arXiv:2603.05419v1 [math.NA] 5 Mar 2026

# 作为非线性方程组的结构化奇异距离

Miryam Gnazzo* Nicola Guglielmi† Federico Poloni‡ Stefano Sicilia§

2026年3月6日

**摘要**

本文研究具有预定线性结构 $\mathcal{S}$ 的非奇异矩阵$A \in \mathbb{C}^{n\times n}$的结构化奇异距离，即使 $A+\Delta$ 变为奇异矩阵所需的最小扰动 $\Delta \in \mathcal{S}$的范数（例如，这种结构可以是稀疏模式或实托普利兹结构）。这是结构化矩阵邻近问题的一个实例：当刻画某个系统性质相对于受特定结构约束的扰动（例如标称系统的结构）的鲁棒性时，控制与系统理论以及数值分析中都会出现这类问题。我们首先强调文献中提出的两种主要工具之间的相似之处：一种是针对特征值泛函的梯度系统方法，它需要求解某些低秩矩阵微分方程（参见 [Guglielmi、Lubich、Sicilia、SINUM 2023]）；另一种是双层优化方法，其中内层线性最小二乘问题可以显式求解（参见 [Usevich、Markovsky、JCAM 2014] 以及 [Gnazzo、Noferini、Nyman、Poloni、FoCM2025]）。特别是，这些文章强调了一个显著性质：$\Delta$ （至少在一般情形下）是秩一矩阵 $uv^*$在该结构 $\mathcal{S}$ 上的正交投影。该性质及上述相似之处启发我们将问题重新表述为关于两个向量未知量 $u, v \in \mathbb{C}^n$的非线性方程组。我们研究了这一新表述，并提出一种利用多元牛顿法直接求解这些非线性方程的算法。我们讨论了如何避免该非线性方程组的奇异性，以及如何确保单调收敛。对于大型矩阵，所得算法比现有算法更快，同时保持了相当的精度。

**关键词：** 结构矩阵邻近问题、秩一扰动、结构化奇异距离、非线性方程

**AMS主题分类：** 15A18, 15A99,65F15

*意大利国家研究委员会“Alessandro Faedo”信息科学与技术研究所，比萨，意大利。email：miryam.gnazzo@isti.cnr.it

†格兰萨索科学研究所数学部，意大利拉奎拉。 email： nicola.guglielmi@gssi.it

‡比萨大学计算机科学系，比萨，意大利。 email： federico.poloni@unipi.it§蒙斯大学数学与运筹学系，蒙斯，比利时。email：Stefano.SICILIA@umons.ac.be

# 1 引言

本文旨在介绍一种计算结构化矩阵非奇异半径的新方法；这是一个经典的矩阵逼近问题。关于相关邻近问题的详尽介绍，请参阅 N. Higham 的开创性论文 [16]。对于给定的非奇异矩阵 $A$，我们希望确定使所考察性质失效、也就是使矩阵变为奇异矩阵的最小加性扰动的大小。在非结构化情形下，这样的距离等于 $A$ 的最小奇异值，正如著名的埃卡特-杨-米尔斯基定理所述。然而，当矩阵 $A$ 具有某种结构时，例如为稀疏矩阵或托普利兹矩阵（这些情形下的结构是一个线性流形 $\mathcal{S}$），在流形 $\mathcal{S}$ 内确定距 $A$ 最近的奇异矩阵便显得更有意义。在这种情况下，埃卡特-杨-米尔斯基定理无济于事，因为它通常给出的是非结构化扰动，而这一数学问题（以及计算问题）颇具挑战性。我们的目标是通过提出一种融合 [6]和 [11]相关思路的新算法，为解决这一重要问题提供思路和计算技术。

对于一般的非奇异复矩阵 $A \in \mathbb{C}^{n\times n}$，我们考虑以下问题：寻找范数最小的扰动 $\Delta \in \mathbb{C}^{n\times n}$，使得 $A+\Delta$ 变为奇异矩阵，且 $\Delta \in \mathcal{S}$，其中 $\mathcal{S}$ 是 $\mathbb{C}^{n\times n}$ 中的一个线性子空间。这种扰动的最小范数称为 $A$ 的 非奇异半径

$$\min_{\Delta\in\mathcal{S}} \|\Delta\|^2 \text{ s.t. } (A+\Delta) \text{ is singular.} \tag{1}$$

本文中，我们在 (1) 中考虑的范数 $\|\cdot\|$ 是弗罗贝尼乌斯范数 $\|\cdot\|_F$，即矩阵元素构成的向量的欧几里得范数：对于 $M \in \mathbb{C}^{n\times n}$，其弗罗贝尼乌斯范数定义为

$$\|M\|_F = \left(\sum_{i,j=1}^{n} |m_{ij}|^2\right)^{1/2}.$$

如果不需要保持任何结构，则离 $A$ 最近的奇异矩阵可直接通过其奇异值分解得到

$$A = \sum_{i=1}^{n} \sigma_i\, u_i v_i^*,$$

其中 $u_i, v_i \in \mathbb{C}^n$ 是与奇异值 $\sigma_i > 0$ 对应的左右奇异向量，$^*$ 表示共轭转置。通过截去奇异值分解展开式的最后一项，即取下式，可得到最近的奇异矩阵

$$\Delta = -\sigma_n\, u_n v_n^*.$$

由于扰动 $\Delta$ 是一个秩一矩阵，因此它也是（算子）2-范数下的极小化点。因此，对于2-范数和弗罗贝尼乌斯范数，奇异距离都等于最小奇异值 $\sigma_n$。这是经典埃卡特–杨定理 [4],的一个特例（秩 $r = n-1$），该定理最初由施密特 [21]提出。然而，当 $A$ 属于某种指定的线性结构时——例如固定的稀疏模式，或托普利兹矩阵、汉克尔矩阵类——情况就发生了根本变化：无法直接通过截断 $A$的奇异值分解，得到同一结构中距离最近的奇异矩阵。最小结构化扰动 $\Delta$ 一般不再是秩

