arXiv:2601.03008v1 [math.OC] 6 Jan 2026

具有二元约束的非光滑非线性优化松弛方法

肖亮海、钱逸天和潘少华

摘要——**我们研究如下形式的二元优化问题：** $\min_{x\in\{-1,1\}^n} f(Ax-b)$ **，其中损失函数** f **可能是非光滑的。遵循提升的秩一半定规划（SDP）方法 [1]，我们针对秩一约束采用差分凸分解，并针对非光滑损失函数采用莫罗包络，从而提出一种主化-最小化算法。我们为所提出的差分凸松弛算法（DCRA）提供全局复杂度保证，并证明该算法能够生成一个近似可行的二进制解，且其最优性间隙具有明确的上界。在合成数据集和真实数据集上的数值实验表明，与现有方法相比，我们的方法具有更高的精度和更好的可扩展性。**

索引词—**二元正交优化问题、全局精确罚函数、松弛方法、语义哈希。**

I. 引言

CO考虑以下二元优化问题 lem:

$$\min_{x\in\{-1,1\}^n} f(Ax-b), \tag{1}$$

其中， $f:\mathbb{R}^r\to\mathbb{R}$ 是一个局部利普希茨连续（可能非光滑）的函数， $A\in\mathbb{R}^{r\times n}$，且 $b\in\mathbb{R}^r$。该模型涵盖了工程、计算机科学和应用数学领域中广泛的重要二元优化问题。下面列举几个具有代表性的实例。

最大 2-SAT。 [2] 考虑 m个逻辑子句，每个子句涉及两个变量。目标是最大化满足的子句数量，等价地说，就是最小化违反的子句数量：

$$\min_{x\in\{-1,1\}^n} \sum_{j=1}^{m} \ell_j(x),$$

其中，每个 $\ell_j(x)$都是一个非光滑损失函数，用于指示子句 j 是否被违反。

二元压缩感知与稳健回归。 [3]，[4] 在稀疏信号恢复中，目标是寻找一个符合线性测量的二元信号：

$$\min_{x\in\{0,1\}^n} \|Ax-b\|_1+\rho\|x\|_1,$$

其中， ℓ_1 范数同时促进稳健性和稀疏度，而通过简单的变量替换可将 $\{0,1\}^n$ 与 $\{-1,1\}^n$ 编码。一个密切相关的特例是绝对值回归：

$$\min_{x\in\{-1,1\}^n} \|Ax-b\|_1,$$

它与（1）直接对应，其中 $f(z)=\|z\|_1$。

通信中的MIMO检测。 [5], [6] 对于接收信号 $y\in\mathbb{R}^r$ 和信道矩阵 $H\in\mathbb{R}^{r\times n}$，符号检测建模为

$$\min_{x\in\{-1,1\}^n} \|Hx-y\|_2^2,$$

这是一个二元最小二乘问题。

哈希与二元嵌入。 [7]–[9] 在大规模检索中，监督哈希通过求解学习二元码

$$\min_{X\in\{-1,1\}^{n\times r},\, W\in\mathbb{R}^{d\times r}} \|B-WX^\top\|_1+\tfrac{\delta}{2}\|W\|_F^2,$$

其中， B 是与数据或标签相关的矩阵， W 是线性投影， X 是二元码矩阵。向量化后，这符合形式（1）。

这些示例表明，形式（1）既涵盖经典的 NP 难组合优化问题（二次指派问题 [10], 2、SAT [2]），也适用于信号处理、统计学、通信和机器学习等现代领域。特别是， ℓ_1 回归和监督哈希等非光滑形式凸显了开发可扩展方法的重要性，以解决超出光滑或二次情形的形式（1）问题。

A. 相关工作

尽管形式（1）具有广泛的适用性，但由于组合约束与非光滑目标函数相互作用，求解此类问题仍面临重大挑战。经典方法包括割平面法和分支割平面算法等精确及近似精确方法 [11]，[12]。割平面法通过添加有效不等式，逐步强化松弛；分支割平面法则将分支与动态生成的割平面相结合，以收紧线性或半定松弛。尽管这些方法能够提供较强的界，并且原则上可以保证全局最优性，但它们的计算成本往往高得难以承受：在最坏情况下，可能需要探索多达 2^n 个子问题，这限制了它们在大规模实例中的实用性。

为提升可扩展性，大量研究聚焦于连续松弛方法，即用凸替代取代二元可行集。线性规划（LP）松弛 [13] 用盒约束$-1\le x\le 1$替代 $\{-1,1\}^n$，随后通过舍入步骤恢复整数解。半正定规划松弛 [14] 则更为强大：对于

前两位作者对本工作的贡献相同。

肖亮海任职于暨南大学信息科学与技术学院，中国广州；钱逸天任职于香港理工大学数据科学与人工智能系，中国香港；潘少华任职于华南理工大学数学学院，中国广州。（电子邮箱：xiaolh@jnu.edu.cn、yitian.qian@polyu.edu.hk、shhpan@scut.edu.cn。通讯作者：钱逸天。）

最大割问题，半正定规划松弛结合随机舍入可达到至少 0.87856 的近似比。后续研究利用割平面不等式 [15]–[17] 和双重非负松弛 [1], [18], [19]强化了半正定规划松弛。这些方法展现了凸松弛的有效性，但其实际性能可能会受到问题结构和舍入方案质量的影响。

另一项重要研究方向采用基于惩罚的方法来施加二元约束。早期研究探讨了光滑惩罚公式 [20]–[22]。随后，研究者提出了基于带均衡约束的数学规划（MPECs） [23], [24] 的策略，并将其应用于二元哈希和子图发现等任务。近年来，研究者开发了增广拉格朗日和惩罚分解框架，用于一般的二元优化 [7]。

尽管取得了这些进展，面向一般非光滑二元优化的通用框架仍相对有限。许多现有方法针对特定的问题类别，尤其是目标函数光滑或为二次函数的问题。惩罚方法 [7], [24] 通常假设目标函数具有可微性，以建立收敛性；而半正定规划（SDP）和双非负等凸松弛在二次目标函数问题上最为有效，因为这类问题的结构可以得到充分利用。当目标 f 非光滑且对 Ax中的扰动高度敏感时，松弛与舍入策略可能会导致解质量显著下降。因此，有必要开发新方法，将半正定规划松弛与精确惩罚技术的优势结合起来，从而在保持可扩展性和理论保证的同时，高效处理非光滑目标函数。

B. 贡献

本文的贡献主要有三点。首先，我们将 [1] 的秩一 DC 惩罚框架扩展到具有一般（可能非光滑）目标函数的 (1) 类问题，从而涵盖了更广泛的应用，例如稳健回归、二元压缩感知和监督哈希。我们发现，当且仅当矩阵

$$X = \begin{bmatrix} 1 & x^\top \\ x & xx^\top \end{bmatrix}$$

为半正定矩阵、秩为一且满足 $\mathrm{diag}(X) = e$时，便有这一性质。基于这一发现，我们可以将 (1) 重新表述为

$$\min_{X \in \mathbb{S}^p_+} \{ f(\mathcal{A}(X) - b) \text{ s.t. } \mathrm{rank}(X) \le 1,\ \mathrm{diag}(X) = e \}, \quad (2)$$

其中 $p := n + 1$和 $\mathcal{A}\colon \mathbb{S}^p \to \mathbb{R}^r$是由 $\mathcal{A}(Z) = \frac{1}{2}\mathcal{A}(z_2 + z_3)$ 对于 $Z = \begin{bmatrix} z_1 & z_2^\top \\ z_3 & z_4 \end{bmatrix}$，且满足$z_1 \in \mathbb{R}$, z_2, $z_3 \in \mathbb{R}^n$, 以及 $z_4 \in \mathbb{R}^{n\times n}$ 所定义的线性算子。我们进一步利用 DC 表示 $\|X\|_* - \|X\| = 0$对秩一约束进行编码，从而得到 DC 约束的半正定规划重构

$$\min_{X \in \mathbb{S}^p_+} \{ f(\mathcal{A}(X) - b), \text{ s.t. } \|X\|_* - \|X\| = 0,\ \mathrm{diag}(X) = e \}. \quad (3)$$

其次，我们提出一种主化–最小化算法（DCRA），该算法利用 f的 Moreau 包络，在子问题中可实现闭式更新，并避免重复进行特征分解。

第三，我们证明，内层迭代方案的最坏情况复杂度为 $O(1/\varepsilon^2)$，以驻点残差衡量；整体惩罚算法则会在有限步内终止。在任何正常终止点，对提升迭代点进行简单的秩一投影，即可得到一个近似可行的二进制解，并获得其最优性间隙的显式可计算上界。

C. 论文结构

本文其余部分安排如下。第二节介绍符号约定和预备知识。第三节构建 DC 惩罚模型及其 Burer–Monteiro分解，并阐明其与提升秩一半定规划的等价性。第四节介绍单位列球面上的平滑MM算法，包括闭式内层更新。第五节给出所提框架的收敛性与复杂度分析，以及带有显式下界间隙的秩一投影保证。第六节报告数值结果，包括参数敏感性研究，以及针对合成问题、二元压缩感知和监督哈希的实验。最后，第七节总结全文。

II. 预备知识

THROUGHOUT 本文使用以下记号和定义。令 $\mathbb{R}^{n_1\times n_2}$ 表示所有实 $n_1 \times n_2$ 矩阵组成的空间，该空间配备迹内积$\langle X, Y\rangle := \mathrm{Tr}(X^\top Y)$及其对应的弗罗贝尼乌斯范数$\|X\|_F := \sqrt{\langle X, X\rangle}$。令 $\mathbb{O}^p$ 表示所有 $p \times p$正交归一矩阵组成的集合。对于正整数 k，记 $[k] :=\{1, 2, \ldots, k\}$。对于矩阵 $X \in \mathbb{R}^{n_1\times n_2}$，分别用 $\|X\|$和 $\|X\|_*$表示其谱范数和核范数。对于指标集 $J \subseteq [n_2]$， X_J 表示 X 中由列 X_j （其中 $j \in J$）组成的子矩阵。以 X 为中心、半径为 $\varepsilon > 0$ 的弗罗贝尼乌斯范数闭球记作 $\mathbb{B}(X, \varepsilon)$。

记为 $\mathbb{S}^p$ 对称矩阵的集合。对于对称矩阵 $X \in \mathbb{S}^p$，我们用 $\mathrm{diag}(X) \in \mathbb{R}^p$ 表示其对角元素组成的向量；对于任意向量 $z \in \mathbb{R}^p$，$\mathrm{Diag}(z) \in \mathbb{S}^p$表示以 z 为对角元素的对角矩阵。我们将 X的特征值记为 $\lambda(X) := (\lambda_1(X) \ldots, \lambda_p(X))$, 并按非增顺序排列；并且， $\Lambda_m(X) :=\mathrm{Diag}(\lambda_1(X) \ldots, \lambda_m(X))$表示由 X的前 m 个特征值构成的对角矩阵。将正交对角化矩阵的集合定义为 $\mathbb{O}(X) := \left\{ P \in \mathbb{O}^p \mid X = P\,\mathrm{Diag}(\lambda(X))P^\top \right\}$。对于 $P \in \mathbb{O}(X)$，我们用 $P_I \in \mathbb{R}^{p\times m}$ 表示 P 中包含其前 m列的子矩阵。单位矩阵、全一向量和全一矩阵分别记为 I、 e和 E，其维度由上下文确定。

给定闭集 $\Delta \subseteq \mathbb{R}^{n_1\times n_2}$，示性函数 δ_Δ 定义为

$$\delta_\Delta(z) := \begin{cases} 0, & \text{if } z \in \Delta, \\ +\infty, & \text{otherwise.} \end{cases}$$

点 $X \in \mathbb{R}^{n_1\times n_2}$ 到 Δ 的距离记为 $\mathrm{dist}(X,\Delta)$，并以弗罗贝尼乌斯范数度量。为方便记号，我们定义：

$$\widetilde{\psi}(V) := \psi(V^\top V),\ \text{for } V\in\mathbb{R}^{m\times p},\ \text{with } \psi(X) := -\|X\|. \quad (4)$$

现在我们回顾 [25]，中正则次微分和（极限）次微分的定义，这些定义将用于分析非光滑函数。

定义 1. 设 $h:\ \mathbb{R}^n \to (-\infty,\ +\infty]$ 和 $x\in\mathbb{R}^n$ 满足 $h(x) < +\infty$。 h 在 x处的正则次微分，记为 $\widehat{\partial}h(x)$，定义为

$$\widehat{\partial}h(x) := \left\{ v\in\mathbb{R}^n \;\middle|\; \liminf_{x'\to x,\, x'\neq x} \frac{h(x') - h(x) - \langle v, x'-x\rangle}{\|x'-x\|} \geq 0 \right\}.$$

h 在 x 处的（极限）次微分定义为

$$\partial h(x) := \left\{ v\in\mathbb{R}^n \;\middle|\; \begin{array}{l} \exists\, x^k \to x \text{ with } h(x^k)\to h(x), \\ v^k \in \widehat{\partial}h(x^k),\ v^k \to v \text{ as } k\to\infty \end{array} \right\}.$$

根据定义可知，$\partial h(x)\subseteq\widehat{\partial}h(x)$，且这两个集合都是闭的；$\partial h(x)$也具有凸性。当 h 为凸函数时，这两个次微分都退化为通常的凸次微分。若点 $\bar{x}$ 满足 $0\in\partial h(\bar{x})$，则称其为 h 的临界点；所有此类点组成的集合记为 crit h。对于示性函数 $h=\delta_\Delta$，这两个次微分分别退化为正则法锥$\widehat{\mathcal{N}}_\Delta(x)$和法锥 $\mathcal{N}_\Delta(x)$。

III. 重构

THE 问题（3）的可行集仍具有组合性质，因为它由具有二元结构的秩一半正定（PSD）矩阵构成。值得注意的是，秩约束可以写成差分凸约束 $\|X\|_* - \|X\| = 0$，但处理差分凸约束通常比处理差分凸目标函数更困难。为解决这一困难，我们考虑将差分凸约束并入目标函数的惩罚重构（3）

$$\min_{X\in\Omega} f\big(\mathcal{A}(X)-b\big) + \rho\big(\|X\|_* - \|X\|\big), \quad (5)$$

其中 $\rho>0$ 是惩罚参数，且 $\Omega := \{X\in\mathbb{S}^p_+ \mid \mathrm{diag}(X) = e\}$.

