



矩阵 0/1 复合优化模型与最优性分析

献给韩继业教授 90 寿辰

张鹏鹤^{1*}, 修乃华², 戚厚铎^{1,3}

1. 香港理工大学应用数学系, 香港 999077;

2. 北京交通大学数学与统计学院, 北京 100044;

3. 香港理工大学数据科学及人工智能学系, 香港 999077

E-mail: penghe.zhang@polyu.edu.hk, nhxiu@bjtu.edu.cn, houduo.qi@polyu.edu.hk

收稿日期: 2024-04-29; 接受日期: 2024-07-26; 网络出版日期: 2024-11-25; * 通信作者

香港研究资助局基金 (批准号: PolyU15309223)、国家重点研发计划 (批准号: 2023YFA1011100) 和国家自然科学基金 (批准号: 12131004) 资助项目

摘要 本文从多分类支持向量机、多标签分类和双边排序等实际应用中抽象凝练出一类矩阵 0/1 复合优化问题, 其目标函数涉及离散的 0/1 损失函数、非光滑的极大值函数以及线性项三重复合结构. 本文给出现有相关工作中 5 种稳定点在这一问题下的具体形式, 分析在特定条件下各稳定点之间以及与该问题解之间的关系, 并进一步建立矩阵 0/1 复合优化问题的二阶必要和充分最优性条件. 尽管矩阵 0/1 复合优化属于非凸非连续问题, 本文的最优性结论均无需约束规范假设.

关键词 矩阵 0/1 复合优化 矩阵 0/1 损失函数 稳定点 二阶最优性条件

MSC (2020) 主题分类 90C46, 90C90

1 引言

本文考虑如下的矩阵 0/1 复合优化问题:

$$\min_{\mathbf{X} \in \mathbb{R}^{n \times s}} h(\mathbf{X}) + \lambda \|\mathbf{A}\mathbf{X} + \mathbf{B}\|_{0/1}^M, \quad (1.1)$$

其中 $h: \mathbb{R}^{n \times s} \rightarrow \mathbb{R}$ 为连续函数, $\lambda > 0$ 为权重参数, $\mathbf{A}: \mathbb{R}^{n \times s} \rightarrow \mathbb{R}^{m \times q}$ 为线性算子, $\mathbf{B} \in \mathbb{R}^{m \times q}$ 为给定矩阵. 对于矩阵 $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_m]^T \in \mathbb{R}^{m \times q}$, 矩阵 0/1 损失函数 $\|\cdot\|_{0/1}^M: \mathbb{R}^{m \times q} \rightarrow \mathbb{R}$ 定义为

$$\|\mathbf{U}\|_{0/1}^M := \sum_{i=1}^m \ell_{0/1}(\max \mathbf{u}_i),$$

英文引用格式: Zhang P H, Xiu N H, Qi H D. Matrixed 0/1 composite optimization model and optimality analysis (in Chinese). Sci Sin Math, 2025, 55: 397–414, doi: 10.1360/SSM-2024-0136

其中 $\max$ 函数取值为向量的最大分量, 而 $\ell_{0/1} : \mathbb{R} \rightarrow \mathbb{R}$ 为经典的一维 0/1 损失函数, 定义为

$$\ell_{0/1}(t) := \begin{cases} 1, & t > 0, \\ 0, & t \leq 0. \end{cases}$$

所以 $\|\mathbf{U}\|_{0/1}^M$ 相当于先对矩阵 $\mathbf{U}$ 的每一行取极大值, 然后将相应的 0/1 损失函数值累加, 这类函数广泛应用于机器学习、统计以及通信等领域, 相应的矩阵 0/1 复合优化 (1.1) 可以涵盖多分类支持向量机 [2,9]、多标签分类 [26,27] 和双边排序 [17] 等实际问题. 然而 (1.1) 目标函数第二部分同时涉及 3 种复合项: 非凸非连续的 $\ell_{0/1}$ 函数、非光滑的 $\max$ 函数以及线性项, 所以矩阵 0/1 复合优化是具有高度复合结构的 NP (non-deterministic polynomial) 困难问题 [7], 与其相关的理论分析工作比较少. 本文的主要工作是直接针对矩阵 0/1 损失函数性质进行研究, 并建立 (1.1) 的一阶和二阶最优性条件.

1.1 文献回顾

下面围绕最优性理论方面回顾与问题 (1.1) 相关的工作, 并阐述本文的研究动机.

一阶最优性理论 当 $s = q = 1$ 时, 矩阵 0/1 损失函数退化为向量 0/1 损失函数 $\|\mathbf{u}\|_{0/1} := \sum_{i=1}^m \ell_{0/1}(u_i)$, 从而 (1.1) 退化为如下的向量 0/1 复合优化问题:

$$\min_{\mathbf{x} \in \mathbb{R}^n} h(\mathbf{x}) + \lambda \|\mathbf{A}\mathbf{x} + \mathbf{b}\|_{0/1}, \quad (1.2)$$

其中矩阵 $\mathbf{A} \in \mathbb{R}^{m \times n}$ 和向量 $\mathbf{b} \in \mathbb{R}^m$ 均为给定数据. Zhou 等 [30] 分别利用 $\|\cdot\|_{0/1}$ 的极限次微分和临近点算子定义了问题 (1.2) 的 KKT (Karush-Kuhn-Tucker) 点和 P 稳定 (proximal stationary) 点, 两种稳定点在适当条件下可以相互转化, 当 h 连续可微时, 局部最优解在矩阵 $\mathbf{A}$ 特定行线性无关条件下可证明是 P 稳定点, 进一步 h 为凸函数时, P 稳定点为局部最优解. Zhang 等 [29] 将 (1.2) 的 KKT 点等价转化为一类带有线性约束的光滑优化问题的 KKT 点, 并进一步与 (1.2) 的局部最优解建立联系, 从而可以在无需线性无关约束规范情形下得到与文献 [30] 相同的一阶最优性结论. 这一结果主要基于 $\ell_{0/1}$ 损失函数的分片线性结构, 而对于一般的下半连续函数和光滑函数构成的非凸复合优化问题, 建立相应的一阶必要最优性条件则需要约束规范 [3,4,23]. 最近, Cui 等 [10] 对一类涉及 $\ell_{0/1}$ 函数与连续函数复合的优化问题, 借助指标集将 $\ell_{0/1}$ 复合项转化为连续函数不等式约束, 进而通过伪 B 稳定点 (pseudo Bouligand stationary point) 建立原问题局部最优解的一阶必要性条件. 对于与文献 [10] 相同的问题, Han 等 [13] 借助其上图化形式提出了一种 E 稳定点 (epi-stationary point), 它要比伪 B 稳定点更强, 且在适当条件下等价于原问题的局部最优解. 由此涉及 $\ell_{0/1}$ 函数的优化问题稳定点种类繁多, 当推广到矩阵 0/1 复合优化问题 (1.1) 时, 这些稳定点的具体形式、相互关系以及适用情形需要进一步研究和讨论.

二阶最优性条件 非凸非连续优化问题的二阶最优性条件是非常具有难度的研究课题, 目前未见问题 (1.1) 的二阶最优性理论工作, 所以本文在此简要回顾一些非凸非连续优化问题的二阶最优性分析相关技术. 对于带有 ℓ_0 稀疏约束, 且其余目标及约束函数均二次连续可微的情形, Pan 等 [22] 利用变分分析方法建立二阶最优性条件, 其中局部最优解的必要条件需要非线性约束集和 ℓ_0 约束集切锥可分离的约束规范. 如果将文献 [22] 中的 ℓ_0 约束替换为 ℓ_0 正则项, 相应的稀疏优化问题可以与一类光滑非线性互补问题稳定点建立联系 [11,15], 由于这类互补问题通常不满足常见的二阶约束规范, 因此要建立其二阶最优性条件, 需要根据问题结构设计新的约束规范. 在向量 0/1 复合优化方面, 文献 [29] 将 (1.2) 与一类带有线性约束的光滑优化问题建立联系, 当 h 二次连续可微时, 无需任何约束规范即

可建立二阶最优性条件. 然而这些技术不能直接应用于 (1.1), 因为矩阵 0/1 损失函数同时包含离散函数 $\ell_{0/1}$ 与非光滑 $\max$ 的复合. 另外, 以上文献中的光滑函数均要求二次可微, 能否推广到更为复杂的情形有待进一步研究.

1.2 方法与主要工作

为了便于后续的理论证明与符号表示, 本文考虑问题 (1.1) 的约束等价形式:

$$\min_{\mathbf{X}, \mathbf{U}} h(\mathbf{X}) + \lambda \|\mathbf{U}\|_{0/1}^M \quad \text{s.t.} \quad \mathbf{A}\mathbf{X} + \mathbf{B} = \mathbf{U}. \quad (1.3)$$

该问题同样称为矩阵 0/1 复合优化.

矩阵 0/1 损失函数的极限次微分显式表达式在研究 (1.3) 各种稳定点关系的过程中发挥着重要作用. 一方面, 它可以在特定条件下转化为临近点算子形式, 从而建立 KKT 点和 P 稳定点的相互转化关系; 另一方面, 它可以保证 KKT 点与相应的乘子满足互补关系, 从而使 (1.3) 的 KKT 点等价转化为一类带有线性约束连续优化问题的 KKT 点. 这类连续优化问题不仅与伪 B 稳定点的定义有着紧密联系, 还可以用于建立二阶最优性条件. 具体来说, 它的目标函数为 h , 约束为带有指标集的线性函数, 一方面可看作问题 (1.3) 在稳定点局部范围内的近似形式; 另一方面它比 (1.3) 的性质更好, 比较直观的一个优点就是仅含线性约束, 可以在无需任何约束规范情形下建立二阶最优性条件. 此外, 可以看到这种技术很好地避免了同时处理函数 $\|\cdot\|_{0/1}^M$ 和 h , 即使问题 (1.3) 中函数 h 不满足二次连续可微性, 也可以建立二阶最优性条件.

本文的主要贡献总结如下:

(i) 建立了多分类支持向量机、多标签分类以及双边分类等实际问题中的矩阵 0/1 复合优化数学模型. 首次计算了矩阵 0/1 损失函数的次微分和临近点算子, 在此基础上定义了问题 (1.3) 的 KKT 点和 P 稳定点, 并证明了它们在临近点算子参数满足特定取值范围条件下可以相互转化.

(ii) 给出了问题 (1.3) 的伪 KKT 点、伪 B 稳定点以及 E 稳定点的具体形式, 并证明了当 h 为光滑或凸函数时, 它们均与 KKT 点等价, 从而特定条件下可以转化为 P 稳定点. 若 h 光滑, 则 P 稳定点为局部最优解的必要条件; 若 h 为凸函数, 则 P 稳定点与局部最优解等价; 若 h 强凸, 则 P 稳定点在临近点参数充分大时可以成为全局最优解.

(iii) 首次建立了 h 为 $C^{1,1}$ 类函数 (光滑且梯度局部 Lipschitz 连续) 情形下矩阵 0/1 复合优化的二阶最优性理论. (1.3) 的局部最优解无需任何约束规范即可满足二阶必要条件; 另外, 若 P 稳定点满足二阶充分条件, 则可以得到 (1.3) 目标函数的二次增长条件.

1.3 论文结构

本文余下内容的安排如下. 第 2 节介绍本文常用符号和基本定义. 第 3 节介绍矩阵 0/1 复合优化的实际应用, 并建立相应的数学模型. 第 4 节给出 $\|\cdot\|_{0/1}^M$ 次微分和临近点算子的显式表达式, 定义问题 (1.3) 的 5 种稳定点, 进而研究它们之间以及与 (1.3) 最优解的关系. 第 5 节建立问题 (1.3) 的二阶必要和充分最优性条件. 第 6 节给出本文研究内容的总结.

2 符号及定义

本文采用大写粗体字母表示矩阵 $\mathbf{U}$, 小写粗体字母表示向量 $\mathbf{u}$, 其中 $\mathbf{u}_i$ 和 $\mathbf{u}^{(k)}$ 分别表示 $\mathbf{U}$ 的第 i 行和第 k 列元素组成的列向量, u_{ik} 表示矩阵 $\mathbf{U}$ 第 i 行第 k 列元素. 此外, 后文还将用 $[u_{ik}]_{m \times q}$ 表示第 i

行第 k 列元素为 u_{ik} 的 $m \times q$ 维矩阵. 记 $[m] := \{1, 2, \dots, m\}$, 给定单元指标集 $\mathcal{S} \subseteq [m]$, $\mathbf{U}_{\mathcal{S}}$ 表示矩阵 $\mathbf{U} \in \mathbb{R}^{m \times q}$ 指标为 $\mathcal{S}$ 的行所组成的子矩阵. 给定二元集 $\mathcal{I} \subseteq [m] \times [q]$, 定义 $\mathbf{U}_{\mathcal{I}} := \{u_{ik} \mid (i, k) \in \mathcal{I}\}$.