一矩阵。不过，对于弗罗贝尼乌斯范数，我们知道，最小扰动可以通过将一个秩一矩阵正交投影到该结构上得到（例如参见 [10], [11]）。

这一观察促成了例如 [10], [11]中提出的两层迭代算法。在内层迭代中，我们求解一个关于两个向量的微分方程组，这两个向量取决于距离参数；这只需要对结构化矩阵进行矩阵向量乘法以及计算向量内积。在外层迭代中，我们求解一个标量非线性方程，以确定结构化奇异距离。梯度系统方法也已在其他涉及矩阵邻近问题的情境中提出：例如在稳定化理论中 [14], ，用于计算互素多项式到公因子整除的距离 [12], ；在神经网络中 [3], ；在矩阵多项式中 [5] ；以及在图论中 [1, 15], ，等等。

解决这一问题的另一种方法基于以下观察：如果我们对给定向量 $v$施加更严格的约束$(A + \Delta)$ $v = 0$ ，那么最优的 $\Delta_*(v)$可以显式求得。因此，只需在 $v$上进行优化即可。Usevich和 Markovsky在一系列论文中对这种方法进行了深入研究 [17, 23], ，尽管这种“变量投影”的思想至少可以追溯到 [7]。最近，这种方法在 [6], 中得到重新审视和扩展，尤其是纳入了正则化技术。这种技术也可应用于上述许多问题，例如稳定化、矩阵多项式和多项式最大公因式问题。

**提纲**

本文组织如下。第 2 节简要介绍两种用于数值近似结构化奇异距离的现有方法，即基于常微分方程的方法和结构化低秩逼近（SLRA）/Riemann-Oracle 方法。此外，我们还列出了这两种优化过程之间的若干相似之处。第 3 节介绍一种用于计算结构化奇异距离的新方法，将该问题重新表述为一个非线性方程组。第 4 节介绍实现细节，第 5 节通过数值实验检验所提方法的表现。最后，第 6 节总结主要结果。

# 2 两种现有方法

## 2.1 梯度系统方法

本段介绍近期多篇文献提出的梯度系统方法（例如，参见 [8, 11, 13, 22]），该方法用于求解矩阵邻近问题。我们考虑其一般框架，并重点说明该方法如何应用于结构化奇异距离的计算问题。

给定矩阵 $A \in \mathbb{C}^{n\times n}$ 以及与 $A$ 的谱相关的性质 $\mathscr{P}$ ，我们寻找优化问题的解

$$\arg\min_{\Delta\in\mathcal{S}}\{\|\Delta\|_F : A + \Delta \text{ does not fulfil the property } \mathscr{P}\}, \tag{2}$$

其中， $\mathcal{S} \subseteq \mathbb{C}^{n\times n}$ 是一个子空间，它对所求扰动 $\Delta$施加结构约束。梯度系统方法采用两层方法，将原问题拆分为两个嵌套的子问题，并分别通过内层迭代和外层迭代求解。在一般情形下，我们考虑一个泛函 $\mathscr{F}\colon \mathcal{S} \to \mathbb{R}$ ，使得 $\mathscr{F}(0) > 0$ ，并且对于所有 $\Delta \in \mathcal{S}$,

$$\mathscr{F}(\Delta) \le 0 \iff A + \Delta \text{ does not fulfil the property } \mathscr{P},$$

这意味着，当且仅当 $\Delta$ 可容许时，泛函 $\mathscr{F}$ 取非正值。我们将扰动 $\Delta = \varepsilon E$改写为如下形式，其中 $\varepsilon > 0$ 是扰动大小， $E$ 的弗罗贝尼乌斯范数为 1，并将泛函 $F_\varepsilon$ 定义为

$$F_\varepsilon(E) := \mathscr{F}(\varepsilon E).$$

当扰动大小 $\varepsilon$ 固定时，内层迭代会最小化泛函 $F_\varepsilon$ ；而外层迭代旨在找到最小值 $\varepsilon_\star$ ，使得对于某个扰动 $\Delta \in \mathcal{S}$ ，满足 $\mathscr{F}(\Delta) = 0$ 成为可能，且该扰动的弗罗贝尼乌斯范数为 $\varepsilon_\star$。两层方法的概要如下：

- 内层迭代：给定固定的 $\varepsilon$，计算矩阵扰动 $E_\star(\varepsilon)$，使得

$$E_\star(\varepsilon) \in \underset{\|E\|_F=1,\ E\in\mathcal{S}}{\arg\min}\ F_\varepsilon(E) = \underset{\|\Delta\|_F=\varepsilon,\ \Delta\in\mathcal{S}}{\arg\min}\ \mathscr{F}(\Delta), \tag{3}$$

- 外层迭代：找到满足以下条件的最小值 $\varepsilon_\star > 0$

$$\phi(\varepsilon) := F_\varepsilon(E_\star(\varepsilon)) = 0. \tag{4}$$

内层迭代是最复杂的过程，而外层迭代在理论上更容易求解，因为它由一个一维求根问题构成，尽管它仍可能具有挑战性。

本文考虑用于计算结构化奇异距离的 [11] 方法，即我们旨在破坏的性质是 $\mathscr{P} = \{$矩阵非奇异$\}$ ，且泛函 $\mathscr{F}$ 取如下形式

$$\mathscr{F}(\Delta) = |\lambda_{\text{target}}(A + \Delta)|^2, \tag{5}$$

其中， $\lambda_{\text{target}}$ 是绝对值最小的特征值。我们还将考虑范围限定为 $\mathcal{S}$ 描述矩阵的稀疏模式的情形。为求解问题（3），内层迭代引入扰动矩阵路径 $E(t)$，其中 $t \geq 0$ ，并对矩阵常微分方程进行积分

$$\dot{E} = -\Pi_{\mathcal{S}} G_\varepsilon(E) + \text{Re}\langle \Pi_{\mathcal{S}} G_\varepsilon(E), E\rangle E, \tag{6}$$