一个关键事实是，惩罚模型（5）是问题（3）的精确惩罚重构。也就是说，存在 $\rho^* = (1+2p)L_f$，使得对于所有 $\rho\geq\rho^*$，根据 [27, 命题2.3和定理3.1]，问题（5）的全局极小点与问题（3）的全局极小点一致，因此也与秩一半正定规划（2）的解一致。此外，对于任意 $\rho>\rho^*$，问题（3）的每个可行点 $X\in\mathcal{F}$ 不仅是问题（3）的局部最优点，也是问题（5）的严格局部极小点（见 [1,命题1(ii)]）。

然而，在实践中，求解问题 (5) 并非易事。即使已知阈值 ρ^*，通常也无法期望在固定的 $\rho\geq\rho^*$ 下获得全局最优解。此外，现有的凸松弛方法在每次迭代时都需要进行完整的特征值分解。这造成了严重的计算瓶颈，阻碍了方法向大规模场景扩展。为克服这一问题，我们采用矩阵分解范式 [21], [28],，该范式在低秩优化和半正定松弛领域重新受到关注（例如，[29], [30]）。具体而言，我们引入如下惩罚因子化形式

$$\min_{V\in\mathcal{S}} \Phi(V) := f\big(\mathcal{A}(V^\top V)-b\big) + \rho\left(\|V\|_F^2 - \|V^\top V\|\right), \quad (6)$$

其中 $\mathcal{S}:=\{V\in\mathbb{R}^{m\times p} \mid \|V_j\|=1 \text{ 对于 } j=1,\ldots,p\}$，$2\leq m<p$，且 V_j 表示 V 的第 j 列。惩罚项$\|V\|_F^2-\|V\|^2$对应于原始秩一正则项的替代项，而按列归一化确保 $\mathrm{diag}(X) = e$。我们要求 $m\geq 2$，因为当 $m=1$时惩罚项为零。尽管对于所有 $V\in\mathcal{S}$，$\|V\|_F^2 = p$，我们仍保留该项，以显式表示对秩替代项的惩罚（$\|X\|_* - \|X\|$）。

根据 [1, 第 3], 节，若 $V^*\in\mathcal{S}$ 是（6）的全局最优解，则$X^* := V^{*\top}V^*$是满足秩$(X^*)\leq m$的（5）的全局最优解。反之，给定（5）的任意全局极小解 X^*，且其秩为 r^*，令 $V^* := \sqrt{\Lambda_m(X^*)}P_I^{*\top}$，即可得到（6）的一个全局最优解。为方便起见，我们将问题（6）称为 *DC* 惩罚矩阵规划，尽管由于可行集 $\mathcal{S}$的非凸性，它并不是真正的 DC 规划。

IV. 算法框架

为 处理（3）中目标函数的非光滑性，我们考虑对其进行 Moreau 包络平滑。对于适当的下半连续函数 $f:\ \mathbb{R}^q\to(-\infty,+\infty]$，其参数为 $\delta>0$ 的 Moreau 包络及相应的近端映射定义为

$$\mathrm{env}_{\delta f}(x) := \min_{y\in\mathbb{R}^q}\left\{ f(y) + \frac{1}{2\delta}\|x-y\|^2 \right\}, \quad \text{and}, \quad (7)$$

$$\mathrm{prox}_{\delta f}(x) := \arg\min_{y\in\mathbb{R}^q}\left\{ f(y) + \frac{1}{2\delta}\|x-y\|^2 \right\}, \quad (8)$$

分别。如果 f是凸的，那么$\mathrm{env}_{\delta f}$连续可微，其梯度由 $\nabla \mathrm{env}_{\delta f}(x) = \frac{1}{\delta}\left(x - \mathrm{prox}_{\delta f}(x)\right)$给出。利用这种平滑技术，我们考虑以下平滑问题：

$$\min_{V\in\mathcal{S}} \Phi_\delta(V) := \widetilde{f}(V) + \rho\left(\|V\|_F^2 + \widetilde{\psi}(V)\right), \quad (9)$$

其中 $\widetilde{f}(V) := \mathrm{env}_{\delta f}(\mathcal{A}(V^\top V)-b)$，且 $\widetilde{\psi}(V)$在式（4）中定义。我们提出以下平滑松弛方法：针对不断增大的 ρ，寻找与之相关的有限个驻点（9）。

算法1 DC R基于模型（9）的A算法（DCRA）

1: 初始化：选择适当的整数 $m\in[2,p]$、较小的 $\epsilon\in(0,1)$以及足够大的 $k_{\max}\in\mathbb{N}$。选取 $\rho_{\max}>0, \rho_0>0, \sigma>1, \delta>0$，以及一个满秩的初始点 $V^0\in\mathcal{S}$。2: **对** $k=0,1,2,\ldots,k_{\max}$ **执行**3: 从 $V^k\in\mathcal{S}$出发，寻找以下问题的近似驻点V^{k+1} : 4: 5:

$$\min_{V\in\mathcal{S}} \Phi_k(V) := \widetilde{f}(V) + \rho_k\left(\|V\|_F^2 + \widetilde{\psi}(V)\right), \quad (10)$$

如果满足 $\|V^{k+1}\|_F^2 - \|V^{k+1}\|^2 \leq \epsilon$，则停止。6: 否则，令 $\rho_{k+1}\leftarrow$ 最小值$\{\sigma\rho_k, \rho_{\max}\}$。7: **结束循环**

接下来，我们考虑子问题 (10)。

定义 2。给定 $\epsilon > 0$，如果存在 ( )，使得对于某个 (,) 满足以下条件，则点 $V^\star \in \mathcal{S}$ 是 (ϵ)-驻点 ˜点 (

10 $\Gamma^\star \in \partial\psi\ U$ $U \in \mathcal{S} \cap$)

$\mathbb{B}\ V\ \epsilon$

$$\mathcal{G}_k(V^\star) := \operatorname{dist}\Big(0,\ \nabla\widetilde{f}(V^\star) + 2\rho_k V^\star + \rho_k\Gamma^\star + N_{\mathcal{S}}(V^\star)\Big) \le \epsilon.$$

˜为计算（10）的一个驻点，我们在当前迭代点 $V^{k,l}$处对 f 和 $\widetilde{\psi}$ 进行线性化，得到以下子问题：

$$\begin{aligned}\min_{V\in\mathcal{S}}\ \widetilde{\Phi}_k(V;V^{k,l}) := \widetilde{f}(V^{k,l}) &+ \langle\nabla\widetilde{f}(V^{k,l}), V - V^{k,l}\rangle \\ &+ \rho_k\left[\|V\|_F^2 + \widetilde{\psi}(V^{k,l}) + \langle\Gamma^{k,l}, V - V^{k,l}\rangle\right] \\ &+ \frac{L}{2}\|V - V^{k,l}\|_F^2. \end{aligned} \tag{11}$$

其中 $\Gamma^{k,l} \in \partial\widetilde{\psi}(V^{k,l})$ 和 $L \ge L_{\widetilde{f}}$。该子问题有闭式解：

$$V^{k,l+1} := \operatorname{proj}_{\mathcal{S}}\left(\frac{1}{2\rho_k + L}\left(LV^{k,l} - \nabla\widetilde{f}(V^{k,l}) - \rho_k\Gamma^{k,l}\right)\right), \tag{12}$$

其中 $\operatorname{proj}_{\mathcal{S}}(U) :=$normalize$(U)$ 按列投影到单位球面上。

算法2 子问题（10）的求解器

Input: current outer iterate $V^{k,0} := V^k \in \mathcal{S}$, $\rho_k > 0$, tolerance $\epsilon_v > 0$, and $l_{\max}$
for $l = 0, 1, 2, \ldots, l_{\max}$ **do**
  Compute $\Gamma^{k,l} \in \partial\widetilde{\psi}(V^{k,l})$.
  Compute:
  $\nabla\widetilde{f}(V^{k,l}) = 2V^{k,l}\mathcal{A}^*\nabla\mathrm{env}_{\delta f}(\mathcal{A}((V^{k,l})^\top V^{k,l}) - b)$.
  Update $V^{k,l+1}$ via (12).
  if $\|V^{k,l+1} - V^{k,l}\|_F \le \epsilon_v$, **then break**
  end if
end for
Output: Approximate solution $V^{k+1} := V^{k,l+1}$.

根据引理1，第 4], 节中我们可以取

$$\Gamma^{k,l} = \Gamma(V^{k,l}) := -2V^{k,l}P_1(V^{k,l})P_1(V^{k,l})^\top,$$

其中， $P_1(V)$是 $V^\top V$的主特征向量，可通过部分奇异值分解获得。由于$V^{k,\ l}$ 的行数较少，因此计算成本很低。

IN V. 收敛性分析

本节将给出我们方法的收敛性结果。定义残差

$$\mathcal{G}(V,\Gamma) := \operatorname{dist}\big(0,\ \nabla\tilde{f}(V) + 2\rho_k V + \rho_k\Gamma + N_{\mathcal{S}}(V)\big).$$

在我们的算法背景下，我们将使用 $\Gamma = \Gamma^{k,\ l} \in \partial\tilde{\psi}(V^{k,\ l})$，并在 $V^{k,\ l+1}$处评估残差。

命题 1. 设 $\{V^{k,l}\}_{l\in\mathbb{N}} \subset \mathcal{S}$ 是固定 $k \in \mathbb{N}$时由算法2生成的序列。则以下陈述成立：

(i) 对所有 $l \ge 0$,

$$\Phi_k(V^{k,l}) - \Phi_k(V^{k,l+1}) \ \ge\ \tfrac{L-L_{\tilde{f}}}{2}\|V^{k,l+1} - V^{k,l}\|_F^2. \tag{13}$$

(ii) 序列 $\{V^{k,l}\} \subset \mathcal{S}$ 有界，且 $\{\Phi_k(V^{k,l})\}$单调不增并收敛。

(iii) 对于所有 $l \in \mathbb{N}$，存在 $C > 0$ 使得

$$\mathcal{G}(V^{k,l+1}, \Gamma^{k,l}) \le C\|V^{k,l+1} - V^{k,l}\|_F. \tag{14}$$

定理1. 设 $\Phi_\delta^\star := \min_{V\in\mathcal{S}}\Phi_\delta(V)$ 为平滑问题 (9) 的最优值。固定外层迭代 $k \in \mathbb{N}$。对于任意满足 $T \in \mathbb{N}$ 的 $l \in [T]$，算法2生成的内层迭代点满足

$$\min_{0\le l<T}\mathcal{G}(V^{k,l+1}, \Gamma^{k,l}) \ \le\ \frac{C}{\sqrt{T}}\sqrt{\frac{2\big(\Phi_k(V^k) - \Phi_\delta^\star\big)}{L - L_{\tilde{f}}}}. \tag{15}$$

特别地，若 $T \ge \left\lceil \frac{2C^2}{L-L_{\tilde{f}}}\frac{\Phi_k(V^k)_{\epsilon^2}}{-\Phi_\delta^\star}\right\rceil$，则内层循环在 $O(\epsilon^{-2})$ 次迭代内达到一个 ϵ-驻点，因此算法2会在有限次内层迭代后终止。

