证明类似。

推论 *3.4*（最后迭代收敛）。设原始-对偶算子 T_{PD} 是极大 $V := blkdiag(\beta_{\mathrm{P}} I_n, \beta_{\mathrm{D}} I_n)$-共单调的，$[\beta_{\mathrm{D}}]_-[\beta_{\mathrm{P}}]_- < \frac{1}{4}$，且 $zer\, T_{\mathrm{PD}}$ 非空，并且 γ 按照 *(3.3)* 选取。考虑由 *DRS* 从 $s^0 \in \mathbb{R}^n$ 开始、步长为 γ 且松弛参数为 $\lambda_k \in (0, 2\alpha)$ 生成的序列 $(u^k, v^k, s^k)_{k\in\mathbb{N}}$，其中对所有 k，$\lambda_k(2\alpha - \lambda_k) \geq \kappa > 0$ 一致成立，且 $\alpha := 1 + \frac{1}{\gamma}\beta_{\mathrm{P}} + \gamma\beta_{\mathrm{D}} > 0$。那么，对于所有 $s^\star \in \{(x^\star - \gamma y^\star) \mid (x^\star, y^\star) \in zer\, T_{\mathrm{PD}}\}$，以下收敛估计成立：

$$\|u^N - v^N\|^2 \leq \frac{\|s^0 - s^\star\|^2}{(N+1)\kappa} \quad and \quad \|u^N - v^N\|^2 = o(1/(N+1)).$$

最后，考虑分段多面体映射这一类，它显然包括分段线性映射，以及所有分段线性二次凸函数的次梯度，例如多面体集的法锥 [77,§12]。

定义 3.5（分段多面体映射）。如果一个集值映射 $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 的图可以表示为有限个多面体集的并，则称其为分段多面体映射。

下面的定理表明，DRS 对于（可能是）非单调分段多面体映射实现了（局部）线性收敛。证明类似于 [48, 引理 IV。4] 并依赖于这样一个事实：分段多面体映射在其图中的所有点处都是度量次正则的 [25, §3]以及定理 2.13和引理 3.1中的先前结果。

定理 *3.6*（分段多面体映射的 *DRS* 线性收敛）。设 A 和 B 是分段多面体映射，$zer\, T_{\mathrm{PD}}$ 是非空闭凸集，并且假设 *II.a3* 对$\mathcal{S}^\star = zer\, T_{\mathrm{PD}}$成立。根据 *(3.3)* 设定步长 γ，并假设预解算子 $J_{\gamma A}$、$J_{\gamma B}$ 具有全定义域且连续。考虑由 *DRS* 生成的序列 $(u^k, v^k, s^k)_{k\in\mathbb{N}}$，其从 $s^0 \in \mathbb{R}^n$ 开始，步长为γ，松弛参数为 $\lambda_k \in (0, 2\alpha)$，并满足 $\liminf_{k\to\infty} \lambda_k(2\alpha - \lambda_k) > 0$，其中 α 按定理 *3.3* 中的定义。则，$(s^k)_{k\in\mathbb{N}}$ R-线性收敛到 $\{(x^\star - \gamma y^\star) \mid (x^\star, y^\star) \in zer\, T_{\mathrm{PD}}\}$ 中的某个元素，并且 $(\|u^k - v^k\|)_{k\in\mathbb{N}}$ R-线性收敛到零。

证明。 参见附录B.1。 □

4 半单调算子

本节中，我们介绍半单调算子类，并给出若干具有半单调次微分的知名函数类示例。关于充分的条件将于DRS 应用于两个半单调算子的和时，确保其收敛的条件将在第 5 节中给出。

我们首先给出（μ，ρ）半单调算子的定义，它可以看作是（弱）单调算子和余弱单调算子的自然推广。此处采用的半单调术语受到 [63, 定义 2] 中引入的半单调性概念的启发（对应于$\mu = \rho$的情形）。

定义 4.1（半单调性）。设 μ，$\rho \in \mathbb{R}$。若算子 $A\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 在 $(x', y') \in$ 图A 处满足以下条件，则称其为（μ，ρ）-半单调的。

$$\langle x - x', y - y' \rangle \geq \mu\|x - x'\|^2 + \rho\|y - y'\|^2, \qquad for\ all\ (x, y) \in \operatorname{gph} A. \tag{4.1}$$

若算子 A 在所有$(x', y') \in$图 A 处均为(μ, ρ)-半单调，则称算子 A 为(μ, ρ)-半单调。若其图未被严格包含在任何其他(μ, ρ)-半单调算子的图中，则称其为极大(μ, ρ)-半单调。

由于这类算子拥有额外的自由度，因此能够涵盖文献中广泛使用的其他算子类。下面列出其中最重要的关联。

注 4.2（与其他类型算子的关系）。

(i) (0, 0)-半单调性等价于A 的单调性（以及其逆A^{-1}的单调性）。

(ii) （μ, 0）-半单调性等价于 μ-单调性，也称为 |μ|-弱单调性[77, 例 12。28] 当 $\mu < 0$ 时，以及当 $\mu > 0$时称为强单调性。

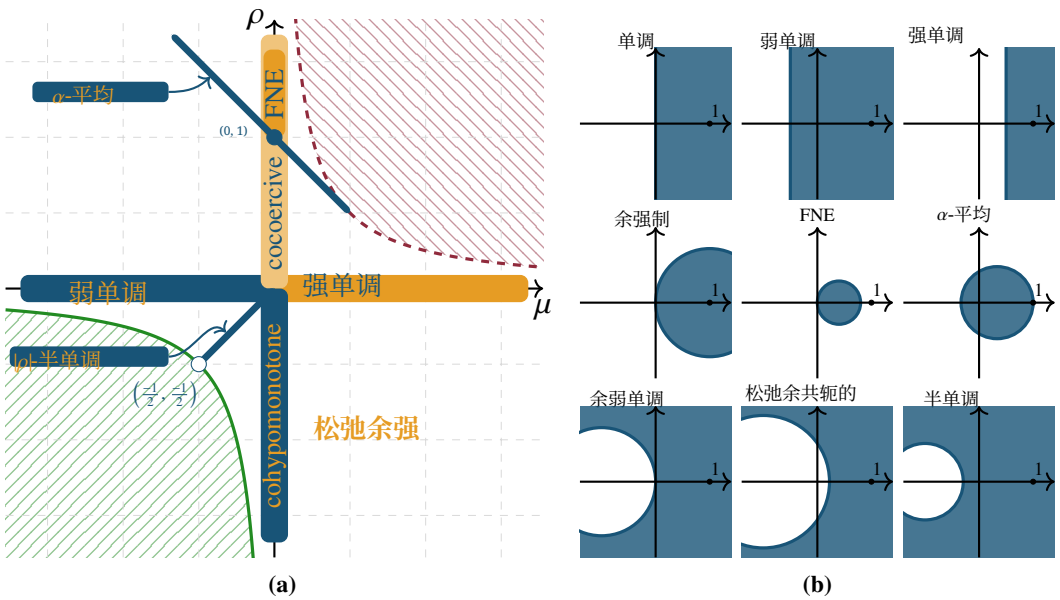

图 5：（左）半单调性与其他类型算子之间的关系。绿色斜线区域表示所有算子都满足 (μ,ρ)-半单调性的区域；红色斜线区域表示不存在任何 (μ,ρ)-半单调算子的区域（见 命题 *4.3*）。（右）不同算子类的缩放相对图（*SRG*）。例如，顶行展示了一个 μ-单调算子的 *SRG*，参见 *[78,* 命题 *3.3* 和定理 *3。5]*。右下图展示了一个 (μ,ρ)-半单调算子，其中 μ 为负，且 ρ。有关 *SRG* 及类似几何示意图的更多详情，请参阅 *[29, 37, 78]*。

(iii) 类似地，$(0,\rho)$-半单调性等价于 ρ-共单调性，也称为当 $\rho<0$ 时的 $|\rho|$-共弱单调性，以及当 $\rho>0$ 时的 ρ-余强单调性。

(iv) 算子 T 在点 $(x^\star, 0)$ 处的 $(0,\rho)$-半单调性 $\in$ 图T 等价于 T 在 $x^\star \in$ 零点集 T 处具有 ρ-弱 Minty 解，有时也称为围绕 $x^\star$ [38, 定义 2.1] 中的 ρ-星余强单调性。

(v) 当 $\mathrm{n}\,\mu = \rho < 0$ 时，我们对 (μ,ρ)-半单调性的定义等价于 $|\rho|$-半调性定义，该定义见 [63, 定义 2]。

(vi) 算子的 α-平均性等价于命题 4.35(iv) [9, 中的 $(1-1/(2-2\alpha), 1/(2-2\alpha))$-半单调性]。此外，算子是 firmly nonexpansive 当且仅当它是 $1/2$-平均的 [9, , 这在任意满足 $\rho \geq 1$ 的 $(0,\rho)$-半单调算子上均成立。],备注 4.34(iii)

(vii) 当 $\mu > 0$ and $\rho < 0$ 时，(μ,ρ)-半单调性的定义与 $(-\rho, \mu)$-松弛余强单调性的定义 [45, 定义 2.3] 相同。□

图 5展示了不同算子类的缩放相对图（SRG） [78]，并给出了它们之间联系的图示。在图 5a中，绿色（分别为红色）阴影区域表示所有算子（分别为没有算子）满足 (μ,ρ)-半单调性的区域，这可由以下命题推出。

命题 4.3（半单调算子的存在性）。设$\mu,\rho \in \mathbb{R}$。以下结论成立。

(i) 若 $[\mu]_-[\rho]_- \geq \frac{1}{4}$，则所有算子 $A: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 均满足(μ, ρ)-半单调 *ity.*

(ii) 若 $[\mu]_+[\rho]_+ > \frac{1}{4}$，则不存在任何满足$(\mu, \rho)$-半单调的算子 $A: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *e.*

(iii) 若 $[\mu]_+[\rho]_+ = \frac{1}{4}$，则算子 $A: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$

1. 若存在某个 c，使得 (μ, ρ)-半单调算子 *iff* 满足 图 $A \subseteq$ 图 T，其中 $T: x \mapsto 2\mu x + c$。
2. 若存在某个 c，使得 $A: x \mapsto 2\mu x + c$，则该算子是极大(μ, ρ)-半单调的 *iff*。

证明。 参见附录 B.2。 □

接下来，我们将考察（极大）半单调算子类的基本性质和微积分法则。例如，它们的逆仍属于同一算子类，但 μ 和 ρ 的作用互换。此外，以下命题分析了半单调算子的缩放和平移。

命题 4.4 （求逆、平移与缩放）。 设算子 $A\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 在(μ,ρ)-半单调 *[于$(x',y') \in$ 图 A]*（极大）。则以下结论成立。

(i) 逆算子 A^{-1} 在(ρ, μ)-半单调 *[于$(y', x') \in$ 图 A^{-1}]*（极大）。

(ii) 对所有 $\alpha \in \mathbb{R}_{++}$和 $u,\ w \in \mathbb{R}^n$，算子 $T(x) := w + \alpha A(x + u)$ 在$(\alpha\mu,\ \alpha^{-1}\rho)$-半单调*[于$(x' - u, w + \alpha y')$]*（极大）。

线性算子是半单调算子的一种特殊情况。对于这些算子，μ 与 ρ之间存在相互作用，以下命题对此进行了概述。该结果推广了 [10, 命题 5。1] 对于μ-单调算子和 ρ-共单调算子。此外，还给出了关于恒等算子的倍数的一个推论，其中包括其极大性。

命题 4.5 （线性算子）。 设 $A \in \mathbb{R}^{n\times n}$ 且$\mu,\rho \in \mathbb{R}$。则，A 是（μ,ρ）-半单调，当且仅当$\frac{1}{2}(A + A^\top) - \rho A^\top A \succeq \mu I$。

证明。考虑 $x \in \mathbb{R}^n$。由 (μ, ρ)-半单调性可得 $\langle x, Ax\rangle \geq \mu\|x\|^2 + \rho\|Ax\|^2 \Leftrightarrow \langle x, (A - \mu I - \rho A^\top A)x\rangle \geq 0 \Leftrightarrow \left\langle x, \left(\frac{1}{2}(A + A^\top) - \mu I - \rho A^\top A\right)x\right\rangle \geq 0 \Leftrightarrow \frac{1}{2}(A + A^\top) - \mu I - \rho A^\top A \succeq 0$。 □

推论 4.6 （单位算子的倍数）。 若 $A = \alpha I_n$ 对某个 $\alpha \in \mathbb{R}$成立，则对于任意 $c \in \mathbb{R}$，A 都是（α（$1 - c\alpha$），c）-半单调的。此外，除非 $\alpha < 0$ 且 $c = 1/2\alpha$，否则 A 是极大的。

证明。 参见附录 B.2。 □

下一个命题讨论两个半单调算子之和及其平行和，A 和B。请注意，所得算子的半单调性参数总是同时涉及A 和B的半单调性参数之和与平行和。不出所料，交换A和B不会影响结果，因此交换性质成立。

命题 4.7 （和与平行和）。 设算子 $A\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 是（μ_A，ρ_A）-半单调的，*[在（x'_A，y'_A）$\in$图 A]* 处；设算子 $B\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 是（μ_B，ρ_B）-半单调的，*[在（x'_B，y'_B）$\in$图 B]*处。则有

(i) *[如果 $x' = x'_A = x'_B$，则]* $A + B$ 是 $(\mu_A + \mu_B, \rho_A \square \rho_B)$-半单调 *[于 $(x', y'_A + y'_B)$]*，如果 $\rho_A + \rho_B > 0$或 $\rho_A = \rho_B = 0$。

(ii) *[若 $y' = y'_A = y'_B$，则]* $A \square B$ 在$(\mu_A \square \mu_B, \rho_A + \rho_B)$-半单调 *[于$(x'_A + x'_B, y')$]*，若$\mu_A + \mu_B > 0$或$\mu_A = \mu_B = 0$。

证明。 参见附录 B.2。 □

结合前述关于和的命题以及推论 4.6中对恒等算子倍数的刻画，可以得到半单调算子与恒等算子倍数之和的半单调性，如命题 4.8*(ii)*所示。然而，在特定条件下，利用现有的恒等算子倍数结构，可以得到更强的等价关系。这一结果见命题 4.8*(i)*，并推广了 [10, 引理 2.6 和 2.8]。

命题 4.8 （与恒等算子的和）。 设 μ_A，$\rho_A \in \mathbb{R}$。考虑算子 $A\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 并定义 $T := A + \alpha I$ 其中 $\alpha \in \mathbb{R}$。则有以下结论。

(i) 若 $1 + 2\rho_A\alpha > 0$，则 T 在$\mu_A + \alpha(1+\rho_A\alpha)1+2\rho_A\alpha$，$\underbrace{\frac{\rho_A}{1+2\rho_A\alpha}}$-半单调 *[于$(x',\ y' + \alpha x') \in$图 T]*（最大）当且仅当 A 在$(x',\ y') \in$图 A*]*处是$(\mu_A,\ \rho_A)$-半单调（最大）。

(ii) 若 $1 + 2\rho_A\alpha \leq 0$，则对于所有 $c > -\rho_A$，只要 A 在$(x',\ y') \in$图 A*]*处是$(\mu_A,\ \rho_A)$-半单调，T 就在$(x',\ y' + \alpha x') \in$图 T*]* 处$(\mu_A + \alpha(1 - c\alpha),\ \rho_A \square c)$-半单调。

证明。 参见附录 B.2。 □

利用前面的结果，我们现在形式化地说明算子 T 与另一（极大）单调算子 A之间的关系，其中前者定义为后者的函数。在实际应用中，可以利用这一结果检验给定算子的（极大）半单调性，或由给定的单调算子生成半单调算子。

推论 4.9 （与单调算子的联系）。 设 $\xi, \nu \in \mathbb{R}$。考虑算子 $A\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$，并定义 $T := (A + \xi\mathrm{id})^{-1} + \nu\mathrm{id}$。则有以下结论。

(i) 若 $1 + 2\xi\nu > 0$，则 T 是（极大） $\left(\frac{\nu(1+\xi\nu)}{1+2\xi\nu}, \frac{\xi}{1+2\xi\nu}\right)$-半单调的，当且仅当 A 是（极大）单调的。

(ii) 若 $1 + 2\xi\nu \leq 0$，则当 A 单调时，对于所有 $c > -\xi$，T 都是 $\big(\nu(1 - c\nu), \xi \square c\big)$-半单调的。

证明。 应用 命题 4.4*(i)*和 4.8 可得，A是（极大）单调的，当且仅当 $A+\xi\mathrm{id}$ 是（极大） $(\xi,0)$-半单调的，当且仅当 $(A+\xi\mathrm{id})^{-1}$ 是（极大） $(0,\xi)$-半单调的。再应用命题 4.8即可得到所述结论。 □

当 $\nu = 0$时，推论 4.9*(i)* 可归结为一个已知的等价关系：算子的共单调性与其 Yosida 正则化的单调性等价，参见 [16,命题 6.9。3]。以下结果是 推论 4.9*(i)* 的直接推论，注意到对于从（4.2）中选取的 ξ 和 ν，有 $1 + 2\xi\nu = 1/\sqrt{1-4\rho\mu} > 0$。 该结果推广了 [63, 定理 2]，，该定理仅涵盖 $\mu = \rho$ 的情形。

推论 4.10. 集合值算子 $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 是（极大）(μ,ρ)-半单调的，且 $\rho\mu < \frac{1}{4}$ 当且仅当 $A := (T^{-1} - \nu\mathrm{id})^{-1} - \xi\mathrm{id}$ 是（极大）单调算子，其中

$$\xi = \frac{\rho}{\sqrt{1-4\mu\rho}} \quad \textit{and} \quad \nu = \frac{2\mu}{1+\sqrt{1-4\mu\rho}}. \tag{4.2}$$