其中，对于两个矩阵 $M$， $N$ ，我们用

$$\langle M, N\rangle = \text{Tr}(M^* N)$$

表示矩阵上的标量积，该标量积诱导出弗罗贝尼乌斯范数；用 $G_\varepsilon(E)$ 表示 $F_\varepsilon$的梯度，用 $\Pi_{\mathcal{S}}$ 表示到 $\mathcal{S}$ 上的正交投影。 $G_\varepsilon(E)$ 的表达式为

$$G_\varepsilon(E) = -\lambda x y^*, \tag{7}$$

其中， $x$ 和 $y$ 分别是与目标特征值 $\lambda = \lambda_{\text{target}}(A + \varepsilon E)$ 对应的单位左、右特征向量，因此 $x^* y > 0$；正交投影 $\Pi_{\mathcal{S}}$ 只需将模式之外的元素置为 0。

可以证明，式 (6) 的驻点对应于 $F_\varepsilon$的局部极小值。为了找到这些驻点，我们对常微分方程 (6) 进行积分，直至到达所求的驻点。由于方程 (6) 是一个梯度系统，积分最终总会到达某个驻点；除非发生非一般情形，这些驻点具有 $E \propto \Pi_{\mathcal{S}} G_\varepsilon(E)$ 的形式。

矩阵 $G_\varepsilon(E)$ 的秩为 1，因此驻点是秩一矩阵在 $\mathcal{S}$ 上的投影。这促使我们考虑另一个常微分方程来求解内层迭代，其轨迹属于秩一流形 $\mathcal{M}_1$：引入一个秩一矩阵路径 $Y(t) \subseteq \mathcal{M}_1$，使得

$$E(t) = \Pi_{\mathcal{S}} Y(t), \qquad t \in [0, +\infty)$$

并考虑常微分方程

$$\dot{Y} = P_Y \left( -G_\varepsilon(\Pi_{\mathcal{S}} Y) + \mathrm{Re}\langle P_Y(G_\varepsilon(\Pi_{\mathcal{S}} Y)), \Pi_{\mathcal{S}} Y \rangle Y \right), \tag{8}$$

其中，$P_Y$ 表示相对于弗罗贝尼乌斯内积投影到 $\mathcal{M}_1$ 在 $Y$中的切空间 $\mathcal{T}_Y\mathcal{M}_1$ 上的正交投影。(6) 的驻点与 (8) 的驻点之间存在显式的一一对应关系（见 [11, 定理 3.1]），这确保了用低秩方程而非满秩方程进行积分时，不会引入或丢失问题 (2) 的解。$Y = \rho \hat{u} v^*$ 的秩一微分方程 (8) 可改写为关于单位范数向量 $\hat{u}$的微分方程，$v$ 以及关于 $\rho$ 的显式公式。

**引理1** (因子的微分方程)。秩一微分方程 (8) 的每个解 $Y(t) \in \mathcal{M}_1$，在满足 $\|\Pi_{\mathcal{S}} Y(t)\|_F = 1$ 时，均可写为 $Y(t) = \rho(t) \hat{u}(t) v(t)^*$，其中，单位范数的 $u(t)$ 和 $v(t)$满足以下微分方程

$$\begin{aligned} \rho \dot{\hat{u}} &= -\tfrac{\mathrm{i}}{2} \operatorname{Im}(\hat{u}^* G v) \hat{u} - (I - \hat{u}\hat{u}^*) G v, \\ \rho \dot{v} &= -\tfrac{\mathrm{i}}{2} \operatorname{Im}(v^* G \hat{u}) v - (I - v v^*) G^* \hat{u}, \end{aligned}$$

其中 $G = G_\varepsilon(E)$ 对于 $E = \Pi_{\mathcal{S}} Y = \rho \Pi_{\mathcal{S}}(\hat{u} v^*)$ 以及 $\rho = 1/\|\Pi_{\mathcal{S}}(\hat{u} v^*)\|_F$。

尽管方程 (8) 不是梯度系统，但它在某种程度上接近于梯度系统：见 [11,定理 4.4]。特别地，当选择一个充分接近某个驻点的合适初始点时，对方程 (8) 积分总会到达该驻点。这只给出了一个局部收敛结果，弱于梯度系统的全局收敛性质。

首先注意，在 (8) 的驻点 $Y$ 处，有 $P_Y G_\varepsilon(E) = G_\varepsilon(E)$，其中$E = \Pi_{\mathcal{S}} Y$。因此，在驻点附近，$P_Y G_\varepsilon(E)$ 会接近 $G_\varepsilon(E)$。事实证明，两者甚至是二次接近的，正如下述引理所述。