接下来，我们证明算法1能够正常终止。引理1确保在适当条件下，差距 $\|V^{k,\ 1}\|_F^2 - \|V^{k,\ 1}\|$ 足够小，而命题2进而保证算法1能够正常退出。尽管本部分的分析沿用了 [1], 的技术，但我们得到的结果有所不同。

引理1. 固定 $k \in \mathbb{N}$。假设存在 $P \in \mathbb{O}\big((V^k)^\top V^k\big)$，使得其主特征向量 P_1 没有零元素。定义 $\mu :=$min $_{1\le j\le p}\ |P_{j1}|\ _> 0.$。令 $\varpi := (L + 2L_{\tilde{f}})$ $\sqrt{p} + \big\|\nabla\widetilde{f}\big(\frac{1}{\sqrt{m}}E\big)\big\|_F$.。取 $\rho_k \ge \rho := \frac{Lp+\varpi}{\mu\, c_0}$，其中 $c_0 \in (0,\ 1)$，则由算法2生成的第一个内层迭代点 $V^{k,\ 1}$满足

$$\|V^{k,1}\|_F^2 - \|V^{k,1}\|^2 \ \le\ c_0^2.$$

命题2. 固定 $k \in \mathbb{N}$，并设 $\{V^{k,l}\}_{l\ge0} \subset \mathcal{S}$ 是算法2在曲率 $L \ge L_{\tilde{f}}$ 和 $V^{k,\ 0} = V^k$ 下生成的内层序列。令 $\varepsilon \in (0,\ c_0^2]$ 为一个容差，其中 $c_0 \in (0,\ 1)$ 与引理 1 $\varpi_{\text{lin}} = 4L_{\tilde{f}}p +$中的常数相 $2\sqrt{p}\left\|\nabla\widetilde{f}\big(\frac{1}{\sqrt{m}}E\big)\right\|_F + \frac{5}{2}Lp\ \eta\ c_0\ p$ 同。定义 : , (,) : $= 2\Big(1 - \frac{1}{p}\Big)\sqrt{p - c_0^2} +$, $\sqrt{2}\sqrt{1 - c_0^2}$ $\widehat{\rho}$ = $\left\{\frac{2\varpi_{\text{lin}}}{\eta\ c_0\ p}\ \ \frac{2\varpi_{\text{lin}}}{\varepsilon}\right\}$. 并设 : max (,) 则有以下结论：

(i) 若 $\rho_k \ge \widehat{\rho}$ $l \in \mathbb{N}$ $\frac{\varepsilon}{2} \le p -$，则对于每个满足的，有 $\|V^{k,l}\|^2 \le c_0^2$

$$\|V^{k,l+1}\|^2 \ \ge\ \|V^{k,l}\|^2 + \tfrac{1}{2}\eta(c_0, p). \tag{16}$$

(ii) 假设存在 $P \in \mathbb{O}\big((V^{k,0})^\top V^{k,0}\big)$，其主特征向量 P_1 没有零元素，并且若 $\rho_k >$max$\{\widehat{\rho}, \rho\}$，且满足引理 (1) 中的 $\rho := \frac{Lp+\varpi}{\mu\, c_0}$，则有

exists $1 \le \bar{l} \le \left\lceil \frac{2c_0^2 - \varepsilon}{\eta(c_0,p)}\right\rceil + 1$ such that for all $l \ge \bar{l}$,

$$\|V^{k,l}\|_F^2 - \|V^{k,l}\|^2 \ =\ p - \|V^{k,l}\|^2 \ \le\ \varepsilon.$$

以下推论可由定理1和命题2轻易得出。

推论1. 假设存在 $P \in \mathbb{O}\big((V^{k,0})^\top V^{k,0}\big)$其主特征向量 P_1 没有零元素，且 ε_v由算法2给出。算法1执行的迭代总次数不超过 $\mathcal{O}\Big(\frac{\log(\rho_\star/\rho_0)}{\varepsilon_v^2}\Big)$，其中 $\rho_\star :=$ max$\{\widehat{\rho}, \rho\}$，而 $\widehat{\rho}$和 ρ分别由命题2和引理1给出。

定理 2指出，算法1的正常输出的秩一投影近似可行且近乎最优。

定理 2. 设 k 是算法1正常终止时的外层索引，此时子问题由算法2求解。设 $X^k = (V^k)^\top V^k$，且 $P \in \mathbb{O}(X^k)$ 对角化 X^k。定义秩一投影 $x^k := \|V^k\| P_1$，其中 P_1 是 X^k 的主特征向量。记 $\hat{} f(X) := \hat{} \mathrm{env}_\delta f(\mathcal{A}(X) - b)$。假设 f 在 Ω 上是利普希茨连续的，模数为 $L_{\hat{f}} > 0$。那么，对于任意 $k^* \in \{0, 1, \ldots, k-1\}$，以下结论成立：

$$\|x^{\bar{k}} \circ x^{\bar{k}} - e\| \le \varepsilon, \tag{17}$$

$$\hat{f}\big(x^{\bar{k}}(x^{\bar{k}})^\top\big) - \widetilde{f}(V^{k^*}) \le \rho_{\bar{k}} \|V^{\bar{k}}\|^2 - \rho_{k^*} \frac{p}{r^*} + \sum_{j=k^*}^{\bar{k}-1} (\rho_j - \rho_{j+1}) \|V^{j+1}\|^2 + L_{\hat{f}}\, \varepsilon, \tag{18}$$

其中 $r^* := \text{秩}(V^{k^*})$。特别地，若对于原问题的最优值 υ^*，有 $\widetilde{f}(V^{k^*}) = \hat{} f\big((V^{k^*})^\top V^{k^*}\big) \le \upsilon^*$，则

$$\hat{f}\big(x^{\bar{k}}(x^{\bar{k}})^\top\big) - \upsilon^* \le \rho_{\bar{k}}(p-1) + \rho_{k^*}\left(1 - \frac{p}{r^*}\right) + (L_{\hat{f}} - \rho_{\bar{k}})\epsilon.$$

VI. 数值实验

为评估所提出的算法在合成数据集和真实世界数据集上的性能，并将其与几种基线方法进行比较。所有实验均使用 MATLAB 实现，并在一台运行 64 位 Windows 的工作站上执行，该工作站配备 Intel(R) Core(TM) i9-12900 CPU @ 2.40 GHz 和 64 GB 随机存取存储器。所有算法的停止准则均为以下两者之一： (i) 达到最大迭代次数，或 (ii) 目标值的变化满足预设容差，以先发生者为准。除非另有说明，算法参数均设为默认值。

A. 参数敏感性分析

由于算法1的性能可能受以下因素影响：(i) 初始化矩阵 V_0中的列数 m，(ii) 初始惩罚参数 ρ，以及 (iii) 子问题构造中的平滑参数 δ，因此我们进行了实验，以考察其性能对这些关键参数的敏感性。对于每个参数，我们考虑两种问题规模：$(n,\ r) = (500,\ 300)$ 和 $(n,\ r) = (1000,\ 600)$。结果取自 50 次独立随机试验的平均值。

对于 m，我们在对数尺度上将其值从10^0 变化到 10^3。对于两种问题规模，随着 m 从较小值增大，目标值迅速改善；当 m 超过约 1–10 后，便趋于稳定。当 $m > 10$ 时，较大问题 $(n,\ r)$ 的目标值稳定性较差 $= (1000,\ 600)$。较小的 m 对应的运行时间增长缓慢，而较大的 m 则会使运行时间急剧增加。

对于 ρ，适中的取值能够以相对较低的计算时间获得稳定且高质量的解。非常小

ρ会导致不稳定的目标值和更长的运行时间，而过大的 ρ 也可能降低性能。

对于 δ，在较宽的取值范围内，性能保持稳定。过大的 δ 可以缩短计算时间，但通常会以牺牲解质量为代价。

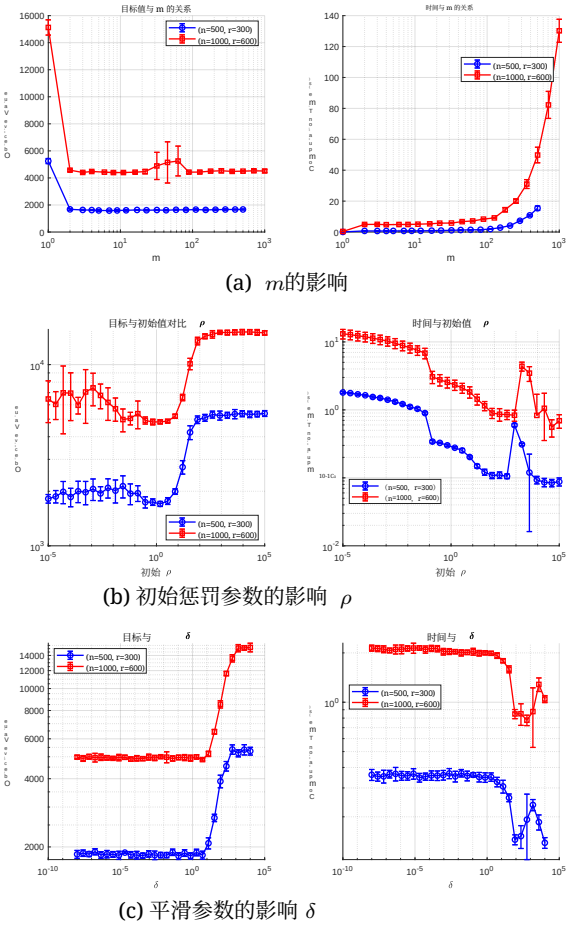


图 1：所提出的算法在两种问题规模下的参数敏感性分析 $(n,\ r) = (500,\ 300)$ 和 $(n,\ r) = (1000,\ 600)$。对于每种参数设置，结果均为 10 次试验的平均值，误差条表示一个标准差。左侧面板：目标值；右侧面板：运行时间。适用时，两个坐标轴均采用对数刻度。

B. 随机生成问题的性能

我们首先在随机生成的实例上评估所提出的算法，以考察其性能和鲁棒性。目标函数为

$$\min_x \|Ax - b\|_1, \tag{19}$$

其中 $A \in \mathbb{R}^{n \times r}$ 且 $b \in \mathbb{R}^n$。对于每种问题规模，我们随机生成 100 个独立实例，其中 A 和 b的每个元素均从标准正态分布中抽取。我们测试的问题规模为 $n \in [100, 3000]$ 和 $d \in [50, 2000]$。DCRA 的参数固定为 $\sigma = 1.2$、$m = 5$和$\rho = 1$。

我们将 DCRA 与三种基线方法进行比较：Gurobi 求解器、MPEC-EPM 方法 [24], 和 L2-boxADMM 方法[34]。通过以下四项指标评估性能：

- DCRA 的胜率：给定方法取得的目标值优于其他方法的实例所占的比例；
- 平均相对差：以下表达式的平均值 $(\mathrm{Obj}_{\mathrm{DCRA}}-\mathrm{Obj}_{\cdot})/\mathrm{Obj}_{\cdot}$，其中负值表示 DCRA 得到的目标值更小；
- 平均目标值：$\|Ax-b\|_1$ 的平均值；以及
- 平均CPU时间：100次运行的平均计算时间。

为了将 Gurobi 应用于我们的二进制优化模型 (19)，我们将其重新表述为一个等价的混合整数线性规划（混合整数线性规划）：

$$\begin{aligned}\min_{x,z}\quad & e^\top z\\ \text{s.t.}\quad & z\ge -(A(2x-1)-b),\\ & z\ge A(2x-1)-b,\\ & x\in\{0,1\}^n, z\in\mathbb{R}^r.\end{aligned}$$

这种表述使我们能够利用 Gurobi 的分支定界框架获得精确的二进制解。

表一报告了结果。在所有测试的问题规模下，DCRA 相对于全部三种基线方法的胜率都接近 100%，表明其目标值始终优于其他方法。所有情况下的平均相对差异均为负值；与其他方法相比，DCRA 的目标值平均低 53%。

在运行时间方面，DCRA 显著快于 Gurobi，但慢于 MPEC-EPM 和 L2-box ADMM。尽管 MPEC-EPM 是所有方法中最快的，但其目标值始终最差。

C. 二元压缩感知的性能 g

二元压缩感知（BCS）数据集通过模拟线性测量模型生成

$$y=Ax_0,$$

其中，$x_0\in\{0,1\}^N$ 是一个稀疏二进制向量，$A\in\mathbb{R}^{M\times N}$ 是一个随机感知矩阵。对于给定的信号维度 N，按照 [35], , 我们选择压缩率 $\alpha=M/N$、稀疏率 ρ以及控制矩阵每个元素均值的偏置参数 μ。A 的每个元素均独立地从高斯分布 $A_{ij}\sim\mathcal{N}\left(\frac{\mu}{N},\frac{1}{N}\right)$, 中抽取，使得各列的期望范数为 1 量级。真实信号 x_0 由伯努利过程生成，其中 $\Pr\{x_{0i}=1\}=\rho$，从而确保约有 ρN 个活跃分量。