鉴于上述与极大单调算子的联系，当 $\rho < \frac{1}{4}\mu$ 时，(μ,ρ)-半单调算子 A 是极大的，若 $(A - \nu\mathrm{id})^{-1}$ 连续（因为任意连续单调映射都是极大单调的）。此外，以下命题表明，极大 (μ,ρ)-半单调算子的图是闭集。

命题 4.11 （半单调算子的外半连续性）。 设 $A\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 是极大(μ, ρ)-半单调算子，且满足 $\rho\mu < \frac{1}{4}$ 或 $[\mu]_+[\rho]_+ = \frac{1}{4}$。则 A 是外半连续的。

证明。 见附录 B.2。 □

本节最后我们将证明，若恰当选择步长，则半单调算子的预解算子具有全定义域且是 Lipschitz连续的。该结果推广了 [10, 命题 2。17] 中关于共单调算子的结果，并推广了 [63, 定理 3],中仅适用于 $\mu = \rho$ 的情形。

命题 4.12 （半单调算子的预解算子）。 设 $A\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 是一个（μ,ρ）-半单调算子，其中 $\rho\mu < \frac{1}{4}$ 或 $[\mu]_+[\rho]_+ = \frac{1}{4}$。设

$$\gamma \in \left(0, \frac{1+\sqrt{1-4\rho\mu}}{2[-\mu]_+}\right). \tag{4.3}$$

那么，以下结论成立。

(i) 预解算子 $J_{\gamma A}$ 具有全定义域，当且仅当 A 是极大的。

(ii) 预解算子 $J_{\gamma A}$ 是 *Lipschitz*连续的，其常数为 $L(\gamma) = \frac{|1+2\rho\gamma^{-1}| + \sqrt{\cdots}}{1-4\mu\rho}$。

证明。 参见附录 B.2。 □

4.1 具有半单调子微分的函数示例

接下来，我们将介绍各种函数类，这些函数的（极限）次微分属于半单调算子类。我们首先研究具有上曲率和/或下曲率的函数类，其定义如下。

定义 4.13（上曲率和下曲率）。设 $\sigma, l \in \mathbb{R}$ 并考虑一个正常下半连续函数 $f\colon \mathbb{R}^n \to \mathbb{R}$。则，

(*i*)（下曲率）若 $\varphi_\sigma := f - \frac{\sigma}{2}\|\cdot\|^2$ 是凸函数，则称 f 具有模数为 σ 的下曲率（记为 $f \in \mathscr{F}_\sigma(\mathbb{R}^n)$）。(*ii*)（上曲率）若 f 取有限值且 $\varphi^\ell := \frac{l}{2}\|\cdot\|^2 - f$ 是凸函数，则称 f 具有模数为 l 的上曲率（记为 $f \in \mathscr{F}^\ell(\mathbb{R}^n)$）。

同时具有上曲率和下曲率的函数记为 $\mathscr{F}^\ell_\sigma(\mathbb{R}^n)$。

许多传统的函数类都属于这一曲率框架。例如，具有正（分别为负）下曲率的函数通常称为强凸（分别为弱凸）函数。另一个值得注意的例子是 l-光滑函数，它们具有上曲率 l 和下曲率 $-l$ [84, 引理 2。1]。

下一个命题根据函数的曲率，给出其次微分的半单调性性质。此外，还给出了一个关于具有上曲率函数的逐点最小值的推论。

命题 4.14（具有上曲率和下曲率的函数）。设 $\sigma, l \in \mathbb{R}$。则以下结论成立。

(i) 若 $f \in \mathscr{F}_\sigma(\mathbb{R}^n)$，即 f 是 σ-凸的，则 ∂f 是 $(\sigma, 0)$-半单调的。

(ii) 若 f 连续且 $f \in \mathscr{F}^l(\mathbb{R}^n)$，其中 $l < 0$，则 ∂f 是 $(0, \frac{1}{l})$-半单调的。

(iii) 若 f 连续且 $f \in \mathscr{F}^l(\mathbb{R}^n)$，其中 $l \geq 0$，则对于所有 $\alpha > l$ 且 $c > \frac{1}{\alpha - l}$，$-\partial f$ 都是 $(\alpha(1-c\alpha), c \square \frac{1}{l})$-半单调的。

若 $f \in \mathscr{F}^l_\sigma(\mathbb{R}^n)$，则，$\sigma \leq l$，$f$ 连续可微，且满足以下性质。

(iv) 若 $l + \sigma > 0$，则 ∇f 是 $(\sigma \square l, \frac{1}{l+\sigma})$-半单调的。

(v) 若 f $l + \sigma \leq 0$，∇f 是 $(\sigma(1-c\sigma),$ 对于满足 $1 + c(l - \sigma) > 0$ 的任意常数 c，均为 $(\sigma(1-c\sigma), c \square \frac{1}{l-\sigma})$-半单调的。

证明。参见附录 B.3。 □

推论 4.15 (逐点最小值)。设 f 是有限值连续函数 $f_i \in \mathscr{F}^{l_i}(\mathbb{R}^n)$ 的逐点最小值，其中 $l_{f_i} \in \mathbb{R}$，即 $f(x) = \min\{f_1(x), \ldots\ldots, f_N(x)\}$。令 $\bar{l} = \max_i l_{f_i}$。则满足以下性质。

(i) If $\bar{l} < 0$, *then* ∂f *is* $\left(0, \frac{1}{\bar{l}}\right)$-*semimonotone.*

(ii) 若 $\bar{l} \geq 0$，则对所有 $\alpha > \bar{l}$，且 $c \geq \frac{1}{\alpha - \bar{l}}$，$\partial f$ 是 $\left(\alpha(1-c\alpha),\ c \square \frac{1}{l}\right)$-半单调的。

证明。参见附录 B.3。 □

函数最基本的运算之一是下确界卷积，接下来的命题将讨论这一主题。为得到该结果，我们首先证明，两个半单调算子的下确界卷积的次微分属于这两个函数的次微分的平行和（见命题 4.16*(i)*），然后使用 命题 4.7*(ii)*。

命题 4.16（下确界卷积）。对于 $i = 1, 2$，设 $f_i\colon \mathbb{R}^n \to \mathbb{R}$ 是正常下半连续函数，其次微分为 (μ_i, ρ_i)-半单调，其中 $\mu_i,\ \rho_i \in \mathbb{R}$ 满足 $\mu_1 + \mu_2 > 0$ 或 $\mu_1 = \mu_2 = 0$。将它们的下确界卷积记为

$$\varphi : s \mapsto \min_{w \in \mathbb{R}^n}\{f_1(w) + f_2(s - w)\}.$$

假设对每个 $r_1 \geq 0$ 和 $\bar{s} \in \mathbb{R}^n$，存在 $r_2 > 0$，使得集合

$$\left\{(w_1, w_2) \in \mathbb{R}^n \;\middle|\; \textstyle\sum_{i=1}^2 f_i(w_i) \leq r_1,\ \left\|\sum_{i=1}^2 w_i - \bar{s}\right\| \leq r_2\right\} \tag{4.4}$$

是有界的。于是，以下结论成立。

(i) $\partial\varphi(\bar{s}) \subseteq (\partial f_1 \square \partial f_2)(\bar{s})$。

(ii) $\partial\varphi$ 是$(\mu_1 \square \mu_2, \rho_1 + \rho_2)$-半单调的。

证明。 见附录 B.3。 □

最后，我们举一个例子，说明在所谓的交互主导条件 [40]下，极小极大优化问题的相关鞍点算子是半单调的。

命题 4.17 (鞍点算子)。 假设 $\varphi: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ 二阶连续可微，且关于某个 $\sigma \in \mathbb{R}$是凸-凹的，即对所有 $z \in \mathbb{R}^{n+m}$，有 $\nabla^2_{xx}\varphi(z) \succeq \sigma \mathrm{I}$ 且 $-\nabla^2_{yy}\varphi(z) \succeq \sigma \mathrm{I}$。此外，对某个 $\alpha > -\frac{1}{\delta}$ 和 $\delta \in (0, 1/[-\sigma]_+)$，$\alpha$ 满足交互主导条件 *[40,* 定义 *1]*，即对所有 $z \in \mathbb{R}^{n+m}$

$$\nabla^2_{xx}\varphi(z) + \nabla^2_{xy}\varphi(z)\Big(\tfrac{1}{\delta}\mathrm{I} - \nabla^2_{yy}\varphi(z)\Big)^{-1}\nabla^2_{yx}\varphi(z) \succeq \alpha \mathrm{I},$$
$$-\nabla^2_{yy}\varphi(z) + \nabla^2_{yx}\varphi(z)\Big(\tfrac{1}{\delta}\mathrm{I} + \nabla^2_{xx}\varphi(z)\Big)^{-1}\nabla^2_{xy}\varphi(z) \succeq \alpha \mathrm{I}.$$

那么，对于相应的鞍点算子 $T_\varphi(z) := (\nabla_x\varphi(z), -\nabla_y\varphi(z))$，以下结论成立。

(i) 若 $\alpha < \frac{1}{\delta}$，则 T_φ 是$\left(\frac{\alpha}{1-\alpha\delta}, \frac{\delta}{1-\alpha\delta}\right)$-半单调。

(ii) 若 $\alpha \geq 0$，则 T_φ 是 $(0, -\delta)$-半单调。

证明。 参见附录 B.3。 □

5 半单调算子的 Douglas–Rachford 分裂

在实际应用中，可能难以确定是否满足DRS 的假设 II 。为解决DRS这一问题，我们给出一组关于两个半单调算子之和的充分条件，以保证DRS 收敛。具体而言，对于问题（*P-I*）中的底层算子*A*和B，我们采用以下假设。

假设 III. 在问题 (P-I)中，存在参数 μ_A，ρ_A，μ_B，$\rho_B \in \mathbb{R}$ 以及一个非空集合 $\mathcal{S}^\star \subseteq$零点集 T_{PD}，使得对每个 $(x^\star, y^\star) \in \mathcal{S}^\star$，以下条件成立。

A1满足以下条件之一：$\mu_A + \mu_B > 0$ 或 $\mu_A = \mu_B = 0$；满足 $\rho_A + \rho_B > 0$ 或 $\rho_A = \rho_B = 0$ 之一；并且$[\mu_A \square \mu_B]_- [\rho_A \square \rho_B]_- < \frac{1}{4}$。

A2 算子 A 在(μ_A, ρ_A)-半单调于$(x^\star, -y^\star) \in$图 A。

A3 算子 B 在(μ_B, ρ_B)-半单调于$(x^\star, y^\star) \in$图 B。

在这些假设下，如下文引理 *5.1*(i)所示，原始-对偶算子T_{PD} 存在斜弱 Minty 解。此外，若算子A 和B如假设 IV所述为极大半单调，则原始-对偶算子为极大余单调。引理 5.1 的证明推迟至附录 B.4。

假设 IV。在问题(P-I)中，零点集$(A+B)$ 非空，且存在参数$\mu_A, \rho_A, \mu_B, \rho_B \in \mathbb{R}$ 满足假设 *III*。A1 使得算子 A 是极大(μ_A, ρ_A)-半单调的，且算子 B 是极大(μ_B, ρ_B)-半单调的。

引理 5.1 (原始-对偶算子)。假设假设 *III* 或假设 *IV* 成立，并令 V 如(3.2)所定义，其中

$$\beta_{\mathrm{P}} = \rho_A \square \rho_B \quad \textit{and} \quad \beta_{\mathrm{D}} = \mu_A \square \mu_B. \tag{5.1}$$

那么，以下结论成立。

(i) 若假设 *III* 成立，则 T_{PD} 在 $\mathcal{S}^\star$处有 V-斜弱 *Minty* 解。

（*ii*） 如果假设 *IV*成立，则 T_{PD} 是极大 V-余单调的。

表 3： 值域： γ 针对*DRS* 与半单调算子。

		$\mu_A = \mu_B = 0$	$\mu_A + \mu_B > 0$	
			$\mu_A\mu_B \geq 0$	$\mu_A\mu_B < 0$
$\rho_A = \rho_B = 0$		$\gamma \in (0, +\infty)$		$\gamma \in \left(0, -\frac{1}{\mu_A \square \mu_B}\right)$
$\rho_A + \rho_B > 0$	$\rho_A\rho_B \geq 0$			$\gamma \in (0, \gamma_+)$
	$\rho_A\rho_B < 0$	$\gamma \in \left(-(\rho_A \square \rho_B), +\infty\right)$	$\gamma \in (\gamma_-, +\infty)$	$\gamma \in (\gamma_-, \gamma_+)$

现在我们可以陈述本节的主要结果：下面给出了使足以保证收敛的充分条件DRS，这些条件通过算子 A 和 B 的半单调性性质给出。证明推迟至 附录 B.4。

定理 5.2 (DRS 在半单调性下的收敛)。假设

(*i*) 要么 假设 *II*。A1，假设 *II*。A2 并且 假设 *III* 成立，

(ii) 或假设 *IV* 成立。

令 γ 根据 表 *3* 进行选择，其中

$$\gamma_- := -\frac{2(\rho_A \square \rho_B)}{1+\sqrt{1-4(\mu_A \square \mu_B)(\rho_A \square \rho_B)}} \quad \textit{and} \quad \gamma_+ := -\frac{1+\sqrt{1-4(\mu_A \square \mu_B)(\rho_A \square \rho_B)}}{2(\mu_A \square \mu_B)}. \tag{5.2}$$

然后， $\alpha := 1 + \frac{1}{\gamma}(\rho_A \square \rho_B) + \gamma(\mu_A \square \mu_B) > 0$。考虑由 *DRS* 生成的序列 $(u^k, v^k, s^k)_{k\in\mathbb{N}}$，其从 $s^0 \in \mathbb{R}^n$ 开始，步长为 γ，松弛参数为 $\lambda_k \in (0, 2\alpha)$，满足 $\liminf_{k\to\infty} \lambda_k(2\alpha - \lambda_k) > 0$。那么，定理 *3.3* 的所有结论均成立。此外，如果假设 *IV* 成立，并且对所有 k 一致地有 $\lambda_k(2\alpha - \lambda_k) \geq \kappa > 0$，那么推论 *3.4* 中的收敛估计也成立。

定理 5.2 为 DRS 的收敛提供了一个通用框架，统一并拓展了关于松弛 Douglas/Peaceman Rachford 分裂方法的现有收敛结果。我们给出几个例子。

注 5.3（与现有理论的比较）。

(i) 当 A 和 B 均为极大（0，0）-半单调算子，即极大单调算子时，可以恢复经典 Douglas–Rachford 分裂的收敛结果，参见例如 [54, 31]。

(ii) 当 A 和 B 分别为 μ_A 和 μ_B 单调时，我们恢复了文献中的若干已有结果。例如，在强单调情形下，其中 $\mu := \mu_A = \mu_B > 0$，我们恢复了 [59]，中的收敛结果，该结果表明 DRS 在 $\lambda \in (0, 2 + \gamma\mu)$ 范围内收敛，而非传统的 $\lambda \in (0,2)$。另一方面，当 $\mu_A + \mu_B > 0$ 时，表 3 中给出的步长范围与 [21, 定理 4.5(ii)] 相符。类似地，我们恢复了 $\rho_A + \rho_B > 0$ 中关于一个 ρ_A 与一个 ρ_B 共单调算子的和的结果，见 [8, 推论 5.5(ii)]。

(iii) 在优化情形下，近期已有研究证明，对于结构化的非光滑非凸优化问题，DRS 可以全局收敛 [53, 84] 。这类问题的目标函数拆分为一个 Lipschitz 可微函数 $f: \mathbb{R}^n \to \mathbb{R}$ 与一个正常下半连续函数 $g: \mathbb{R}^n \to \mathbb{R}$之和，即 $A = \nabla f$ 和 $B = \partial g$。这些结果依赖于证明迭代点在所谓的 Douglas-Rachford 包络 [65], 上单调下降；确定可接受步长的范围时，仅利用了 Lipschitz 可微函数 f 的性质。例如，考虑 f 为 l-光滑且 σ-凸的情形，其中 $l > 0$ 且 $\sigma \in [-l, 0)$。此时， [84] 证明，只要极小点集合 $\arg\min_x f(x) + g(x)$ 非空，当 $\lambda \in (0, 2)$ 且 $\gamma < \min\{\frac{\lambda-2}{2\sigma}, \frac{1}{l}\} < -\frac{1}{\sigma}$ 时，DRS 就会全局收敛。

注意，当给定 g 的额外半单调性性质时，可以通过定理 5.2 为 γ 获得更大的步长范围。为此，首先注意在此情形下，对于某个 $\mu_A < 0$ 和 $\rho_A \geq 0$，A 是 (μ_A, ρ_A)-半单调的（见命题 4.14(iv)）。因此，假设 ∂g 在 $(x^\star, y^\star)$ 处是 (μ_B, ρ_B)-半单调的，且 $\in \operatorname{gph} \partial g$，其中 $\mu_A + \mu_B > 0$ 且 $\rho_B \geq 0$，则由定理 5.2 可知，DRS 对 $\gamma \in (0, \gamma_+)$ 收敛。这是一个更大的步长范围，因为 $\gamma_+ > -\frac{1}{\sigma}$。 □

6 示例问题

本节给出若干超越标准单调情形的DRS 示例问题。本节所呈现示例的详细内容可见于附录C。

我们首先将 DRS 应用于一个简单的线性包含问题。由于两个具体原因，这种设定尤为值得关注。首先，该示例展示了定理 3.3 中给出的步长范围的紧致性（见示例 6.1*(ii)*）。其次，它说明我们的理论能够涵盖原始包含、对偶包含和原始-对偶包含均非单调的情形（见示例 6.1*(iii)*）。