定义 $\mathbf{1}$ 和 $\mathbf{0}$ 分别为所有元素为 1 和 0 且具有适当维度的矩阵. $\mathbf{e}_k$ 表示第 k 个元素等于 1、其余元素为 0 的适当维度列向量, 在此基础上 $\bar{\mathbf{e}}_k = \mathbf{1} - \mathbf{e}_k$. 令 $\mathbb{R}_{\leq}^q := \{\mathbf{u} \in \mathbb{R}^q \mid \mathbf{u} \leq \mathbf{0}\}$, 它的补集记作 $\mathbb{R}_{>}^q$. 给定向量 $\mathbf{u} = (u_1, \dots, u_q)^\top$, 则 $\mathbf{u}_+ := (\max(u_1, 0), \dots, \max(u_q, 0))^\top$ 且 $\mathbf{u}_- := -(-\mathbf{u})_+$. 给定矩阵 $\mathbf{U}, \mathbf{Z} \in \mathbb{R}^{m \times q}$, 矩阵内积 $\langle \mathbf{U}, \mathbf{Z} \rangle := \sum_{i=1}^m \sum_{k=1}^q u_{ik} z_{ik}$, 符号 $\mathbf{U} \perp \mathbf{Z}$ 表示两矩阵对应元素至少一个为 0, 记 $N(\mathbf{U}, \delta) := \{\mathbf{Z} \in \mathbb{R}^{m \times q} \mid \|\mathbf{Z} - \mathbf{U}\| < \delta\}$. $\|\cdot\|$ 用于表示向量的二范数或矩阵的 Frobenius 范数. 给定线性算子 $\mathcal{A}$, 其伴随算子记作 $\mathcal{A}^\top$. 给定表达式 π , 符号 $\llbracket \pi \rrbracket = 1$ 表示 π 成立, $\llbracket \pi \rrbracket = 0$ 表示 π 不成立.

下面介绍一些基本数学定义, 可参见文献 [23]. 给定函数 $\varphi: \mathbb{R}^{m \times q} \rightarrow \mathbb{R}$, 该函数在 $\bar{\mathbf{U}}$ 的正则次微分定义为

$$\hat{\partial}\varphi(\bar{\mathbf{U}}) := \left\{ \mathbf{V} \in \mathbb{R}^{m \times q} \mid \liminf_{\substack{\mathbf{U} \rightarrow \bar{\mathbf{U}} \\ \mathbf{U} \neq \bar{\mathbf{U}}}} \frac{\varphi(\mathbf{U}) - \varphi(\bar{\mathbf{U}}) - \langle \mathbf{V}, \mathbf{U} - \bar{\mathbf{U}} \rangle}{\|\mathbf{U} - \bar{\mathbf{U}}\|} \geq 0 \right\},$$

该函数在 $\bar{\mathbf{U}}$ 的极限次微分定义为

$$\partial\varphi(\bar{\mathbf{U}}) := \{\mathbf{V} \in \mathbb{R}^{m \times q} \mid \mathbf{U} \rightarrow \bar{\mathbf{U}}, \varphi(\mathbf{U}) \rightarrow \varphi(\bar{\mathbf{U}}), \mathbf{Z} \in \hat{\partial}\varphi(\mathbf{U}), \mathbf{Z} \rightarrow \mathbf{V}\},$$

函数 φ 在 $\bar{\mathbf{U}}$ 的关于参数 $\alpha > 0$ 的临近点算子定义为

$$\text{Prox}_{\alpha\varphi}(\bar{\mathbf{U}}) := \arg \min_{\mathbf{U} \in \mathbb{R}^{m \times q}} \frac{1}{2\alpha} \|\mathbf{U} - \bar{\mathbf{U}}\|^2 + \varphi(\mathbf{U}).$$

如果给定 $\sigma > 0$, 对于任意 $\mathbf{U}, \mathbf{V} \in \mathbb{R}^{m \times q}$ 以及 $\mu \in [0, 1]$, 都有

$$\varphi(\mu\mathbf{U} + (1 - \mu)\mathbf{V}) \leq \mu\varphi(\mathbf{U}) + (1 - \mu)\varphi(\mathbf{V}) - \frac{\sigma}{2}\mu(1 - \mu)\|\mathbf{U} - \mathbf{V}\|^2.$$

则称 φ 为 σ -强凸函数.

函数 φ 在 $\bar{\mathbf{U}}$ 关于 $\mathbf{D}$ 的方向导数定义为

$$\varphi'(\bar{\mathbf{U}}; \mathbf{D}) := \lim_{t \downarrow 0} \frac{\varphi(\bar{\mathbf{U}} + t\mathbf{D}) - \varphi(\bar{\mathbf{U}})}{t}.$$

如果函数 φ 在所有点处局部 Lipschitz 连续且方向导数存在, 则称其为 B 可微 (Bouligand differentiable) 函数.

集合 $\Omega \subseteq \mathbb{R}^{m \times q}$ 在 $\bar{\mathbf{U}}$ 的切锥定义为

$$\mathcal{T}_{\Omega}(\bar{\mathbf{U}}) := \{\mathbf{V} \in \mathbb{R}^{m \times q} \mid \mathbf{U} \in \Omega, \mathbf{U} \rightarrow \bar{\mathbf{U}}, t \downarrow 0, (\mathbf{U} - \bar{\mathbf{U}})/t \rightarrow \mathbf{V}\}.$$

若矩阵函数 $\Phi: \mathbb{R}^{n \times s} \rightarrow \mathbb{R}$ 在 $\mathbf{X}$ 处可微, 则它的梯度为

$$\nabla\Phi(\mathbf{X}) := \left[\frac{\partial\Phi(\mathbf{X})}{\partial x_{ij}} \right]_{n \times s}.$$

如果 Φ 在 $\mathbf{X}$ 处二次可微, 本文将其 Hessian 矩阵记作如下线性算子 [18]:

$$\nabla^2\Phi(\mathbf{X}) := \{[\langle \nabla\phi_{ij}(\mathbf{X}), \cdot \rangle]_{n \times s} \mid \phi_{ij}(\mathbf{X}) = [\nabla\Phi(\mathbf{X})]_{ij}, i \in [n], j \in [s]\},$$

特别地, $\|\nabla^2\Phi(\mathbf{X})\| := \sqrt{\sum_{i=1}^n \sum_{j=1}^s \|\nabla\phi_{ij}(\mathbf{X})\|^2}$.

如果函数 Φ 连续可微, 且梯度满足局部 Lipschitz 连续, 则记作 $\Phi \in C^{1,1}$. 在这一条件下, 根据 Rademacher 定理, 函数 Φ 在 $\mathbb{R}^{n \times s}$ 上几乎处处二次可微, 它在 $\mathbf{X}$ 处的广义 Hessian 矩阵定义如下:

$$\partial^2 \Phi(\mathbf{X}) := \text{co}\{\mathcal{M} \mid \exists \mathbf{X}^\nu \rightarrow \mathbf{X}, \Phi \text{ 在 } \mathbf{X}^\nu \text{ 处二次可微}, \nabla^2 \Phi(\mathbf{X}^\nu) \rightarrow \mathcal{M}\},$$

其中 co 表示集合的凸包.

3 应用实例

本节展示矩阵 0/1 损失函数在实际领域的应用, 它可以准确表示相关问题的评价指标, 刻画问题的本质, 进而建立矩阵 0/1 复合优化模型 (1.3).

(i) 多分类支持向量机. 考虑含有 q 个类别的多分类问题, 令 $\{(\mathbf{c}_i, v_i)\}_{i=1}^m$ 为训练样本, 其中 $\mathbf{c}_i \in \mathbb{R}^n$ 为样本特征, $v_i \in [q]$ 为类别标签. 给定待预测样本特征 $\mathbf{c}$, 可定义 $\langle \mathbf{c}, \mathbf{x}^{(k)} \rangle$ 为 $\mathbf{c}$ 关于第 k 个类别的评分, 多分类器 $f_{\mathbf{X}}: \mathbb{R}^n \rightarrow [q]$ 通常将 $\mathbf{c}$ 归为最大评分所在的类别 [9]:

$$f_{\mathbf{X}}(\mathbf{c}) := \arg \max_{k \in [q]} \langle \mathbf{x}^{(k)}, \mathbf{c} \rangle.$$

多分类学习的主要目标就是估计以上的参数矩阵 $\mathbf{X} = [\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(q)}] \in \mathbb{R}^{n \times q}$, 从而使如下分类准确率 (accuracy) 尽可能高:

$$\frac{1}{m} \sum_{i=1}^m \mathbb{I}[f_{\mathbf{X}}(\mathbf{c}_i) = v_i].$$

为达成这一目的, 可令如下不等式成立数量尽可能多:

$$\max(\mathbf{X}^\top \mathbf{c}_i - \mathbf{1} \langle \mathbf{x}^{(v_i)}, \mathbf{c}_i \rangle + \bar{\mathbf{e}}_{v_i}) \leq 0, \quad i \in [m]. \quad (3.1)$$

这里的 $\bar{\mathbf{e}}_{v_i}$ 可以保证样本实际所属类别的评分比其他类别至少高一分, 从而有利于提升模型的容错率. 借助 $\ell_{0/1}$ 损失函数, 极大化 (3.1) 中不等式成立数量可转化为如下的优化问题:

$$\min_{\mathbf{X} \in \mathbb{R}^{m \times q}} h(\mathbf{X}) + \lambda \sum_{i=1}^m \ell_{0/1}(\max(\mathbf{X}^\top \mathbf{c}_i - \mathbf{1} \langle \mathbf{x}^{(v_i)}, \mathbf{c}_i \rangle + \bar{\mathbf{e}}_{v_i})),$$

以上问题可看作特定参数设置下的矩阵 0/1 复合优化. 这里函数 h 为正则项, 用于降低过拟合风险或特征选择, 可取作 ℓ_2 范数平方 [9] $\|\mathbf{X}\|^2$ 、 ℓ_1 范数 [25] $\|\mathbf{X}\|_1 := \sum_{k=1}^q \|\mathbf{x}^{(k)}\|_1$ 和 $\ell_{2,1}$ 范数 [2] $\|\mathbf{X}\|_{2,1} := \sum_{j=1}^n \|\mathbf{x}_j\|$. 此外, 其他参数可采用如下设置:

$$s = q, \quad \mathcal{A}(\cdot) = [\langle \mathbf{A}^{(i,k)}, \cdot \rangle]_{m \times q}, \quad \mathbf{A}^{(i,k)} = \mathbf{c}_i(\mathbf{e}_k - \mathbf{e}_{v_i})^\top, \quad \mathbf{B} = [\bar{\mathbf{e}}_{v_1}, \dots, \bar{\mathbf{e}}_{v_m}]^\top.$$

(ii) 多标签分类. 对于含有 q 个类别的多标签分类问题, 训练样本为 $\{(\mathbf{c}_i, \mathbf{v}_i)\}_{i=1}^m$, 其中 $\mathbf{c}_i \in \mathbb{R}^n$ 为样本特征, $\mathbf{v}_i \in \{-1, 1\}^q$ 为类别标签, $v_{ik} = 1$ 表示样本属于第 k 类, 否则不属于, 所以这类问题允许样本同时属于多个类别. 基于二元相关性 [26] (binary relevance) 的多标签分类器 $f_{\mathbf{X}}: \mathbb{R}^n \rightarrow \{-1, 1\}^q$ 通常采用如下方式预测样本 $\mathbf{c}$ 的类别标签:

$$f_{\mathbf{X}}(\mathbf{c}) := [\text{sgn}(\langle \mathbf{x}^{(1)}, \mathbf{c} \rangle), \dots, \text{sgn}(\langle \mathbf{x}^{(q)}, \mathbf{c} \rangle)]^\top,$$

其中 sgn 为符号函数, 变量为正数, 相应取值为 1, 否则为 -1 . 多标签分类器的一个重要评价指标为子集准确率^[27] (subset accuracy), 定义如下:

$$\frac{1}{m} \sum_{i=1}^m \mathbb{I}[f_{\mathbf{X}}(\mathbf{c}_i) = \mathbf{v}_i].$$