通过改变 α、ρ 和 μ，可以构造出压缩率和稀疏度不同的一系列实例，正如 中所做的那样。每个实例均由 (N,α,ρ,μ) 和随机种子完全确定。二元压缩感知的原始形式为

$$\min_{x\in\{0,1\}^N}\|Ax-y\|_1+\lambda\|x\|_1, \tag{20}$$

其中，$A\in\mathbb{R}^{M\times N}$ 是感知矩阵，$y\in\mathbb{R}^M$ 是测量向量，$\lambda>0$ 是正则化参数。项 $\|x\|_1$ 促进稀疏性，因为 $x\in\{0,1\}^N$。

由于我们的算法是针对$\{-1,1\}$中的二元变量设计的，我们采用变量变换 $x=(z+e)/2$，其中 $z\in\{-1,1\}^N$。代入（20）后，得到一个等价问题：

$$\min_{z\in\{-1,1\}^N}\left\|\tfrac{1}{2}Az-\left(y-\tfrac{1}{2}Ae\right)\right\|_1+\tfrac{\lambda}{2}e^\top z. \tag{21}$$

定义 $\tilde{A}=\frac{1}{2}A$、$\tilde{b}=y-\frac{1}{2}Ae$ 和 $c=(\lambda/2)e$，其紧凑形式为

$$\min_{z\in\{-1,1\}^N}\|\tilde{A}z-\tilde{b}\|_1+c^\top z.$$

线性项 $c_z^\top z$ 通常无法仅通过 $u=A'z-b'$ 表示。为了符合 $\min_z f(Az-b)$ 的模板，我们引入一种增广形式：

$$\bar{A}=\begin{bmatrix}A'\\ I_N\end{bmatrix},\quad \bar{b}=\begin{bmatrix}b'\\ 0\end{bmatrix},\quad \bar{u}=\bar{A}z-\bar{b}=\begin{bmatrix}A'z-b'\\ z\end{bmatrix}. \tag{22}$$

接下来定义

$$\bar{f}(\bar{u})=\|\bar{u}_{1:M}\|_1+c_z^\top\bar{u}_{M+1:M+N}. \tag{23}$$

通过这一构造，

$$\bar{f}(\bar{A}z-\bar{b})=\|A'z-b'\|_1+c_z^\top z.$$

增广损失 f 可按两个分块分离。在前 M 个 f 坐标上，ℓ_1化为 ℓ_1-范数。其近端算子为软阈值化：

$$\mathrm{prox}_{\gamma\|\cdot\|_1}(x)=\mathrm{sign}(x)\cdot\max\{|x|-\gamma,0\}.$$

相应的 Moreau 包络为

$$\mathrm{env}_{\gamma\|\cdot\|_1}(x)=\sum_{i=1}^{M}\begin{cases}\frac{1}{2\gamma}x_i^2, & |x_i|\le\gamma,\\ |x_i|-\frac{\gamma}{2}, & |x_i|>\gamma.\end{cases}$$

在最后 N 个坐标上，$\bar{f}$ 是线性的：$c_z^\top u$。其近端算子是一个简单的平移：

$$\mathrm{prox}_{\gamma\langle c_z,\cdot\rangle}(y)=y-\gamma c_z,$$

而 Moreau 包络为

$$\mathrm{env}_{\gamma\langle c_z,\cdot\rangle}(y)=\langle c_z,y\rangle-\tfrac{\gamma}{2}\|c_z\|_2^2.$$

因此，结合这两个模块即可高效计算 f 的近端算子和包络。

为求解 (21)，我们比较以下算法：MPEC-EPM、DCRA、MPEC-ADM 和 GUROBI。每个求解器都会生成一个向量 z，我们通过 $x=(z+e)/2$将其映射回 $\{0,1\}^N$。为进行评估，我们计算原始目标函数值 (20)、均方误差 (MSE) $\|x-x_0\|_2^2/N$、x 与 x_0之间的汉明距离，以及实际运行时间。

为了使用 Gurobi 求解 (20)，我们还需要将其重新表述为一个等价的混合整数线性规划：

$$\begin{aligned}\min_{x,z}\quad & e^\top z+\lambda e^\top x\\ \text{s.t.}\quad & z\ge -(Ax-b),\\ & z\ge Ax-b,\\ & x\in\{0,1\}^N, z\in\mathbb{R}^M.\end{aligned}$$

图2展示了四种求解器在三个偏置水平（$\mu=0$、2、10）下，目标值随参数 ρ和 α 变化的相图。表 II 报告了四种求解器在不同偏置 μ 和稀疏度 ρ下求解二进制压缩感知问题的平均运行时间和目标值。DCRA 始终取得最低的目标值，证实了其更强的优化能力，而 MPEC-EPM 和 MPEC-ADM 的运行速度最快。

表一：随机生成问题的性能比较

n	d	DCRA 的胜率，对比对象：			平均CPU时间			
		Gurobi	MPEC-EPM	L2-box ADMM	DCRA	Gurobi	MPEC-EPM	L2-box ADMM
100	50	100%	100%	100%	0.0170	0.0737	0.0031	0.0060
100	100	100%	100%	100%	0.0145	0.0807	0.0034	0.0063
100	200	100%	100%	100%	0.0203	0.1116	0.0054	0.0090
200	200	100%	100%	100%	0.0340	0.1638	0.0075	0.0112
300	300	100%	100%	100%	0.0551	0.3080	0.0114	0.0150
300	500	100%	100%	100%	0.1218	0.6111	0.0151	0.0183
500	500	100%	100%	100%	0.9115	2.7283	0.1432	0.1552
300	1000	100%	100%	100%	0.2258	1.9576	0.0215	0.0249
500	1000	100%	100%	100%	1.2865	8.5689	0.2341	0.2424
1000	1000	100%	100%	100%	1.8787	5.4533	0.1957	0.1989
300	2000	100%	100%	100%	0.4629	6.8323	0.0635	0.0720
500	2000	100%	100%	100%	1.0622	12.8283	0.1595	0.1609
1000	2000	100%	100%	100%	3.1180	19.6319	0.4112	0.4206
2000	2000	100%	100%	99%	7.3039	29.8193	1.4853	1.6716
3000	2000	100%	100%	88%	13.4699	37.2612	3.6330	3.7831

n	d	平均相对差异			平均目标值			
		Gurobi	MPEC-EPM	L2-box ADMM	DCRA	Gurobi	MPEC-EPM	L2-box ADMM
100	50	-64%	-77%	-59%	1.44^{02}	3.96^{02}	6.33^{02}	3.53^{02}
100	100	-59%	-73%	-56%	3.30^{02}	8.06^{02}	1.21^{03}	7.54^{02}
100	200	-47%	-62%	-60%	8.50^{02}	1.60^{03}	2.27^{03}	2.12^{03}
200	200	-61%	-74%	-51%	8.89^{02}	2.27^{03}	3.46^{03}	1.81^{03}
300	300	-61%	-75%	-45%	1.62^{03}	4.20^{03}	6.37^{03}	2.97^{03}
300	500	-51%	-66%	-51%	3.43^{03}	6.94^{03}	1.00^{04}	7.02^{03}
500	500	-61%	-74%	-36%	3.52^{03}	8.98^{03}	1.38^{04}	5.54^{03}
300	1000	-35%	-52%	-51%	9.00^{03}	1.38^{04}	1.87^{04}	1.85^{04}
500	1000	-46%	-62%	-48%	9.60^{03}	1.79^{04}	2.54^{04}	1.85^{04}
1000	1000	-60%	-74%	-22%	1.01^{04}	2.52^{04}	3.88^{04}	1.29^{04}
300	2000	-22%	-38%	-38%	2.16^{04}	2.77^{04}	3.47^{04}	3.48^{04}
500	2000	-31%	-48%	-47%	2.46^{04}	3.57^{04}	4.73^{04}	4.67^{04}
1000	2000	-46%	-62%	-37%	2.72^{04}	5.04^{04}	7.20^{04}	4.33^{04}
2000	2000	-60%	-74%	-11%	2.87^{04}	7.15^{04}	1.10^{05}	3.22^{04}
3000	2000	-67%	-79%	-4%	2.92^{04}	8.75^{04}	1.40^{05}	3.06^{04}

注：上标 k 表示 10^k的尺度。

D. 监督哈希代码学习的性能

在本实验中，我们评估了所提出方法在监督离散哈希问题上的性能。该问题是基于哈希的检索系统中的一项关键任务。监督哈希模型遵循 [7], [9], [36] 中的形式，并采用平均绝对误差（MAE）损失:

$$\min_{X\in\{-1,1\}^{n\times r},\, W\in\mathbb{R}^{d\times r}} \left\|B - WX^\top\right\|_1 + \frac{\delta}{2}\|W\|_F^2, \qquad (24)$$

其中， $B \in \mathbb{R}^{d\times n}$ 是数据矩阵， $W \in \mathbb{R}^{d\times r}$ 是实值投影矩阵， $X \in \{-1,1\}^{n\times r}$ 是二进制哈希码矩阵。由于对 X 施加了二进制约束，且使用了非光滑的 ℓ_1-范数，该目标函数是非凸的。

我们在三个基准数据集上进行了实验：MNIST[1],CIFAR-10[2], 和MIRFLICKR-25K[3]。

为求解问题 (24)，我们采用交替最小化方案。具体而言，对于 W-子问题，我们固定二进制矩阵 X ，并使用梯度下降法最小化目标函数的平滑版本:

$$\min_{W\in\mathbb{R}^{d\times r}} h_\mu\left(B - W(X^{(k)})^\top\right) + \frac{\delta}{2}\|W\|_F^2, \qquad (25)$$

其中， $h_\mu(\cdot)$表示参数为$\mu > 0$的 Huber 损失函数。

令 $R = B - WX^\top$，平滑后的˜目标的梯度为 $\nabla_W f(W) = -H'_\mu(R)X + \delta W$，其中$H'_\mu$ 逐元素应用 Huber 导数。

我们使用固定步长 $\alpha = 1/L$，其中 Lipschitz 常数估计为 $L = \frac{nr}{\mu} + \delta.$。对于 X-子问题，我们固定 W 并求解:

$$\min_{X\in\{-1,1\}^{n\times r}} \left\|B - WX^\top\right\|_1,$$

使用的算法包括 DCRA、MPEC-EDM、 L_2-box ADMM 和次梯度下降法。MPEC-EDM 和 L_2-box ADMM 的参数采用默认设置。

我们沿用 [7], [36], [37]中的实现设置，并使用线性哈希函数将查询编码为二元码。我们使用三种标准指标评估各方法的性能：平均精度均值（MAP）、精确率–召回率（PR）曲线以及归一化折损累计增益（NDCG）。MAP 衡量平均精确率

[1]MNIST 数据集：https://pjreddie.com/projects/mnist-in-csv [2]CIFAR-10 数据集：https://www.cs.toronto.edu/~kriz/cifar.html [3]MIRFLICKR-25K 数据集：https://press.liacs.nl/mirflickr/mirdownload. html