示例 6.1 (鞍点问题)。 考虑寻找线性包含问题的驻点，其中的特定结构自然出现在鞍点问题中 [33, §1.4。1]

$$0 \in T_{\mathrm{P}}x = \begin{bmatrix} b & a \\ -a & b \end{bmatrix} x = \underbrace{\begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}}_{A} x + \underbrace{\begin{bmatrix} b & 0 \\ 0 & b \end{bmatrix}}_{B} x, \tag{6.1}$$

其中 a、$b \in \mathbb{R} \setminus \{0\}$。特别地，包含问题$0 \in T_{\mathrm{P}}x$的任一解都是 $f(x_1, x_2) := ax_1x_2 + b2(x_1^2 - x_2^2)$ 的极小极大解（当 $b > 0$ 时），而当 $b < 0$ 时则是极大极小解。对于这个问题，以下断言成立。

(i) 通过考察算法算子的谱半径可知，由固定松弛参数 λ 下的 DRS 生成的序列 $(u^k, v^k, s^k)_{k\in\mathbb{N}}$ 收敛，当且仅当落在区间内

$$0 < \lambda < \bar{\lambda} := 2\frac{(a^2\gamma + b)(1 + \gamma b)}{\gamma(a^2 + b^2)}, \tag{6.2}$$

该集合非空，当且仅当$b > 0$ 或

$$b < 0,\ a^2 \neq b^2 \text{ and } \gamma \in \left(\min\{-\tfrac{1}{b}, -\tfrac{b}{a^2}\}, \max\{-\tfrac{1}{b}, -\tfrac{b}{a^2}\}\right).$$

验证这一收敛结果的数值实验见
见 图 6。

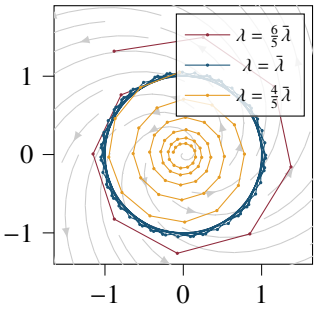

图 6： 通过对 例 *6.1* 应用 *DRS* 生成的序列 u^k，其中 $= 2$，$b = -1$，步长 $\gamma = \frac{1}{2}$。

(ii) 定理 3.3 的结果是紧的，因为当在 (2.1) 中将向量 v限制为 $\mathcal{R}(P)$（见 定义 2.1 下方的注释）时，松弛参数 λ 在 (6.2) 中的整个取值范围均被覆盖。

*(iii)*a 与 b 的值域　例 6.1*(i)* 中满足以下条件的情形：DRS 收敛，而原问题、对偶问题和原始—对偶问题均不具有单调性。

此外，附录C 对定理 5.2的紧致性进行了分析。在该分析中，根据命题 4.5$_A$，$_B$，ρ_A，ρ_B 选取半单调性参数，使得定理 5.2中的松弛参数达到最大值。请注意，尽管A 显然是 (0，0)-半单调的，但最大松弛参数通常在其他μ_A 以及 ρ_A取值下才能达到。

□

接下来，我们将关注一个非光滑优化问题。该例凸显了基本假设的普适性，因为据作者所知，先前关于DRS 的研究均未涵盖该问题中DRS的收敛性。

例 6.2 （非光滑优化）. 考虑以下可分离的非光滑优化问题

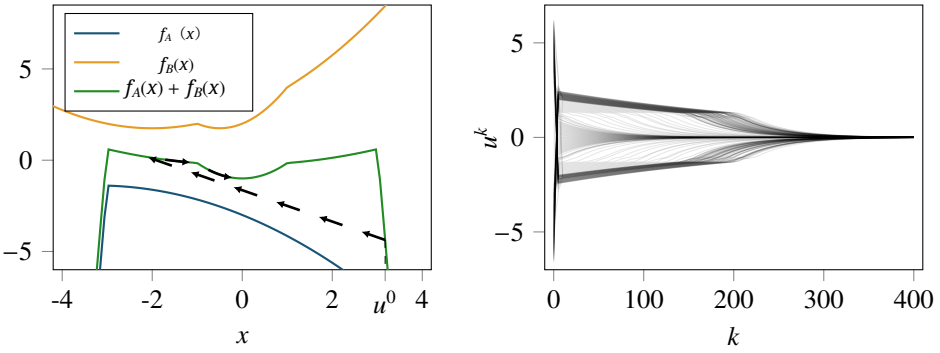

图 8：（左）例 6.4 中所涉及函数的可视化。此外，还使用黑色箭头标出了序列 $(u^k)^{400}_{k=0}$，该序列通过对 $0 \in \partial f_A(x) + \partial f_B(x)$ 应用 DRS 获得，步长为 $\gamma = 11/60 \in (1/6, 1/5)$，松弛参数为 $\lambda = 9/10(2 - 4/15\gamma - 12\gamma/5) \approx 0.095$。在计算（多值）预解算子 $J_{\gamma A}$ 和 $J_{\gamma B}$ 时，以相同概率对每个元素进行采样。（右）对 200 次此类实验所得序列 $(u^k)^{400}_{k=0}$ 的可视化，其中初始值 s^0 在区间 $[-5, 5]$ 内均匀分布。

$$\underset{x \in \mathbb{R}}{\text{minimize}} \quad f_A(x) + f_B(x)$$

其可视化结果见图 7，其中

$$f_A(x) := \begin{cases} 3x^2 - 3x + 8, & \text{if } x \in [-1, 1], \\ -\frac{1}{2}x^2 - 3x + \frac{23}{2}, & \text{otherwise}, \end{cases}$$

$$f_B(x) := \begin{cases} 2x^2 - x, & \text{if } x < -1, \\ x^2 + 3x + 5, & \text{if } x \in [-1, 2], \\ x^2 + 15x - 19, & \text{if } x > 2. \end{cases}$$

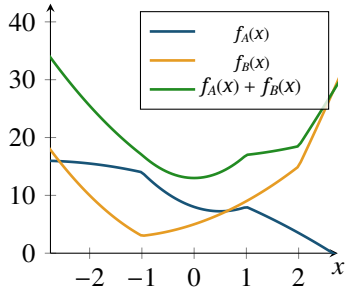

图 7： 示例 6.2 中涉及的（非光滑）函数的可视化。

对于包含关系 $0 \in T_P x = \partial f_A(x) + \partial f_B(x)$，以下结论成立。

(i) 假设 II。A1 和假设 II。A2 在 $\gamma \neq 1$ 时成立。
(ii) 假设 III 对 $\mathcal{S}^\star = \{(0,3)\}$ 以及 $\mu_A = -1$成立。2，$\rho_A = 0.2$，$\mu_B = 1.6$ 和 $\rho_B = 0.1$。

(iii) 定理 3.3 中的结论对由步长为 $\gamma \in$(0、0。2614) 且松弛参数为 $\lambda \in (0、2 + 2/15\gamma - 48\gamma/5)$ 的 DRS 生成的序列 $(u^k, v^k, s^k)_{k \in \mathbb{N}}$ 成立。 □

备注 6.3 （与先前工作的比较）。

(i) 由于 f_A 和 f_B 均非光滑，且 f_A 是非凸的，因此无法应用 DRS在优化情形下的收敛结果，例如 [53, 84]。

(ii) 由于 $\rho_A > 0$和 $\rho_B > 0$，根据定义，∂f_A 在 $(x^\star, -y^\star)$ 处是 μ_A-单调的，而 ∂f_B 在 $(x^\star, y^\star)$ 处是 μ_B-单调的。然而，由于 A 和 B 不具备全局单调性，因此[21]中的结果在此处不适用。 □

接下来，考虑以下示例，我们将 DRS应用于可分非光滑函数的次微分。对于此示例，我们将说明对 DRS 的分析涵盖了其向驻点收敛的行为，并在 图 8中给出数值结果。请注意，迭代点会被局部最小值吸引；相比之下，下降法等方法则会迅速发散。

例 6.4 (驻点)。设 $\mu_A = -0。3$，$\rho_A = -0。1$，$\mu_B = 0。4$ 和 $\rho_B = 0。4$。考虑如下问题：求 $f_A(x) + f_B(x)$ 的驻点，该问题如 图 8 所示，其中

$$f_A(x) := \begin{cases} \frac{1 - \sqrt{1 - 4\mu_A\rho_A}}{4\rho_A} x^2 - x - 3, & \text{if } x \in [-3, 3], \\ \frac{1 + \sqrt{1 - 4\mu_A\rho_A}}{4\rho_A} x^2 - x - \frac{9\sqrt{1 - 4\mu_A\rho_A}}{2\rho_A} - 3, & \text{otherwise}, \end{cases}$$

$$f_B(x) := \begin{cases} x^2 + x + 2, & \text{若 } x \in [-1, 1], \\ \frac{1}{4}x^2 + x + \frac{11}{4}, & \text{否则。} \end{cases}$$

以下包含关系对 $0 \in T_{\mathrm{P}}x = \partial f_A(x) + \partial f_B(x)$ 成立。

(i) 假设 II.A1 和 假设 II.A2 对 $\gamma \in (0, \frac{-2\rho_A}{1+\sqrt{1-4\mu_A\rho_A}}) \cup (\frac{-2\rho_A}{1+\sqrt{1-4\mu_A\rho_A}}, 1/5)$成立。

(ii) 假设 III 对 $\mathcal{S}^\star = \{(0, -1)\}$ 以及给定的 $\mu_A, \rho_A, \mu_B, \rho_B$成立。

(iii) 定理 3.3 中的结论 对由 *DRS* 生成的序列 $(u^k, v^k, s^k)_{k\in\mathbb{N}}$ 成立，其中 DRS 的步长为 $\gamma \in (1/6, 1/5)$，固定松弛参数为 $\lambda \in (0, 2 - 4/15\gamma - 12\gamma/5)$ (参见 图 8 中的数值结果)。 □

最后，考虑以下带有盒约束和线性等式约束的二次规划（QP）问题。当 Q 为半正定矩阵时，即处于凸情形下，DRS 对例 6.5 的收敛性可直接由 A 和 B的单调性得出。下面的例子说明，利用第 4 节中的微积分法则，我们的结果能够拓展到单调情形之外。请注意，在此例中，f 和 g 均非光滑，因此诸如 [53, 84] 这样的 DRS 收敛结果并不适用于此处。

例 6.5 （带约束的 QP）。 考虑以下二次规划

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & \tfrac{1}{2}x^\top Qx + q^\top x \\ \text{subject to} \quad & Hx = h, \\ & l \le x \le u, \end{aligned} \tag{6.3}$$

其中 $Q \in \mathbb{S}^n$，$q \in \mathbb{R}^n$，$H \in \mathbb{R}^{m\times n}$ 具有满行秩，$h \in \mathbb{R}^m$，标量 l、$u \in \mathbb{R}$ 满足 $l < u$。该问题可等价地表述为最小化 $f(x) + g(x)$，其中

$$f(x) := \tfrac{1}{2}x^\top Qx + q^\top x + \delta_{\{s \mid Hs=h\}}(x) \quad \text{and} \quad g(x) := \delta_{\{s \mid l\le s\le u\}}(x),$$

其对应的一阶最优性条件为 $0 \in \mathrm{A(x)} + \mathrm{B(x)} := \partial \mathrm{f}(x) + \partial \mathrm{g(x)}$，将二次规划问题 (6.3) 分解为 f 和 g 后，可以通过闭式形式高效计算所得的 DRS 迭代。具体而言，f 对应一个带等式约束的二次规划问题，其 prox 可通过相应的 KKT 系统高效求解 [62,§16.1], 而计算 ∂g 的预解算子则对应于投影到一个盒集上。

例如，设 $Q = \mathrm{diag}(-1, 1/2), q = [-4\ -1]^\top, H = [1\ 1], h = 1, l = 0$ 且 $u = 1$。那么，(6.3) 的唯一极小化解为 $x^\star = [1\ 0]^\top$，并且以下断言成立。

(i) 算子 A 是 $(-1, 0)$-半单调。

(ii) 算子 B 在 $(x^\star, y^\star) = ([1\ 0]^\top, [2\ -2]^\top) \in$图 B处是 $(2, 0)$-半单调的。

(iii) 序列 $(u^k, 1/\gamma(u^k - s^k))_{k\in\mathbb{N}}$ 由 DRS 采用固定松弛参数 λ 生成，收敛于零点集 T_{PD} 当 $\gamma \in (0, 1/2)$ 和 $\lambda \in (0, 2 - 4\gamma)$。

证明。 根据 命题 4.4*(ii)* 和 4.5，可知映射 $x \mapsto Qx + q$ 是 $(-1,0)$-半单调的。通过观察法锥 A，即可得到 $N_{\{s \mid Hs=h\}}$ 是单调的 [9,例 20。26]，并应用 命题 4.7*(i)*。最后，B的半单调性由 命题 A.7 得出，而 DRS 的收敛结果由定理 5.2得出。 □

7 结论

本文全面研究了适用于满足斜弱 Minty 条件的一类非单调问题的（松弛）预条件近端点算法（PPPA）和（松弛）Douglas-Rachford 分裂（DRS）方法。为了使所得到的 DRS 结果能够轻松应用于实践，本文引入了半单调算子类，并提出了保证 DRS 对两个半单调算子之和收敛的充分条件。文中还提供了多个相关示例问题，展示了所得到的收敛结果的紧确性。

未来的研究方向包括探讨 PPPA 的非精确变体、分析非单调情形下的其他分裂方法，以及研究其在非合作博弈论和非凸分布式优化中的应用。

A 辅助引理

引理 A.1。 设 $P \in \mathbb{S}_+^n$ 为对称半正定矩阵，U 为 $\mathcal{R}(P)$ 的标准正交基，并定义 $Q := P + \Pi_{\mathcal{N}(P)}$。则以下性质成立。

(i) P 的值域上的投影由 $\Pi_{\mathcal{R}(P)} = UU^\top$ 给出。

(ii) $P = P\,\Pi_{\mathcal{R}(P)} = \Pi_{\mathcal{R}(P)}\,P = Q\,\Pi_{\mathcal{R}(P)}$.

(iii) $Q \succ 0$。*(iv)* $U^\top PU = U^\top QU \succ 0$。

*Pr*证明。 命题 A.1*(i)* 由 [57, §5.13]和 A.1*(ii)* 根据 P 和 Q的定义可直接得出。

♠ A.1(iii)：设 $\mathrm{x} \in \mathbb{R}^n \backslash \{0\}$，并考虑将其分解为 $\mathrm{x} = \mathrm{x}_1 + \mathrm{x}_2$，其中 $\mathrm{x}_1 \in \mathcal{R}(\mathrm{P})$ 且 $\mathrm{x}_2 \in \mathcal{N}(\mathrm{P})$。设 $\mathrm{U}_1\Lambda_1\mathrm{U}_1^\top$ 是 P 的紧致特征分解，使得 $\mathrm{U}_1\mathrm{U}_1^\top = \Pi_{\mathcal{R}(P)}$ 且 $\mathrm{p}_{\min} := \lambda_{\min}(\Lambda_1) > 0$。那么，

$$\begin{aligned}\langle x, Qx\rangle &= \langle x, Px\rangle + \langle x, \Pi_{\mathcal{N}(P)}\,x\rangle = \langle x_1, U_1\Lambda_1 U_1^\top x_1\rangle + \|x_2\|^2 \\ &\geq p_{\min}\langle x_1, \Pi_{\mathcal{R}(P)}\,x_1\rangle + \|x_2\|^2 = p_{\min}\|x_1\|^2 + \|x_2\|^2 \geq \min\{1, p_{\min}\}\|x\|^2 > 0.\end{aligned}$$

♠ A.1*(iv)*：等价关系 $U^\top PU = U^\top QU$ 由以下结果得出：引理 A.1*(i)* 和 A.1*(ii)*，并利用 $U^\top U = \mathrm{I}$。正定性由引理 A.1*(iii)*以及 U 具有满列秩这一事实得出 [11, 命题 8.1.2(xiii)]。 □

命题 A.2 （共单调算子的预解算子）。 设 $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 是一个 V-余单调算子，并假设 P 根据假设 *I.*A4 选取。那么，预条件化预解算子 $(P+T)^{-1} \circ P$ 的定义域是全空间，当且仅当 T 是极大的。

证明。由于 T 是 V-余单调的，根据定义，$\mathrm{T}^{-1} - \mathrm{V}$ 是单调的，因此 $\tilde{\mathrm{T}} := \mathrm{U}^\top\mathrm{T}^{-1}\mathrm{U} - \mathrm{U}^\top\mathrm{VU}$ 也单调。此外，根据假设 I.a4 和注 2.2，有 $(\mathrm{U}^\top\mathrm{PU})^{-1} + \mathrm{U}^\top\mathrm{VU} \succ 0$。记 $\hat{\mathrm{T}} := \mathrm{U}^\top\mathrm{T}^{-1}\mathrm{U} + (\mathrm{U}^\top\mathrm{PU})^{-1}$。由 $(\mathrm{U}^\top\mathrm{PU})^{-1} + \mathrm{U}^\top\mathrm{VU} \succ 0$ 以及 Minty 定理[9, 定理 21.1]，算子

$$((U^\top PU)^{-1} + U^\top VU)^{-1}\hat{T} = ((U^\top PU)^{-1} + U^\top VU)^{-1}\tilde{T} + \mathrm{id}$$

当且仅当 $(U^\top PU)^{-1} + U^\top VU)^{-1}\,\tilde{T}$ 是极大单调的时，具有满值域。由于 $(U^\top PU)^{-1} + U^\top VU \succ 0$，这意味着 $\hat{T}$ 具有满值域，当且仅当 $\tilde{T}$ 是极大单调的；也就是说，当且仅当T 是极大V-共单调的。