为了极大化子集准确率, 实际当中可考虑求解 $\mathbf{X}$ 使得如下不等式成立数量尽可能多:

$$\max(\mathbf{1} - \mathbf{v}_i \odot (\mathbf{X}^\top \mathbf{c}_i)) \leq 0, \quad i \in [m],$$

其中 $\odot$ 为 Hadamard 乘积. 可转化为求解如下优化问题:

$$\min_{\mathbf{X} \in \mathbb{R}^{m \times q}} h(\mathbf{X}) + \lambda \sum_{i=1}^m \ell_{0/1}(\max(\mathbf{1} - \mathbf{v}_i \odot (\mathbf{X}^\top \mathbf{c}_i))).$$

以上问题可看作模型 (1.1) 的一个特例, 其中函数 h 为正则项, 在多标签分类领域相关工作中可取作 ℓ_2 范数平方^[26] $\|\mathbf{X}\|^2$ 、 $\ell_{2,1}$ 范数^[20] $\|\mathbf{X}\|_{2,1} := \sum_{i=1}^m \|\mathbf{x}_i\|$ 和特征图函数^[6] $\langle \mathbf{X}, \mathbf{LX} \rangle$ ($\mathbf{L} \in \mathbb{R}^{n \times n}$ 为给定 Laplace 矩阵) 等. 此外, 其他数据取值如下:

$$s = q, \quad \mathcal{A}(\cdot) = [\langle \mathbf{A}^{(i,k)}, \cdot \rangle]_{m \times q}, \quad \mathbf{A}^{(i,k)} = -v_{ik} \mathbf{c}_i \mathbf{e}_k^\top, \quad \mathbf{B} = \mathbf{1}.$$

(iii) 双边排序. 令 $\mathbf{C}^+ = [\mathbf{c}_1^+, \dots, \mathbf{c}_m^+]^\top \in \mathbb{R}^{m \times n}$ 和 $\mathbf{C}^- := [\mathbf{c}_1^-, \dots, \mathbf{c}_q^-] \in \mathbb{R}^{q \times n}$ 分别为正类和负类样本矩阵. 给定样本 $\mathbf{c} \in \mathbb{R}^n$, 排序问题通常将 $f_{\mathbf{x}}(\mathbf{c}) = \langle \mathbf{x}, \mathbf{c} \rangle$ 作为样本的评分, 同时要求正类样本评分要高于负类样本, 其中顶端准确率 (accuracy at the top) 考虑了每个正类样本评分与最大负类样本评分的大小关系, 定义为

$$\frac{1}{m} \sum_{i=1}^m \mathbb{I}[f_{\mathbf{x}}(\mathbf{c}_i^+) > \max_{k \in [q]} f_{\mathbf{x}}(\mathbf{c}_k^-)].$$

为了让顶端准确率尽可能大, 可考虑求解 $\mathbf{x}$ 使得如下不等式成立数量尽可能多:

$$\max(\mathbf{C}^- \mathbf{x} - \mathbf{1} \langle \mathbf{c}_i^+, \mathbf{x} \rangle + \mathbf{1}) \leq 0, \quad i \in [m].$$

进一步转化为如下优化问题:

$$\min_{\mathbf{x} \in \mathbb{R}^n} h(\mathbf{x}) + \lambda \sum_{i=1}^m \ell_{0/1}(\max(\mathbf{C}^- \mathbf{x} - \mathbf{1} \langle \mathbf{c}_i^+, \mathbf{x} \rangle + \mathbf{1})),$$

以上模型可看作问题 (1.1) 的特例, 其中可取函数 $h(\mathbf{x}) = \|\mathbf{x}\|^2/2$ 来增强模型的泛化性, 其他参数设置如下:

$$s = 1, \quad \mathcal{A}(\cdot) = [\langle \mathbf{a}^{(i,k)}, \cdot \rangle]_{m \times q}, \quad \mathbf{a}^{(i,k)} = \mathbf{c}_k^- - \mathbf{c}_i^+, \quad \mathbf{B} = \mathbf{1}.$$

(iv) 向量 0/1 复合优化. 当 $s = q = 1$ 时, 矩阵 0/1 复合优化退化为向量 0/1 复合优化 (1.2). 文献 [30] 对这类优化模型的应用进行了详细的介绍, 可以涵盖支持向量机^[8]、1-bit 压缩感知^[5] 以及最大序相关^[12] 等问题.

4 一阶最优性理论

本节给出非凸复合最优性分析中的几类稳定点来刻画问题 (1.1) 的局部最优解. 这些稳定点适用于不同的情形, 它们之间的关系将会被重点研究. 在定义这些稳定点之前, 需要给出矩阵 0/1 损失函数的一些基本性质.

4.1 次微分与临近点算子

矩阵 0/1 损失函数是下半连续且有下界的, 它的次微分和临近点算子具有显示表达式.

定理 4.1 (矩阵 0/1 损失次微分) 给定 $\mathbf{U} \in \mathbb{R}^{m \times q}$, 矩阵 0/1 损失函数在 $\mathbf{U}$ 处的正则和极限次微分相等, 且可以表示为

$$\hat{\partial}\|\mathbf{U}\|_{0/1}^M = \partial\|\mathbf{U}\|_{0/1}^M = \Theta := \left\{ \mathbf{Z} \in \mathbb{R}^{m \times q} \mid \begin{cases} \mathbf{z}_i = \mathbf{0}, & \text{若 } \mathbf{u}_i \in \overline{\mathbb{R}}_{\leq}^q, \\ \mathbf{z}_i \geq \mathbf{0}, \quad \mathbf{z}_i \perp \mathbf{u}_i, & \text{若 } \mathbf{u}_i \in \mathbb{R}_{<}^q, \end{cases} \quad \forall i \in [m] \right\}. \quad (4.1)$$

证明 首先来证明 $\hat{\partial}\|\mathbf{U}\|_{0/1}^M \supseteq \Theta$. 任取 $\mathbf{Z} \in \Theta$, 当序列 $\mathbf{U}^\nu \rightarrow \mathbf{U}$ 时, 对于任意 $i \in [m]$, 考虑如下表达式的取值情形:

$$\Delta_i^\nu := \ell_{0/1}(\max \mathbf{u}_i^\nu) - \ell_{0/1}(\max \mathbf{u}_i) - \langle \mathbf{z}_i, \mathbf{u}_i^\nu - \mathbf{u}_i \rangle.$$

若 $\mathbf{u}_i \in \overline{\mathbb{R}}_{\leq}^q$, 则 $\mathbf{z}_i = \mathbf{0}$ 且 $\max \mathbf{u}_i > 0$. 由于 $\max$ 函数连续, 当 $\mathbf{U}^\nu$ 充分靠近 $\mathbf{U}$ 时, $\max \mathbf{u}_i^\nu > 0$, 从而 $\Delta_i^\nu = 1 - 1 - 0 = 0$.

若 $\mathbf{u}_i \in \mathbb{R}_{<}^q$ 且 $\max \mathbf{u}_i < 0$, 则 $\mathbf{z}_i = \mathbf{0}$. 同上可根据 $\max$ 函数连续性推导出 $\Delta_i^\nu = 0$.

当 $\mathbf{u}_i \in \overline{\mathbb{R}}_{\leq}^q$ 且 $\max \mathbf{u}_i = 0$ 时, 不妨假定 $\mathbf{U}^\nu$ 充分靠近 $\mathbf{U}$, 从而使得 $1 - \langle \mathbf{z}_i, \mathbf{u}_i^\nu - \mathbf{u}_i \rangle \geq 0$. 如果 $\max \mathbf{u}_i^\nu > 0$, 则 $\Delta_i^\nu = 1 - 0 - \langle \mathbf{z}_i, \mathbf{u}_i^\nu - \mathbf{u}_i \rangle \geq 0$. 如果 $\max \mathbf{u}_i^\nu \leq 0$, 则 $\mathbf{u}_i^\nu \leq \mathbf{0}$, 再根据 $\mathbf{z}_i \perp \mathbf{u}_i$ 可得 $\Delta_i^\nu = 0 - 0 - \langle \mathbf{z}_i, \mathbf{u}_i^\nu \rangle \geq 0$.

综合以上 3 种情形可知 $\mathbf{U}^\nu$ 充分靠近 $\mathbf{U}$ 时, 对于任意 $i \in [m]$, 都有 $\Delta_i^\nu \geq 0$, 从而

$$\liminf_{\mathbf{U}^\nu \rightarrow \mathbf{U}, \mathbf{U}^\nu \neq \mathbf{U}} \frac{\sum_{i=1}^m \Delta_i^\nu}{\|\mathbf{U}^\nu - \mathbf{U}\|} \geq 0.$$

这意味着 $\mathbf{Z} \in \hat{\partial}\|\mathbf{U}\|_{0/1}^M, \hat{\partial}\|\mathbf{U}\|_{0/1}^M \supseteq \Theta$.

接下来证明 $\hat{\partial}\|\mathbf{U}\|_{0/1}^M \subseteq \Theta$. 对于任意 $\mathbf{Z} \in \hat{\partial}\|\mathbf{U}\|_{0/1}^M$, 都有

$$\eta := \liminf_{\mathbf{U}^\nu \rightarrow \mathbf{U}, \mathbf{U}^\nu \neq \mathbf{U}} \frac{\|\mathbf{U}^\nu\|_{0/1}^M - \|\mathbf{U}\|_{0/1}^M - \langle \mathbf{Z}, \mathbf{U}^\nu - \mathbf{U} \rangle}{\|\mathbf{U}^\nu - \mathbf{U}\|} \geq 0.$$

取定序号 $(i_0, k_0) \in [m] \times [q]$, 下面选取不同序列 $\{\mathbf{U}^\nu\}$ 来逼近 $\mathbf{U}$.

当 $\mathbf{u}_{i_0} \in \overline{\mathbb{R}}_{\leq}^q$ 时, 任取 $k_0 \in [m]$, $z_{i_0 k_0} = 0$ 一定成立, 否则令

$$u_{ik}^\nu = \begin{cases} u_{ik}, & \text{若 } i \neq i_0 \text{ 或 } k \neq k_0, \\ u_{ik} + t z_{ik}, & \text{若 } i = i_0 \text{ 且 } k = k_0. \end{cases} \quad (4.2)$$

显然 $t \downarrow 0$ 时, $\mathbf{U} \neq \mathbf{U}^\nu \rightarrow \mathbf{U}$, 并且 $\mathbf{U}^\nu$ 充分靠近 $\mathbf{U}$ 时, $\max \mathbf{u}_{i_0}^\nu > 0$ 且 $\max \mathbf{u}_{i_0} > 0$, 进一步得到

$$0 \leq \eta = \liminf_{t \downarrow 0} -\frac{t z_{i_0 k_0}^2}{t |z_{i_0 k_0}|} = -|z_{i_0 k_0}|. \quad (4.3)$$

这意味着 $z_{i_0 k_0} = 0$, 从而推导出矛盾.

当 $\mathbf{u}_{i_0} \in \mathbb{R}_{\leq}^q$ 且 $u_{i_0 k_0} < 0$ 时, $z_{i_0 k_0} = 0$ 一定成立. 否则按照 (4.2) 取序列 $\mathbf{U}^\nu$ 充分靠近 $\mathbf{U}$ 时, $u_{i_0 k_0}^\nu < 0$, 从而 $\|\mathbf{U}^\nu\|_{0/1}^M = \|\mathbf{U}\|_{0/1}^M$, 仍然可以推导出矛盾 (4.3).

当 $\mathbf{u}_{i_0} \in \mathbb{R}_{\leq}^q$ 且 $u_{i_0 k_0} = 0$ 时, 令

$$u_{ik}^\nu = \begin{cases} u_{ik}, & \text{若 } i \neq i_0 \text{ 或 } k \neq k_0, \\ u_{ik} - t, & \text{若 } i = i_0 \text{ 且 } k = k_0. \end{cases}$$

显然 $t \downarrow 0$ 时, $\mathbf{U} \neq \mathbf{U}^\nu \rightarrow \mathbf{U}$, 并且 $u_{ik}^\nu < 0$, 从而 $\|\mathbf{U}^\nu\|_{0/1}^M = \|\mathbf{U}\|_{0/1}^M$. 进一步可推导出

$$0 \leq \eta = \liminf_{t \downarrow 0} \frac{t z_{i_0 k_0}}{t},$$

从而 $z_{i_0 k_0} \geq 0$.