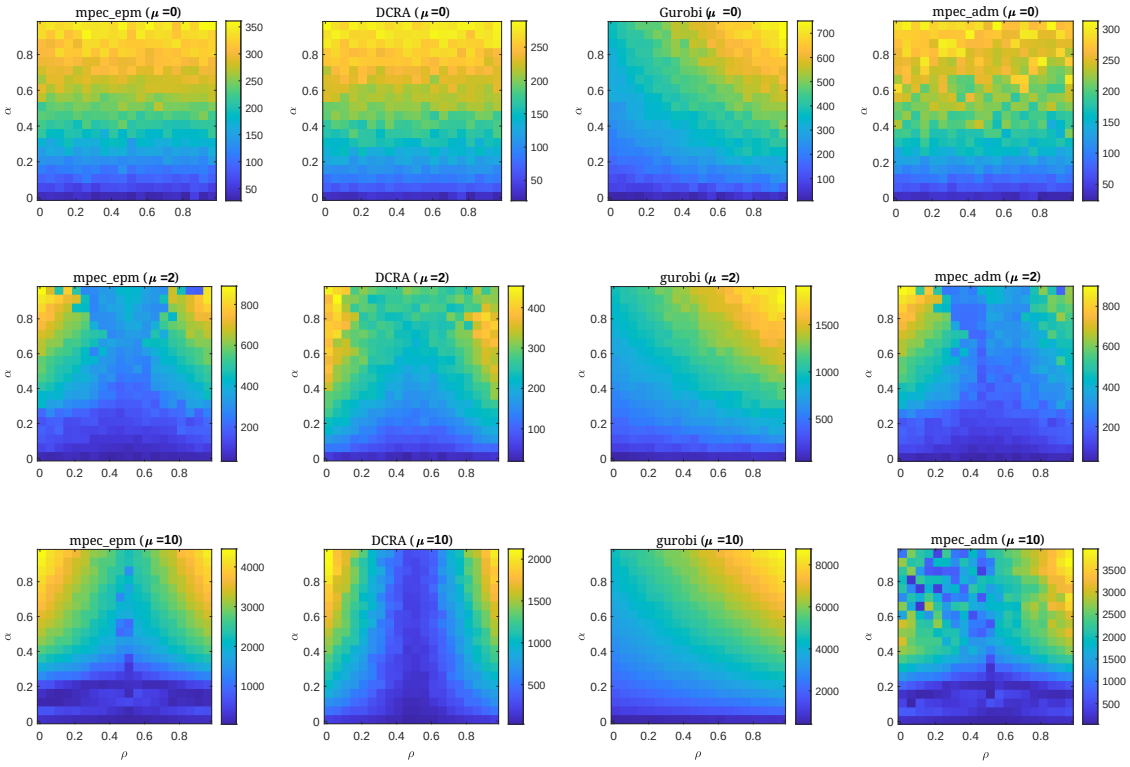


图2： $\mu = 0$，2，10 个案例的相图。

算法3 平滑监督哈希的交替最小化

1: **输入：** 数据矩阵 $B \in \mathbb{R}^{d\times n}$，正则化参数$\delta > 0$，初始矩阵 $W^{(0)} \in \mathbb{R}^{d\times r}$, $X^{(0)} \in \{-1,1\}^{n\times r}$2: **对于** $k = 0$ 到 $K-1$ **执行**3: 固定 $X^{(k)}$，通过使用步长 $\alpha = \frac{1}{\frac{nr}{\mu}+\delta}$ 的梯度下降求解 (25)，更新 $W^{(k+1)}$ 。4: 固定 $W^{(k+1)}$，通过求解以下问题更新 $X^{(k+1)}$： 5: **结束循环**6: **输出：** 最终哈希码 $X^{(K)}$ 和投影 $\min_{X\in\{-1,1\}^{n\times r}} \left\| B - W^{(k+1)} X^{\top} \right\|_1$.

每个查询的检索项目，并提供排序质量的单值概述。精确率-召回率曲线展示了不同检索阈值下精确率与召回率之间的权衡。归一化折损累计增益通过衡量检索项目的累计增益（按其秩位置折损），并将其归一化为理想排序增益，来评估排序质量。这些指标全面评估了检索性能。

尽管与其他方法相比，DCRA 的计算成本更高（见图 8），但在所有其他质量指标上，它始终表现更优。如图 3 所示，在不同检索深度下，DCRA 均保持了具有竞争力的精确率。16 位和 32 位配置的精确率-召回率曲线分别见图 4 和图 5。图 6 进一步表明，在所有测试数据集上、16 至 96 位范围内，DCRA 的 MAP 得分均更优。在 NDCG 指标方面（图 7），DCRA 在所有数据集上的性能也相当或更优。

VII. 结论

我们提出了一种新颖的矩阵松弛方法，用于求解非光滑二元优化问题，重点关注绝对值回归。我们从秩约束半正定规划重构出发，为秩约束引入了差分凸代理函数，并推导出一种精确惩罚形式，使得能够在光滑流形上进行连续优化。为高效求解由此得到的惩罚问题，我们提出了一种基于矩阵分解的算法（DCRA），并为其提供了可证明的收敛保证。我们的理论分析确立了惩罚形式与原始形式的驻点之间的对应关系，并证明了在温和的谱间隙条件下内层迭代的收敛性。在合成数据和监督哈希问题上的大量数值结果证实，该方法能够以良好的计算效率获得高质量的二元解。未来工作可探索将该方法扩展到约束二元规划、其他类型的非光滑目标函数，以及用于调整惩罚参数的自适应策略。

表 II：二元压缩感知性能比较

μ = 0	ρ	1%	11%	21%	31%	41%	51%	61%	71%	81%	91%
平均时间	MPEC EPM	0.0739	0.0723	0.0688	0.0706	0.0688	0.0698	0.0719	0.0717	0.0725	0.0710
	DCRA	0.4965	0.4979	0.5001	0.5012	0.5006	0.4836	0.4981	0.5066	0.4950	0.5084
	Gurobi	0.4618	0.4531	0.4459	0.4485	0.4437	0.4599	0.4484	0.4568	0.4535	0.4525
	MPEC ADM	0.0751	0.0748	0.0692	0.0721	0.0723	0.0730	0.0734	0.0736	0.0722	0.0724
平均目标值	MPEC EPM	217.20	219.85	215.11	216.59	221.54	213.83	216.29	218.81	216.38	221.65
	DCRA	182.33	182.95	184.21	179.41	182.23	184.25	182.88	184.01	185.24	180.05
	Gurobi	241.34	275.05	303.96	329.82	359.00	386.72	413.01	440.86	453.59	474.44
	MPEC ADM	197.98	189.91	192.16	190.88	192.02	189.28	185.33	190.88	185.62	179.29
μ = 2	ρ	0.01	0.11	0.21	0.31	0.41	0.51	0.61	0.71	0.81	0.91
平均时间	MPEC EPM	0.0658	0.0655	0.0654	0.0656	0.0649	0.0651	0.0662	0.0653	0.0640	0.0654
	DCRA	0.4836	0.5208	0.5019	0.5264	0.4960	0.4901	0.5020	0.5116	0.5004	0.4969
	Gurobi	0.3955	0.4035	0.3953	0.3907	0.3939	0.3935	0.3906	0.3855	0.3921	0.3826
	MPEC ADM	0.0658	0.0661	0.0666	0.0651	0.0664	0.0666	0.0660	0.0666	0.0659	0.0672
平均目标值	MPEC EPM	500.00	433.93	375.50	269.88	237.53	252.43	244.96	305.22	384.34	420.96
	DCRA	315.32	271.81	227.10	201.53	186.98	182.20	190.28	203.84	229.08	280.08
	Gurobi	641.19	711.16	759.10	833.88	892.97	975.99	1,035.97	1,091.03	1,158.40	1,222.42
	MPEC ADM	503.83	438.47	378.77	261.84	204.37	233.10	227.19	252.34	322.03	344.80
μ = 10	ρ	0.01	0.11	0.21	0.31	0.41	0.51	0.61	0.71	0.81	0.91
平均时间	MPEC EPM	0.0650	0.0652	0.0662	0.0672	0.0651	0.0659	0.0661	0.0652	0.0664	0.0665
	DCRA	0.6348	0.6440	0.6601	0.6736	0.6848	0.7283	0.6903	0.6648	0.6580	0.6374
	Gurobi	0.0801	0.0796	0.0786	0.0823	0.0806	0.0886	0.0903	0.0898	0.0892	0.0906
	MPEC ADM	0.0658	0.0660	0.0646	0.0664	0.0658	0.0673	0.0645	0.0664	0.0666	0.0666
平均目标值	MPEC EPM	2,614.78	2,319.62	2,061.92	1,795.14	1,574.88	1,158.58	1,567.82	1,825.19	2,093.73	2,368.07
	DCRA	1,432.58	1,133.98	848.24	541.57	298.18	186.64	306.21	595.97	895.64	1,190.97
	Gurobi	2,984.03	3,276.34	3,569.60	3,881.01	4,133.50	4,471.29	4,749.66	5,057.28	5,365.27	5,652.05
	MPEC ADM	1,761.00	1,423.93	1,337.58	1,106.91	1,107.50	975.05	1,440.93	1,703.56	1,901.69	2,177.99

致谢

本研究得到国家自然科学基金项目（项目编号：12371299）的资助。

附录A 命题 1 的证明

证明：*(i)* 根据 $L_{\tilde{f}}$-光滑性，

$$\tilde{f}(V^{k,l+1}) \le \tilde{f}(V^{k,l}) + \langle \nabla \tilde{f}(V^{k,l}), \Delta^{k,l} \rangle + \tfrac{L_{\tilde{f}}}{2}\|\Delta^{k,l}\|_F^2,$$

其中 $\Delta^{k,\,l} := V^{k,\,l+1} - V^{k,\,l}$。根据 $V^{k,\,l+1}$ 在 (11) 中的最优性以及 $V^{k,\,l}$的可行性，

$$\begin{aligned}&\langle \nabla \tilde{f}(V^{k,l}), \Delta^{k,l} \rangle + \rho_k \langle \Gamma^{k,l}, \Delta^{k,l} \rangle \\ &+ \rho_k\big(\|V^{k,l+1}\|_F^2 - \|V^{k,l}\|_F^2\big) + \tfrac{L}{2}\|\Delta^{k,l}\|_F^2 \;\le\; 0.\end{aligned}$$

由于 $\|V\|_F^2$ $\mathcal{S}$ $\|V^{k\,l+1}\|_F^2$ $\|V^{k\,l}\|_F^2$ 在 上为常数，项 (. .) 相互抵消。将上述两式相加，并利用 $\tilde{\psi}(V^{k,\,l+1}) \le \tilde{\psi}(V^{k,\,l}) + \langle \Gamma^{k,\,l},\ \Delta^{k,\,l}\rangle$，得到

$$\Phi_k(V^{k,l+1}) - \Phi_k(V^{k,l}) \;\le\; -\tfrac{L-L_{\tilde{f}}}{2}\|\Delta^{k,l}\|_F^2.$$

(ii) 由于 $\Phi_k(V^{k,\,l})$ 单调不增且下有界，因此它收敛。$\{V^{k,l}\}$ 的有界性成立，因为 $\mathcal{S}$ 是闭且有界的。

(iii) (11) 的最优性条件给出，对于某个$W^{k,\,l+1} \in N_{\mathcal{S}}(V^{k,\,l+1})$,

$$0 = \nabla \tilde{f}(V^{k,l}) + \rho_k \Gamma^{k,l} + 2\rho_k V^{k,l+1} + L\Delta^{k,l} + W^{k,l+1}.$$

因此，

$$\begin{aligned}&\nabla \tilde{f}(V^{k,l+1}) + 2\rho_k V^{k,l+1} + \rho_k \Gamma^{k,l} + W^{k,l+1} \\ =&\big(\nabla \tilde{f}(V^{k,l+1}) - \nabla \tilde{f}(V^{k,l})\big) - L\Delta^{k,l}.\end{aligned}$$

˜取范数并利用 ∇f 的 $L_{\tilde{f}}$-利普希茨连续性，即可得到 (14)。 ■

附录B 定理1的证明

证明：令 $\Delta^{k,\,l} := V^{k,\,l+1} - V^{k,\,l}$。由命题1（iii），对 (14) 两边在 $l = 0, \ldots, T-1$ 上取最小值，得到

$$\min_{0\le l<T} \mathcal{G}\big(V^{k,l+1}, \Gamma^{k,l}\big) \;\le\; C \min_{0\le l<T} \|\Delta^{k,l}\|_F. \tag{26}$$

对于任意非负序列 $\{a_l\}$ $0\le l<T}\, a_l \le$，我们有 min 。将此应用于 $a_l = \|\Delta^{k\,l}\|_F$ ，得到

$$\min_{0\le l<T} \|\Delta^{k,l}\|_F \;\le\; \sqrt{\frac{1}{T}\sum_{l=0}^{T-1} \|\Delta^{k,l}\|_F^2}. \tag{27}$$

由命题 1(i)，

$$\sum_{l=0}^{T-1} \|\Delta^{k,l}\|_F^2 \le\; \frac{2}{L-L_{\tilde{f}}}\Big(\Phi_k(V^{k,0}) - \Phi_k(V^{k,T})\Big). \tag{28}$$

由于 (11) 是 (9) 的主化，因此对于所有 $V \in \mathcal{S}$,$\Phi_\delta^\star = \inf_{V\in\mathcal{S}} \Phi_\delta(V) \le \Phi_\delta(V) \le \Phi_k(V),$, 进而$\Phi_k(V^{k,\,T}) \le \Phi_k(V^{k,\,0}) - \Phi_\delta^\star$. 结合 (26)、(27) 和 (28) 得到 (15)，其中 $V^k = V^{k,\,0}$由算法2的步骤1得到。特别地，若 $T \ge \left\lceil \frac{2C^2}{L-L_{\tilde{f}}} \frac{\Phi_k(V^k)}{\epsilon^2} - \Phi_\delta^\star \right\rceil$, , 则 $\min_{0\le l<T} \mathcal{G}(\cdot) \le \epsilon$, 即内层循环经过 $\mathcal{O}(\epsilon^{-2})$ 次迭代后达到一个 ϵ-驻点。