因此，只需证明 $\hat{T}$具有满值域，当且仅当 $\mathrm{dom}\,(P+T)^{-1} \circ P = \mathbb{R}^n$。首先假设 $\hat{T}$具有满值域。令m 表示P 的值域维数，并固定 $\hat{y} \in \mathbb{R}^m$。于是，根据 $\hat{T}$的满值域假设，存在某个 $\hat{x} \in \mathbb{R}^m$，使得 $\hat{y} \in \hat{T}\hat{x}$。令 $\hat{z} = \hat{y} - (U^\top PU)^{-1}\hat{x}$，这意味着

$$\hat{y} \in U^\top T^{-1}U\hat{x} + (U^\top PU)^{-1}\hat{x} \iff \hat{z} \in U^\top T^{-1}U(U^\top PU(\hat{y} - \hat{z})) \tag{A.1}$$

两边乘以 U，可得 $\mathrm{U}\hat{\mathrm{z}} \in \Pi_{\mathcal{R}(P)}\mathrm{T}^{-1}(\mathrm{PU}(\hat{\mathrm{y}} - \hat{\mathrm{z}}))$。定义 $\mathrm{x}' = \mathrm{U}\hat{\mathrm{x}}$、$\mathrm{y}' = \mathrm{U}\hat{\mathrm{y}}$ 和 $\mathrm{z}' = \mathrm{U}\hat{\mathrm{z}}$，这对应于 $\mathrm{z}' \in \Pi_{\mathcal{R}(P)}\mathrm{T}^{-1}(\mathrm{P}(\mathrm{y}' - \mathrm{z}'))$。因此，$\mathrm{T}^{-1}(\mathrm{P}(\mathrm{y}' - \mathrm{z}'))$非空，并且存在一个 $\mathrm{z}'' \in \mathcal{N}(\mathrm{P}) \subseteq \mathbb{R}^n$，使得 $\mathrm{z}'' \in \Pi_{\mathcal{N}(P)}\mathrm{T}^{-1}(\mathrm{P}(\mathrm{y}' - \mathrm{z}'))$。将两个包含关系相加，得到

$$z' + z'' \in T^{-1}(P(y' - z')) \iff P(y' - z') \in T(z' + z'') \iff z' + z'' \in (P+T)^{-1}Py' \tag{A.2}$$

由于 y' 在 $\mathcal{R}(P)$ 中的选择是任意的，$(P+T)^{-1}P$ 具有全定义域。

逆命题类似可得。假设 $\mathrm{dom}\,(\mathrm{P}+\mathrm{T})^{-1}\mathrm{P} = \mathbb{R}^n$。于是，对于任意固定的 $\hat{\mathrm{y}} \in \mathbb{R}^m$ 以及 $\mathrm{y}' = \mathrm{U}\hat{\mathrm{y}}$，都存在某个 $\mathrm{z} \in \mathbb{R}^n$ 使得 $\mathrm{z} \in (\mathrm{T}+\mathrm{P})^{-1}\mathrm{Py}$。不妨将 $\mathrm{z} := \mathrm{U}\hat{\mathrm{z}} + \mathrm{z}''$分解，其中 $\hat{\mathrm{z}} \in \mathbb{R}^m$ 且 $\mathrm{z}'' \in \mathcal{N}(\mathrm{P})$。由 (A.2) 可知，$\mathrm{U}\hat{\mathrm{z}} + \mathrm{z}'' \in \mathrm{T}^{-1}(\mathrm{PU}(\hat{\mathrm{y}} - \hat{\mathrm{z}}))$。两边乘以 $\mathrm{U}^\top$，得到 $\hat{\mathrm{z}} \in \mathrm{U}^\top\mathrm{T}^{-1}\mathrm{U}(\mathrm{U}^\top\mathrm{PU}(\hat{\mathrm{y}} - \hat{\mathrm{z}}))$。定义 $\hat{\mathrm{x}} = \mathrm{U}^\top\mathrm{PU}(\hat{\mathrm{y}} - \hat{\mathrm{z}})$，并利用 (A.1)，可得 $\hat{\mathrm{y}} \in \hat{\mathrm{T}}\hat{\mathrm{x}}$，从而证明 $\hat{\mathrm{T}}$ 具有满值域。 □

事实 A.3 (二次不等式的解)。 设 a,b,$c \in \mathbb{R}$，并假设当 $ac \geq 0$时，$b > -2\sqrt{ac}$。那么，下列结论成立。

表 4： 使预解算子 $J_{\gamma A}$ 和 $J_{\gamma B}$ 具有全定义域的 γ 取值范围。

		$\mu_A=\mu_B=0$	$\mu_A+\mu_B>0$	
			$\mu_A\mu_B\geq 0$	$\mu_A\mu_B<0$
$\rho_A=\rho_B=0$		$\gamma\in(0,+\infty)$		$\gamma\in\left(0,-\frac{1}{\{A\text{最小值}\mu,\ \mu_B\}}\right)$
$\rho_A+\rho_B>0$	$\rho_A\rho_B\geq 0$			$\gamma\in(0,\bar{\gamma}_+)$
	$\rho_A\rho_B<0$	$\gamma\in(-\min\{\rho_A,\rho_B\},+\infty)$	$\gamma\in(\bar{\gamma}_-,+\infty)$	$\gamma\in(\bar{\gamma}_-,\bar{\gamma}_+)$

(i) 存在一个 $\gamma>0$ 满足 $a\gamma^2+b\gamma+c>0$，当且仅当满足

$$b>2\sqrt{ac}\quad when\quad a<0\ and\ c<0. \tag{A.3}$$

(ii) 若(A.3)成立，则 $\gamma>0$ 满足 $a\gamma^2+b\gamma+c>0$ 当且仅当 $\gamma\in\left(\frac{2[-c]_+}{b+\sqrt{b^2-4ac}},\frac{2[-a]_+}{b+\sqrt{b^2-4ac}}\right)$，且该集合非空。

事实 A.4. 设 a、$b\in\mathbb{R}$ 满足 $a+b>0$。则 $a\square b<\min\{a,b\}$。

引理 A.5 （假设 II的推论）。若假设 *II* 成立，则假设 *I.*A1 至*I.*A3 对算子 T 成立$_{\text{PD}}$ 其中 （PD-I）以及式 （(3.1)） 中的预条件子 P。

证明。首先，T_{PD} 的外半连续性由 A 和 B 的外半连续性得出 [77, 定理 5.7(a)],，从而证明 I.a1。其次，由于 $J_{\gamma A}$ 和 $J_{\gamma B}$ 具有全定义域，预条件化预解算子 $(P+T_{PD})^{-1}P$ 具有全定义域，因此 I.a2 成立；这是由 [77, 引理 12.14] 得出的。最后，陈述 I.a3 显然成立，证明完毕。 □

引理 A.6 （假设 IV的推论）。设假设 *IV* 成立，且 γ 按照表 *3*选择，其中 γ_- 和 γ_+ 如 （*5.2*） 中所定义。则假设 *II.*A1 和*II.*A2 成立，且预解算子 $J_{\gamma A}$ 和 $J_{\gamma B}$ 均为 *Lipschitz*连续。

证明。首先，注意到 $[\mu_A\square\mu_B]_-[\rho_A\square\rho_B]_-<\frac{1}{4}$，由事实 A.4 和命题 4.3(ii) 可知，这意味着$[\mu_A]_-[\rho_A]_-<\frac{1}{4}$ 且 $[\mu_B]_-[\rho_B]_-\prec\frac{1}{4}$。因此，由命题4.11 可知，A 和 B 是外半连续的，这表明假设 II.a1 成立。此外，由命题 4.12(i) 可知，预解算子 $J_{\gamma A}$ 和 $J_{\gamma B}$ 是 Lipschitz连续的，并且当 γ 符合表 4 时具有全定义域，其中

$$\bar{\gamma}_-=\begin{cases}\frac{-2\rho_A}{1+\sqrt{1-4\mu_A\rho_A}}, & \text{if } \rho_A<0,\\ \frac{-2\rho_B}{1+\sqrt{1-4\mu_B\rho_B}}, & \text{otherwise,}\end{cases}\quad\text{and}\quad \bar{\gamma}_+=\begin{cases}-\frac{1+\sqrt{1-4\mu_A\rho_A}}{2\mu_A}, & \text{if } \mu_A<0,\\ -\frac{1+\sqrt{1-4\mu_B\rho_B}}{2\mu_B}, & \text{otherwise.}\end{cases} \tag{A.4}$$

由于 γ 是根据 表 3选取的，只需证明这些值域是 γ 在 表 4中给出的值域的子集，这可直接利用 事实 A.4验证。 □

命题 *A.7*（盒的法锥）。*n* 维盒 *C x_l^i, . . . , x_u^i, i = 1, . . . , n* 的法锥算子 *N min v_u^l. . . , v_u^l,* -半单调于 *(x, v)* ∈ 图 *N*。 —— ——

证明。 根据法锥的单调性 N_C [9, 例 20。26], 可得

$$\langle\tilde{v},\tilde{x}-x\rangle=\sum_{i=1}^{n}|\tilde{v}_i||\tilde{x}_i-x_i|\geq\sum_{i=1}^{n}\frac{|\tilde{v}_i|}{u_i-l_i}|\tilde{x}_i-x_i|^2\geq\min\left\{\frac{|\tilde{v}_1|}{u_1-l_1},\dots,\frac{|\tilde{v}_n|}{u_n-l_n}\right\}\|\tilde{x}-x\|^2,$$

其中第一个不等式成立，是因为对于所有$i\in\{1,\dots,n\}$，$|\tilde{}x_i-x_i|\leq u_i-l_i$。 □

B 省略的证明

B.1 Douglas-Rachford 分裂

证明 引理 ***3.1*** *(DRS 与 PPPA等价性)*。对带有前置条件子（3.1）的（PD-I）应用更新规则（PPPA），可等价地写为

$$\bar{x}^k \in J_{\gamma A}(x^k - \gamma y^k) \tag{B.1a}$$

$$\bar{y}^k \in (\gamma \mathrm{id} + B^{-1})^{-1}(2\bar{x}^k - x^k + \gamma y^k) \Leftrightarrow 2\bar{x}^k - x^k + \gamma y^k - \gamma \bar{y}^k \in J_{\gamma B}(2\bar{x}^k - x^k + \gamma y^k) \tag{B.1b}$$

$$x^{k+1} = x^k + \lambda_k(\bar{x}^k - x^k) \tag{B.1c}$$

$$y^{k+1} = y^k + \lambda_k(\bar{y}^k - y^k), \tag{B.1d}$$

其中，(B.1b) 中用到了 [77, 引理 12。14] 我们通过归纳法进行证明。基例 $k = 0$ 根据初始化 $s^0 = x^0 - \gamma y^0$, , 并利用 (B.1a) 和 (B.1b) 可直接满足；注意，对于任意 $\bar{x}^0, \bar{y}^0$ ，存在 $u^0 = \bar{x}^0 \in J_{\gamma A}(s^0)$, $v^0 = 2u^0 - s^0 - \gamma \bar{y}^0 \in J_{\gamma B}(2u^0 - s^0)$（反之亦然）。现在假设该结论对某个 $k \geq 0$成立。特别地，$s^k = x^k - \gamma y^k$, $u^k = \bar{x}^k$ ，且 $v^k = 2u^k - s^k - \gamma \bar{y}^k$。根据更新规则（DRS），有如下结论。

$$s^{k+1} = s^k + \lambda_k(v^k - u^k) = s^k + \lambda_k(\bar{x}^k - s^k - \gamma \bar{y}^k) = x^k - \gamma y^k + \lambda_k(\bar{x}^k - \gamma \bar{y}^k - x^k + \gamma y^k) = x^{k+1} - \gamma y^{k+1},$$

其中前两个等式使用了归纳法假设，最后一个等式使用了 (B.1c) 和 (B.1d)。在确立了 $s^{k+1} = x^{k+1} - \gamma y^{k+1}$ 后，关于 $\bar{x}^{k+1}$、$\bar{y}^{k+1}$ 与u^{k+1}、v^{k+1} 之间对应关系的断言，可如基例中的论证那样由 (B.1a) 和 (B.1b) 得出。 □

引理 3.2 的证明（原始算子和对偶算子的弱明蒂性质）。令 $(x_A, y_A) \in A$ 的图，且令 $(x_B, y_B) \in B$ 的图。于是，根据原始-对偶算子的定义，并利用 $(y_B, x_B) \in B^{-1}$ 的图，有 $(z, v) := \big((x_A, y_B), (y_A + y_B, x_B - x_A)\big) \in T$ 的图$_{PD}$。因此，根据假设 II.a3，对所有 $z^\star = (x^\star, y^\star) \in \mathcal{S}^\star$，都有$\langle v, z - z^\star \rangle \geq q_V(v)$，即，

$$\big\langle y_A + y_B, x_A - x^\star \big\rangle + \big\langle x_B - x_A, y_B - y^\star \big\rangle \geq \beta_P \|y_A + y_B\|^2 + \beta_D \|x_B - x_A\|^2. \tag{B.2}$$

设 $x_A = x_B =: x_P$。那么，由 (B.2) 可得 $\langle y_A + y_B, x_P - x^\star \rangle \geq \beta_P \|y_A + y_B\|^2$ 对所有 $x^\star \in \mathcal{S}_P^\star$ 都成立。由于 $(x_P, y_A + y_B) \in$属于 T 的图$_P$，容易验证 T_P 在 $\mathcal{S}_P^\star$ 处具有 β_P-弱 Minty 解。

相似地，设 $-y_A = y_B =: y_D$。于是，由于 (B.2) $\langle x_B - x_A$，对于所有 $y^\star \in \mathcal{S}_D^\star$，有 $y_D - y^\star \rangle \geq \beta_D \|x_B - x_A\|^2$，而关于 T_D 的结论可类似地由 $(y_D, x_B - x_A) \in$图T_D 这一事实推出，证明完成。□

证明定理 ***3.6*** *(DRS 对分段多面体映射的线性收敛)*。
由于算子 B 是分段多面体映射，其逆也是如此，因此映射 $\hat{T} = A \times B^{-1}$ 也是如此。于是，T_{PD} 是 $\hat{T}$ 与一个线性映射的和，因此也是分段多面体映射。由于分段多面体映射的逆也是分段多面体映射，由 [25, 3H.1 和 3H.3]可知，对于任意满足 $(z, v) \in T$ 的图$_{PD}$的 v，T_{PD} 在任意 z 处都是度量次正则的。由定理 2.13 以及关系 $U_1^\top z^k = s^k$ 和 (3.5)，可立即得到 DRS 的所述收敛速率。 □

B.2 半单调算子

证明 命题 ***4.3*** *(半单调算子的存在性)*。考虑 $(x, y), (x', y') \in A$ 的图 A 并令 $\epsilon > 0$。于是，由 Fenchel–Young 不等式可得

$$-\epsilon \|x - x'\|^2 - \epsilon^{-1} \|y - y'\|^2 \leq 2\langle x - x', y - y' \rangle \leq \epsilon \|x - x'\|^2 + \epsilon^{-1} \|y - y'\|^2, \tag{B.3}$$

♠ 4.3*(i)*：不等式（B.3）中的第一个不等式意味着（4.1）对于 $\mu \leq -\frac{\epsilon}{2}$ 且 $\rho \leq -\frac{1}{2\epsilon}$ 始终成立，即对于$[\mu]_- [\rho]_- \geq \frac{1}{4}$，从而证明完毕。

♠ 4.3*(ii)*：不等式（B.3）中的第二个不等式意味着（4.1）在 $\mu > \frac{\epsilon}{2}$ 且 $\rho > \frac{1}{2\epsilon}$ 时永远无法成立，即当 $[\mu]_+[\rho]_+ > \frac{1}{4}$ 时。

♠ 4.3*(iii)*：不等式（B.3）的第二个不等式意味着，只有当 4.1）取等号时，才能满足（$\mu = \frac{\epsilon}{2} > 0$ 且 $\rho = \frac{1}{2\epsilon} > 0$，即

$$\langle x - x', y - y' \rangle = \mu \|x - x'\| + \tfrac{1}{4\mu}\|y - y'\|^2, \qquad \text{for all } (x, y), (x', y') \in \operatorname{gph} A,$$

这等价于

$$\mu \left\| x - x' - \tfrac{1}{2\mu}(y - y') \right\|^2 = 0, \qquad \text{for all } (x, y), (x', y') \in \operatorname{gph} A.$$

根据构造，只有当存在某个 $c \in \mathbb{R}^n$ 使得图 $A \subseteq$ 图T 时，这才可能成立。由于图T无法再加入其他点（$\hat{x}$，$\hat{y}$）并同时保持（μ，ρ）-半单调性，因此T是极大（μ，ρ）-半单调的，证明完毕。 □

推论 4.6 的证明（单位算子的倍数）。根据命题 4.5，并利用 $A = \alpha \mathrm{I}_n$，可得 A 是（μ，ρ）-半单调的，当且仅当 $\alpha \geq \mu + \rho\alpha^2$；对于任意 $c \in \mathbb{R}$，当（μ，ρ）=（α（$1 - c\alpha$），c）时，这显然成立。现在，令（$\hat{x}$，$\hat{y}$）$\in \mathbb{R}^n \times \mathbb{R}^n$，并假设

$$\langle x - \hat{x}, \alpha x - \hat{y} \rangle \geq \alpha(1 - c\alpha)\|x - \hat{x}\|^2 + c\|\alpha x - \hat{y}\|^2, \qquad \forall x \in \mathbb{R}^n. \tag{B.4}$$

为了说明 A 何时是极大（$\alpha(1 - c\alpha)$，c）半单调的，我们必须说明在什么条件下可以保证 $(\hat{x}, \hat{y}) \in$图 A，也就是说，$\hat{y} = \alpha\hat{x}$。通过初等代数，可以证明 (B.4) 等价于

$$0 \geq (1 - 2c\alpha)\langle x, \hat{y} - \alpha\hat{x} \rangle + \alpha(1 - c\alpha)\|\hat{x}\|^2 - \langle \hat{x}, \hat{y} \rangle + c\|\hat{y}\|^2, \qquad \forall x \in \mathbb{R}^n. \tag{B.5}$$