**引理 2** (驻点附近的投影梯度)。设 $Y_\star \in \mathcal{M}_1$ 的弗罗贝尼乌斯范数为 $E_\star = \Pi_{\mathcal{S}} Y_\star \in \mathcal{S}$。设 $Y_\star$ 是秩-*1* 投影微分方程(8)的一个驻点，其中$\bar{}$是 $A + \varepsilon E_\star$ 的一个相应目标特征值 $\lambda$，且该特征值是单的。那么，存在 $\delta > 0$ 和一个实数 $C\bar{}$，使得对于所有正的 $\delta \le \delta$ 以及所有满足 $\|Y - Y_\star\| \le \delta$ 且相应 $E = \Pi_{\mathcal{S}} Y$ 具有单位范数的 $Y \in \mathcal{M}_1$，我们有

$$\|P_Y G_\varepsilon(E) - G_\varepsilon(E)\| \le C\delta^2. \tag{9}$$

作为这一引理的直接推论，比较微分方程 (8) 和 (6) 可知，当 $\delta$接近驻点时，泛函沿着方程 (8) 的解单调递减，直到 $O(\delta^2)$；而且在达到$O(\delta^2)$之前，其导数甚至与梯度流 (6) 的导数相同。注意，在强局部极小值的 $\delta$邻域内，泛函的导数与 $-\delta$ 成正比。Guglielmi, Lubich & Sicilia [11] 利用引理 2 证明了如下局部收敛结果：当 $t \to \infty$ 趋近于泛函 $F_\varepsilon$ 的强局部极小值时，对于与秩一微分方程 (8) 的解 $Y(t)$相关联的单位弗罗贝尼乌斯范数 $E(t)$ 的 $= \Pi_{\mathcal{S}} Y(t)$。

数值实验表明，这足以使该方法在实践中奏效。通过这种方式，将 (8) 纳入其中，可以利用矩阵逼近问题固有的低秩特征，从而在数值计算中获益。

注1。 在某些框架中，特征值优化问题 (5) 被奇异值优化问题所取代

$$\widetilde{\mathscr{F}}(\Delta) = \sigma_{\text{target}}(A + \Delta),$$

其中，我们在 $\mathscr{F}$ 的定义中用 $\sigma_{\text{target}}$ 替换 $\lambda^2_{\text{target}}$。这一思路与此前在 [9]，中采用的方法类似。本文将考虑这一特定情形，尤其关注最小奇异值，即 $\sigma_{\text{target}} = \sigma_{\min}$，因此

$$\arg\min_{\Delta \in \mathcal{S}} \widetilde{\mathscr{F}}(A + \Delta)$$

给出了 $A$ 的结构化奇异距离。在这种情况下，也可以研究与式 (6) 类似的相应梯度系统，即

$$\dot{E} = -\Pi_{\mathcal{S}} \widetilde{G}_\varepsilon(E) + \operatorname{Re}\langle \Pi_{\mathcal{S}} \widetilde{G}_\varepsilon(E), E\rangle\, E, \tag{10}$$

其中，$G = uv^*$ 是与函数 $F_\varepsilon(E) := \mathscr{F}(\varepsilon E)$ 相关联的梯度（对应于特征值情形下的 $F_\varepsilon(E)$），而 $u$ 和 $v$ 是与 $\sigma_{\min}$ 相关联的左右奇异向量

## 2.2 结构化低秩逼近/黎曼预言机方法

本节介绍另一种框架。Markovsky 和 Usevich 在一系列论文 [17,23,24]，中以结构化低秩逼近之名研究了这一框架，随后又将其扩展至更一般的邻近问题，并在 [6]，中以黎曼预言机之名进行了更深入的研究。就本文而言，我们简要介绍该方法在计算给定非奇异矩阵的结构化奇异距离这一特定情形下的应用。我们再次考虑矩阵 $A \in \mathbb{C}^{n\times n}$ 和一个维数为 $p$ 、包含允许的扰动$\Delta \in \mathbb{C}^{n\times n}$ 的线性子空间 $\mathcal{S}$。设 $P^{(1)}, \ldots, P_{(p)} \in \mathbb{C}^{n\times n}$ 是 $\mathcal{S}$ 的一组标准正交基，即

$$\mathcal{S} = \left\{ \Delta \in \mathbb{C}^{n\times n} \colon \Delta = \sum_{i=1}^{p} P^{(i)} \delta_i, \quad \delta = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_p \end{bmatrix} \in \mathbb{C}^p \right\}, \quad \operatorname{Tr}((P^{(i)})^* P^{(j)}) = \delta_{ij}.$$

等价地，$\operatorname{vec}(\Delta) = \mathcal{P}\delta$，其中

$$\mathcal{P} = \begin{bmatrix} \operatorname{vec} P^{(1)} & \ldots & \operatorname{vec} P^{(p)} \end{bmatrix} \in \mathbb{C}^{n^2\times p} \tag{11}$$

具有正交归一列。

该方法基于如下观察：若我们固定一个向量$v \in \mathbb{C}$，$v \neq 0$ 且该向量必须位于 $A + \Delta$ 的核中（在 [6], 的表述中，该目标向量由一个预言机提供）。为了求解

$$\min \|\Delta\|_F^2 \text{ s.t. } \Delta \in \mathcal{S}, (A + \Delta)v = 0, \tag{12}$$

我们可以通过将 (12) 改写为关于向量 $\delta$ 的形式来消除约束 $\Delta \in \mathcal{S}$，使得 $\operatorname{vec}\Delta = \mathcal{P}\delta$；于是 (12) 变为

$$\min \|\delta\|^2 \text{ s.t. } M\delta = r, \tag{13}$$

其中 $r = r(v) = -Av \in \mathbb{C}^n$，且

$$M = M(v) = (v^\top \otimes I_n)\mathcal{P} = \begin{bmatrix} P^{(1)}v & P^{(2)}v & \dots & P^{(p)}v \end{bmatrix} \in \mathbb{C}^{n\times p}. \tag{14}$$

这是一个经典问题，其（唯一）解为 $\delta_* = M^+ r$，其中 $M^+$表示 Moore-Penrose 伪逆（假设可行域非空，即 $r \in$ 值域$(M)$）。