综合以上 3 种情形可以推导出 $\mathbf{Z} \in \Theta$. 至此已证得 $\widehat{\partial}\|\mathbf{U}\|_{0/1}^M = \Theta$.

下面证明 $\partial\|\mathbf{U}\|_{0/1}^M = \Theta$. 由文献 [23, 定理 8.6] 可以推导出 $\Theta = \widehat{\partial}\|\mathbf{U}\|_{0/1}^M \subseteq \partial\|\mathbf{U}\|_{0/1}^M$, 所以接下来只需要证明 $\Theta \supseteq \partial\|\mathbf{U}\|_{0/1}^M$. 任取 $\mathbf{Z} \in \partial\|\mathbf{U}\|_{0/1}^M$, 根据极限次微分定义, 存在序列 $\mathbf{U}^\nu \rightarrow \mathbf{U}$ 且 $\|\mathbf{U}^\nu\|_{0/1}^M \rightarrow \|\mathbf{U}\|_{0/1}^M$, 使得存在 $\mathbf{Z}^\nu \in \widehat{\partial}\|\mathbf{U}^\nu\|_{0/1}^M$, 满足 $\mathbf{Z}^\nu \rightarrow \mathbf{Z}$. 当 $\mathbf{U}^\nu$ 充分靠近 $\mathbf{U}$ 时, 由于 $\|\cdot\|_{0/1}^M$ 仅在有限集 $\{0, 1, \dots, m\}$ 中取值, 所以有

$$\|\mathbf{U}^\nu\|_{0/1}^M = \|\mathbf{U}\|_{0/1}^M, \quad (4.4)$$

除此之外以下集合的包含关系成立:

$$\{i \in [m] \mid \max \mathbf{u}_i^\nu > 0\} \supseteq \{i \in [m] \mid \max \mathbf{u}_i > 0\}, \quad (4.5a)$$

$$\{(i, k) \in [m] \times [q] \mid u_{ik}^\nu = 0\} \subseteq \{(i, k) \in [m] \times [q] \mid u_{ik} = 0\}, \quad (4.5b)$$

$$\{(i, k) \in [m] \times [q] \mid u_{ik}^\nu < 0\} \supseteq \{(i, k) \in [m] \times [q] \mid u_{ik} < 0\}. \quad (4.5c)$$

表达式 (4.4) 和 (4.5a) 实际上意味着 (4.5a) 中的两个集合相等, 再结合 (4.5b) 和 (4.5c), 可以推导出 $\mathbf{Z}^\nu \in \widehat{\partial}\|\mathbf{U}^\nu\|_{0/1}^M \subseteq \Theta$. 由于 Θ 为闭集, 所以 $\mathbf{Z}^\nu \rightarrow \mathbf{Z} \in \Theta$. 至此完成了 (4.1) 的证明. $\square$

从表达式 (4.1) 中可以看出, 如果 $\mathbf{Z} \in \partial\|\mathbf{U}\|_{0/1}^M$, 则一定有互补性质 $\mathbf{U} \perp \mathbf{Z} \geq \mathbf{0}$ 成立. 同时对于任意 $\lambda > 0$ 都有 $\partial\lambda\|\mathbf{U}\|_{0/1}^M = \lambda\partial\|\mathbf{U}\|_{0/1}^M = \partial\|\mathbf{U}\|_{0/1}^M$.

定理 4.2 (矩阵 0/1 损失临近点算子) 给定 $\mathbf{Y} \in \mathbb{R}^{m \times q}$, $\lambda > 0$ 和 $\alpha > 0$, 以下结论成立:

(i) 在点 $\mathbf{Y}$ 处, 函数 $\lambda\|\cdot\|_{0/1}^M$ 关于参数 $\alpha > 0$ 的临近点算子可表示为

$$(\text{Prox}_{\alpha\lambda\|\cdot\|_{0/1}^M}(\mathbf{Y}))_i = \begin{cases} \mathbf{y}_i, & \text{若 } \|(\mathbf{y}_i)_+\| \in (\sqrt{2\alpha\lambda}, +\infty), \\ \{\mathbf{y}_i, (\mathbf{y}_i)_-\}, & \text{若 } \|(\mathbf{y}_i)_+\| \in \{0, \sqrt{2\alpha\lambda}\}, \\ (\mathbf{y}_i)_-, & \text{若 } \|(\mathbf{y}_i)_+\| \in (0, \sqrt{2\alpha\lambda}). \end{cases} \quad (4.6)$$

(ii) 若矩阵 $\mathbf{U}, \mathbf{Z} \in \mathbb{R}^{m \times q}$ 满足 $\mathbf{U} \in \text{Prox}_{\alpha\lambda\|\cdot\|_{0/1}^M}(\mathbf{U} + \alpha\mathbf{Z})$, 则对于任意 $i \in [m]$, 向量 $\mathbf{u}_i$ 和 $\mathbf{z}_i$ 的取值范围满足如下两种情形之一:

$$\begin{aligned} \|(\mathbf{u}_i)_+\| &\in [\sqrt{2\alpha\lambda}, +\infty), \quad \mathbf{z}_i = \mathbf{0}, \\ \mathbf{0} &\geq \mathbf{u}_i \perp \mathbf{z}_i \geq \mathbf{0}, \quad \|\mathbf{z}_i\| \in [0, \sqrt{2\lambda/\alpha}]. \end{aligned} \quad (4.7)$$

证明 (i) 根据 $\|\cdot\|^2$ 和 $\|\cdot\|_{0/1}^M$ 的可分离性以及临近点算子定义, 对于任意 $i \in [m]$, 都有

$$(\text{Prox}_{\alpha\lambda\|\cdot\|_{0/1}^M}(\mathbf{Y}))_i = \arg \min_{\mathbf{u} \in \mathbb{R}^q} \frac{1}{2\alpha} \|\mathbf{u} - \mathbf{y}_i\|^2 + \lambda \ell_{0/1}(\max \mathbf{u}). \quad (4.8)$$

将以上问题的目标函数定义为 φ , 取值如下:

$$\varphi(\mathbf{u}) = \begin{cases} \frac{1}{2\alpha} \|\mathbf{u} - \mathbf{y}_i\|^2, & \text{若 } \mathbf{u} \in \mathbb{R}_{\leq}^q, \\ \frac{1}{2\alpha} \|\mathbf{u} - \mathbf{y}_i\|^2 + \lambda, & \text{若 } \mathbf{u} \in \overline{\mathbb{R}}_{\leq}^q. \end{cases}$$

将函数 φ 在区域 $\mathbb{R}_{\leq}^q$ 和 $\overline{\mathbb{R}}_{\leq}^q$ 上的函数最小取值分别记作 val_1 和 val_2 , 相应的点记作 sol_1 和 sol_2 . 下面根据 $\mathbf{y}_i$ 取值考虑两种情形:

若 $\mathbf{y}_i \in \mathbb{R}_{\leq}^q$, 即 $\|(\mathbf{y}_i)_+\| = 0$, 则显然 $0 = \text{val}_1 < \text{val}_2$. 从而问题 (4.8) 的解为 $\text{sol}_1 = \mathbf{y}_i = (\mathbf{y}_i)_-$.

若 $\mathbf{y}_i \in \overline{\mathbb{R}}_{\leq}^q$, 即 $\|(\mathbf{y}_i)_+\| > 0$, 则 $\text{sol}_1 = (\mathbf{y}_i)_-$, $\text{val}_1 = \|\mathbf{y}_i - (\mathbf{y}_i)_-\|^2/(2\alpha) = \|(\mathbf{y}_i)_+\|^2/(2\alpha)$. 同时, $\text{sol}_2 = \mathbf{y}_i$ 且 $\text{val}_2 = \lambda$. 接下来只需要比较 val_1 和 val_2 的大小来决定 (4.8) 的形式. 当 $\text{val}_1 > \text{val}_2$ 时, $\|(\mathbf{y}_i)_+\| > \sqrt{2\alpha\lambda}$, (4.8) 的解为 $\text{sol}_2 = \mathbf{y}_i$. 当 $\text{val}_1 = \text{val}_2$ 时, $\|(\mathbf{y}_i)_+\| = \sqrt{2\alpha\lambda}$, (4.8) 的解为 $\{\text{sol}_1, \text{sol}_2\} = \{(\mathbf{y}_i)_-, \mathbf{y}_i\}$. 当 $\text{val}_1 < \text{val}_2$ 时, $\|(\mathbf{y}_i)_+\| < \sqrt{2\alpha\lambda}$, (4.8) 的解为 $\text{sol}_1 = (\mathbf{y}_i)_-$.

(ii) 如果 $\mathbf{U} \in \text{Prox}_{\alpha\lambda\|\cdot\|_{0/1}^M}(\mathbf{U} + \alpha\mathbf{Z})$ 成立, 则按照如下几种情形考虑对于任意 $i \in [m]$, $\mathbf{u}_i$ 和 $\mathbf{z}_i$ 的取值范围:

若 $\|(\mathbf{u}_i + \alpha\mathbf{z}_i)_+\| \in (\sqrt{2\alpha\lambda}, +\infty)$, 则 $\mathbf{u}_i = \mathbf{u}_i + \alpha\mathbf{z}_i$, 从而 $\mathbf{z}_i = \mathbf{0}$ 且 $\|(\mathbf{u}_i)_+\| \in (\sqrt{2\alpha\lambda}, +\infty)$.

若 $\|(\mathbf{u}_i + \alpha\mathbf{z}_i)_+\| = 0$, 则 $\mathbf{u}_i + \alpha\mathbf{z}_i \leq \mathbf{0}$ 且 $\mathbf{u}_i = \mathbf{u}_i + \alpha\mathbf{z}_i$, 从而 $\mathbf{z}_i = \mathbf{0}$ 且 $\mathbf{u}_i \leq \mathbf{0}$.

若 $\|(\mathbf{u}_i + \alpha\mathbf{z}_i)_+\| = \sqrt{2\alpha\lambda}$, 则 $\mathbf{u}_i \in \{\mathbf{u}_i + \alpha\mathbf{z}_i, (\mathbf{u}_i + \alpha\mathbf{z}_i)_-\}$. 如果 $\mathbf{u}_i = \mathbf{u}_i + \alpha\mathbf{z}_i$, 则 $\mathbf{z}_i = \mathbf{0}$ 且 $\|(\mathbf{u}_i)_+\| = \sqrt{2\alpha\lambda}$; 如果 $\mathbf{u}_i = (\mathbf{u}_i + \alpha\mathbf{z}_i)_-$, 则 $\mathbf{0} \geq \mathbf{u}_i \perp \mathbf{z}_i \geq \mathbf{0}$ 且 $\|\mathbf{z}_i\| = \|(\mathbf{z}_i)_+\| = \sqrt{2\lambda/\alpha}$.

若 $\|(\mathbf{u}_i + \alpha\mathbf{z}_i)_+\| \in (0, \sqrt{2\alpha\lambda})$, 则 $\mathbf{u}_i = (\mathbf{u}_i + \alpha\mathbf{z}_i)_-$, 从而 $\mathbf{0} \geq \mathbf{u}_i \perp \mathbf{z}_i \geq \mathbf{0}$ 且 $\|\mathbf{z}_i\| = \|(\mathbf{z}_i)_+\| \in (0, \sqrt{2\lambda/\alpha})$.

综合以上讨论, $\mathbf{u}_i$ 和 $\mathbf{z}_i$ 满足取值范围 (4.7).

下面证明相反的结论. 如果 $\mathbf{u}_i$ 和 $\mathbf{z}_i$ 满足取值范围 (4.7), 考虑如下情形: 若 $\|(\mathbf{u}_i)_+\| \in [\sqrt{2\alpha\lambda}, +\infty)$ 且 $\mathbf{z}_i = \mathbf{0}$, 则 $\|(\mathbf{u}_i + \alpha\mathbf{z}_i)_+\| \in [\sqrt{2\alpha\lambda}, +\infty)$, 从而 $\mathbf{u}_i = \mathbf{u}_i + \alpha\mathbf{z}_i \in (\text{Prox}_{\alpha\lambda\|\cdot\|_{0/1}^M}(\mathbf{U} + \alpha\mathbf{Z}))_i$.