由于第一项取决于 x，因此必然有 $\hat{y} = \alpha\hat{x}$ 或 $c = \frac{1}{2\alpha}$。将 $\hat{y} = \alpha\hat{x}$ 代入（B.5），得到 $0 \geq 0$（正如预期）。将 $c = \frac{1}{2\alpha}$ 代入（B.5），得到

$$0 \geq \tfrac{1}{2}\alpha\|\hat{x}\|^2 - \langle \hat{x}, \hat{y} \rangle + \tfrac{1}{2\alpha}\|\hat{y}\|^2 = \tfrac{1}{2\alpha}\|\alpha\hat{x} - \hat{y}\|^2, \qquad \forall x \in \mathbb{R}^n.$$

因此，(B.5) 成立，当且仅当 以下条件之一成立：(i) $\hat{y} = \alpha\hat{x}$；或 (ii) $c = \frac{1}{2\alpha}$ 且 $\alpha < 0$，证明完毕。 □

证明命题 ***4.7*** *(*和与平行和*)*。

♠ 4.7*(i)*：首先，考虑 A 和B 仅在单点处为半单调的情形。令（x，y_A）$\in$图 A 以及（x，y_B）$\in$图B。则有

$$\begin{aligned}\langle x - x', y_A + y_B - (y'_A + y'_B) \rangle &= \langle x - x', y_A - y'_A \rangle + \langle x - x', y_B - y'_B \rangle \\ &\geq (\mu_A + \mu_B)\|x - x'\|^2 + \rho_A\|y_A - y'_A\|^2 + \rho_B\|y_B - y'_B\|^2.\end{aligned}$$

因此，根据定义，$A + B$ 在（x'，$y'_A + y'_B$）处为（$\mu_A + \mu_B$，ρ）-半单调，当且仅当

$$\rho\|y_A + y_B - (y'_A + y'_B)\|^2 = \rho\|y_A - y'_A\|^2 + 2\rho\langle y_A - y'_A, y_B - y'_B \rangle + \rho\|y_B - y'_B\|^2 \leq \rho_A\|y_A - y'_A\|^2 + \rho_B\|y_B - y'_B\|^2,$$

这等价于线性矩阵不等式 $\begin{pmatrix} \rho_A - \rho & -|\rho| \\ -|\rho| & \rho_B - \rho \end{pmatrix} \geq 0$。当且仅当行列式和迹均为非负时，该条件成立，即，

$$\rho_A\rho_B - \rho(\rho_A + \rho_B) \geq 0 \ \text{ and } \ \rho_A + \rho_B - 2\rho \geq 0 \quad \Leftrightarrow \quad \rho \leq \begin{cases} \frac{\rho_A\rho_B}{\rho_A + \rho_B}, & \text{if } \rho_A + \rho_B > 0, \\ 0, & \text{if } \rho_A = \rho_B = 0. \end{cases} \tag{B.6}$$

根据定义，满足 ρ（B.6）的最大可能值由 $\rho_A \square \rho_B$ 给出（如果 $\rho_A + \rho_B > 0$ 或$\rho_A = \rho_B = 0$），证明的第一部分即告完成。

如果A 和B 是半单调的，那么根据定义，A 在所有（x'_A，y'_A）$\in \operatorname{gph} A$ 处都是（μ_A，ρ_A）-半单调的，并且B在所有（x'_B，y'_B）$\in \operatorname{gph} B$处都是（μ_B，ρ_B）-半单调的。因此，利用前面的结果，我们知道$A + B$ 在集合中的所有点处都是（$\mu_A + \mu_B$，$\rho_A \square \rho_B$）-半单调的

$$\{(x', v'_A + v'_B) \mid x' \in \operatorname{dom} A \cap \operatorname{dom} B,\ v'_A \in Ax',\ v'_B \in Bx'\}.$$

由于此集合等于 $\operatorname{gph}(A + B)$，证明完成。

♠ 4.7*(ii)*：这是平行和定义的直接推论，这归功于 命题 4.4*(i)*和 命题 4.7*(i)*。

□

证明命题 ***4.8*** *(*与恒等算子的和*)*。

♠ 4.8(i)：首先，考虑半单调性仅在单个点处成立的命题。算子 A 在 $(x', y') \in \operatorname{gph} A$ 处是 (μ_A, ρ_A)-半单调的，当且仅当 $\langle x - x', y - y' \rangle \geq \mu_A \|x - x'\|^2 + \rho_A \|y - y'\|^2$ 对所有 $(x, y) \in \operatorname{gph} A$ 均成立。定义 $p := y + \alpha x$ 和 $p' := y' + \alpha x'$，使得 $(x, p), (x', p') \in \operatorname{gph} T$，则这等价于

$$(1 + 2\rho_A\alpha)\langle x - x', p - p' \rangle \geq (\mu_A + \alpha(1 + \rho_A\alpha))\|x - x'\|^2 + \rho_A\|p - p'\|^2, \quad \forall (x, p) \in \operatorname{gph} T.$$

两边都除以 $1 + 2\rho_A\alpha > 0$ 后，即可得出 A 在 (x', y') 处的 (μ_A, ρ_A)-半单调性与 T 在 $(x', y' + \alpha x')$ 处的 $\left(\frac{\mu_A + \alpha}{(1+\rho_A\alpha)1+2\rho_A\alpha}, \frac{\rho_A}{1+2\rho_A\alpha}\right)$-半单调性之间所声称的等价关系。通过考虑所有点 $(x', y') \in$ 图 A，即可类似地得出（全局）半单调性的结论。

最后，极大性的断言成立，因为如果不破坏（μ_A, ρ_A）-半单调性，则 A 的图 A 可以添加点 $(\hat{x}, \hat{y})$，当且仅当 T 的图 T可以添加点 $(\hat{x}, \hat{y} + \alpha\hat{x})$ 而不违反 $\left(\mu_A + \alpha(1+\rho_A\alpha)1+2\rho_A\alpha, \frac{\rho_A}{1+2\rho_A\alpha}\right)$- 半单调性。

♠ 4.8*(ii)*：该结果可由 命题 4.7*(i)*和 推论 4.6直接推出。 □

命题 4.11（半单调算子的外半连续性）的证明。首先，设 $\rho\mu < \frac{1}{4}$，并令 ξ 和 ν 如 (4.2) 中所定义。令 $(x^k, y^k) \in A$ 的图，其中 $x^k \to \bar{x}$ 且 $y^k \to \bar{y}$。根据构造，有 $x^k - \xi(y^k - \nu x^k) \in \left[(A^{-1} - \nu\mathrm{id})^{-1} - \xi\mathrm{id}\right](y^k - \nu x^k)$。由推论 4.10 可知，$(A^{-1} - \nu\mathrm{id})^{-1} - \xi\mathrm{id}$ 是极大单调的，因而也外半连续 [77, Ex. 12.8]。因此可得 $\bar{x} - \xi(\bar{y} - \nu\bar{x}) \in \left[(A^{-1} - \nu\mathrm{id})^{-1} - \xi\mathrm{id}\right](\bar{y} - \nu\bar{x})$，进而推出 $\bar{x} \in A\bar{y}$，证明了 A 的外半连续性。最后，当 $[\mu]_+[\rho]_+ = 1/4$，时，结合命题 4.3(iii)-2 和 4.4(ii) 以及推论 4.6，可得 A 的极大单调性，由此推出其外半连续性 [77,Ex. 12.8]。 □

证明 命题 ***4.12***（半单调算子的预解算子）。

♠ 4.12 *(i)* ：首先，假设 $\mu\rho < \frac{1}{4}$。定义 $\theta := \sqrt{1 - 4\rho\mu}$，并令 $\nu := \frac{2\mu}{1+\theta}$、$\xi := \frac{\rho}{\theta}$和 $\omega := \frac{\gamma}{1+\gamma\nu}$。利用 (4.3) 中的上界 $\gamma \leq \frac{2[-\mu]_+}{1+\theta}$，可得 $1 + \gamma\nu = 1 + \gamma\frac{2\mu}{1+\theta} \geq 1 - \frac{2\gamma[-\mu]_+}{1+\theta} \geq 0$。因此，$\omega + \xi > 0$ 因为

$$\frac{\gamma}{1+\gamma\nu} + \frac{\rho}{\theta} > 0 \overset{1+\gamma\nu>0}{\iff} \gamma + \left(1 + \gamma\frac{2\mu}{1+\theta}\right)\frac{\rho}{\theta} > 0 \overset{\theta>0}{\iff} \left(2\theta + \frac{4\mu\rho}{1+\theta}\right)\gamma + 2\rho > 0 \iff (1+\theta)\gamma + 2\rho > 0$$

这由 (4.3) 中的下界 $\gamma > \frac{2[-\rho]_+}{1+\theta}$ 得到满足。记 $\hat{A} := (A - \nu\mathrm{id})^{-1} + \omega\mathrm{id}$。由于$\omega + \xi > 0$ 以及 Minty 定理 [9, 定理 21.1]，，算子 $\frac{1}{\omega+\xi}\hat{A} = \frac{1}{\omega+\xi}\left((A - \nu\mathrm{id})^{-1} - \xi\mathrm{id}\right) + \mathrm{id}$ 具有满值域，当且仅当 $\frac{1}{\omega+\xi}((A - \nu\mathrm{id})^{-1} - \xi\mathrm{id})$ 是极大单调的。由于 $\omega + \xi > 0$，这意味着 $\hat{A}$ 具有满值域，当且仅当算子 $\tilde{A} := (A - \nu\mathrm{id})^{-1} - \xi\mathrm{id}$ 是极大单调的。

另一方面，根据定义 $1 + 2\xi\nu = 1 + \frac{1-\theta^2}{\theta 1+\theta}(\quad) = \frac{1}{\theta} \geq 0$。因此，利用推论 4.9(i)，算子 $\tilde{A} := (A - \nu\mathrm{id})^{-1} - \xi\mathrm{id}$ 是极大单调的，当且仅当 A 是极大 $\left(\frac{\nu(1+\xi\nu)}{1+2\xi\nu}, \frac{\xi}{1+2\xi\nu}\right) = (\mu, \rho)$ 半单调的，其中等式是通过代入 ξ 和 ν 得出的。因此，只需证明 $\hat{A}$ 具有满值域，当且仅当 $\mathcal{R}(\mathrm{id} + \gamma A) = \mathbb{R}^n$，即预解算子具有全定义域。首先，假设 $\hat{A}$ 具有满值域。固定 $y \in \mathbb{R}^n$。那么，根据 $\hat{A}$ 的满值域假设，存在某个 $x \in \mathbb{R}^n$，使得 $\frac{\omega}{\gamma}y \in (A - \nu\mathrm{id})^{-1}x + \omega x$。令 $z = \omega(\frac{1}{\gamma}y - x)$，我们有

$$\frac{1}{\gamma}y - \frac{1}{\omega}z \in (A - \nu\mathrm{id})z \iff \frac{1}{\gamma}y \in (A + (\frac{1}{\omega} - \nu)\mathrm{id})z \iff y \in (\gamma A + \mathrm{id})z, \tag{B.7}$$

其中，在最后一个等价变形中，我们使用了如下事实：$\gamma = \frac{\omega}{1-\omega\nu}$。由于 y 的选取是任意的，$J_{\gamma A}$ 具有全定义域。逆命题的证明类似。假设 $\mathcal{R}(I + \gamma A) = \mathbb{R}^n$。对于任意固定的 $y \in \mathbb{R}^n$，存在某个 $z \in \mathbb{R}^n$，使得 $y \in \gamma Az + z$。因此，根据 (B.7)，并令 $x = \frac{1}{\gamma}y - \frac{1}{\omega}z$，我们得到 $\frac{\omega}{\gamma}y \in \hat{A}x$，从而证明了 $\mathcal{R}(\hat{A}) = \mathbb{R}^n$。

最后，考虑 $[\mu]_+[\rho]_+ = \frac{1}{4}$ 的情形。此时，(4.3) 中步长的取值范围缩小为 $\gamma \in (0, +\infty)$。首先，假设 $J_{\gamma A}$ 具有全定义域，即 $\mathcal{R}(\mathrm{id} + \gamma A) = \mathbb{R}^n$。固定 $v \in \mathbb{R}^n$。于是，存在 $(x, y) \in \operatorname{gph} A$，使得 $v = x + \gamma y$。根据命题 4.3(iii)-1.，$y = 2\mu x + c$，其中某个 $c \in \mathbb{R}$ 以及某个 $x \in \operatorname{dom} A$，因此，$v = (1 + 2\gamma\mu)x + c\gamma$。由于 $1 + 2\gamma\mu > 0$，且 v 的选取是任意的，必然有 $\operatorname{dom} A = \mathbb{R}^n$，因此 A 是极大 (μ, ρ)-半单调的，见命题 4.3(iii)-2.。现在，假设 A 是极大 (μ, ρ)- 半单调的。于是，根据命题 4.3(iii)-2.，有 $A : x \mapsto 2\mu x + c$，其中某个 $c \in \mathbb{R}^n$，这意味着 A 是极大单调的。因此，根据 Minty 定理 [9, 定理 21.1]，预解算子 $J_{\gamma A}$ 具有全定义域。

♠ 4.12(ii)：取 (u, v)、$(u', v') \in$ 图 $J_{\gamma A}$。由 A 的半单调性，并利用 $\frac{1}{\gamma}(u - v) \in Av, \frac{1}{\gamma}(u' - v') \in Av'$，我们有

$$\mu\|v - v'\|^2 + \rho\gamma^{-2}\|(u - v) - (u' - v')\|^2 \le \gamma^{-1}\langle (u - v) - (u' - v'), v - v'\rangle.$$

展开左侧的最后一项并重新整理，得到

$$\begin{aligned}\left(1 + \mu\gamma + \rho\gamma^{-1}\right)\|v - v'\|^2 + \rho\gamma^{-1}\|u - u'\|^2 &\le \left(1 + 2\rho\gamma^{-1}\right)\langle u - u', v - v'\rangle \\ &\le \left|1 + 2\rho\gamma^{-1}\right|\|u - u'\|\|v - v'\|.\end{aligned} \tag{B.8}$$

根据 事实 A.3*(ii)* 可知，(4.3) 给出了 γ 关于 $1 + \mu\gamma + \rho\gamma^{-1} > 0$ 的所有严格正解。因此，如果 $u = u'$，那么必然有 $v = v'$，此时 Lipschitz 不等式显然成立。否则，将两边同时除以 $\|u - u'\|^2$，得到

$$\left(1 + \mu\gamma + \rho\gamma^{-1}\right)\omega^2 - \left|1 + 2\rho\gamma^{-1}\right|\omega + \rho\gamma^{-1} \le 0,$$

其中 $\omega := \frac{\|v - v'\|}{\|u - u'\|}$。对该不等式求解 ω，即可得到所述 Lipschitz 模数。 □

B.3 半单调次微分的函数示例

证明 命题 ***4.14***（具有上曲率和下曲率的函数）。

♠ 4.14*(i)*：根据定义可得。♠ 4.14*(ii)*：根据 [77, 命题 12。60] 以及 $-f$ 的强凸性，共轭函数 $(-f)^*$ 是 $-\ell^{-1}$- 光滑的，即

$$\langle \nabla(-f)^*(x) - \nabla(-f)^*(x'), x - x'\rangle \le -\tfrac{1}{\ell}\|x - x'\|^2, \quad \forall x, x' \in \mathbb{R}^n. \tag{B.9}$$

此外，由于 $-f$ 是连续凸函数，根据 [77, 命题 11。3] 可知 $\nabla(-f)^* = (\partial(-f))^{-1}$，因此（B.9）等价于

$$\langle y - y', x - x'\rangle \le -\tfrac{1}{\ell}\|y - y'\|^2, \quad \forall (x, y), (x', y') \in \operatorname{gph} \partial(-f). \tag{B.10}$$

最后，由于 $-f$ 是有限凸函数，根据 [77, 推论 9。21] 可知 $\partial f(\bar{x}) \subseteq -\partial(-f)(\bar{x}), \ \forall \bar{x} \in \mathbb{R}^n$，因此

$$\langle y - y', x - x'\rangle \ge \tfrac{1}{\ell}\|y - y'\|^2, \quad \forall (x, y), (x', y') \in \operatorname{gph} \partial f \subseteq -\operatorname{gph} \partial(-f).$$

♠ 4.14*(iii)*：考虑函数 $\varphi = f - \frac{\alpha}{2}\|\cdot\|^2$。那么，根据定义 $\varphi \in \mathscr{F}^{\ell-\alpha}(\mathbb{R}^n)$，这意味着 $\partial\varphi$ 由于 命题 4.14~~*(ii)*~~ 而是 $(0, 1\,1-\alpha)$-半单调的，因为 $\alpha > 1$。由于 $\partial f = \partial\varphi + \alpha\mathrm{id}$，[72, 例 8.8(c)], 可由命题 4.8 得出，对于 $c > \frac{1}{\alpha - 1}$，所声称的半单调性成立。

♠ 4.14(iv), 4.14(v)：假设 $1 > \sigma$，因为如果 $1 = \sigma$，那么 $\nabla f = \mathrm{id}$，结论显然成立。考虑函数 $\varphi = f - \frac{\sigma}{2}\|\cdot\|^2 \in \mathscr{F}_0^{\ell-\sigma}(\mathbb{R}^n)$。由于 φ 是有限凸函数，根据 [84, 定理 23.4]，其次微分 $\partial\varphi(x)$ 对所有 $x \in \mathbb{R}^n$ [72, 均非空。因此，根据 [84, 引理 2.1]，φ 是连续可微的；并且根据 [60, 定理 2.1.10]，$\nabla\varphi$ 是 $(0, 1\,1-\sigma)$-半单调的。因此，根据命题 4.4(i)，存在单调映射 $\tilde{A}$，使得 $(\nabla\varphi)^{-1} = \tilde{A} + \frac{1}{1-\sigma}$；由此，根据 $\nabla f = \nabla\varphi + \sigma\mathrm{id}$ [77, 例 8.8(c)]，可得

$$\nabla f = \nabla\varphi + \sigma\mathrm{id} = (\tilde{A} + \tfrac{1}{\ell - \sigma}\mathrm{id})^{-1} + \sigma\mathrm{id}.$$