因此，我们知道如何以闭式求解关于 $\Delta$ 的内层最小化子问题，并对给定的候选核向量 $v$计算最优的 $\|\Delta\|_F$；该核向量可取自 $\mathbb{C}^n$中的单位球面，此处记为 $\mathbb{S}_1$。因此，为了针对弗罗贝尼乌斯范数求解原问题 (1)，我们求解外层最小化问题

$$\min_{v\in\mathbb{S}_1} f(v), \quad f(v) = \|M(v)^+ r(v)\|^2,$$

即通过对流形 $\mathbb{S}_1$进行黎曼优化。

在 [23, 第 3.1], 节中，这些公式是在 $M$ 具有列满秩这一附加假设下得到的。然而，正如 [6, 第 2.1], 节所指出的，该假设存在问题：当 $M(v)$ 的秩下降时，函数 $f(v)$ 具有可去间断；而$f(v)$ 的全局极小值可能出现在该间断点处。在这种情况下，针对 $f(v)$ 的经典基于导数的优化方法会错过该间断，转而返回一个次优的局部极小值。

在 [6], 中，这个问题通过正则化过程来解决：我们最小化目标泛函 $f(v)$ 的一个松弛版本，即

$$f_\varepsilon(v) = \min_{\Delta\in\mathcal{S}} \|\Delta\|_F^2 + \varepsilon^{-1}\|(A+\Delta)v\|^2, \tag{15}$$

对于给定的 $\varepsilon > 0$。仿照上述推导，可以得到最小值的闭式表达式 m

$$f_\varepsilon(v) = r^*(MM^* + \varepsilon I)^{-1} r$$

以及相应的极小点

$$\delta_* = M^*(MM^* + \varepsilon I)^{-1} r = (M^*M + \varepsilon I)^{-1} M^* r, \quad \operatorname{vec}\Delta_* = \mathcal{P}\delta_*. \tag{16}$$

当 $\varepsilon \to 0$ 时， $f_\varepsilon(v)$ 的最小值趋于 $f(v)$ 的最小值 [6, 定理 2.11]。

此外，按照 [20] 中针对类似问题所采用的方法，我们注意到 $\mathcal{P}\mathcal{P}^*$ 是 vec $\mathcal{S}$ 上的正交投影矩阵，因此若令 $u = (MM^* + \varepsilon I)^{-1} r \in \mathbb{C}^n$，则有$\delta_* = M^* u$ 以及

$$\operatorname{vec}(\Delta_*) = \mathcal{P}\delta_* = \mathcal{P}M^* u = \mathcal{P}\mathcal{P}^*(\overline{v}\otimes u) = \operatorname{vec}\Pi_{\mathcal{S}}(uv^*), \tag{17}$$

也就是说，每次迭代计算得到的矩阵 $\Delta_*$ 是秩一矩阵 $uv^*$ 在结构 $\mathcal{S}$ 上的投影。

式（15）的欧几里得梯度为

$$\nabla_v f_\varepsilon = (A+\Delta)^* u, \quad \Delta = \Pi_{\mathcal{S}}(uv^*).$$

在单位球面上的驻点 $v_*$ of $f_\varepsilon$ 处，欧几里得梯度 $\nabla_{v_*} f$ must be a multiple of$v_*$。在极限 $\varepsilon \to 0$ 下，函数 $f(v)$ 具有尺度不变性，即对于每个 $\alpha \in \mathbb{C}$， $v \in \mathbb{C}^n$， $f(v\alpha) = f(v)$，因此梯度的径向分量为零，且必有 $\nabla_{v_*} f = 0$。

如前所述，在这项工作中，我们主要关注稀疏结构。在这种情况下， $MM^*$变为对角矩阵（见 [6, 节 5] 以及下文引理 8 的证明），从而使方法更加简单、快速。

### 2.3 两种方法的相似之处

值得注意的是，这两种方法之间存在相似之处。

| **常微分方程方法** | **黎曼–Oracle 方法** |
|---|---|
| 内层迭代：对于给定的 $\varepsilon > 0$，使用常微分方程积分器计算 $\Delta_* = \Pi_{\mathcal{S}}(uv^*)$，其中 $u = \varepsilon\dot{\rho}u$，它满足 $\dot{Y} = 0$。 | 内层迭代：对于给定的 $\varepsilon > 0$，使用黎曼优化计算 $\Delta_* = \Pi_{\mathcal{S}}(uv^*)$，使其满足 $\nabla_v f_\varepsilon = v\mu$。 |
| 外层迭代：通过牛顿-二分法计算最优的 $\varepsilon$ 以获得 $\lambda = 0$（单变量最小化-问题）。 | 外层迭代：令 $\varepsilon \to 0$。 |
| 在每次内层迭代中：$(A + \Delta_*)v = v\lambda$，$(A + \Delta_*)^* u = u\overline{\lambda}$。 | 在每次内层迭代中：$\|\Delta\|_F^2 + \varepsilon^{-1}\|(A+\Delta)v\|^2$ 达到最小值且为 $(A + \Delta_*)^* u = v\mu$。 |
| 在外层迭代收敛时：$(A + \Delta_*)v = 0$，$u^*(A + \Delta_*) = 0$。 | |

虽然这两种方法并不相同，但它们的结构非常相似：它们在每一步最小化不同的函数，但在嵌套迭代结束时，它们计算出一对 $(u, v)$，该对满足同一个非线性方程组。

## 3 重新表述为非线性方程组

这两种表述促使我们直接寻找满足这两个所需方程的向量 $u$，$v \in \mathbb{C}^n$。

问题：寻找满足非线性方程组的 $u$，$v \in \mathbb{C}^n$

$$(A + \Pi_{\mathcal{S}}(uv^*))v = (A + \Pi_{\mathcal{S}}(uv^*))^* u = 0. \tag{18}$$