若 $\mathbf{0} \geq \mathbf{u}_i \perp \mathbf{z}_i \geq \mathbf{0}$ 且 $\|\mathbf{z}_i\| \in [0, \sqrt{2\lambda/\alpha}]$, 则 $\|(\mathbf{u}_i + \alpha\mathbf{z}_i)_+\| \in [0, \sqrt{2\alpha\lambda}]$, 从而 $\mathbf{u}_i = (\mathbf{u}_i + \alpha\mathbf{z}_i)_- \in (\text{Prox}_{\alpha\lambda\|\cdot\|_{0/1}^M}(\mathbf{U} + \alpha\mathbf{Z}))_i$. 综合以上两种情形可得 $\mathbf{U} \in \text{Prox}_{\alpha\lambda\|\cdot\|_{0/1}^M}(\mathbf{U} + \alpha\mathbf{Z})$. 至此完成了结论 (ii) 的证明. $\square$

关于定理 4.1 和 4.2, 有如下讨论.

注 4.1 (i) 首先给出一个计算矩阵 0/1 损失函数次微分和临近点算子的实例. 考虑

$$\mathbf{U}^* = \begin{bmatrix} 3, & 0 \\ 2, & -1 \\ 0, & -1 \end{bmatrix}, \quad \lambda = 1, \quad \alpha = 2.$$

显然 $\mathbf{u}_1^*, \mathbf{u}_2^* \notin \mathbb{R}_{\leq}^2$, 而 $\mathbf{u}_3^* \in \mathbb{R}_{\leq}^2$, 从而

$$\partial\|\mathbf{U}^*\|_{0/1}^M = \left\{ \begin{bmatrix} 0, & 0 \\ 0, & 0 \\ \beta, & 0 \end{bmatrix} \mid \beta \geq 0 \right\}.$$

另外, 可以计算 $\sqrt{2\lambda\alpha} = 2$, $\|(\mathbf{u}_1^*)_+\| = 3$, $\|(\mathbf{u}_2^*)_+\| = 2$, $\|(\mathbf{u}_3^*)_+\| = 0$, 从而

$$\text{Prox}_{\alpha\lambda\|\cdot\|_{0/1}^M}(\mathbf{U}^*) = \left\{ \begin{bmatrix} 3, & 0 \\ 2, & -1 \\ 0, & -1 \end{bmatrix}, \begin{bmatrix} 3, & 0 \\ 0, & -1 \\ 0, & -1 \end{bmatrix} \right\}.$$

(ii) 通过比较表达式 (4.1) 和 (4.7), 如果 $\mathbf{U}, \mathbf{Z} \in \mathbb{R}^{m \times q}$ 满足 $\mathbf{U} \in \text{Prox}_{\alpha\lambda\|\cdot\|_{0/1}^M}(\mathbf{U} + \alpha\mathbf{Z})$, 则 $\mathbf{Z} \in \partial\|\mathbf{U}\|_{0/1}^M$ 一定成立, 这种情形下不仅互补性质 $\mathbf{U} \perp \mathbf{Z} \geq \mathbf{0}$ 成立, 而且 $\mathbf{U}$ 和 $\mathbf{Z}$ 相应分量的取值范围还会被 λ 和 α 限制.

(iii) 当 $s = q = 1$ 时, $\|\cdot\|_{0/1}^M$ 退化为向量 0/1 损失函数 $\|\cdot\|_{0/1}$, 相应次微分和临近点算子表达式为

$$\begin{aligned} \partial\|\mathbf{u}\|_{0/1} &= \left\{ \mathbf{z} \in \mathbb{R}^m \mid z_i \begin{cases} = 0, & \text{若 } u_i \neq 0, \\ \geq 0 & \text{若 } u_i = 0 \end{cases} \right\}, \\ [\text{Prox}_{\alpha\lambda\|\cdot\|_{0/1}}(\mathbf{y})]_i &= \begin{cases} y_i, & \text{若 } y_i \in (-\infty, 0) \cup (\sqrt{2\alpha\lambda}, +\infty), \\ \{y_i, 0\}, & \text{若 } y_i \in \{0, \sqrt{2\alpha\lambda}\}, \\ 0, & \text{若 } y_i \in (0, \sqrt{2\alpha\lambda}), \end{cases} \quad i \in [m]. \end{aligned}$$

从而得到文献 [30] 的相关结果.

4.2 稳定点关系

矩阵 0/1 复合优化 (1.3) 可行集定义为 $\mathcal{F} := \{(\mathbf{X}, \mathbf{U}) \mid \mathcal{A}\mathbf{X} + \mathbf{B} = \mathbf{U}\}$, 它的增广 Lagrange 函数可写作

$$\mathcal{L}(\mathbf{W}, \mathbf{Z}) := h(\mathbf{X}) + \lambda\|\mathbf{U}\|_{0/1}^M + \langle \mathbf{Z}, \mathcal{A}\mathbf{X} + \mathbf{B} - \mathbf{U} \rangle,$$

其中 $\mathbf{W} := (\mathbf{X}, \mathbf{U})$ 为原始变量, 后续将沿用该简记形式, $\mathbf{Z}$ 为乘子. 基于 Lagrange 函数的临界点, 可以定义问题 (1.1) 的 KKT 点. 这类稳定点在文献 [3] 中用于研究非凸复合优化问题的一阶必要性条件, 而文献 [30] 利用这种稳定点建立了向量 0/1 复合优化的一阶必要以及充分最优性条件.

定义 4.1 (KKT 点) 若点 $\mathbf{W}^* := (\mathbf{X}^*, \mathbf{U}^*) \in \mathcal{F}$ 和 $\mathbf{Z}^*$ 满足

$$\mathbf{0} \in \partial_{\mathbf{W}} \mathcal{L}(\mathbf{W}^*, \mathbf{Z}^*), \quad (4.9)$$

则称它们分别为问题 (1.3) 的 KKT 点和 Lagrange 乘子, $(\mathbf{W}^*, \mathbf{Z}^*)$ 称作该问题的 KKT 对.

表达式 (4.9) 实际上等价于如下的 KKT 系统:

$$\begin{cases} \partial h(\mathbf{X}^*) + \mathcal{A}^\top \mathbf{Z}^* \ni \mathbf{0}, \\ -\mathbf{Z}^* + \partial\lambda\|\mathbf{U}^*\|_{0/1}^M \ni \mathbf{0}, \\ \mathcal{A}\mathbf{X}^* + \mathbf{B} - \mathbf{U}^* = \mathbf{0}. \end{cases} \quad (4.10)$$

注 4.1 表明使用临近点算子可以对 (4.10) 中的 $\mathbf{Z}^* \in \partial\lambda\|\mathbf{U}^*\|_{0/1}^M$ 进行更加细致的刻画. 临近点算子通常用于设计迭代算法 [24], 因此相应的 P 稳定点 (proximal stationary point) 会更加有利于算法性质分析. 文献 [28, 30] 针对向量 0/1 复合优化的算法收敛性分析均采用了 P 稳定点, 本文在此将这一定义推广到矩阵 0/1 复合优化 (1.3).

定义 4.2 (P 稳定点) 如果点 $\mathbf{W}^* := (\mathbf{X}^*, \mathbf{U}^*)$ 和 $\mathbf{Z}^*$ 满足如下系统:

$$\begin{cases} \partial h(\mathbf{X}^*) + \mathcal{A}^\top \mathbf{Z}^* \ni \mathbf{0}, \\ \mathbf{U}^* \in \text{Prox}_{\alpha\lambda\|\cdot\|_{0/1}^M}(\mathbf{U}^* + \alpha\mathbf{Z}^*), \\ \mathcal{A}\mathbf{X}^* + \mathbf{B} - \mathbf{U}^* = \mathbf{0}, \end{cases} \quad (4.11)$$

则称它们分别为问题 (1.3) 的 P 稳定点和 P 稳定乘子, $(\mathbf{W}^*, \mathbf{Z}^*)$ 称作该问题的 P 稳定对.

稳定点系统 (4.10) 和 (4.11) 的第二行均可保证互补条件 $\mathbf{U}^* \perp \mathbf{Z}^* \geq \mathbf{0}$ 成立, 这说明它们与带有不等式约束的优化问题可以建立联系. 文献 [29] 利用这种思想将向量 0/1 复合优化问题与特定的光滑线性约束优化的 KKT 点建立等价关系, 从而在最优性分析中无需引入任何约束规范假设. 本文接下来将这类等价关系推广到矩阵 0/1 复合优化. 给定问题 (1.3) 的可行点 $\mathbf{W}^* = (\mathbf{X}^*, \mathbf{U}^*)$, 定义如下指标集:

$$\mathcal{S}_{\leq}^* := \{i \in [m] \mid \max \mathbf{u}_i^* \leq 0\}, \quad \mathcal{S}_{>}^* := \{i \in [m] \mid \max \mathbf{u}_i^* > 0\}.$$

基于指标集 $\mathcal{S}_{\leq}^*$, 定义如下线性约束优化问题:

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{U}} \quad & h(\mathbf{X}) \\ \text{s.t.} \quad & \mathcal{A}\mathbf{X} + \mathbf{B} - \mathbf{U} = \mathbf{0}, \\ & \mathbf{U}_{\mathcal{S}_{\leq}^*} \leq \mathbf{0} \end{aligned} \quad \Bigg\} =: \mathcal{F}_K^* \text{ (可行集)}. \quad (4.12)$$

该问题的 Lagrange 函数可写作

$$\mathcal{L}^*(\mathbf{W}, \mathbf{Z}, \mathbf{Q}) = h(\mathbf{X}) + \langle \mathbf{Z}, \mathcal{A}\mathbf{X} + \mathbf{B} - \mathbf{U} \rangle + \langle \mathbf{Q}, \mathbf{U}_{\mathcal{S}_{\leq}^*} \rangle,$$

其中 $(\mathbf{Z}, \mathbf{Q})$ 为问题 (4.12) 的 Lagrange 乘子.

定义 4.3 (伪 KKT 点) 若 $\mathbf{W}^* \in \mathcal{F}_K^*$ 和 $(\mathbf{Z}^*, \mathbf{Q}^*)$ 分别为问题 (4.12) 的 KKT 点和 Lagrange 乘子, 即满足

$$\begin{cases} \mathbf{0} \in \partial_{\mathbf{W}} \mathcal{L}^*(\mathbf{W}^*, \mathbf{Z}^*, \mathbf{Q}^*), \\ \mathbf{0} \leq \mathbf{Q}^* \perp \mathbf{U}_{\mathcal{S}_{\leq}^*}^* \leq \mathbf{0}, \end{cases} \quad (4.13)$$

则称它们为问题 (1.3) 的伪 KKT 点和伪 Lagrange 乘子, 此外称 $(\mathbf{W}^*, \mathbf{Z}^*, \mathbf{Q}^*)$ 为伪 KKT 对.

定义 4.3 的命名方式参考了文献 [10] 中的伪 B 稳定点 (pseudo Bouligand stationary point). 这一工作考虑了 0/1 损失函数和 B 可微函数复合的优化问题, 因此可以包含本文的 0/1 损失函数和 max 函数复合情形. 为了给出问题 (1.3) 的伪 B 稳定点定义, 需要引入如下涉及指标集 $\mathcal{S}_{\leq}^*$ 和 $\mathcal{S}_{>}^*$ 的优化问题:

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{U}} \quad & h(\mathbf{X}) \\ \text{s.t.} \quad & \mathcal{A}\mathbf{X} + \mathbf{B} - \mathbf{U} = \mathbf{0}, \\ & \max \mathbf{u}_i \geq 0, \quad i \in \mathcal{S}_{>}^*, \\ & \max \mathbf{u}_j \leq 0, \quad j \in \mathcal{S}_{\leq}^* \end{aligned} \quad \Bigg\} =: \mathcal{F}_B^* \text{ (可行集)}. \quad (4.14)$$

定义 4.4 (伪 B 稳定点) 若 $\mathbf{W}^* \in \mathcal{F}_B^*$ 为问题 (4.14) 的 B 稳定点, 即满足

$$h'(\mathbf{X}^*; \mathbf{D}^X) \geq 0, \quad \forall (\mathbf{D}^X, \mathbf{D}^U) \in \mathcal{T}_{\mathcal{F}_B^*}(\mathbf{W}^*),$$

则称其为问题 (1.3) 的伪 B 稳定点.

优化问题 (4.12) 的可行集 $\mathcal{F}_K^*$ 为凸集, 而相比之下 (4.14) 的可行集 $\mathcal{F}_B^*$ 增加了约束 $\max \mathbf{u}_i \geq 0$, $i \in \mathcal{S}_>^*$, 所以 $\mathcal{F}_B^*$ 非凸. 实际上, 对于点 $\mathbf{W}^*$, 这些约束可以严格成立, 属于非积极约束, 在后续分析中不必考虑. 所以相比之下, 问题 (4.12) 的性质更好且形式简洁, 它在第 5 节还可以用于建立 (1.3) 的二阶最优性条件.