对于任意 $\epsilon_v > 0$，选择 $T \ge \frac{\Phi_k\ (V^k)}{c\,\epsilon_v^2} - \Phi_\delta^\star$可保证存在一个索引 $\ell < T$，使得 $\|V^{k,\,\ell+1} - V^{k,\,\ell}\|_F \le \epsilon_v$，因此算法2终止。 ■

附录 C 引理1的证明

证明：设 $V^{k,\,0} := V^l$和 $U^{k,\,0} := V^{k,\,0} \in \mathcal{S}$。取 $\Gamma^{k,\,0} := -2\,V^{k,\,0}\,P_1 P_1^\top \in \partial\widetilde{\psi}(V^{k,\,0})$，其中 $P_1 \in \mathbb{R}^p$是

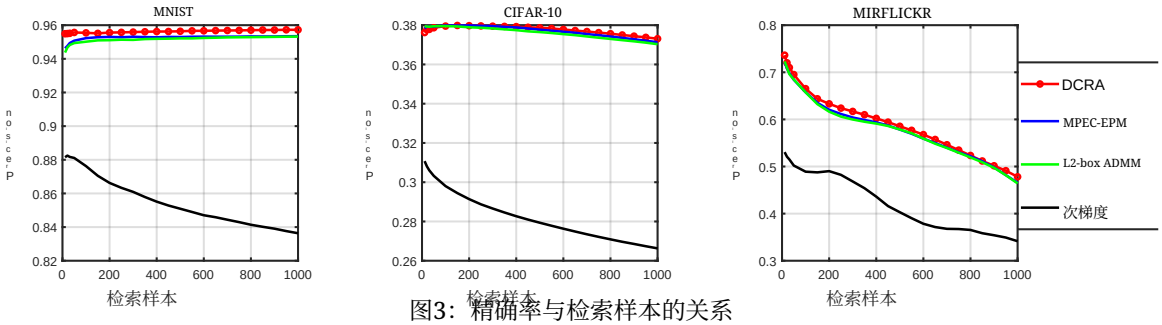


图3：精确率与检索样本的关系

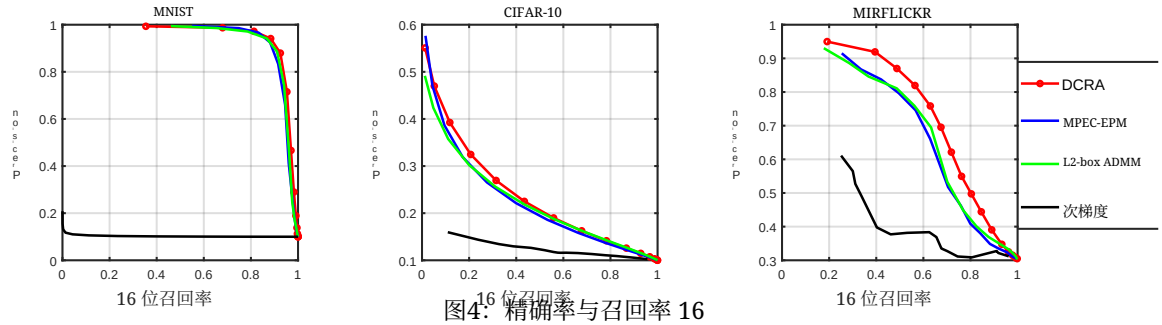


图4：精确率与召回率 16

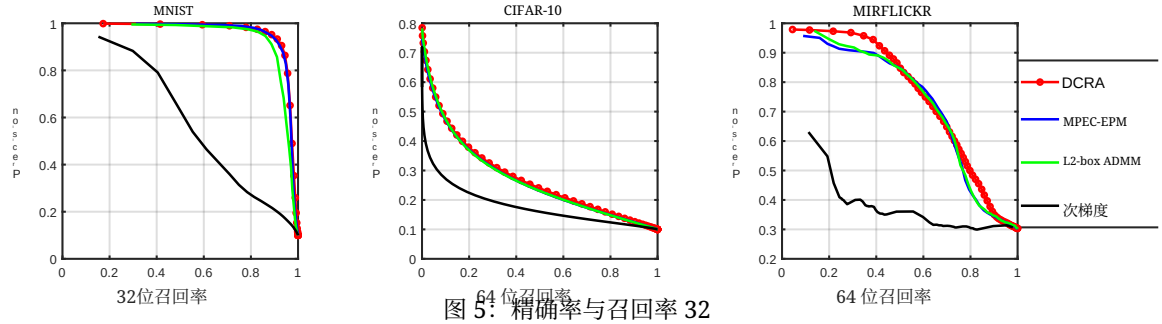


图 5：精确率与召回率 32

$(V^{k,0})^\top V^{k,0}$的主特征向量。对于每个 $j \in [p]$，$\Gamma^{k,0}$ 的第 j 列为

$$\begin{aligned}\Gamma_j^{k,0} &= -2\,V^{k,0}P_1\,(P_1)_j \Longrightarrow \\ \|\Gamma_j^{k,0}\| &= 2\|V^{k,0}P_1\||P_{j1}| = 2\,\|V^{k,0}\|\,|P_{j1}| \geq 2|P_{j1}|,\end{aligned}$$

其中我们利用了 P_1 是与 $\|V^{k,0}\|$ 相关联的右奇异向量，且 $\|V^{k,0}\| \geq \|V^{k,j}{}^0\| = 1$ 对所有 j 均成立。因此

$$\|\Gamma_j^{k,0}\| \;\geq\; 2\mu \text{ for all } j \in [p],\ \mu := \min_{1\leq j\leq p}|P_{j1}| > 0. \tag{29}$$

根据（12），定义

$$G^{k,0} := \frac{1}{L+2\rho_k}\left(LV^{k,0} - \nabla\widetilde{f}(V^{k,0}) - \rho_k\Gamma^{k,0}\right). \tag{30}$$

利用 ∇f的利普希茨连续性，其常数为$L_{\widetilde{f}}$，且$\|V^{k,0}\|_F = \sqrt{p}$, $\|E/\sqrt{m}\|_F = \sqrt{p}$，我们得到

$$\begin{aligned}\left\|LV^{k,0} - \nabla\widetilde{f}(V^{k,0})\right\|_F &\leq L\|V^{k,0}\|_F \\ &+ L_{\widetilde{f}}\|V^{k,0} - \tfrac{1}{\sqrt{m}}E\|_F + \left\|\nabla\widetilde{f}(\tfrac{1}{\sqrt{m}}E)\right\|_F \leq \varpi,\end{aligned} \tag{31}$$

其中 $\varpi := (L + 2L_{\widetilde{f}})\sqrt{p} + \|\nabla\widetilde{f}(\frac{1}{\sqrt{m}}E)\|_F$。取 (30) 中的第 j列，利用三角不等式以及 (29)–(31)，可得

$$\begin{aligned}\|G_j^{k,0}\| &\geq \frac{1}{L+2\rho_k}\left(\rho_k\|\Gamma_j^{k,0}\| - \left\|[LV^{k,0} - \nabla\widetilde{f}(V^{k,0})]_j\right\|\right) \\ &\geq \frac{2\rho_k\mu - \varpi}{L+2\rho_k}.\end{aligned} \tag{32}$$

在 $\rho_k \geq \overline{\rho} = (Lp + \varpi)\,/\,(\mu c_0)$ 且 $c_0 \in (0,\ 1)$ 时，我们有$\rho_k\mu \geq \varpi$，因此由（32）可得

$$\|G_j^{k,0}\| \;\geq\; \frac{\rho_k\mu}{L+2\rho_k} \;>\; 0 \qquad \forall j. \tag{33}$$

令 $D := \mathrm{Diag}\big(1/\|G^{k,1}{}^0\|,\ \ldots,\ 1/\|G^{k,p}{}^0\|\big)$。则 $V^{k,1} = \mathrm{Proj}_{\mathcal{S}}(G^{k,0}) = G^{k,0}D$。由（33）

$$\|D\| \;\leq\; \frac{L+2\rho_k}{\rho_k\mu}, \|D\|_F \leq \sqrt{p}\,\|D\| \leq \sqrt{p}\,\frac{L+2\rho_k}{\rho_k\mu}. \tag{34}$$

定义 $\Gamma^{k,0} := -\frac{1}{2}\Gamma^{k,0}D$，根据 $\Gamma^{k,0}$ 30 $V^{k,1} = G^{k,0}D$ $V^{k,1}$ 的定义，其秩 Γ^k $LV^{k,0} - \nabla\widetilde{f}\,V^{k,0}$ 为一。利用 ( ) 和 ， ， ， $= \frac{1}{L+2\rho_k}\left(\frac{L}{2}\Gamma^{k,0}D + (\ \ (\ \ )\ )\ D\right)$. 取弗罗贝尼乌斯范数并利用 $\|\Gamma^{k,0}\| \leq 2\|V^{k,0}\| \leq 2\|V^{k,0}\|_F =$

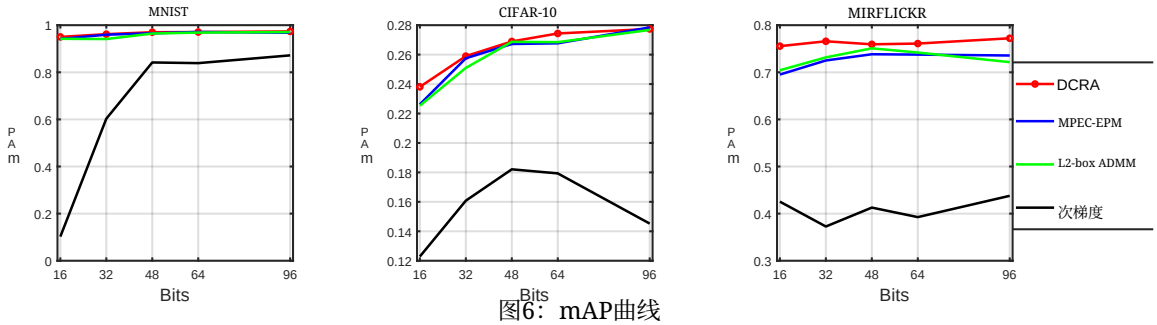


图6：mAP曲线

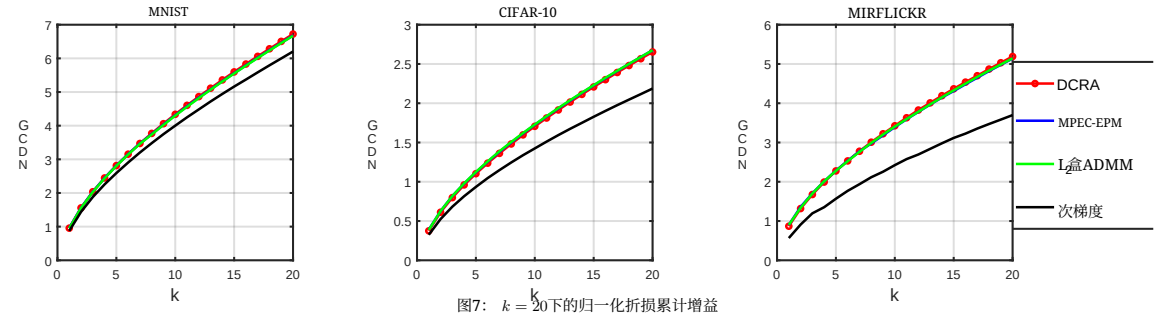


图7：$k=20$下的归一化折损累计增益

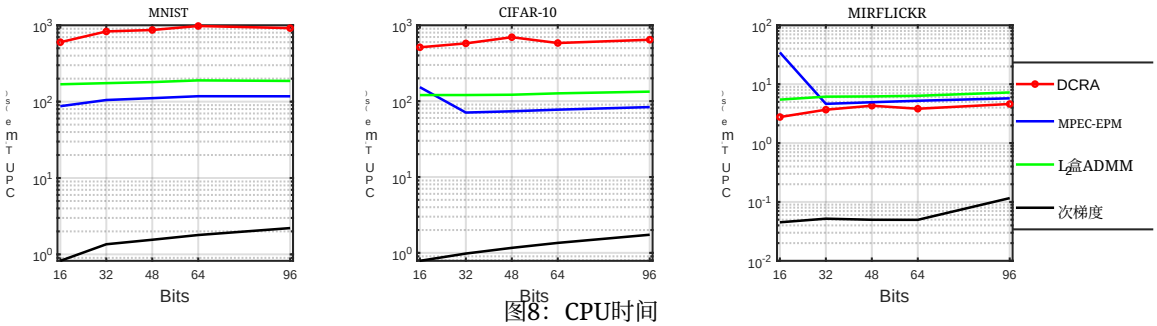


图8：CPU时间

$2\sqrt{p}$ 结合（31）和（34），可得

$$\begin{aligned}\operatorname{dist}(V^{k,1},\overline{\mathcal{R}}) &\le \|V^{k,1}-\overline{\Gamma}^{k,0}\|_F \\ &\le \frac{1}{L+2\rho_k}\left(\frac{L}{2}\|\Gamma^{k,0}\|\,\|D\|_F+\|LV^{k,0}-\nabla\widetilde{f}(V^{k,0})\|_F\,\|D\|\right) \\ &\le \frac{1}{L+2\rho_k}\left(L\sqrt{p}\cdot\frac{L+2\rho_k}{\rho_k\mu}+\varpi\cdot\frac{L+2\rho_k}{\rho_k\mu}\right)=\frac{Lp+\varpi}{\rho_k\mu}.\end{aligned}$$