当这两个方程（18）成立时，$(u, v)$ 是常微分方程方法和预言机方法中优化问题的一个临界点。

我们证明如下结论：常微分方程方法计算得到的向量 $u$，$v$ 经过适当缩放后，是问题（18）的一个解。

**定理 3**（固定 $\varepsilon$时的驻点）。设 $\mathscr{F}(\Delta) = \sigma_{\min}(A + \Delta)$，并固定 $\varepsilon > 0$。假设常微分方程方法所考察的轨迹上，$\sigma_{\min}(A+\varepsilon E)$ 为正且是单重的。那么，约束流中任一满足 $\|E(\varepsilon)\|_F = 1$ 的驻点 $E(\varepsilon) \in \mathcal{S}$，在乘以一个实数后满足：

$$E(\varepsilon) \propto \Pi_{\mathcal{S}}\big(u(\varepsilon)v(\varepsilon)^*\big), \tag{19}$$

其中，$u(\varepsilon)$ $v(\varepsilon)$ 是与 $\sigma(\varepsilon) := \sigma_{\min}(A+\varepsilon E(\varepsilon)$ 相关联的单位范数左右奇异向量。特别地，在采用归一化

$$E(\varepsilon) = \frac{\Pi_{\mathcal{S}}(u(\varepsilon)v(\varepsilon)^*)}{\|\Pi_{\mathcal{S}}(u(\varepsilon)v(\varepsilon)^*)\|_F}, \tag{20}$$

奇异向量关系可写为

$$(A+\varepsilon E(\varepsilon))v(\varepsilon)=\sigma(\varepsilon)u(\varepsilon),\qquad (A+\varepsilon E(\varepsilon))^{*}u(\varepsilon)=\sigma(\varepsilon)v(\varepsilon). \tag{21}$$

~证明。 </style}令 $\Delta(t)=\varepsilon E(t)$，其中 $\|E(t)\|_F\equiv 1$。与注 1 中一样，我们定义泛函 $F_\varepsilon(E)$ $=\sim\mathscr{F}(\varepsilon E)$ $=\sigma_{\min}(A+\varepsilon E)$。当 $\sigma_{\min}$ 为单重且严格为正时，成立标准微分恒等式

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{F}_\varepsilon(E(t))=\varepsilon\ \mathrm{Re}\langle\dot{E}(t),\,u(t)v(t)^{*}\rangle,$$

其中 $u(t)$ $v(t)$ 是相应的左右奇异向量。因此， $F_\varepsilon$ 在 $E$ 处的（欧氏）梯度~ ~为 $G_\varepsilon(E)=uv^{*}$。

$S$ 上的 Frobenius 范数单位球面上的约束梯度流为

$$\dot{E}=-\Pi_{\mathcal{S}}\widetilde{G}_\varepsilon(E)+\mathrm{Re}\langle\Pi_{\mathcal{S}}\widetilde{G}_\varepsilon(E),E\rangle\,E,$$

~ ~因此，在驻点 $E(\varepsilon)$处，我们有 $-\Pi_{\mathcal{S}}G_\varepsilon(E(\varepsilon))$ $+\mathrm{Re}\langle\Pi_{\mathcal{S}}G_\varepsilon(E(\varepsilon))$ $E(\varepsilon)\rangle E$ $(\varepsilon)$ $=0$，这蕴含式 (19)。采用归一化 (20) 可得 (21)。 □

请注意，在前一个定理中我们必须假设 $\sigma_{\min}>0$ 。在下一个定理中，我们将其延伸至极限，从而得到关于 $\sigma_{\min}=0$ 的结果。

**定理 4 (**在 $\varepsilon_\star$ 处的极限刻画)。 假设对于每个 $\varepsilon\in(0,\varepsilon_\star)$ ，都存在一个弗罗贝尼乌斯范数为单位的驻点 $E(\varepsilon)$ $\in\mathcal{S}$ $\|E(\varepsilon)\|_F=1$，使得最小奇异值

$$\sigma(\varepsilon):=\sigma_{\min}(A+\varepsilon E(\varepsilon))$$

为正且是单重的，并且相应的左右奇异向量 $u(\varepsilon)$ $v(\varepsilon)$ （单位范数）满足

$$(A+\varepsilon E(\varepsilon))v(\varepsilon)=\sigma(\varepsilon)u(\varepsilon),\qquad (A+\varepsilon E(\varepsilon))^{*}u(\varepsilon)=\sigma(\varepsilon)v(\varepsilon), \tag{22}$$

*and*

$$E(\varepsilon)=\frac{\Pi_{\mathcal{S}}(u(\varepsilon)v(\varepsilon)^{*})}{\|\Pi_{\mathcal{S}}(u(\varepsilon)v(\varepsilon)^{*})\|_F}. \tag{23}$$

此外，假设

$$\lim_{\varepsilon\nearrow\varepsilon_\star}\sigma(\varepsilon)=0,$$

并且存在一个常数 $c>0$ ，使得

$$\|\Pi_{\mathcal{S}}(u(\varepsilon)v(\varepsilon)^{*})\|_F\geq c\qquad \textit{for all } \varepsilon\in(0,\varepsilon_\star). \tag{24}$$

^那么存在向量 $\widehat{u}$、 $\widehat{v}\in\mathbb{C}^n$，满足 $\|\widehat{v}\|_2=1$，以及一个矩阵 $\Delta:=\Pi_{\mathcal{S}}(\widehat{u}\widehat{v}^{*})$ ，使得

$$(A+\widehat{\Delta})\widehat{v}=0,\qquad \textit{and}\qquad (A+\widehat{\Delta})^{*}\widehat{u}=0, \tag{25}$$