最近, Han 等^[13] 针对 0/1 损失和一般函数复合的优化问题提出了一种 E 稳定点 (epi-stationary point). 为了给出问题 (1.3) 的 E 稳定点, 需要考虑其等价的上图化形式:

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{U}, r} \quad & r \\ \text{s.t.} \quad & \mathcal{A}\mathbf{X} + \mathbf{B} - \mathbf{U} = \mathbf{0}, \\ & h(\mathbf{X}) + \lambda \|\mathbf{U}\|_{0/1}^M \leq r \end{aligned} \quad \Bigg\} =: \mathcal{F}_E \text{ (可行集).} \quad (4.15)$$

给定点 $\mathbf{W}^*$, 为简便起见, 接下来定理中将记 $r^* := h(\mathbf{X}^*) + \lambda \|\mathbf{U}^*\|_{0/1}^M$.

定义 4.5 (E 稳定点) 若 $(\mathbf{W}^*, r^*) \in \mathcal{F}_E$ 为问题 (4.15) 的 B 稳定点, 即满足

$$d^r \geq 0, \quad \forall (\mathbf{D}^W, d^r) \in \mathcal{T}_{\mathcal{F}_E}(\mathbf{W}^*),$$

则称 $\mathbf{W}^*$ 为问题 (1.3) 的 E 稳定点.

下面将给出定义 4.1–4.5 五种稳定点之间的关系, 由于 P 稳定点的定义方式引入了参数 $\alpha > 0$, 因此它和其他稳定点的关系会单独讨论.

定理 4.3 问题 (1.3) 的稳定点具有如下关系:

- (i) 若 h 为光滑或凸函数, 则 KKT 点、伪 KKT 点、伪 B 稳定点以及 E 稳定点均相互等价.
- (ii) 若 $(\mathbf{W}^*, \mathbf{Z}^*)$ 为 P 稳定对, 则它为 KKT 对; 反之, 若 $(\mathbf{W}^*, \mathbf{Z}^*)$ 为 KKT 对, 则它为 P 稳定对, 且参数 $\alpha \in (0, \min\{\alpha_U^*, \alpha_Z^*\})$, 其中

$$\alpha_U^* := \begin{cases} +\infty, & \text{若 } \mathbf{U}^* \leq \mathbf{0}, \\ \min_{\mathbf{u}_i^* \in \mathbb{R}_{\leq}^q, i \in [m]} \frac{\|(\mathbf{u}_i^*)_+\|^2}{2\lambda}, & \text{其他,} \end{cases} \quad \alpha_Z^* := \begin{cases} +\infty, & \text{若 } \mathbf{Z}^* = \mathbf{0}, \\ \min_{\mathbf{z}_i^* \neq \mathbf{0}, i \in [m]} \frac{2\lambda}{\|\mathbf{z}_i^*\|^2}, & \text{其他.} \end{cases}$$

证明 (i) KKT 点 $\Leftrightarrow$ 伪 KKT 点. 伪 KKT 点对 $(\mathbf{W}^*, \mathbf{Z}^*, \mathbf{Q}^*)$ 满足的系统 (4.13) 等价于

$$\begin{cases} \partial h(\mathbf{X}^*) + \mathcal{A}^\top \mathbf{Z}^* \ni \mathbf{0}, \\ \mathbf{0} \leq \mathbf{Z}_{\mathcal{S}_{\leq}^*} \perp \mathbf{U}_{\mathcal{S}_{\leq}^*} \leq \mathbf{0}, \quad \mathbf{Z}_{\mathcal{S}_{>}^*} = \mathbf{0}, \\ \mathcal{A}\mathbf{X}^* + \mathbf{B} - \mathbf{U}^* = \mathbf{0}, \quad \mathbf{Q}^* = \mathbf{Z}_{\mathcal{S}_{\leq}^*}^*, \end{cases} \quad (4.16)$$

相比于 (4.10), 仅在第二行不同. 两者的等价性由矩阵 0/1 损失函数的极限次微分表达式 (4.1) 可直接验证.

伪 KKT 点 $\Leftrightarrow$ 伪 B 稳定点. 令 $\mathbf{W}^*$ 为伪 B 稳定点, 首先来说明当邻域半径 δ_1 足够小时, 可保证 $\mathcal{F}_B^* \cap N(\mathbf{W}^*, \delta_1) = \mathcal{F}_K^* \cap N(\mathbf{W}^*, \delta_1)$. 由于 $\mathcal{F}_B^* \subseteq \mathcal{F}_K^*$, 故只需要证明 $\mathcal{F}_B^* \cap N(\mathbf{W}^*, \delta_1) \supseteq \mathcal{F}_K^*$

$\cap N(\mathbf{W}^*, \delta_1)$. 由于对于任意 $i \in \mathcal{S}_>$ 都有 $\max \mathbf{u}_i^* > 0$, 所以当 δ_1 足够小时, 根据 $\max$ 函数的连续性, 任取 $\mathbf{W} \in \mathcal{F}_K^* \cap N(\mathbf{W}^*, \delta_1)$, 对于任意 $i \in \mathcal{S}_>$, $\max \mathbf{u}_i > 0$ 均成立, 所以 $\mathbf{W} \in \mathcal{F}_B^* \cap N(\mathbf{W}^*, \delta_1)$. 当 $\mathcal{F}_B^* \cap N(\mathbf{W}^*, \delta_1) = \mathcal{F}_K^* \cap N(\mathbf{W}^*, \delta_1)$ 成立时, 由切锥定义可得 $\mathcal{T}_{\mathcal{F}_B^*}(\mathbf{W}^*) = \mathcal{T}_{\mathcal{F}_K^*}(\mathbf{W}^*)$, 从而 $\mathbf{W}^*$ 为伪 B 稳定点等价于

$$h'(\mathbf{X}^*; \mathbf{D}^X) \geq 0, \quad \forall (\mathbf{D}^X, \mathbf{D}^U) \in \mathcal{T}_{\mathcal{F}_K^*}(\mathbf{W}^*). \quad (4.17)$$

当 f 为光滑或凸函数时, 可由文献 [23, 定理 6.12] 和 [19, 定理 4.14] 证明 $\mathbf{W}^*$ 等价于问题 (4.12) 的 KKT 点, 即 (1.3) 的伪 KKT 点.

伪 KKT 点 $\Leftrightarrow$ E 稳定点. 令 $\mathbf{W}^*$ 为 E 稳定点且 $r^* := h(\mathbf{X}^*) + \lambda \|\mathbf{U}^*\|_{0/1}^M$, 则 $(\mathbf{W}^*, r^*)$ 同时为如下优化问题的可行点:

$$\left. \begin{array}{l} \min_{\mathbf{X}, \mathbf{U}, r} \quad r \\ \text{s.t.} \quad \mathbf{A}\mathbf{X} + \mathbf{B} - \mathbf{U} = \mathbf{0}, \\ h(\mathbf{X}) + \lambda \|\mathbf{U}^*\|_{0/1}^M \leq r, \\ \mathbf{U}_{\mathcal{S}_\leq^*} \leq \mathbf{0} \end{array} \right\} =: \mathcal{F}^* \text{ (可行集)}. \quad (4.18)$$

令邻域半径 δ_2 足够小从而使 $\mathbf{W} \in N(\mathbf{W}^*, \delta_2)$ 满足如下条件:

$$\delta_2 < \lambda/4, \quad |h(\mathbf{X}) - h(\mathbf{X}^*)| \leq \lambda/2, \quad \mathcal{S}_\leq^* \subseteq \mathcal{S}_> := \{i \in [m] \mid \max \mathbf{u}_i > 0\}, \quad (4.19)$$

其中第二个不等式由于 h 为光滑或有限凸函数从而满足连续性, 集合包含关系由 $\max$ 函数连续性得到. 接下来将通过左右集合互相包含证明 $\mathcal{F}_E \cap N((\mathbf{W}^*, r^*), \delta_2) = \mathcal{F}^* \cap N((\mathbf{W}^*, r^*), \delta_2)$.

“ $\supseteq$ ”. 对于任意 $(\mathbf{W}, r) \in \mathcal{F}^* \cap N((\mathbf{W}^*, r^*), \delta_2)$, 均满足 $\mathbf{U}_{\mathcal{S}_\leq^*} \leq \mathbf{0}$, 结合 (4.19) 中的 $\mathcal{S}_\leq^* \subseteq \mathcal{S}_>$ 可推导出 $\|\mathbf{U}\|_{0/1}^M = \|\mathbf{U}^*\|_{0/1}^M$, 从而 $(\mathbf{W}, r) \in \mathcal{F}_E \cap N((\mathbf{W}^*, r^*), \delta_2)$.

“ $\subseteq$ ”. 对于任意 $(\mathbf{W}, r) \in \mathcal{F}_E \cap N((\mathbf{W}^*, r^*), \delta_2)$, 均满足

$$r \geq h(\mathbf{X}) + \lambda \|\mathbf{U}\|_{0/1}^M. \quad (4.20)$$

为证明 $(\mathbf{W}, r) \in \mathcal{F}^* \cap N((\mathbf{W}^*, r^*), \delta_2)$, 只需要证明 $\mathcal{S}_\leq^* = \mathcal{S}_>$ 即可. 假设这一结论不成立, 则 $\mathcal{S}_\leq^* \subsetneq \mathcal{S}_>$, 必然有 $\|\mathbf{U}\|_{0/1}^M \geq \|\mathbf{U}^*\|_{0/1}^M + 1$. 结合 (4.20) 和 (4.19), 可以得到

$$r \geq h(\mathbf{X}) + \lambda \|\mathbf{U}^*\|_{0/1}^M + \lambda \geq h(\mathbf{X}^*) + \lambda \|\mathbf{U}^*\|_{0/1}^M + \lambda/2 = r^* + \lambda/2.$$

显然这与 $|r - r^*| < \delta_2 \leq \lambda/4$ 矛盾.

当 $\mathcal{F}_E \cap N((\mathbf{W}^*, r^*), \delta_2) = \mathcal{F}^* \cap N((\mathbf{W}^*, r^*), \delta_2)$ 时, 由切锥定义可得 $\mathcal{T}_{\mathcal{F}_E}(\mathbf{W}^*) = \mathcal{T}_{\mathcal{F}^*}(\mathbf{W}^*)$, 从而 $\mathbf{W}^*$ 为 E 稳定点等价于

$$d^r \geq 0, \quad \forall (\mathbf{D}^W, d^r) \in \mathcal{T}_{\mathcal{F}^*}(\mathbf{W}^*). \quad (4.21)$$

根据文献 [13, 命题 5], 可以证明 (4.21) 实际上等价于 (4.17). 同上一条等价关系证明, 可以得到 $\mathbf{W}^*$ 等价于 (1.3) 的伪 KKT 点.

(ii) 比较表达式 (4.11) 和 (4.10), 根据定理 4.2(ii), 只需要证明

$$\mathbf{Z}^* \in \partial \|\mathbf{U}^*\|_{0/1}^M \Leftrightarrow \text{对于任意 } i \in [m], \mathbf{u}_i^* \text{ 和 } \mathbf{z}_i^* \text{ 满足表达式 (4.7),}$$

其中“ $\Leftarrow$ ”通过比较 (4.7) 和 (4.1) 可直接验证. 接下来证明“ $\Rightarrow$ ”. 任取 $\mathbf{Z}^* \in \partial\|\mathbf{U}^*\|_{0/1}^M$, 对于任意 $i \in [m]$, 考虑如下 3 种情形:

若 $\mathbf{u}_i^* \in \mathbb{R}_{\leq}^q$ 且 $\mathbf{z}_i^* = \mathbf{0}$, 则 $\|(\mathbf{u}_i^*)_+\|^2/(2\lambda) \geq \alpha_U^* > \alpha$, 从而 $\|(\mathbf{u}_i^*)_+\| \in [\sqrt{2\alpha\lambda}, +\infty)$.

若 $\mathbf{u}_i^* \leq \mathbf{0}$ 且 $\mathbf{z}_i^* = \mathbf{0}$, 则显然两者取值满足 (4.7) 第二条.