所声称的结果来自 命题 4.14*(iv)* 以及 4.14*(v)*，随后由 推论 4.9 得到，其中令 $\xi = \frac{1}{1-\sigma}$ 和 $\nu = \sigma$。这是因为 $1 - \sigma > 0$，因此 $1 + 2\xi\nu > 0$（分别为 ≤ 0）当且仅当 $\sigma + 1 > 0$（分别为 ≤ 0）。 □

推论 4.15 的证明（逐点最小值）。由于 $f_i \in \mathscr{F}(\mathbb{R}^n)$，因此 $\frac{1}{2}\|\cdot\|^2 - f_i$ 是凸的。因此，$\frac{1}{2}\|x\|^2 - f(x) = \max\left\{\frac{1}{2}\|x\|^2 - f_1(x), \dots, \frac{1}{2}\|x\|^2 - f_N(x)\right\}$ 作为凸函数的逐点最大值是凸的，故 $f \in \mathscr{F}(\mathbb{R}^n)$。所述结果随即由命题 4.14(ii) 和 4.14(iii) 得出。

□

证明命题 ***4.16***（下确界卷积）。

♠ 4.16(i)：设 $f(w) = f_1(w_1) + f_2(w_2)$，并令 L 为线性映射 $Lw = w_1 + w_2$。那么，下确界卷积可以视为关于变量 w 对 $f(w, s) = f(w) + \delta_{\{0\}}(Lw - s)$ 进行的参数化最小化，即 $\varphi(s) = \min_w f(w, s)$。令 $M(s) := \operatorname{极小点集合}_{(w_1,w_2)}\{f_1(w_1) + f_2(w_2) \mid w_1 + w_2 = s\}$。那么，由于 (4.4)，根据 [77, 定理 10.13]，对所有满足 $s \in \mathbb{R}^n$ 的情形，均有

$$\partial\varphi(\bar{s}) \subseteq \bigcup_{\bar{x}\in M(\bar{s})} \left\{y \in R^n \mid (0, y) \in \partial\tilde{f}(\bar{x}, \bar{s})\right\}. \tag{B.11}$$

考虑 $f(w, u) = f(w) + \delta_{\{0\}}(u)$ 以及线性变换 $K\colon (w, s) \mapsto (w, Lw - s)$。那么，$f = f \circ K$。由于 K 是满秩的，根据 [77, 例 10.7]，对于任意满足 $x \in M(s)$ 的数对 (x, s)，有：

$$\partial\tilde{f}(\bar{x}, \bar{s}) = (K^\top \circ \partial\bar{f} \circ K)(\bar{x}, \bar{s}) \quad \text{where} \quad \partial\bar{f}(w, u) = (\partial f(w), N_{\{0\}}(u)).$$

将此代入（B.11）可得

$$\partial\varphi(\bar{s}) \subseteq \bigcup_{\bar{x}\in M(\bar{s})} \left\{y \in R^n \mid (0, y) \in \left(\partial f(\bar{x}) + L^\top N_{\{0\}}(L\bar{x} - \bar{s}), -N_{\{0\}}(L\bar{x} - \bar{s})\right)\right\} = \bigcup_{(\bar{x}_1,\bar{x}_2)\in M(\bar{s})} \partial f_1(\bar{x}_1) \cap \partial f_2(\bar{x}_2).$$

因此，所述结果由 [9,命题 25.30(i)]得出。

♠ 4.16*(ii)*：由命题 4.16*(i)*和命题 4.7*(ii)*立即可得。 □

命题 ***4.17*** 的证明（鞍点算子）。考虑鞍包络

$$\varphi_\delta(x, y) := \min_{u\in\mathbb{R}^n} \max_{v\in\mathbb{R}^m} \{\varphi(u, v) + \tfrac{1}{2\delta}\|u - x\|^2 - \tfrac{1}{2\delta}\|v - y\|^2\},$$

以及相应的预解算子 $J_{\delta T_\varphi} := (\mathrm{id} + \delta T_\varphi)^{-1}$。对于任意 $(z, w) \in$ 算子 $J_{\delta T_\varphi}$ 的图，其中分解为 $z = (x, y)$ 和 $w = (u, v)$，根据构造可知

$$0 = \nabla_x\varphi(w) + \tfrac{1}{\delta}(u - x) = \nabla_x\varphi(w) - \nabla_x\varphi_\delta(z) \quad \text{and} \quad 0 = \nabla_y\varphi(w) - \tfrac{1}{\delta}(v - y) = \nabla_y\varphi(w) - \nabla_y\varphi_\delta(z).$$

因此，

$$T_\varphi(w) = (\nabla_x\varphi(w), -\nabla_y\varphi(w)) = (\nabla_x\varphi_\delta(z), -\nabla_y\varphi_\delta(z)) =: T_{\varphi_\delta}(z). \tag{B.12}$$

另一方面，由 [40, 命题 1]可知

$$\tfrac{1}{\delta}\mathrm{I} \succeq \nabla_{xx}\varphi_\delta(z) \succeq \tfrac{\alpha}{1+\delta\alpha}\mathrm{I}, \quad \tfrac{1}{\delta}\mathrm{I} \succeq -\nabla_{yy}\varphi_\delta(z) \succeq \tfrac{\alpha}{1+\delta\alpha}\mathrm{I}, \quad \forall z, \tag{B.13}$$

因此可知，算子 T_{φ_δ} 是 $\frac{\alpha}{1+\delta\alpha}$-单调的。设 $w_1, w_2 \in \mathbb{R}^n$ 且 $z_i = w_i + \delta T_\varphi w_i$，其中 $i = 1、2$。于是，对于 $i = 1、2$，$(z_i, w_i) \in$ 属于 $J_{\delta T_\varphi}$ 的图，并且我们有

$$\begin{aligned}
\langle T_\varphi w_1 - T_\varphi w_2, w_1 - w_2\rangle &= \langle T_\varphi w_1 - T_\varphi w_2, z_1 - z_2\rangle - \delta\|T_\varphi w_1 - T_\varphi w_2\|^2 \\
&\overset{\text{(B.12)}}{=} \langle T_{\varphi_\delta} z_1 - T_{\varphi_\delta} z_2, z_1 - z_2\rangle - \delta\|T_\varphi w_1 - T_\varphi w_2\|^2 \\
&\overset{\text{(B.13)}}{\geq} \tfrac{\alpha}{1+\alpha\delta}\|z_1 - z_2\|^2 - \delta\|T_\varphi w_1 - T_\varphi w_2\|^2.
\end{aligned} \tag{B.14}$$

第二个断言由上述不等式得出，注意到 $1 + \alpha\delta > 0$。此外，将 $z_i = w_i + \delta T_\varphi w_i$ 代入（B.14），再乘以（$1 + \alpha\delta$），得到

$$(1 - \alpha\delta)\langle T_\varphi w_1 - T_\varphi w_2, w_1 - w_2\rangle \geq \alpha\|w_1 - w_2\|^2 - \delta\|T_\varphi w_1 - T_\varphi w_2\|^2,$$

从而证明了第一个断言。 □

B.4 适用于半单调算子的 Douglas-Rachford 分裂

引理 5.1（原始-对偶算子）的证明。令 $(\mathrm{x}_A, \mathrm{y}_A) \in \mathrm{A}$ 的图，且 $(\mathrm{x}_B, \mathrm{y}_B) \in \mathrm{B}$ 的图。此外，如果假设 IV 成立，令 $(\mathrm{x}'_A, \mathrm{y}'_A) \in \mathrm{A}$ 的图，且 $(\mathrm{x}'_B, \mathrm{y}'_B) \in \mathrm{B}$ 的图。否则，如果假设 III 成立，令 $(\mathrm{x}^\star, \mathrm{y}^\star) \in \mathcal{S}^\star$ 并定义 $(\mathrm{x}'_A, \mathrm{y}'_A) := (\mathrm{x}^\star, -\mathrm{y}^\star) \in \mathrm{A}$ 的图，且 $(\mathrm{x}'_B, \mathrm{y}'_B) := (\mathrm{x}^\star, \mathrm{y}^\star) \in \mathrm{B}$ 的图。于是，根据 (PD-I) 中原始-对偶算子的定义，有 马会

$$(z, v) := \big((x_A, y_B), (y_A + y_B, x_B - x_A)\big) \in \operatorname{gph} T_{\mathrm{PD}} \quad \text{and} \quad (z', v') := \big((x'_A, y'_B), (y'_A + y'_B, x'_B - x'_A)\big) \in \operatorname{gph} T_{\mathrm{PD}}.$$

由于半单调性假设以及 命题 4.4*(i)*，这意味着

$$\begin{aligned} \langle v - v', z - z' \rangle &= \langle y_A + y_B - (y'_A + y'_B), x_A - x'_A \rangle + \langle x_B - x_A - (x'_B - x'_A), y_B - y'_B \rangle \\ &= \langle y_A - y'_A, x_A - x'_A \rangle + \langle x_B - x'_B, y_B - y'_B \rangle \\ &\geq \mu_A \|x_A - x'_A\|^2 + \rho_A \|y_A - y'_A\|^2 + \rho_B \|y_B - y'_B\|^2 + \mu_B \|x_B - x'_B\|^2. \end{aligned} \tag{B.15}$$

另一方面，由柯西–施瓦茨不等式可得

$$\begin{aligned} \mathrm{q}_V(v - v') &= \beta_{\mathrm{P}} \|y_A + y_B - (y'_A + y'_B)\|^2 + \beta_{\mathrm{D}} \|x_B - x_A - (x'_B - x'_A)\|^2 \\ &\leq \beta_{\mathrm{P}} \|y_A - y'_A\|^2 + \beta_{\mathrm{P}} \|y_B - y'_B\|^2 + 2|\beta_{\mathrm{P}}| \|y_A - y'_A\| \|y_B - y'_B\| \\ &\quad + \beta_{\mathrm{D}} \|x_B - x'_B\|^2 + \beta_{\mathrm{D}} \|x_A - x'_A\|^2 + 2|\beta_{\mathrm{D}}| \|x_B - x'_B\| \|x_A - x'_A\|. \end{aligned} \tag{B.16}$$

结合 (B.15) 和 (B.16) 可知，若假设 III 成立且 $\mathrm{X} := \begin{pmatrix} \mu_A - \beta_{\mathrm{D}} & -|\beta_{\mathrm{D}}| \\ -|\beta_{\mathrm{D}}| & \mu_B - \beta_{\mathrm{D}} \end{pmatrix} \succeq 0$，$\mathrm{Y} := \begin{pmatrix} \rho_B - \beta_{\mathrm{P}} & -|\beta_{\mathrm{P}}| \\ -\beta_{\mathrm{P}} & \rho_A - |\beta_{\mathrm{P}}| \end{pmatrix} \succeq 0$，则原始-对偶算子 T_{PD} 具有位于 $\mathcal{S}^\star$ 的 V-斜弱 Minty 解；若假设 IV 成立，则该算子是 V-余单调。如前文在命题 4.7(i) 的证明中所述，满足 $\mathrm{X} \succeq 0$ 的 β_{D} 的最大可能值为 $\mu_A \square \mu_B$，前提是 $\mu_A + \mu_B > 0$ 或 $\mu_A = \mu_B = 0$（类似地，满足 $\mathrm{Y} \succeq 0$ 的 β_{P} 的最大值为 $\rho_A \square \rho_B$，前提是 $\rho_A + \rho_B > 0$ 或 $\rho_A = \rho_B = 0$），由此完成引理 5.1(i) 的证明。为证明引理 5.1(ii)，只需证明极大性。为此，考虑任意一个符合表 3 的步长 γ，其中 γ_- 和 γ_+ 按 (5.2) 定义。于是，根据引理 A.6，预解算子 $\mathrm{J}_{\gamma A}$ 和 $\mathrm{J}_{\gamma B}$ 具有全定义域，这又意味着预条件化预解算子 $(\mathrm{P} + \mathrm{T}_{\mathrm{PD}})^{-1}\mathrm{P}$ 具有全定义域，其中 P 由 (3.1) [77, 引理 12.14] 给出。结论随即由命题 A.2 得出。 □

证明 定理 ***5.2***（半单调性下 ***DRS*** 的收敛）。考虑以下蕴含关系 s.

♠ 对于 定理 5.2*(i)*，由 假设 III 和 引理 5.1*(i)*可知 假设 II.A3 成立。

♠ 对于 定理 5.2*(ii)*，由 假设 IV 和 引理 A.6 可知 假设 II.A1和 II.A2成立；由 假设 IV 和 引理 5.1*(ii)* 可知 T_{PD} 是极大 V_--共单调，其中 $V_- := \operatorname{blkdiag}(\beta_{\mathrm{P}}\mathrm{I}_n, \beta_{\mathrm{D}}\mathrm{I}_n)$，$[\beta_{\mathrm{D}}]_- [\beta_{\mathrm{P}}]_- < \frac{1}{4}$。

此外，步长 γ 满足 (3.3，因为 γ 的表 3 所给界是将 (5.1) 代入 (3.3) 得到的。基于这些观察，结论可直接由 定理 3.3 和 推论 3.4得出，证明完毕。 □

C 示例

详情：例 ***2.6***（玩具示例）。

♠ 2.6*(i)*-1。：由于T等于线性映射$T = \begin{bmatrix} b & a \\ -a & b \end{bmatrix}$，因此它是$\rho$共单调的，当且仅当$\frac{1}{2}(T + T^\top) \preceq \rho T^\top T$，这在$\rho \leq b/(a^2+b^2)$时成立。

♠ 2.6*(i)*-2.：为了说明定理 2.4*(v)*中的结果是紧的，定义

$$H := \mathrm{I}_2 + \lambda\Big((\mathrm{I}_2+\gamma T)^{-1} - \mathrm{I}_2\Big) = \mathrm{I}_2 + \lambda\left(\begin{bmatrix} 1+\gamma b & \gamma a \\ -\gamma a & 1+\gamma b \end{bmatrix}^{-1} - \mathrm{I}_2\right) = \mathrm{I}_2 + \lambda\left(\tfrac{1}{(1+\gamma b)^2+\gamma^2 a^2}\begin{bmatrix} 1+\gamma b & -\gamma a \\ \gamma a & 1+\gamma b \end{bmatrix} - \mathrm{I}_2\right),$$

使得采用固定松弛参数的近端点算法的迭代 λ 可表示为线性动力系统$z^{k+1} = Hz^k$。当且仅当H的谱半径 $\sqrt{1 - \frac{\lambda}{(2\bar{\lambda}-\lambda)(a^2+b^2)\gamma^2}\frac{(a^2+b^2)\gamma^2+2b\gamma+1}{}}$严格小于 1 时，该系统全局渐近稳定；而这在当且仅当 (2.18)成立时满足。

♠ 2.6*(ii)*-1.：回想一下，对于给定的 $\rho \in \mathbb{R}$，该问题的 ρ-弱 Minty 解构成如下集合$\mathcal{S}^\star \subseteq$ 零点集T，满足

$$\langle T(z), z - z^\star\rangle \geq \rho\|T(z)\|^2, \quad \forall z \in \mathbb{R}^2, z^\star \in \mathcal{S}^\star. \tag{C.1}$$

图 3 中所定义的f的解析表达式为

$$f\big(\|z\|\big) := \begin{cases} \|z\|, & \text{if } \|z\| \leq 2/5, \\ 4/5 - \|z\|, & \text{if } \|z\| \in (2/5, 4/5), \\ 0, & \text{if } \|z\| \in [4/5, 1], \\ 5/2(\|z\| - 1), & \text{if } \|z\| \in (1, 7/5), \\ 1, & \text{if } \|z\| \geq 7/5. \end{cases} \tag{C.2}$$

根据T的定义，有零点集$T = \left\{z \in \mathbb{R}^2 \mid \|z\| \in \{0\} \cup [4/5, \quad 1]\right\}$并且 (C.1)对应于 to 因此

$$f\big(\|z\|\big)\Big(b\|z\|^2 - b(z_1 z_1^\star + z_2 z_2^\star) + a(z_1 z_2^\star - z_1^\star z_2)\Big) \geq \rho(a^2+b^2) f\big(\|z\|\big)^2\|z\|^2, \qquad \forall z \in \mathbb{R}^2, \ z^\star \in \mathcal{S}^\star。 \tag{C.3}$$

显然，对于所有 $z \in$零点集 T，该结论都成立。对于 $z \notin$零点集 T，我们区分以下两种情况。

◇ $\|z^\star\| = 0$：此时，(C.3) 成立，当且仅当对于所有 $z \notin$零点集 T，都有 $\rho \leq b(a^2+b^2)f(\|z\|)$。由于 $f(\|z\|) \in (0, 1]$, 该条件对于 $\rho = b/a^2+b^2$ 得到满足。

◇ $\|z^\star\| \in [4/5, 1]$：令 $z = \epsilon z^\star$，其中 $\epsilon \in (0, 2/5)$。于是，$f(\|z\|) = \epsilon\|z^\star\|$，并且 (C.3) 成立，当且仅当 $\rho \leq b(1-\epsilon^{-1})\epsilon(a^2+b^2)\|z^\star\|$。令 ϵ从上方趋于零，得到 $\lim_{\epsilon\to 0+} b(1-\epsilon^{-1})\epsilon(a^2+b^2)\|z^\star\| = -\infty$，因此对于任意 $\rho \in \mathbb{R}$，(C.3) 都不可能成立。 □