即，$(\widehat{u},\widehat{v})$ 满足非线性方程组 $(A+\Pi_{\mathcal{S}}(\widehat{u}\widehat{v}^{*}))\widehat{v}=(A+\Pi_{\mathcal{S}}(\widehat{u}\widehat{v}^{*}))^{*}\widehat{u}=0.$

证明。 定义 $\Delta(\varepsilon) := \varepsilon E(\varepsilon)$。由 (23) 可改写为

$$\Delta(\varepsilon) = \Pi_{\mathcal{S}}(\alpha(\varepsilon)\, u(\varepsilon) v(\varepsilon)^*)\,, \qquad \text{where} \qquad \alpha(\varepsilon) := \frac{\varepsilon}{\|\Pi_{\mathcal{S}}(u(\varepsilon)v(\varepsilon)^*)\|_F}.$$

设 $\tilde{u}(\varepsilon) := \alpha(\varepsilon)\, u(\varepsilon)$。则

$$\Delta(\varepsilon) = \Pi_{\mathcal{S}}(\tilde{u}(\varepsilon) v(\varepsilon)^*). \tag{26}$$

由 (24)，标量 $\alpha(\varepsilon)$ 在 $(0, \varepsilon_\star)$ 上有界，因此 $\tilde{u}(\varepsilon)$ 也有界，因为 $\|u(\varepsilon)\|_2 = 1$。此外，对于所有 $\varepsilon$，都有 $\|v(\varepsilon)\|_2 = 1$。因此，根据单位球面的紧致性，存在一个序列 $\varepsilon_k \nearrow \varepsilon_\star$，使得

$$v(\varepsilon_k) \to \widehat{v}, \qquad \tilde{u}(\varepsilon_k) \to \widehat{u}$$

对于某个 $\widehat{u}, \widehat{v}$，满足 $\|\widehat{v}\|_2 = 1$。由于 $\|\Delta(\varepsilon)\|_F = \varepsilon$，序列 $\Delta(\varepsilon_k)$ 有界，因此（取子序列后） $\Delta(\varepsilon_k) \to \Delta \in \mathcal{S}.$。在 (26) 中取极限并利用 $\Pi_{\mathcal{S}}$ 的连续性，得到

$$\widehat{\Delta} = \Pi_{\mathcal{S}}(\widehat{u}\widehat{v}^*). \tag{27}$$

No我们考虑 (22) 中的第一个关系：$(A + \Delta(\varepsilon_k))v(\varepsilon_k) = \sigma(\varepsilon_k)u(\varepsilon_k)$。取范数得 s

$$\|(A + \Delta(\varepsilon_k))v(\varepsilon_k)\|_2 = \sigma(\varepsilon_k),$$

并且根据假设 $\sigma(\varepsilon_k) \to 0$。由于

$$A + \Delta(\varepsilon_k) \to A + \widehat{\Delta}, \qquad v(\varepsilon_k) \to \widehat{v},$$

^我们得到 $(A + \Delta)\widehat{v} = 0$。

对于伴随关系，将 (22) 中的第二个方程乘以 $\alpha(\varepsilon_k)$：

$$(A + \Delta(\varepsilon_k))^* \tilde{u}(\varepsilon_k) = \alpha(\varepsilon_k)\sigma(\varepsilon_k)\, v(\varepsilon_k).$$

右端趋于零，因为 $\alpha(\varepsilon_k)$ 有界，且 $\sigma(\varepsilon_k) \to 0$。取极限^，得到 $(A + \Delta)^*\widehat{u} = 0$。这证明了 (25)。 □

若假设 $\exists \lim_{\varepsilon \nearrow \varepsilon_\star} {}^{u}(\varepsilon) = \widehat{u}$ 且 $\lim_{\varepsilon \nearrow \varepsilon_\star} {}^{v}(\varepsilon) = \widehat{v}$，则证明可以简化。

我们还可以证明极限矩阵具有明确的驻定性。

^**定理 5** （Clarke 驻点性于 $\Delta$）。 在定理 *4* 的假设下，设

$$\mathcal{S}_{\varepsilon_\star} := \{\Delta \in \mathcal{S} : \ \|\Delta\|_F = \varepsilon_\star\}, \qquad T_{\widehat{\Delta}}\mathcal{S}_{\varepsilon_\star} = \{Z \in \mathcal{S} : \ \mathrm{Re}\langle Z, \widehat{\Delta}\rangle = 0\}$$

*and*

$$\widetilde{\mathscr{F}}(\Delta) := \sigma_{\min}(A + \Delta), \qquad \Delta \in \mathcal{S}.$$

^则 $\Delta$ 在 $\mathcal{S}_{\varepsilon_\star}$ 上是 ***Clarke*** 驻点的（更多细节见 *[2]*）。等价地，以 $\Pi$ 为到切空间上的投影算子，

$$0 \in \Pi_{T_{\widehat{\Delta}}\mathcal{S}_{\varepsilon_\star}}\big(\partial^C \widetilde{\mathscr{F}}(\widehat{\Delta})\big),$$

即存在 $G \in \partial^C \sigma_{\min}(A + \widehat{\Delta})$ （***Clarke*** 次梯度），使得 $\Pi_{\mathcal{S}}(G)$ 与 $\widehat{\Delta}$ 共线。

证明。使用定理 4 的相同论证，我们令 $\Delta_k = \varepsilon_k E_k$。由于 $\sigma_k = \sigma_{\min}(A+\Delta_k) > 0$是简单的， $X \mapsto \sigma_{\min}(X)$ 在 $X_k$ 处可微：$= A + \Delta_k$，其梯度为 $u_k v_k^*$。$E_k$ 在单位球面上的驻点性意味着其梯度与 $E_k$ 共线。