利用$\rho_k \ge \rho = (Lp+\varpi)/(\mu c_0)$，可得 $\operatorname{dist}(V^{k,1},\mathcal{R}) \le c_0$。最后，根据埃克特–杨–米尔斯基定理，定义$\mathcal{R} := \{V \in \mathbb{R}^{m\times p} \mid \operatorname{rank}(V) \le 1\}$，

$$\begin{aligned}\operatorname{dist}(V^{k,1},\overline{\mathcal{R}})^2 &= \|V^{k,1}\|_F^2-\sigma_1(V^{k,1})^2 \\ &= \|V^{k,1}\|_F^2-\|V^{k,1}\|^2.\end{aligned}$$

■

Appendix D
Proof of Proposition 2

证明：(i) 固定 $l \in \mathbb{N}$ 使 $\frac{\varepsilon}{2} \le p-\|V^{k,l+1}\|^2 \le c_0^2$，其满足。根据，在线性化子问题中的最优性，

对于任意 $V \in \mathcal{S}$，$\widetilde{\Phi}_k(V^{k,l+1};V^{k,l}) \le \widetilde{\Phi}_k(V;V^{k,l})$，因此

$$\begin{aligned}\rho_k\left\langle\Gamma^{k,l},V^{k,l+1}-V\right\rangle &\le \left\langle\nabla\widetilde{f}(V^{k,l}),V-V^{k,l+1}\right\rangle \\ &\quad+\frac{L}{2}\left(\|V-V^{k,l}\|_F^2-\|V^{k,l+1}-V^{k,l}\|_F^2\right) \\ &\le 2\sqrt{p}\left(L_{\widetilde{f}}\left\|V^{k,l}-\tfrac{1}{\sqrt{m}}E\right\|_F+\left\|\nabla\widetilde{f}\left(\tfrac{1}{\sqrt{m}}E\right)\right\|_F\right) \\ &\quad+\frac{L}{2}\|V\|_F^2+L\|V-V^{k,l+1}\|_F\|V^{k,l}\|_F \\ &\le 2\sqrt{p}\left(L_{\widetilde{f}}\cdot 2\sqrt{p}+\left\|\nabla\widetilde{f}\left(\tfrac{1}{\sqrt{m}}E\right)\right\|_F\right)+\frac{L}{2}p+2Lp \\ &= 4L_{\widetilde{f}}p+2\sqrt{p}\left\|\nabla\widetilde{f}\left(\tfrac{1}{\sqrt{m}}E\right)\right\|_F+\tfrac{5}{2}Lp:=\varpi_{\mathrm{lin}},\end{aligned}$$

~其中，我们使用了 ∇f 的利普希茨连续性以及以下事实：$\|V\|_F=\|V^{k,l}\|_F=\sqrt{p}$ 且 $\|V-V^{k,l+1}\|_F \le 2\sqrt{p}$，对于$V \in \mathcal{S}$。重新排列可得，对于任意 $V \in \mathcal{S}$，

$$\begin{aligned}-\left\langle\Gamma^{k,l},V\right\rangle &\le \rho_k^{-1}\varpi_{\mathrm{lin}}+\left\langle\Gamma^{k,l},V^{k,l+1}\right\rangle \\ &\le \rho_k^{-1}\varpi_{\mathrm{lin}}+\|\Gamma^{k,l}\|_*\,\|V^{k,l+1}\| \\ &= \rho_k^{-1}\varpi_{\mathrm{lin}}+2\|V^{k,l}\|\,\|V^{k,l+1}\|, \qquad (35)\end{aligned}$$

其中我们使用了谱选择 $\Gamma^{k,l}=-2V^{k,l}u_1u_1^\top$以及秩一恒等式 $\|\Gamma^{k,l}\|_*=2\|V^{k,l}\|$。现在进行奇异值分解 $V^{k,l}=Q\operatorname{Diag}(\sigma)\,U^\top$，其顶部奇异向量为 q_1，u_1。令 $\widehat{u}_1 \in \mathbb{R}^p$ 由 $(\widehat{u}_1)_j = u_{j1}/|u_{j1}|$ 定义，并设

$\widehat{V} = q_1\widehat{u}_1^\top \in \mathcal{S}$ $-\langle \Gamma^{k,l}\ \widehat{V}\rangle = 2\langle V^{k,l}u_1u_1^\top\ q_1\widehat{u}_1^\top\rangle =:$. 然后, , , , . 代入 ( ) $2\|V^{k,l}\|\sum_{j=1}^p|u_{j1}|$. $V=\widehat{V}$ 35 得到

$$2\|V^{k,l}\|\sum_{j=1}^p|u_{j1}| \ \le\ \rho_k^{-1}\varpi_{\rm lin} + 2\|V^{k,l}\|\,\|V^{k,l+1}\|. \quad (36)$$

由 $\|V^{k,\ j\ l}\|=1$ 对每一列成立以及奇异值分解恒等式,

$$\sigma_1(V^{k,l})^2\,u_{j1}^2 = 1-\sum_{i\ge 2}\sigma_i(V^{k,l})^2\,u_{ji}^2$$
$$\ge 1-\sum_{i\ge 2}\sigma_i(V^{k,l})^2 = 1-\big(\|V^{k,l}\|_F^2-\|V^{k,l}\|^2\big).$$

由于 $p-\|V^{k,\ l}\|^2=\|V^{k,\ l}\|_F^2-\|V^{k,\ l}\|^2\le c_0^2$, 我们有

$$|u_{j1}| \ \ge\ \frac{\sqrt{1-c_0^2}}{\|V^{k,l}\|}\quad \forall j. \quad (37)$$

根据 $\sum_{j=1}^p u_{j1}^2=1$, 存在一个指标 $\hat{}\jmath$, 使得 $u_{\hat{j}1}^2\le 1/p$。对 $x\in[0,1]$ 使用 $x\ge x^2$,

$$\sum_{j=1}^p|u_{j1}| \ \ge\ |u_{\hat{j}1}|+\sum_{j\ne\hat{j}}u_{j1}^2 \ \ge\ \frac{\sqrt{1-c_0^2}}{\|V^{k,l}\|}+\Big(1-\frac{1}{p}\Big). \quad (38)$$

将 (38) 代入 (36) 得

$$2\|V^{k,l}\|\,\|V^{k,l+1}\| \ \ge\ 2\sqrt{1-c_0^2}+2\Big(1-\frac{1}{p}\Big)\|V^{k,l}\| - \rho_k^{-1}\varpi_{\rm lin}.$$

由于 $p-\|V^{k,\ l}\|^2\le c_0^2$ 蕴含 $\|V^{k,\ l}\|\ge\sqrt{p-c_0^2}$, 利用 $\|V^{k,\ l+1}\|^2+\|V^{k,\ l}\|^2\ge 2\|V^{k,\ l}\|\,\|V^{k,\ l+1}\|$, 可得 $\|V^{k,l+1}\|^2 \ge \|V^{k,l}\|^2+2\Big(\Big(1-\frac{1}{p}\Big)\sqrt{p-c_0^2}+\sqrt{1-c_0^2}\Big)-\frac{\varpi_{\rm lin}}{\rho_k}$.

令 $\eta(c_0,p):=2\Big(1-\frac{1}{p}\Big)\sqrt{p-c_0^2}+2\sqrt{1-c_0^2}$。因此, 选择 $\rho_k\ge\dfrac{2\,\varpi_{\rm lin}}{\eta\ c_0\ p}$(,) 可保证 (16)。

(ii) 令 $g_\ell:=p-\|V^{k,\ell}\|^2$。反设对所有 $\ell=1,\dots,\widehat{l}:=\lceil 2(c_0^2-\varepsilon)/\eta(c_0,p)\rceil+1$ 均有 $g_\ell>\varepsilon$。那么对于 $\ell=1,\dots,\widehat{l}-1$, 有 $g_\ell\in[\varepsilon/2,\,c_0^2]$, 因此可应用条目 (i) 得到 $\|V^{k\ \ell+1}\|^2\ge\|V^{k\ \ell}\|^2+\frac{1}{2}\eta\ c_0\ p \iff g_{\ell+1}\le$,, $g_\ell-\frac{1}{2}\eta\ c_0\ p$. $\widehat{l}-1$ (,) (,)。对这些不等式求和可得 $g_{\widehat{l}}\le$ $c_0^2-\frac{1}{2}\widehat{l}-1\ \eta\ c_0\ p\ \le\varepsilon$ ( ) (,), 矛盾。因此, 存在 $\nearrow 1\le l\le l$, 使得 $g_l\le\varepsilon$。

由引理1和 $\rho_k\ge\rho$, 我们得到 $g_1\le c_0^2$。若 $g_1\le\varepsilon$, 取$l=1$ 即可。否则 $\varepsilon<g_1\le c_0^2$。

若 $g_j\in[\varepsilon/2,\varepsilon]$, 再次应用条目 (i) 得到 $g_{j+1}\le g_j-\frac{1}{2}\eta(c_0,p)\le\varepsilon$。若 $g_j<\varepsilon/2$, 利用 (35) 并取 $V=V^{k,\ j}$, 得到

$$2\|V^{k,j}\|\|V^{k,j+1}\| \ \ge\ 2\|V^{k,j}\|^2-\rho_k^{-1}\varpi_{\rm lin}$$
$$\Longrightarrow\|V^{k,j+1}\|^2\ge\|V^{k,j}\|^2-\rho_k^{-1}\varpi_{\rm lin}.$$

因此 $g_{j+1}\le g_j+\rho_k^{-1}\varpi_{\rm lin}<\varepsilon/2+\varepsilon/2=\varepsilon$, 因为 $\rho_k\ge 2\,\varpi_{\rm lin}/\varepsilon$。所以, 对所有 $l\ge l$, 都有 $g_l\le\varepsilon$。

(iii) 由 (37) 和 $l=1$, 并在每个 $l\ge 1$ 处重复相同的论证, 我们得到

$$|u_{1j}|\ge\frac{\sqrt{1-c_0^2}}{\sigma_1}=\frac{\sqrt{1-c_0^2}}{\|V^{k,l}\|}\ge\frac{\sqrt{1-c_0^2}}{\sqrt{p}}>0,\ \forall j\in[p]. \quad (39)$$

根据 $P_1(V^{k,l})=u_1(V^{k,l})$ 这一事实, $(V^{k,l})^\top V^{k,l}$ 的主特征向量$P_1(V^{k,l})$ 对每个 $l\ge 1$ 均没有零元素。 ■

附录 E 定理 2 的证明

证明: 定义 $\Delta^{k,l}:=V^{k,l+1}-V^{k,l}$。对于固定的 k, 取 $L\ge L_{\widetilde{f}}$ 后, 更新式 (12) 给出

$$\rho_k\|V^{k,l+1}\|_F^2\le\rho_k\|V^{k,l}\|_F^2-\langle\nabla\widetilde{f}(V^{k,l})+\rho_k\Gamma^{k,l},\Delta^{k,l}\rangle-\frac{L}{2}\|\Delta^{k,l}\|_F^2.$$

结合 $\widetilde{f}$ 的下降引理, $\widetilde{f}(V^{k,\ l+1})$ $\rho_k\|V^{k\ l+1}\|^2\ \le\ \widetilde{f}\ V^{k\ l}\ -\rho_k\|V^{k\ l}\|^2$. $\widetilde{f}\cdot\ -\rho_k\|\cdot\|^2$, (,), 因此, ( ) 单调不增。由于 $V^{k+1}\in\{V^{k,l}\}$ 和 $V^{k,\ 0}=V^k$, 可得

$$\widetilde{f}(V^{k+1})-\rho_k\|V^{k+1}\|^2 \ \le\ \widetilde{f}(V^k)-\rho_k\|V^k\|^2. \quad (40)$$