若 $\mathbf{u}_i^* \leq \mathbf{0}$ 且 $\mathbf{z}_i^* \neq \mathbf{0}$, 则 $2\lambda/\|\mathbf{z}_i^*\|^2 \geq \alpha_Z^* > \alpha$, 从而 $\|\mathbf{z}_i^*\| \in (0, \sqrt{2\lambda/\alpha}]$.

以上 3 种情形均满足 (4.7), 至此完成定理证明. $\square$

注 4.2 通过观察 (4.16) 可以发现, 若 $(\mathbf{Z}^*, \mathbf{Q}^*)$ 为 (1.3) 的伪 Lagrange 乘子, 则一定有 $\mathbf{Q}^* = \mathbf{Z}_{S_{\leq}}^*$. 此外, $(\mathbf{W}^*, \mathbf{Z}^*, \mathbf{Z}_{S_{\leq}}^*)$ 为伪 KKT 对当且仅当 $(\mathbf{W}^*, \mathbf{Z}^*)$ 为 KKT 对.

由于定理 4.3 给出了各稳定点相互等价关系, 下面统一采用 P- 稳定点来建立问题 (1.3) 的一阶最优性条件. 值得注意的是, 由于 P- 稳定点引入了参数 $\alpha > 0$, 这一特点可以使它在特定条件下与问题 (1.3) 的全局最优解建立联系.

定理 4.4 (一阶最优性条件) 对于问题 (1.3), 给定可行点 $\mathbf{W}^*$, 以下结论成立:

- (i) 若函数 h 连续可微, 则局部最优解 $\mathbf{W}^*$ 一定是 P 稳定点;
- (ii) 若 h 为凸函数, 则 $\mathbf{W}^*$ 为局部最优解当且仅当它为 P 稳定点;
- (iii) 若 h 为 σ_h - 强凸函数, $\mathbf{W}^*$ 为 P 稳定点且相应的参数 $\alpha \geq \|\mathbf{A}\|^2/\sigma_h$, 则 $\mathbf{W}^*$ 为全局最优解.

证明 (i) 由文献 [13, 命题 3] 可知问题 (1.3) 的局部最优解 $\mathbf{W}^*$ 一定是 E 稳定点, 再根据定理 4.3 可推导出 $\mathbf{W}^*$ 为 P 稳定点.

(ii) 若 $\mathbf{W}^*$ 为 P 稳定点, 由定理 4.3 可知它一定是 E 稳定点, 根据文献 [13, 推论 3] 可证明 $\mathbf{W}^*$ 与局部最优解等价.

(iii) 若 $\mathbf{W}^*$ 为 P 稳定点, 根据 (4.11) 和临近点算子定义, 对于任意 $\mathbf{U} = \mathbf{A}\mathbf{X} + \mathbf{B}$, 都有

$$\frac{1}{2\alpha}\|\mathbf{U} - (\mathbf{U}^* + \alpha\mathbf{Z}^*)\|^2 + \lambda\|\mathbf{U}\|_{0/1}^M \leq \frac{1}{2\alpha}\|\alpha\mathbf{Z}^*\|^2 + \lambda\|\mathbf{U}^*\|_{0/1}^M,$$

化简可得

$$\begin{aligned} \lambda\|\mathbf{U}\|_{0/1}^M - \lambda\|\mathbf{U}^*\|_{0/1}^M &\geq \langle \mathbf{Z}^*, \mathbf{U} - \mathbf{U}^* \rangle - \frac{1}{2\alpha}\|\mathbf{U} - \mathbf{U}^*\|^2 \\ &= \langle \mathbf{A}^\top \mathbf{Z}^*, \mathbf{X} - \mathbf{X}^* \rangle - \frac{1}{2\alpha}\|\mathbf{A}(\mathbf{X} - \mathbf{X}^*)\|^2, \end{aligned}$$

其中上式第二行由 $\mathbf{U}^* = \mathbf{A}\mathbf{X}^* + \mathbf{B}$ 得到. 由 (4.11) 第一行可知, 存在 $\mathbf{V}^* \in \partial h(\mathbf{X}^*)$, 使得 $-\mathbf{V}^* = \mathbf{A}^\top \mathbf{Z}^*$, 从而

$$\lambda\|\mathbf{U}\|_{0/1}^M - \lambda\|\mathbf{U}^*\|_{0/1}^M \geq -\langle \mathbf{V}^*, \mathbf{X} - \mathbf{X}^* \rangle - \frac{\|\mathbf{A}\|^2}{2\alpha}\|\mathbf{X} - \mathbf{X}^*\|^2. \quad (4.22)$$

当函数 h 满足 σ_h - 强凸时, 根据文献 [1, 定理 5.24], 可得

$$h(\mathbf{X}) - h(\mathbf{X}^*) \geq \langle \mathbf{V}^*, \mathbf{X} - \mathbf{X}^* \rangle + \frac{\sigma_f}{2}\|\mathbf{X} - \mathbf{X}^*\|.$$

将该不等式与 (4.22) 相加, 可知对于任意 $\mathbf{W} \in \mathcal{F}$ 都有

$$h(\mathbf{X}) + \lambda\|\mathbf{U}\|_{0/1}^M - (h(\mathbf{X}^*) + \lambda\|\mathbf{U}^*\|_{0/1}^M) \geq \left(\frac{\sigma_f}{2} - \frac{\|\mathbf{A}\|^2}{2\alpha}\right)\|\mathbf{X} - \mathbf{X}^*\|^2 \geq 0.$$

从而 $\mathbf{W}^*$ 为 (1.3) 全局最优解. $\square$

5 二阶最优性理论

给定问题 (1.3) 的 P 稳定对 $(\mathbf{W}^*, \mathbf{Z}^*)$, 定理 4.3 和注 4.2 表明 $(\mathbf{W}^*, \mathbf{Z}^*, \mathbf{Z}_{S_{\leq}^*}^*)$ 一定是问题 (4.12) 的 KKT 对. 相比于问题 (1.3), 问题 (4.12) 目标函数去掉了非凸非连续的 $\|\cdot\|_{0/1}^M$ 且仅含有线性约束, 非常有利于最优性分析.

基于以上观察, 本节的核心思想是利用简洁的线性约束连续优化问题 (4.12) 来建立 (1.3) 二阶最优性条件. 这种技术可以使考虑更具一般性的函数 h , 例如, 利用文献 [14, 16] 中的结果, 可以在函数 $h \in C^{1,1}$ 且无需任何约束规范情形下建立 (1.3) 二阶最优性条件. 在此之前需要介绍临界锥的相关概念. 定义如下指标集:

$$\mathcal{I}_0^* := \{(i, k) \in [m] \times [q] \mid u_{ik}^* = 0, z_{ik}^* = 0\}, \quad \mathcal{I}_{>}^* := \{(i, k) \in [m] \times [q] \mid u_{ik}^* = 0, z_{ik}^* > 0\}.$$

由文献 [21, 第 12.5 小节] 可知, 问题 (4.12) 在乘子 $(\mathbf{W}^*, \mathbf{Z}^*, \mathbf{Z}_{S_{\leq}^*}^*)$ 处的临界锥表达式为

$$\mathcal{C}(\mathbf{W}^*, \mathbf{Z}^*) := \{\mathbf{D} = (\mathbf{D}^X, \mathbf{D}^U) \mid \mathcal{A}\mathbf{D}^X - \mathbf{D}^U = \mathbf{0}, \mathbf{D}_{\mathcal{I}_0^*}^U \leq \mathbf{0}, \mathbf{D}_{\mathcal{I}_{>}^*}^U = \mathbf{0}\},$$

可进一步简化为

$$\mathcal{C}^* = \{\mathbf{D}^X \mid (\mathcal{A}\mathbf{D}^X)_{\mathcal{I}_0^*} \leq \mathbf{0}, (\mathcal{A}\mathbf{D}^X)_{\mathcal{I}_{>}^*} = \mathbf{0}\}.$$

定理 5.1 (二阶必要条件) 若 $h \in C^{1,1}$ 且 $\mathbf{W}^*$ 为问题 (1.3) 的局部最优解, 则存在相应的 P 稳定乘子 $\mathbf{Z}^*$. 那么对每个 $\mathbf{D}^X \in \mathcal{C}^*$, 都存在线性算子 $\mathcal{M}^* \in \partial^2 h(\mathbf{X}^*)$, 使得 $\langle \mathbf{D}^X, \mathcal{M}^* \mathbf{D}^X \rangle \geq 0$.

证明 首先来说明 $\mathbf{W}^*$ 是问题 (4.12) 的局部最优解. 由于 $\mathbf{W}^*$ 为 (1.3) 的局部最优解, 所以存在 $\delta > 0$, 使得

$$h(\mathbf{X}) + \lambda \|\mathbf{U}\|_{0/1}^M \geq h(\mathbf{X}^*) + \lambda \|\mathbf{U}^*\|_{0/1}^M, \quad \forall \mathbf{W} \in N(\mathbf{W}^*, \delta) \cap \mathcal{F}. \quad (5.1)$$

若进一步 $\mathbf{W} \in N(\mathbf{W}^*, \delta) \cap \mathcal{F}_K^*$, 则 $\mathbf{U}_{S_{\leq}^*} \leq \mathbf{0}$, 从而 $\|\mathbf{U}\|_{0/1}^M \leq \|\mathbf{U}^*\|_{0/1}^M$. 再根据 (5.1) 可得

$$h(\mathbf{X}) \geq h(\mathbf{X}^*), \quad \forall \mathbf{W} \in N(\mathbf{W}^*, \delta) \cap \mathcal{F}_K^*,$$

即 $\mathbf{W}^*$ 为 (4.12) 局部最优解. 再根据文献 [14, 推论 3.1], 可证明对于任意 $\mathbf{D} \in \mathcal{C}(\mathbf{W}^*, \mathbf{Z}^*)$, 都存在 $\widetilde{\mathcal{M}}^* \in \partial_{\mathbf{W}, \mathbf{W}}^2 \mathcal{L}^*(\mathbf{W}^*, \mathbf{Z}^*, \mathbf{Z}_{S_{\leq}^*}^*)$, 使得 $\langle \mathbf{D}, \widetilde{\mathcal{M}}^* \mathbf{D} \rangle \geq 0$. 根据函数 $\mathcal{L}^*(\cdot, \mathbf{Z}^*, \mathbf{Z}_{S_{\leq}^*}^*)$ 定义可以验证

$$\partial_{\mathbf{W}, \mathbf{W}}^2 \mathcal{L}^*(\mathbf{W}^*, \mathbf{Z}^*, \mathbf{Z}_{S_{\leq}^*}^*) = \begin{pmatrix} \partial^2 h(\mathbf{X}^*) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}. \quad (5.2)$$

代入以上表达式即可得到最终结论. □

下面用一个例子来验证矩阵 0/1 复合优化稳定点关系和二阶必要最优性条件.

例 5.1 定义一维函数 $g(t) := t^3 - t$, 考虑问题 (1.3) 满足如下参数设置:

$$m = n = s = 2, \quad q = 3, \quad h(\mathbf{X}) = \frac{1}{2} \sum_{i=1}^2 \sum_{j=1}^2 [g(x_{ij})]_+^2, \quad \lambda = 1,$$

$$\mathcal{A}\mathbf{X} = \begin{bmatrix} -x_{11} & x_{12} & -x_{11} + x_{12} \\ x_{11} + x_{21} & x_{12} + x_{22} & x_{22} \end{bmatrix}, \quad \mathbf{B} = \mathbf{1},$$

此时有 $h \in C^{1,1}$. 下面考虑点 $\mathbf{X}^* = [1, -1; -1, 1]$, $\mathbf{U}^* = \mathbf{A}\mathbf{X}^* + \mathbf{B} = [0, 0, -1; 1, 1, 2]$. 可以验证对于任意 $\mathbf{W} = (\mathbf{X}, \mathbf{U}) \in N(\mathbf{W}^*, 1)$, 都有

$$h(\mathbf{X}) \geq h(\mathbf{X}^*) = 0 \quad \text{且} \quad \|\mathbf{U}\|_{0/1}^M \geq \|\mathbf{U}^*\|_{0/1}^M = 1,$$

从而 $\mathbf{W}^* = (\mathbf{X}^*, \mathbf{U}^*)$ 为 (1.3) 的局部最优解. 令 $\mathbf{Z}^* = \mathbf{0}$, $\mathbf{Q}^* = \mathbf{0}$, 则 $(\mathbf{W}^*, \mathbf{Z}^*)$ 满足 (4.10) 且 $(\mathbf{W}^*, \mathbf{Z}^*, \mathbf{Q}^*)$ 满足 (4.16), 从而 $\mathbf{W}^*$ 同时为 (1.3) 的 KKT 点和伪 KKT 点. 此外, 还可以验证 $\mathbf{W}^*$ 满足 P- 稳定点系统 (4.11), 其中满足参数取值范围 $\alpha \in (0, 3)$. 不仅如此, 由于 $h'(\mathbf{X}^*; \mathbf{D}^X) = 0$, $\forall \mathbf{D}^X \in \mathbb{R}^{2 \times 2}$ 恒成立, 所以 $\mathbf{X}^*$ 还是伪 B 稳定点. 另外, 根据文献 [14, 例 2.2], 可以求解出

$$\partial^2 h(\mathbf{X}^*) = \{\mathbf{V} \odot (\cdot) \mid v_{ij} \in [0, 4], i, j \in \{1, 2\}\},$$

其中 $\odot$ 为 Hadamard 乘积. 在这种情形下, 对于任意 $\mathbf{D}^X \in \mathcal{C}^*$ 和 $\mathcal{M}^* \in \partial^2 h(\mathbf{X}^*)$, 都有 $\langle \mathbf{D}^X, \mathcal{M}^* \mathbf{D}^X \rangle \geq 0$, 因此二阶必要性条件成立.