例例**2.7** （冯·诺伊曼经济均衡模型）。我们首先为集合 $C := \Delta_2 \times \Delta_2$ 以及算子 F 和 N_C建立三个中间结论。首先，根据单纯形的定义可知

$$z \in C \iff z = (s, 1-s, t, 1-t) \text{ for some } s, t \in [0, 1]. \tag{C.4}$$

其次，回顾一下，

$$F(z) = (F_1(z), F_2(z), F_3(z), F_4(z)) = \Big(\nabla_x f(x,y), -\nabla_y f(x,y)\Big) = \left(\tfrac{\langle x, Sy\rangle Ry - \langle x, Ry\rangle Sy}{\langle x, Sy\rangle^2}, \tfrac{\langle x, Ry\rangle S^\top x - \langle x, Sy\rangle R^\top x}{\langle x, Sy\rangle^2}\right).$$

因此，利用（C.4）中的参数化，对于任意 $z \in C$，都有

$$\langle F(z), z - z^\star\rangle = \tfrac{\epsilon(s+t) - st(\epsilon+s)}{1/2(2-s)^2}, \quad F_1(z) - F_2(z) = \tfrac{\epsilon - t(2+\epsilon/2)}{1/2(2-s)^2} \quad \text{and} \quad F_3(z) - F_4(z) = \tfrac{\epsilon/2 + s}{1/2(2-s)}. \tag{C.5}$$

最后，将维数为 n 的全 1 向量记为 $\mathbf{1}_n$，由 [43, 例5.2.6(c)] 可知

$$N_C(z) = \begin{cases} \Big\{\alpha \otimes \mathbf{1}_2 - \beta \,\Big|\, \alpha \in \mathbb{R}^2, \beta \in \mathbb{R}^4_+, \langle \beta, z\rangle = 0\Big\}, & \text{if } z \in C, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{C.6}$$

现在我们开始论述主要论证。在 [22, 命题 2(1)], 中已证明，$z^\star = (0,1,0,1)$ 是（2.19）的唯一纳什均衡（注意到$f(x,y) - 2 = \langle x, \begin{bmatrix} -1 & \epsilon/-\epsilon \\ 2/2 & 0 \end{bmatrix} y/\langle x, Sy\rangle\big)$ 。该纳什均衡是一个 ρ_ϵ-弱

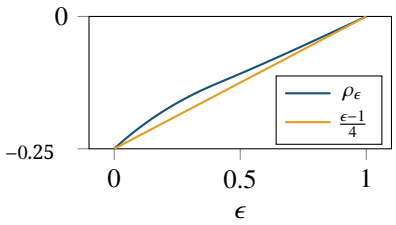

图 9： 最优弱 *Minty* 常数与其下界的比较 ρ_ϵ $\frac{\epsilon-1}{4}$ 对于例2.7。 —

T 的明蒂解，T 前提是 ρ_ϵ = 下确界$_{z\in C}\varrho_\epsilon(z)$，其中 $\varrho_\epsilon(z)$ = 下确界$_{v\in N_C(z)}\langle F(z)+v, z-z^\star\rangle/\|F(z)+v\|^2$。从现在起，我们不加说明地假设 $z \in C$。于是，由于（C.6），有

$$\varrho_\epsilon(z) = \inf\left\{\frac{\langle F(z),z-z^\star\rangle+\langle\alpha\otimes\mathbf{1}_2-\beta,z-z^\star\rangle}{\|F(z)+\alpha\otimes\mathbf{1}_2-\beta\|^2} \mid \alpha\in\mathbb{R}^2,\beta\in\mathbb{R}^4_+,\ \langle\beta,z\rangle=0\right\}.$$

由于参数化来自（C.4），且$z^\star = (0,\ 1,\ 0,\ 1)$，因此有 $\langle\alpha\otimes\mathbf{1}_2,\ z-z^\star\rangle = 0$ 且$\langle\beta,\ z^\star\rangle = \beta_2+\beta_4$，因此

$$\varrho_\epsilon(z) = \inf\left\{\frac{\langle F(z),z-z^\star\rangle+\beta_2+\beta_4}{\|F(z)+\alpha\otimes\mathbf{1}_2-\beta\|^2} \mid \alpha\in\mathbb{R}^2,\beta\in\mathbb{R}^4_+,\ \langle\beta,z\rangle=0\right\}.$$

注意，根据构造，$\varrho_\epsilon(z)$ 的符号与 $\langle F(z),z-z^\star\rangle$ 的符号相同：如果 $\langle F(z),z-z^\star\rangle$ 为正，这显然成立，因为 $\beta_i \geq 0$；如果为负，则可通过取候选解 $\beta_2 = \beta_4 = 0$ 看出。因此，由 (C.5) 可知，$\varrho_\epsilon(z) < 0$ 当且仅当 $\epsilon(s+t) - st(\epsilon+s) < 0$，也就是说，当且仅当 z 属于以下非空集合

$$D := \left\{(s,1-s,t,1-t) \mid s\in(\sqrt{\epsilon},1],\ t\in(\bar{t}(s),1]\right\} \quad \text{where} \quad \bar{t}(s) := \frac{\epsilon s}{s(\epsilon+s)-\epsilon}.$$

因此，$\rho_\epsilon = \inf_{z\in C}\varrho_\epsilon(z) = \inf_{z\in D}\varrho_\epsilon(z)$。考虑$z \in D \subset C$。那么，根据构造，$\varrho_\epsilon(z)$为负，这意味着对于使分母最小的 α，$\varrho_\epsilon(z)$取得最小值，即

$$\arg\min_{\alpha\in\mathbb{R}^2}\|F(z)+\alpha\otimes\mathbf{1}_2-\beta\|^2 = \left(\tfrac{1}{2}(F_1(z)-F_2(z)-\beta_1+\beta_2),\tfrac{1}{2}(F_3(z)-F_4(z)-\beta_3+\beta_4)\right),$$

这意味着

$$\varrho_\epsilon(z) = \inf\left\{\frac{2\langle F(z),z-z^\star\rangle+2\beta_2+2\beta_4}{\left(F_1(z)-\beta_1-F_2(z)+\beta_2\right)^2+\left(F_3(z)-\beta_3-F_4(z)+\beta_4\right)^2} \mid \beta\in\mathbb{R}^4_+,\ \langle\beta,z\rangle=0\right\}.$$

代入 参数化式z（来自C.4），利用（C.5）并引入$\bar{\beta} := {}^1\!/_2\,(2-s)^2\beta$，可得 at

$$\varrho_\epsilon(s,t) = \inf\left\{\frac{4(2-s)^2\left(\epsilon(s+t)-st(\epsilon+s)+\bar{\beta}_2+\bar{\beta}_4\right)}{\left(2\epsilon-t(4+\epsilon)-2\bar{\beta}_1+2\bar{\beta}_2\right)^2+\left((\epsilon+2s)(2-s)-2\bar{\beta}_3+2\bar{\beta}_4\right)^2} \mid \bar{\beta}\in\mathbb{R}^4_+,\ \langle\bar{\beta},(s,1-s,t,1-t)\rangle=0\right\},$$

我们声称，对于任意$s \in (\sqrt{\epsilon}$以及$t \in (\bar{}t(s), \bar{}1]$，其在下确界处成立。对于 $\bar{\beta}_1$ 和 $\bar{\beta}_3$，这由等式约束以及s 和t 严格为正这一事实立即得到，因此

$$\varrho_\epsilon(s,t) = \inf\left\{\frac{4(2-s)^2\left(\epsilon(s+t)-st(\epsilon+s)+\bar{\beta}_2+\bar{\beta}_4\right)}{\left(2\epsilon-t(4+\epsilon)+2\bar{\beta}_2\right)^2+\left((\epsilon+2s)(2-s)+2\bar{\beta}_4\right)^2} \mid (\bar{\beta}_2,\bar{\beta}_4)\in\mathbb{R}^2_+,\ \langle(\bar{\beta}_2,\bar{\beta}_4),(1-s,1-t)\rangle=0\right\}.$$

若$s \in (\sqrt{\epsilon},\ 1)$ 且$t \in (\bar{}t(s),\ 1)$，则对于 $\bar{\beta}_2$ 和 $\bar{\beta}_4$ 的断言也可由等式约束得到。否则，考虑以下情形。

♠ 若 $s = 1$ 和 $t \in (\bar{}t(1), 1) = (\epsilon, 1)$，则 $\beta_4 = 0$ 由于等式约束以及

$$\varrho_\epsilon(1,t) = \inf_{\bar{\beta}_2\in\mathbb{R}_+}\frac{4(\epsilon-t+\bar{\beta}_2)}{(2\epsilon-t(4+\epsilon)+2\bar{\beta}_2)^2+(\epsilon+2)^2} = \inf_{\bar{\beta}_2\in[0,t-\epsilon)}\frac{4(\epsilon-t+\bar{\beta}_2)}{(2\epsilon-t(4+\epsilon)+2\bar{\beta}_2)^2+(\epsilon+2)^2}.$$

- ¯由于该函数在 $\beta_2 \in [0, t-\epsilon)$ 上单调递增，因此下确界在 $\beta_2 =$ 处取得 0.

♠ 如果$s \in (\sqrt{\epsilon}, 1)$且$t = 1$，那么$\bar{\beta}_2 = 0$由于等式约束和

$$\varrho_\epsilon(s,1) = \inf_{\bar{\beta}_4 \in \mathbb{R}_+} \frac{4(2-s)^2(\epsilon - s^2 + \bar{\beta}_4)}{(\epsilon-4)^2 + \big((\epsilon+2s)(2-s) + 2\bar{\beta}_4\big)^2} = \inf_{\bar{\beta}_4 \in [0, s^2-\epsilon)} \frac{4(2-s)^2(\epsilon - s^2 + \bar{\beta}_4)}{(\epsilon-4)^2 + \big((\epsilon+2s)(2-s) + 2\bar{\beta}_4\big)^2}.$$

由于该函数在$\bar{\beta}_4 \in [0, s^2 - \epsilon)$上单调递增，因此下确界在$\bar{\beta}_4 =$处取得 0.

♠ 如果$s = 1$且$t = 1$，那么

$$\varrho_\epsilon(1,1) = \inf_{(\bar{\beta}_2, \bar{\beta}_4) \in \mathbb{R}^2_+} \frac{4(\epsilon - 1 + \bar{\beta}_2 + \bar{\beta}_4)}{(\epsilon - 4 + 2\bar{\beta}_2)^2 + (\epsilon + 2 + 2\bar{\beta}_4)^2},$$

其全局最小值在$\beta_2 = \beta_4 = 0$处取得。

既然已确定下确界处的$\beta = 0$，便可得

$$\rho_\epsilon = \inf\left\{ \varrho_\epsilon(s,t) \mid s \in (\sqrt{\epsilon}, 1],\ t \in \left(\frac{\epsilon s}{s(\epsilon+s) - \epsilon}, 1\right] \right\} \quad \text{where} \quad \varrho_\epsilon(s,t) = \frac{4(2-s)^2\big(\epsilon(s+t) - st(\epsilon+s)\big)}{\big(2\epsilon - t(4+\epsilon)\big)^2 + (\epsilon+2s)^2(2-s)^2}. \tag{C.7}$$

最后，通过数值求解此参数优化问题，得到下界$\frac{\epsilon-1}{4} \le \rho_\epsilon$（另见 图 9）。

□

例***6.1*** 的细节 （鞍点问题）。

♠ 6.1(i)：通过定义$\bar{\mathrm{H}} := (\mathrm{I} + \gamma \mathrm{A})^{-1}$，$\widehat{\mathrm{H}} := (\mathrm{I} + \gamma \mathrm{B})^{-1}(2\mathrm{H} - \mathrm{I})$和$\mathrm{H} := \mathrm{I} + \lambda(\widehat{\mathrm{H}} - \mathrm{H})$，DRS 可表示为线性动力系统$s^{k+1} = \mathrm{H}s^k$。该系统全局渐近稳定，当且仅当 H 的谱半径严格小于 1；通过初等代数可以证明，这一条件成立当且仅当满足 (6.2)。

♠ 6.1*(ii)*：原始-对偶算子及其逆为

$$T_{\mathrm{PD}} = \begin{bmatrix} A & \mathrm{I}_2 \\ -\mathrm{I}_2 & B^{-1} \end{bmatrix} \quad \text{and} \quad T_{\mathrm{PD}}^{-1} = \begin{bmatrix} (A+B)^{-1} & -(A+B)^{-1}B \\ B(A+B)^{-1} & B - B(A+B)^{-1}B \end{bmatrix}$$

其中我们使用了舒尔补引理。因此，在 (2.1) 中将向量v限制为$\mathcal{R}(P)$时，我们也可以等价地表述 假设II。A3 如下

$$z^\top \left(\frac{T_{\mathrm{PD}} + T_{\mathrm{PD}}^\top}{2} - T_{\mathrm{PD}}^\top V T_{\mathrm{PD}} \right) z \ge 0, \qquad \text{for all } z \in \mathbb{R}^n : z \in T_{\mathrm{PD}}^{-1}\mathcal{R}(P), \tag{C.8}$$

其中V定义如 (3.2) 所示。利用$U = \begin{bmatrix} \mathrm{I} \\ -\gamma \mathrm{I}_2 \end{bmatrix}$是$\mathcal{R}(P)$的一组正交基这一点，可知 (C.8) 成立，当且仅当

$$(T_{\mathrm{PD}}^{-1}U)^\top \left(\frac{T_{\mathrm{PD}} + T_{\mathrm{PD}}^\top}{2} - T_{\mathrm{PD}}^\top V T_{\mathrm{PD}} \right) T_{\mathrm{PD}}^{-1} U \ge 0 \quad \Longleftrightarrow \quad \tfrac{1}{\gamma}\beta_{\mathrm{P}} + \gamma\beta_{\mathrm{D}} \le \tfrac{b(a^2\gamma+1)}{\gamma(a^2+b^2)}.$$

因此，由λ定理 3.3给出的上界，记为$\lambda_{\max}$，满足

$$\lambda_{\max} := 2\left(1 + \tfrac{1}{\gamma}\beta_{\mathrm{P}} + \gamma\beta_{\mathrm{D}}\right) \le 2 + 2\tfrac{b(a^2\gamma+1)}{\gamma(a^2+b^2)}.$$

由于该界与式（λ 6.2）给出的上界相同，证明完毕。

♠ 6.1*(iii)*：例如，考虑一个$= 2$，$b = -1$，其迹T_{P}、T_{D}和T_{PD}均等于 -2 .

♠ 最后，我们将讨论λ定理 5.2所涵盖的松弛参数范围。为此，需要选择半单调性参数μ_A，ρ_A、μ_B，ρ_B使松弛参数最大化，即

$$\lambda_{\max}^{\mathrm{semi}} := \max_{\mu_A, \rho_A, \mu_B, \rho_B} \ 2\left(1 + \tfrac{1}{\gamma}(\rho_A \square \rho_B) + \gamma(\mu_A \square \mu_B)\right)$$
$$\text{subject to} \quad \mu_A + a^2\rho_A \le 0 \quad \text{and} \quad \mu_B + b^2\rho_B - b \le 0,$$

其中约束条件通过采用命题 4.5 并对算子(μ_A, ρ_A)-和(μ_B, ρ_B)-半单调性A和B分别施加约束而得到。请注意，该优化问题可以编码为凸半定规划。对$a = 10$、$b = -1$进行数值求解后，得到可视化于图 10 中的结果。因此，在分别通过A和B的半单调性性质，而不是通过组合原始-对偶算子的（斜）弱 Minty 条件，来刻画该问题的非单调性时，不可避免地会引入一些松弛。

□

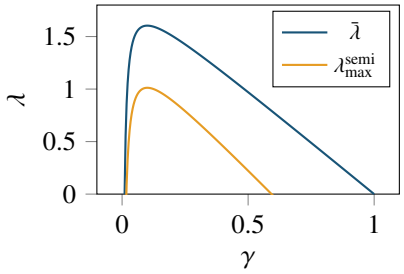

图 10: 松弛参数上界的比较，作为步长 λ 的函数，适用于 例 6.1 ，其中 $= 10$，$b = -1$。 $\bar{\lambda}$ 是由 (6.2) 给出的上界，而 $\lambda^{\text{semi}}_{\max}$ 是定理 5.2所隐含的上界。

关于例 **6.2** （非光滑优化）的详细信息。相应的极限次微分ff为

$$\partial f_A(x) = \begin{cases} -x-3, & \text{if } x < -1, \\ \{-2,-9\}, & \text{if } x = -1, \\ 6x-3, & \text{if } x \in (-1,1), \\ \{-4,3\}, & \text{if } x = 1, \\ -x-3, & \text{if } x > 1, \end{cases} \quad \text{and} \quad \partial f_B(x) = \begin{cases} 4x-1, & \text{if } x < -1, \\ [-5,1] & \text{if } x = -1, \\ 2x+3, & \text{if } x \in (-1,2), \\ [7,19], & \text{if } x = 2, \\ 2x+15, & \text{if } x > 2. \end{cases} \tag{C.9}$$

♠ 假设 II。A1：由极限次微分的定义可得ff [72,定义 8.3(b)]。

♠ 假设 II。A2：容易验证，对于 $\gamma \neq 1$ 成立 $\mathcal{R}(1+\gamma\partial f_A) = \mathcal{R}(1+\gamma\partial f_B) = \mathbb{R}$。因此，预解算子 $J_{\gamma\partial f_A}$ 和 $J_{\gamma\partial f_B}$ 具有全定义域。

♠ 假设 III：设 $(x^\star, y^\star) = (0,3)$。那么，可以直接验证，次微分 ∂f_A 和∂f_B 分别在 $(x^\star, -y^\star)$ 和 $(x^\star, y^\star)$ 处是半单调的，其参数分别为 $\mu_A = -1.2, \rho_A = 0.2$、$\mu_B = 1.6$ 和 $\rho_B = 0.1$。