在（40）的两边加上 $(\rho_k-\rho_{k+1})\|V^{k+1}\|^2$, 并从 $k=k^*$ 到 $k-1$ 求和消去中间项, 可得

$$\widetilde{f}(V^{\bar{k}})-\rho_{\bar{k}}\|V^{\bar{k}}\|^2\le\widetilde{f}(V^{k^*})-\rho_{k^*}\|V^{k^*}\|^2$$
$$+\sum_{j=k^*}^{\bar{k}-1}\big(\rho_j-\rho_{j+1}\big)\|V^{j+1}\|^2. \quad (41)$$

写出 $V^k=Q\mathrm{Diag}(\sigma_1,\dots,\sigma_r)U^\top$ 的奇异值分解, 以及 $X^k=U\mathrm{Diag}(\lambda_1,\dots,\lambda_r,0,\dots,0)U^\top$ 的特征分解。由于 $X^k=(V^k)^\top V^k=\sum_{i=1}^p\lambda_iP_iP_i^\top$ 且 $x^k:=\|V^k\|P_1$, 我们有 $x^k(x^k)^\top=\lambda_1P_1P_1^\top$。于是,

$$\big\|X^{\bar{k}}-x^{\bar{k}}(x^{\bar{k}})^\top\big\|_F=\Big(\sum_{i=2}^p\lambda_i^2\Big)^{\frac{1}{2}}\le\sum_{i=2}^p\sigma_i^2=\|V^{\bar{k}}\|_F^2-\|V^{\bar{k}}\|^2\le\varepsilon.$$

取对角线可得 $\|x^k\circ x^k-e\|\le\|x^k(x^k)^\top-X^{\bar{k}}\|_F\le\varepsilon$。由于 $\hat{}\,f$ 在 Ω 上是利普希茨连续的, 模数为 $L_{\hat{f}}$, 且 $X^k\in\Omega$,

$$\hat{f}(X^{\bar{k}})\ge\ \hat{f}\big(x^{\bar{k}}(x^{\bar{k}})^\top\big)-L_{\hat{f}}\big\|X^{\bar{k}}-x^{\bar{k}}(x^{\bar{k}})^\top\big\|_F$$
$$\ge\ \hat{f}\big(x^{\bar{k}}(x^{\bar{k}})^\top\big)-L_{\hat{f}}\,\varepsilon. \quad (42)$$

由于 $\|V^{k^*}\|\ \ge\ \|V^{k^*}\|_F/\sqrt{\text{秩}(V^{k^*})}=\sqrt{p/r^*}$, 利用 $\widetilde{f}(V^k)=\hat{}f\big((V^k)^\top V^k\big)=\hat{}f(X^k))$ 和 (41), 我们得到

$$\hat{f}(X^{\bar{k}})-\widetilde{f}(V^{k^*})\le\ \rho_{\bar{k}}\|V^{\bar{k}}\|^2\ -\ \rho_{k^*}\frac{p}{r^*}$$
$$+\ \sum_{j=k^*}^{\bar{k}-1}\big(\rho_j-\rho_{j+1}\big)\,\|V^{j+1}\|^2.$$

将上述结果与 (42) 结合, 得到 (18)。根据命题2(ii), 令 $g_k:=p-\|V^k\|^2\in[0,\epsilon]$, 使得 $\|V^k\|^2=p-g_k$。由于对于所有 j, $\rho_{j+1}\ge\rho_j$ 和 $\|V^{j+1}\|^2\ge 1$,

$$\sum_{j=k^*}^{\bar{k}-1}\big(\rho_j-\rho_{j+1}\big)\,\|V^{j+1}\|^2\ \le\ \sum_{j=k^*}^{\bar{k}-1}\big(\rho_j-\rho_{j+1}\big)\ =\ \rho_{k^*}-\rho_{\bar{k}}.$$

因此,

$$\hat{f}\big(x^{\bar{k}}(x^{\bar{k}})^\top\big)-\widetilde{f}(V^{k^*})\le\rho_{\bar{k}}\|V^{\bar{k}}\|^2-\rho_{k^*}\frac{p}{r^*}+\big(\rho_{k^*}-\rho_{\bar{k}}\big)+L_{\hat{f}}\,g_{\bar{k}}$$
$$=\rho_{\bar{k}}\,(p-g_{\bar{k}})-\rho_{k^*}\frac{p}{r^*}+\rho_{k^*}-\rho_{\bar{k}}+L_{\hat{f}}\,g_{\bar{k}}$$
$$=\rho_{\bar{k}}(p-1)+\rho_{k^*}\Big(1-\frac{p}{r^*}\Big)+\big(L_{\hat{f}}-\rho_{\bar{k}}\big)\,g_{\bar{k}}$$
$$\le\rho_{\bar{k}}(p-1)+\rho_{k^*}\Big(1-\frac{p}{r^*}\Big)+\big(L_{\hat{f}}-\rho_{\bar{k}}\big)\,\epsilon.$$

■

参考文献

[1] Y. Qian、S. Pan 和 S. Bi，“无约束二元多项式规划的矩阵非凸松弛方法”，《计算与优化应用》，第84卷第3期，875–919页，2023年。[2] B. Aspvall、M. F. Plass 和 R. E. Tarjan，“用于检验某些量化布尔公式真值的线性时间算法”，《信息处理快报》，第8卷第3期，121–123页，1979年。[3] B. Chevalier、W. Roga 和 M. Takeoka，“利用量子近似优化算法增强压缩感知”，《物理评论A》，第110卷第6期，062410页，2024年。[4] M. Lotfi 和 M. Vidyasagar，“利用维度近乎最优的二元矩阵进行压缩感知”，《IEEE信号处理汇刊》，第68卷，3008–3021页，2020年。[5] L. V. Nguyen、A. L. Swindlehurst 和 D. H. Nguyen，“采用单比特模数转换器的大规模MIMO系统的线性接收机与深度神经网络接收机”，《IEEE无线通信汇刊》，第20卷，7345，2021年11月，7333页。[6] M. Shao、W.-K. Chen、C.-Y. Yu、Y.-F. Liu 和 W.-K. Ma，“单比特MIMO检测：从全局最大似然检测器到幅度恢复方法”，arXiv预印本 arXiv:2407.11079，2024年。[7] H. Xiong、M. Yu、L. Liu、F. Zhu、J. Qin、F. Shen 和 L. Shao，“一种二元优化通用方法：收敛性分析与应用”，《IEEE模式分析与机器智能汇刊》，第44卷第9期，4524–4543页，2021年。[8] L. Xiao、Y. Qian 和 S. Pan，“基于流形梯度法的二元正交优化问题松弛方法及其应用”，《信号处理》，第238卷，110187页，2026年。[9] J. Gui、T. Liu、Z. Sun、D. Tao 和 T. Tan，“带松弛的监督式离散哈希”，《IEEE神经网络与学习系统汇刊》，第29卷第2期，608–617页，2018年。[10] R. Burkard，“二次指派问题”，载于《组合优化手册》，2741–2814页，Springer，纽约，2013年。[11] Q. H. Wu、Y. Wang 和 Z. P. L¨u，“基于禁忌搜索的最大割问题混合进化算法”，《应用软计算》，第34卷，827–837页，2015年。[12] F. Glover、Z. P. L¨u 和 J. K. Hao，“以多样化为导向的无约束二元二次问题禁忌搜索”，《运筹学研究》，第4卷，253页，2010年8月，239页。[13] N. Komodakis 和 G. Tziritas，“基于线性规划、通过图割进行近似标记”，《IEEE模式分析与机器智能汇刊》，第29卷第8期，1436–1453页，2007年。[14] M. X. Goemans 和 D. P. Williamson，“利用半定规划改进最大割与可满足性问题的近似算法”，《计算机协会学报》，第42卷，1145页，1995年6月，1115页。[15] C. Helmberg 和 F. Rendl，“利用半定规划和切平面求解二次（0, 1）问题”，《数学规划》，第82卷，291–315页，1998年。[16] F. Rendl、G. Rinaldi 和 A. Wiegele，“通过交叉结合半定松弛与多面体松弛，将最大割问题求至最优”，《数学规划》，第121卷第2期，307–335页，2010年。[17] N. Krislock、J. Malick 和 F. Roupin，“改进的半定界定程序：将最大割问题求至最优”，《数学规划》，第143卷第1期，61–86页，2014年。[18] P. Wang、C. Shen 和 A. Van Den Hengel，“求解二元二次问题的快速半定方法”，载于《IEEE计算机视觉与模式识别会议论文集》，2013年，1312–1319页。[19] T. R. Fu、D. D. Ge 和 Y. Y. Ye，“关于双重正半定规划松弛”，《计算数学学报》，第36卷第3期，391–403页，2018年。[20] M. F. Anjos 和 H. Wolkowicz，“通过二次提升强化最大割问题的半定松弛”，《离散应用数学》，第119卷第1–2期，79–106页，2002年。[21] S. Burer、R. D. C. Monteiro 和 Y. Zhang，“用于最大割及其他二元二次规划的秩二松弛启发式方法”，《SIAM优化杂志》，第12卷第2期，503–521页，2001年。[22] P. Chardaire 和 A. Sutter，“求解二次零一规划的分解法”，《管理科学》，第41卷第4期，704–712页，1995年。[23] S. Bi、X. Liu 和 S. Pan，“基于数学规划均衡约束公式的零范数最小化精确罚分解法”，《SIAM科学计算杂志》，第36卷第4期，A1451–A1477页，2014年。[24] G. Yuan 和 B. Ghanem，“基于数学规划均衡约束公式的二元优化精确罚函数法”，载于《AAAI人工智能会议论文集》，2017年，2867–2875页。[25] R. T. Rockafellar 和 R. J.-B. Wets，《变分分析》。德国柏林：Springer，1998年。[26] A. S. Lewis，“特征值的非光滑分析”，数学规划，第84卷，第1期，第1–24页，1999年。[27] S. Bi 和 S. Pan，“秩约束优化问题的误差界及其应用”，运筹学快报，第44卷，341，2016年 3，336第期，第–页。[28] S. Burer 和 R. D. C. Monteiro，“通过低秩分解求解半定规划的非线性规划算法”，数学规划，第95卷，第2期，第329–357页，2003年。[29] R. Y. Sun 和 Z. Q. Luo，“通过非凸分解保证矩阵补全”，*IEEE*信息论汇刊，第62卷，6579，2016年 11，6535第期，第–页。[30] Q. W. Li、Z. H. Zhu 和 G. G. Tang，“低秩矩阵优化的非凸几何”，信息与推断，第8卷，96，2019年 1，51第期，第–页。[31] T. Pham Dinh 和 H. A. Le Thi，“DC规划的凸分析方法：理论、算法与应用”，越南数学学报，第22卷，第289–355页，1997年。[32] T. Pham Dinh 和 H. A. Le Thi，“DC规划与DCA的最新进展”，计算智能汇刊第十三卷，第8342卷，第1–37页，2014年。[33] T. X. Liu、T. K. Pong 和 A. Takeda，“PDCA$_e$ 的精细收敛性分析及其在同时稀疏恢复和异常值检测中的应用”，计算与优化应用，第73卷，第1期，第69–100页，2019年。[34] B. Wu 和 B. Ghanem，“ℓ_p-盒 ADMM：一种通用的整数规划框架”，*IEEE*模式分析与机器智能汇刊，第41卷，第7期，第1695–1708页，2018年。[35] M. Doi 和 M. Ohzeki，“基于 L_1-范数最小化的二元压缩感知中的相变”，日本物理学会刊，第93卷，第8期，第084003页，2024年。[36] J. Gui、T. Liu、Z. Sun、D. Tao 和 T. Tan，“快速监督式离散哈希”，*IEEE*模式分析与机器智能汇刊，第40卷，第2期，第490–496页，2018年。[37] F. Shen、C. Shen、W. Liu 和 H. T. Shen，“监督式离散哈希”，载于*IEEE*计算机视觉与模式识别会议论文集，2015年，第37–45页。[38] A. D. Ioffe 和 J. V. Outrata，“次微分演算中的度量与平静性资格条件”，集值分析，第16卷，第3期，第199–227页，2008年。[39] D. R. Luke，“秩约束集的近端正则性及其对算法的影响”，数学成像与视觉期刊，第47卷，第3期，第231–248页，2013年。[40] R. T. Rockafellar，凸分析。美国新泽西州普林斯顿：普林斯顿大学出版社，1970年。

F 个人简介部分

肖亮海 于2020年获得英国伯明翰大学数学学院博士学位。此后，他曾在中国华南理工大学数学学院从事博士后研究。目前，他是中国暨南大学信息科学与技术学院的讲师。

钱逸天 于2018年和2023年分别获得中国华南理工大学数学学院学士学位和博士学位。目前，他是中国香港香港理工大学数据科学与人工智能系的博士后研究员。

潘少华 于2003年获得中国大连理工大学数学学院博士学位。目前，她是中国华南理工大学数学学院的教授。她的主要研究方向包括低秩与稀疏优化，以及非凸与非光滑优化。