定理 5.2 (二阶充分条件) 令 $h \in C^{1,1}$ 且 $(\mathbf{W}^*, \mathbf{Z}^*)$ 为问题 (1.3) 的 P 稳定对, 若对每个 $\mathcal{M}^* \in \partial^2 h(\mathbf{X}^*)$, 都有如下条件成立:

$$\langle \mathbf{D}^X, \mathcal{M}^* \mathbf{D}^X \rangle > 0, \quad \forall \mathbf{D}^X \in \mathcal{C}^* \setminus \{\mathbf{0}\}, \quad (5.3)$$

则 $\mathbf{W}^*$ 为问题 (1.3) 的严格局部最优解, 同时存在 $\delta^* > 0$ 和 $\gamma > 0$ 使得

$$h(\mathbf{X}) + \lambda \|\mathbf{U}\|_{0/1}^M \geq h(\mathbf{X}^*) + \lambda \|\mathbf{U}^*\|_{0/1}^M + \gamma \|\mathbf{W} - \mathbf{W}^*\|^2, \quad \forall \mathbf{W} \in N(\mathbf{W}^*, \delta^*) \cap \mathcal{F}. \quad (5.4)$$

证明 只需要证明 (5.4) 即可说明 $\mathbf{W}^*$ 为严格局部最优解. 不妨令 δ^* 足够小从而使得对于任意 $\mathbf{W} \in N(\mathbf{W}^*, \delta^*)$ 都有

$$\delta^* < \sqrt{\lambda/2\gamma}, \quad |h(\mathbf{X}) - h(\mathbf{X}^*)| \leq \lambda/2, \quad \mathcal{S}_>^* \subseteq \mathcal{S}_> := \{i \in [m] \mid \max \mathbf{u}_i > 0\}. \quad (5.5)$$

根据表达式 (5.2) 可知, 二阶充分条件 (5.2) 等价于对每个 $\widetilde{\mathcal{M}}^* \in \partial_{\mathbf{W}, \mathbf{W}}^2 \mathcal{L}^*(\mathbf{W}^*, \mathbf{Z}^*, \mathbf{Z}_{\mathcal{S}_<}^*)$, 都有

$$\langle \mathbf{D}, \widetilde{\mathcal{M}}^* \mathbf{D} \rangle > 0, \quad \forall \mathbf{D} \in \mathcal{C}(\mathbf{W}^*, \mathbf{Z}^*) \setminus \{\mathbf{0}\}.$$

当 δ^* 充分小时, 根据文献 [16, 定理 1], 可知存在 $\gamma > 0$ 使得

$$h(\mathbf{X}) - h(\mathbf{X}^*) \geq \gamma \|\mathbf{W} - \mathbf{W}^*\|^2, \quad \forall \mathbf{W} \in N(\mathbf{W}^*, \delta^*) \cap \mathcal{F}_K^*. \quad (5.6)$$

下面将分为两部分证明 (5.4).

若 $\mathbf{W} \in N(\mathbf{W}^*, \delta^*) \cap \mathcal{F}_K^*$, 由 (5.5) 中的 $\mathcal{S}_>^* \subseteq \mathcal{S}_>$ 可以证明 $\|\mathbf{U}\|_{0/1}^M \geq \|\mathbf{U}^*\|_{0/1}^M$, 再与 (5.6) 相加即可证明 (5.4).

若 $\mathbf{W} \in N(\mathbf{W}^*, \delta^*) \cap (\mathcal{F} \setminus \mathcal{F}_K^*)$, 结合 $\mathcal{S}_>^* \subseteq \mathcal{S}_>$, 则一定有 $\|\mathbf{U}\|_{0/1}^M \geq \|\mathbf{U}^*\|_{0/1}^M + 1$, 再根据 (5.5) 可得

$$h(\mathbf{X}) + \lambda \|\mathbf{U}\|_{0/1}^M - (h(\mathbf{X}^*) + \lambda \|\mathbf{U}^*\|_{0/1}^M) \geq \lambda - \lambda/2 \geq \gamma \|\mathbf{W} - \mathbf{W}^*\|^2.$$

至此完成本定理的证明. □

注 5.1 本文最优性理论无需任何约束规范的主要原因在于构造的连续优化问题 (4.12) 仅含线性约束, 这种约束集的切锥等于线性化可行方向锥 (参见文献 [21, 引理 12.7]). 所以, 如果在模型 (1.1) 基础上增加线性约束, 相应的最优性分析依然无需任何约束规范. 而对于矩阵 0/1 损失函数和一般光滑函数复合的情形, 利用本文方法转化后的优化问题涉及非线性约束, 相应的最优性理论则需要引入特定的约束规范.

6 结论

本文研究的矩阵 0/1 复合优化问题具有广泛的应用背景, 但是它是具有高度复合结构的 NP (non-deterministic polynomial) 困难问题, 直接对其进行最优性分析的文献非常少. 本文计算了矩阵 0/1 损失函数的极限次微分和临近点算子, 然后借鉴相关文献中的思想, 定义了矩阵 0/1 复合优化的 5 种稳定点, 研究了它们之间以及和该问题解的关系. 此外, 本文还建立了该问题的二阶必要和充分最优性条件. 如何准确高效地计算矩阵 0/1 复合优化的稳定点是重要的课题, 这需要在未来进一步细致研究.

参考文献

- 1 Beck A. First-Order Methods in Optimization. Philadelphia: SIAM, 2017
- 2 Blondel M, Seki K, Uehara K. Block coordinate descent algorithms for large-scale sparse multiclass classification. *Mach Learn*, 2013, 93: 31–52
- 3 Bolte J, Sabach S, Teboulle M. Nonconvex Lagrangian-based optimization: Monitoring schemes and global convergence. *Math Oper Res*, 2018, 43: 1210–1232
- 4 Boş I, Csetnek E R, Nguyen D K. A proximal minimization algorithm for structured nonconvex and nonsmooth problems. *SIAM J Optim*, 2019, 29: 1300–1328
- 5 Boufounos P T, Baraniuk R G. 1-bit compressive sensing. In: 2008 42nd Annual Conference on Information Sciences and Systems. Princeton: IEEE, 2008, 16–21
- 6 Cai Z, Zhu W. Multi-label feature selection via feature manifold learning and sparsity regularization. *Int J Mach Learn Cyber*, 2018, 9: 1321–1334
- 7 Chen C, Mangasarian O L. Hybrid misclassification minimization. *Adv Comput Math*, 1996, 5: 127–136
- 8 Cortes C, Vapnik V. Support-vector networks. *Mach Learn*, 1995, 20: 273–297
- 9 Crammer K, Singer Y. On the algorithmic implementation of multiclass kernel-based vector machines. *J Mach Learn Res*, 2001, 2: 265–292
- 10 Cui Y, Liu J, Pang J-S. The minimization of piecewise functions: Pseudo stationarity. *J Convex Anal*, 2023, 30: 793–834
- 11 Feng M, Mitchell J E, Pang J-S, et al. Complementarity formulations of ℓ_0 -norm optimization problems. *Pac J Optim*, 2018, 14: 273–305
- 12 Han A K. Non-parametric analysis of a generalized regression model. *J Econom*, 1987, 35: 303–316
- 13 Han S, Cui Y, Pang J-S. Analysis of a class of minimization problems lacking lower semicontinuity. <https://optimization-online.org/?p=24391>, 2023
- 14 Hiriart-Urruty J B, Strodio J J, Nguyen V H. Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data. *Appl Math Optim*, 1984, 11: 43–56
- 15 Kanzow C, Schwarz A, Weiß F. The sparse(st) optimization problem: Reformulations, optimality, stationarity, and numerical results. *arXiv:2210.09589*, 2022
- 16 Klatte D, Tammek K. On second-order sufficient optimality conditions for $C^{1,1}$ -optimization problems. *Optimization*, 1988, 19: 169–179
- 17 Li N, Jin R, Zhou Z-H. Top rank optimization in linear time. In: *Advances in Neural Information Processing Systems*. Montreal: Curran Associates, 2014, 1–9
- 18 Lütkepohl H. *Handbook of Matrices*. New York: John Wiley & Sons, 1997
- 19 Mordukhovich B S, Nam N M. *An Easy Path to Convex Analysis and Applications*. San Rafael: Morgan & Claypool Publishers, 2013
- 20 Nie F, Huang H, Cai X, et al. Efficient and robust feature selection via joint $\ell_{2,1}$ -norms minimization. In: *Advances in Neural Information Processing Systems*. Vancouver: Curran Associates, 2010, 1–9
- 21 Nocedal J, Wright S. *Numerical Optimization*. New York: Springer, 2006
- 22 Pan L L, Xiu N H, Fan J. Optimality conditions for sparse nonlinear programming. *Sci China Math*, 2017, 60: 759–776
- 23 Rockafellar R T, Wets R J-B. *Variational Analysis*. Berlin: Springer, 1998
- 24 Themelis A, Stella L, Patrinos P. Forward-backward envelope for the sum of two nonconvex functions: Further properties and nonmonotone linesearch algorithms. *SIAM J Optim*, 2018, 28: 2274–2303
- 25 Wang L, Shen X. On ℓ_1 -norm multiclass support vector machines. *J Am Statistical Assoc*, 2007, 102: 583–594
- 26 Wu G, Zhu J. Multi-label classification: Do Hamming loss and subset accuracy really conflict with each other? In: *Advances in Neural Information Processing Systems*. New York: Curran Associates, 2020, 3130–3140
- 27 Zhang M L, Zhou Z H. A review on multi-label learning algorithms. *IEEE Trans Knowl Data Eng*, 2013, 26: 1819–1837

- 28 Zhang P, Xiu N, Luo Z. Zero-one composite optimization: Lyapunov exact penalty and a globally convergent inexact augmented Lagrangian method. *Math Oper Res*, 2024, 49: 2049–2802
- 29 Zhang P, Xiu N, Qi H-D. iNALM: An inexact Newton augmented Lagrangian method for zero-one composite optimization. *arXiv:2306.08991*, 2023
- 30 Zhou S, Pan L, Xiu N, et al. Quadratic convergence of smoothing Newton’s method for 0/1 loss optimization. *SIAM J Optim*, 2021, 31: 3184–3211

Matrixed 0/1 composite optimization model and optimality analysis

Penghe Zhang, Naihua Xiu & Houduo Qi

Abstract In this paper, a novel matrixed 0/1 composite optimization problem is abstracted from practical applications including multi-class support vector machine, multi-label learning, and bipartite ranking. Its objective function involves three layers of composition: the discrete 0/1 loss function, the nonsmooth maximum function, and a linear term. Inspired by some related works, we define five types of stationary points for this problem and examine the interrelations among these stationary points and their connection to minimizers of the problem. Additionally, second-order necessary and sufficient optimality conditions for matrixed 0/1 composite optimization problems are established. Although the matrixed 0/1 composite optimization essentially belongs to nonconvex and discrete composite problems, the optimality results derived in this paper do not require any constraint qualification.

Keywords matrixed 0/1 composite optimization, matrixed 0/1 loss function, stationary point, second-order optimality condition

MSC(2020) 90C46, 90C90

doi: 10.1360/SSM-2024-0136