♠ 由应用 定理 5.2 并利用 $\mu_A \square \mu_B = -24/5$ 以及 $\rho_A \square \rho_B = 2/30$。 □

例 6.4（驻点）的详细内容。定义 $g_1(x) := (1-\sqrt{1-4\mu_A\rho_A}/2\rho_A)x - \overline{1,\ g_2}(x) := (1+\sqrt{1-4\mu_A\rho_A}/2\rho_A)x - \overline{1,\ h_1}(x) := 2x+1$ 以及 $h_2(x) := \frac{1}{2}x + 1$。那么，$f_A$ 和 f_B 的极限次微分为

$$\partial f_A(x) = \begin{cases} \{g_1(x), g_2(x)\}, & \text{if } x \in \{-3,3\}, \\ g_1(x), & \text{if } x \in (-3,3), \\ g_2(x), & \text{otherwise}, \end{cases} \quad \text{and} \quad \partial f_B(x) = \begin{cases} \{h_1(x), h_2(x)\}, & \text{if } x \in \{-1,1\}, \\ h_1(x), & \text{if } x \in (-1,1), \\ h_2(x), & \text{otherwise}. \end{cases}$$

♠ 假设 II。A1：由极限次微分的定义可得ff [72,定义 8.3(b)]。

♠ 假设 II。A2：预解算子 $J_{\gamma\partial f_A}$和 $J_{\gamma\partial f_B}$具有全定义域，当且仅当 $\mathcal{R}(1+\gamma\partial f_A) = \mathcal{R}(1+\gamma\partial f_B) = \mathbb{R}$。容易验证，对于所有 $\gamma \in \mathbb{R}_+$， $\mathcal{R}(1+\gamma\partial f_B) = \mathbb{R}$。另一方面，我们有 玩彩神争霸

$$(1+\gamma\partial f_A)(x) = \begin{cases} x - \gamma + \left\{\frac{1-\sqrt{1-4\mu_A\rho_A}}{2\rho_A}\gamma x, \frac{1+\sqrt{1-4\mu_A\rho_A}}{2\rho_A}\gamma x\right\}, & \text{if } x \in \{-3,3\}, \\ x - \gamma + \frac{1-\sqrt{1-4\mu_A\rho_A}}{2\rho_A}\gamma x, & \text{if } x \in (-3,3), \\ x - \gamma + \frac{1+\sqrt{1-4\mu_A\rho_A}}{2\rho_A}\gamma x, & \text{otherwise} \end{cases}$$

由此可得，对于所有 $\gamma \in \mathbb{R}_+$， $\mathcal{R}(1+\gamma\partial f_A) = \mathbb{R}$ 当且仅当 γ 按照给定的值域选取。

♠ 假设 III：设 $(x^\star, y^\star) = (0,-1)$。不难验证，对于给定的参数 $\mu_A, \rho_A, \mu_B, \rho_B$，次微分 ∂f_A 和∂f_B 分别在 $(x^\star, -y^\star)$ 和 $(x^\star, y^\star)$ 处半单调。此外，这些参数满足假设 III.a1。

♠ 通过应用定理 5.2并利用$\mu_A \square \mu_B = -6/5$ 以及 $\rho_A \square \rho_B = -4/30$。 □

参考文献

[1] A. Alacaoglu、A. B¨ohm 和 Y. Malitsky，《变分不等式算法中超越黄金比例》，《机器学习研究杂志》，24（2023年），第1–33页。[2] A. Alacaoglu、D. Kim 和 S. J. Wright，《利用余拟单调性拓展非凸极小极大问题的一阶算法适用范围》，2024年，arXiv 预印本 arXiv:2402.05071。[3] E. Allevi、A. Gnudi 和 I. V. Konnov，《非单调变分不等式的近端点方法》，《运筹学数学方法》，63（2006年），第553–565页。[4] F. J. Arag´on Artacho 和 M. H. Geoffroy，《次微分度量正则性的刻画》，《凸分析杂志》，15（2008年），第365–380页。[5] F. J. Arag´on Artacho、A. Dontchev 和 M. Geoffroy，《度量正则映射的近端点方法的收敛性》，ESAIM会议录，17（2007年），第1–8页，https://doi.org/10.1051/ proc:071701。[6] H. Attouch 和 M. Th´era，《两个算子之和的一般对偶原理》，《凸分析杂志》，3（1996年），第1–24页。[7] A. Auslender 和 M. Teboulle，《优化与变分不等式中的渐近锥与函数》，施普林格科学与商业媒体，柏林，2006年。[8] S. Bartz、M. N. Dao 和 H. M. Phan，《不动点算法的锥平均性与收敛性分析》，《全球优化杂志》，82（2022年），第351–373页。[9] H. H. Bauschke 和 P. L. Combettes，《希尔伯特空间中的凸分析与单调算子理论》，CMS数学丛书，施普林格国际出版，纽约，第2版，2017年。[10] H. H. Bauschke、W. M. Moursi 和 X. Wang，《广义单调算子及其平均预解式》，《数学规划》，189（2021年），第55–74页。[11] D. S. Bernstein，《矩阵数学：理论、事实与公式》，普林斯顿大学出版社，普林斯顿，第2版，2009年。[12] A. B¨ohm，《求解具有弱 Minty 解的非凸非凹极小极大问题》，2022年，arXiv 预印本 arXiv:2201.12247。[13] K. Bredies、E. Chenchene、D. A. Lorenz 和 E. Naldi，《退化预条件近端点算法》，《SIAM优化杂志》，32（2022年），第2376–2401页。[14] H. Br´ezis 和 P. L. Lions，《预解算子的无穷乘积》，《以色列数学杂志》，29（1978年），第329–345页。[15] M. N. Bui 和 P. L. Combettes，《单调包含的扭曲近端迭代》，《数学分析与应用杂志》，491（2020年），第124315页。[16] R. S. Burachik 和 A. N. Iusem，《集值映射与单调算子的扩张》，施普林格优化及其应用丛书，施普林格，2008年6月。[17] J. V. Burke 和 M. Qian，《单调算子的变度量近端点算法》，《SIAM控制与优化杂志》，37（1999年），第353–375页。[18] A. Chambolle 和 T. Pock，《一种用于凸问题的一阶原始-对偶算法及其在成像中的应用》，《数学成像与视觉杂志》，40（2011年），第120–145页。[19] P. L. Combettes 和 T. Pennanen，《余拟单调算子的近端方法》，《SIAM控制与优化杂志》，43（2004年），第731–742页。[20] L. Condat，《一种涉及 Lipschitz 连续、可近端计算及线性复合项的凸优化原始-对偶分裂方法》，《优化理论与应用杂志》，158（2013年），第460–479页。

[21] *M. N. Dao*和*H. M. Phan*，两个算子之和的自适应*Douglas–Rachford*分裂算法，《*SIAM*优化杂志》，*29*（*2019*年），第*2697–2724*页。[22] *C. Daskalakis*、*D. J. Foster*和*N. Golowich*，竞争性强化学习中的独立策略梯度方法，《神经信息处理系统进展》，*33*（*2020*年），第*5527–5540*页。[23] *D. Davis*和*W. Yin*，若干分裂方案的收敛速率分析，*2015*年，*arXiv*预印本 *arXiv:1406.4834*。[24]*J. Diakonikolas*、*C. Daskalakis*和*M. I. Jordan*，结构化非凸-非凹极小极大优化的高效方法，载于《人工智能与统计国际会议》，*2021*年，第*2746–2754*页。[25] *A. L. Dontchev*和*R. T. Rockafellar*，《隐函数与解映射》，*Springer* 数学专著。施普林格，*208*（*2009*年）。[26] *J. Douglas*和*H. H. Rachford*，关于两个和三个空间变量热传导问题的数值解，《美国数学学会会刊》，*82*（*1956*年），第*421–439*页。[27] *D. Drusvyatskiy*和*A. S. Lewis*，近端方法的误差界、二次增长与线性收敛，《运筹学数学》，*43*（*2018*年），第*919–948*页。[28] *J. Eckstein*，*Lions–Mercier* 分裂算法和交替方向法都是近端点算法的实例，技术报告，*LIDS-P-1769*，*1988*年。[29] *J. Eckstein*，《单调算子的分裂方法及其在并行优化中的应用》，博士论文，麻省理工学院，*1989*年。[30] *J. Eckstein*，使用*Bregman* 函数的非线性近端点算法及其在凸规划中的应用，《运筹学数学》，*18*（*1993*年），第*202–226*页。[31] *J. Eckstein*和*D. P. Bertsekas*，关于极大单调算子的*Douglas–Rachford* 分裂法和近端点算法，《数学规划》，*55*（*1992*年），第*293–318*页。[32] *E. Esser*、*X. Zhang*和*T. F. Chan*，成像科学中凸优化一类一阶原始-对偶算法的通用框架，《*SIAM*影像科学杂志》，*3*（*2010*年），第*1015–1046*页。[33] *F. Facchinei*和*J.-S. Pang*，《有限维变分不等式与互补问题》第一卷，*Springer* 运筹学丛书，施普林格，纽约，*2003*年。[34] *N. Farouq*，伪单调变分不等式：近端方法的收敛性，《优化理论与应用杂志》，*109*（*2001*年），第*311–326*页。[35] *F. Giannessi*，关于*Minty* 变分原理，载于《数学规划新趋势：献给*Steven Vajda*》，*Springer US*，波士顿，*MA*，*1998*年，第*93–99*页。[36] *P. Giselsson*，带投影校正的非线性前向-后向分裂，《*SIAM*优化杂志》，*31*（*2021*年），第*2199–2226*页。[37] *P. Giselsson*和*S. Boyd*，*Douglas-Rachford* 分裂和*ADMM*的线性收敛与度量选择，《*IEEE* 自动控制汇刊》，*62*（*2017*年），第*532–544*页。[38] *E. Gorbunov*、*N. Loizou*和*G. Gidel*，外梯度法：$\mathcal{O}(1/k)$ 单调变分不等式的最后迭代收敛及其与余强单调性的联系，载于《人工智能与统计国际会议》，*2022*年，第*366–402*页。[39] *E. Gorbunov*、*A. Taylor*、*S. Horv´ath*和*G. Gidel*，超越单调性的近端点法和基于外梯度方法的收敛性：负共单调情形，载于《机器学习国际会议》，*PMLR*，*2023*年，第*11614–11641*页。

[40] B. Grimmer、H. Lu、P. Worah和V. Mirrokni，非凸–非凹极小极大优化中近端点方法的研究现状，《数学规划》，（2022年），第1–35页。[41] K. Guo、D. Han和X. Yuan，“强弱”凸规划的Douglas–Rachford分裂法收敛性分析，《SIAM数值分析杂志》，55（2017年），第1549–1577页。+ B. He和X. Yuan，鞍点问题原始-对偶算法的收敛性分析：压缩映射视角，《SIAM影像科学杂志》，5（2012年），第119–149页。[42] J.-B. Hiriart-Urruty和C. Lemarechal，《凸分析与最小化算法I：基础》，数学科学基础丛书305，施普林格，柏林，第二次校订版，1996年。[43] R. A. Horn和C. R. Johnson，《矩阵分析》，剑桥大学出版社，剑桥，2012年。[44] Z. Huang和M. A. Noor，针对具有不同（[45]，r）-余强制映射的非线性变分不等式组的显式投影法，《应用数学与计算》，190（2007年），第356–361页。γA. N. Iusem、T. Pennanen和B. F. Svaiter，无单调性条件下近端点算法的非精确变体，《SIAM优化杂志》，13（2003年），第1080–1097页。[46] I. V. Konnov，一类求解变分不等式的组合迭代方法，《优化理论与应用杂志》，94（1997年），第677–693页。[47] P. Latafat、N. M. Freris和P. Patrinos，一种用于分布式优化的新型随机块坐标原始-对偶近端算法，《IEEE 自动控制汇刊》，64（2019年），第4050–4065页。[48] P. Latafat和P. Patrinos，用于求解涉及三个算子的单调包含问题的非对称前向–后向–伴随分裂，《计算优化与应用》，68（2017年），第57–93页。[49] S. Lee和D. Kim，光滑结构化非凸-非凹极小极大问题的快速外梯度法，《神经信息处理系统进展》，34（2021年），第22588–22600页。[50] D. Leventhal，度量次正则性与近端点方法，《数学分析与应用杂志》，360（2009年），第681–688页。[51] G. Li、T. Liu和T. K. Pong，针对一类非凸优化问题的Peaceman–Rachford分裂，《计算优化与应用》，68（2017年），第407–436页。[52] G. Li和T. K. Pong，Douglas–Rachford分裂法及其在非凸可行性问题中的应用，《数学规划》，159（2016年），第371–401页。[53] P. L. Lions和B. Mercier，两个非线性算子之和的分裂算法，《SIAM数值分析杂志》，16（1979年），第964–979页。[54] J. Luque，非线性近端点算法，载于《第26届IEEE决策与控制会议》，第26卷，1987年，第816–817页。[55] B. Martinet，通过逐次逼近对变分不等式进行正则化，《法国计算机与运筹学评论》。红色系列，4（1970年），第154–158页。[56]C. D. Meyer，《矩阵分析与应用线性代数》，第71卷，工业与应用数学学会，费城，2000年。[57] G. J. Minty，希尔伯特空间中的单调（非线性）算子，《杜克数学杂志》，29（1962年），第341–346页。[58] R. D. C. Monteiro和C.-K. Sim，两个极大强单调算子之和的松弛Peaceman–Rachford分裂法的复杂度，《计算优化与应用》，70（2018年），第763–790页。

[60] Y. Nesterov，《凸优化导论：基础课程》，《应用优化》第87卷，Springer US，波士顿，马萨诸塞州，2004年。[61] J. v. Neumann，一般经济均衡模型，《经济研究评论》，13（1945），第1–9页。[62] J. Nocedal 和 S. Wright，《数值优化》，《运筹学与金融工程丛书》，Springer New York，纽约，纽约州，第2版，2006年。[63] R. G. Otero 和 A. Iusem，半单调算子的正则性结果，《计算与应用数学》，30（2011），第13页。[64] N. Parikh 和 S. Boyd，近端算法，《优化基础与趋势》，1（2014），第127–239页。[65] P. Patrinos、L. Stella 和 A. Bemporad，Douglas–Rachford 分裂：复杂度估计与加速变体，载于《第53届 IEEE 决策与控制会议》，2014年，第4234–4239页。[66] D. W. Peaceman 和 H. H. Rachford, Jr.，抛物型与椭圆型微分方程的数值解法，《工业与应用数学学会期刊》，3（1955），第28–41页。[67] T. Pennanen，无需单调性的近端点算法与乘子方法的局部收敛，《运筹学数学》，27（2002），第170–191页。[68] T. Pethick、O. Fercoq、P. Latafat、P. Patrinos 和 V. Cevher，在不增加批量大小的情况下求解随机弱 Minty 变分不等式，2023年，arXiv 预印本 arXiv:2302.09029。[69] T. Pethick、P. Latafat、P. Patrinos、O. Fercoq 和 V. Cevher，逃离极限环：约束非凸非凹极小极大问题的全局收敛，载于《国际学习表征会议》，2021年。[70] B. Polyak，《优化导论》，Optimization Software，纽约，1987年。[71] G. Pontus 和 W. M. Moursi，关于 Lipschitz 连续算子的特殊情形之复合，《科学与工程中的不动点理论与算法》，2021年（2021）。[72] R. T. Rockafellar，《凸分析》，普林斯顿大学出版社，普林斯顿，1970年。[73] R. T. Rockafellar，增广拉格朗日函数及近端点算法在凸规划中的应用，《运筹学数学》，1（1976），第97–116页。[74] R. T. Rockafellar，单调算子与近端点算法，《SIAM 控制与优化杂志》，14（1976），第877–898页。[75] R. T. Rockafellar，在具有可揭示凸性或单调性的优化与变分不等式中逐步解耦关联关系，《集值与变分分析》，27（2019），第863–893页。[76] R. T. Rockafellar 和 J. Sun，通过渐进对冲求解单调随机变分不等式与互补问题，《数学规划》，174（2019），第453–471页。[77] R. T. Rockafellar 和 R. J.-B. Wets，《变分分析》，Springer 科学与商业媒体，柏林，2009年。[78] E. K. Ryu、R. Hannah 和 W. Yin，缩放相对图：通过二维欧几里得几何研究非扩张算子，《数学规划》（2021），第1–51页。[79] L. Shen 和 S. Pan，广义近端点算法及若干用于复合包含问题的分裂方法的线性收敛，2016年，arXiv 预印本 arXiv:1508.05156。[80] M. V. Solodov 和 B. F. Svaiter，一种混合投影近端点算法，《凸分析杂志》，6（1999），第59–70页。

[81] M. V. Solodov 和 P. Tseng，单调变分不等式的修正投影型方法，《SIAM 控制与优化杂志》，34（1996），第1814–1830页。[82] J. E. Spingarn，次单调映射与近端点算法，《数值泛函分析与优化》，4（1982），第123–150页。[83] B. F. Svaiter，关于 Douglas–Rachford 方法的弱收敛，《SIAM 控制与优化杂志》，49（2011），第280–287页。[84] A. Themelis 和 P. Patrinos，非凸优化的 Douglas-Rachford 分裂与 ADMM：紧确收敛结果，《SIAM优化杂志》，30（2020），第149–181页。[85]Q. Tran-Dinh，具有 $\mathcal{O}(1/k)$ 最后迭代收敛速率的外梯度型方法，用于余假单调包含问题，2023年，arXiv 预印本 arXiv:2302.04099。[86] P. Tseng，极大单调映射的修正前向-后向分裂方法，《SIAM 控制与优化杂志》，38（2000），第431–446页。[87] Z. Zhou、P. Mertikopoulos、N. Bambos、S. Boyd 和 P. W. Glynn，变分相干优化问题中的随机镜像下降，《神经信息处理系统进展》，30（2017）。