由于 $\|u_k v_k^*\|_F = 1$，序列 $\{u_k v_k^*\}$ 有界；考虑一个收敛子列使得 $u_k v_k^* \to G$。我们得到

$$\Pi_{\mathcal{S}}(u_k v_k^*) = \|\Pi_{\mathcal{S}}(u_k v_k^*)\|_F \, E_k \implies \Pi_{\mathcal{S}}(G) = \eta \, \widehat{E}, \qquad \widehat{E} := \widehat{\Delta}/\varepsilon_\star,$$

$\widehat{}$对某个 $\eta > 0$ 成立（由非退化性假设(24)），故 $\Pi_{\mathcal{S}}(G)$与 $\Delta$共线。映射$X \mapsto \sigma_{\min}(X)$是局部 Lipschitz 的，其 Clarke 次微分是外半连续的。因此，由于$X_k \to \widehat{X} := A + \widehat{\Delta}$ 且 $u_k v_k^* = \nabla \sigma_{\min}(X_k)$， $u_k v_k^*$ 的任一极限 $G$ 属于 $\partial^C \sigma_{\min}(\widehat{X})$。于是$G \in \partial^C \widetilde{\mathscr{F}}(\widehat{\Delta})$，且 $\Pi_{\mathcal{S}}(G)$与 $\widehat{\Delta}$共线，从而证明了 Clarke 驻点性这一陈述。

□

注记 2. 若 $(u, v)$ 是方程 (18) 的一个解，则

$$(u\alpha^{-1}, v\overline{\alpha}) \tag{28}$$

对每个 $\alpha \in \mathbb{C}, \alpha \neq 0$，它也是另一个解；因此，解仅在归一化因子的意义下确定。不失一般性，我们可以假设 $\|v\| = 1$。

## 3.1 微分、梯度和海森矩阵

我们设

$$G(u, v) = \begin{bmatrix} (A+\Delta)v \\ (A+\Delta)^* u \end{bmatrix}, \quad \Delta = \Pi_{\mathcal{S}}(uv^*),$$

于是 (18) 变为 $G(u, v) = 0$。值得注意的是，当 $A \in \mathcal{S}$ 时，这个量 $G(u, v)$ 是某个标量函数 $F(u, v)$ 的梯度。我们将在下面的引理中证明这一点。

**引理 6.** 考虑函数 $F(u, v) = \frac{1}{2}\|A + \Pi_{\mathcal{S}}(uv^*)\|_F^2$。那么，$F$ 关于 $u$ 和 $v$ 的两个偏梯度为

$$\nabla_u F(u, v) = (\Pi_{\mathcal{S}}(A) + \Delta)v, \quad \nabla_v F(u, v) = (\Pi_{\mathcal{S}}(A) + \Delta)^* u, \quad \Delta = \Pi_{\mathcal{S}}(uv^*). \tag{29}$$

证明。 注意

$$\langle A + \Delta, A + \Delta \rangle = \|A\|_F^2 + 2\langle \Pi_{\mathcal{S}}(A), uv^* \rangle + \|\Pi_{\mathcal{S}}(uv^*)\|_F^2.$$

我们计算微分

$$\begin{aligned} dF(u, v) &= d\frac{1}{2}\left(\|A\|_F^2 + 2\langle \Pi_{\mathcal{S}}(A), uv^* \rangle + \|\Pi_{\mathcal{S}}(uv^*)\|_F^2\right) \\ &= \langle \Pi_{\mathcal{S}}(A), (du)v^* + u(dv)^* \rangle + \langle \Pi_{\mathcal{S}}((du)v^* + u(dv)^*), \Pi_{\mathcal{S}}(uv^*) \rangle \\ &= \langle \Pi_{\mathcal{S}}(A) + \Delta, (du)v^* + u(dv)^* \rangle \\ &= \langle (\Pi_{\mathcal{S}}(A) + \Delta)v, du \rangle + \langle dv, (\Pi_{\mathcal{S}}(A) + \Delta)^* u \rangle, \end{aligned}$$

其中我们用到了 $\langle \cdot, \cdot \rangle$ 的定义、 $\Pi_{\mathcal{S}}$ 是一个正交投影这一事实，以及迹的循环性质。从最后一行可以直接读出这两个梯度。 □

当 $A \in \mathcal{S}$, $\Pi_{\mathcal{S}}(A) = A$ 时，(29) 因而与 $G(u,v)$ 的两个块一致。这个结果表明 (18) 是一个梯度系统。

该 $F(u,v)$ 与原始极小化问题 (1) 之间的关系并非显而易见；一般而言，全局距离极小化器 $\Delta_*$ 既不是泛函 $F$ 的极小值，也不是其极大值：见下文例 9。

我们可以给出 $G(u,v)$ 的微分 $dG(u,v)$ 的显式公式，它是一个 $\mathbb{R}$-线性算子 $H : \mathbb{C}^{2n} \to \mathbb{C}^{2n}$。我们使用字母 $H$，是因为在 $G$ 为$F(u,v)$ 的梯度的情形下，算子 $H$是 $F(u,v)$ 的海森矩阵。

我们注意到，即使当 $u, v$ 为复向量时， $H$ 也仅保证是 $\mathbb{R}$-线性的，因为其表达式中会出现共轭。为固定记号，令 $x$ 表示标量、向量或矩阵 $x$ 的（逐项）共轭；而 $x^\top$ 表示 $x$ 的转置。

**引理 7.** $G(u,v)$ 的微分是 $\mathbb{R}$-线性算子 $H$，使得对每个 $\left[\begin{smallmatrix} du \\ dv \end{smallmatrix}\right] \in \mathbb{C}^{2n}$

$$H \begin{bmatrix} du \\ dv \end{bmatrix} = \begin{bmatrix} MM^*\, du + (A+\Delta) dv + MN^\top \overline{dv} \\ (A+\Delta)^* du + NM^\top \overline{du} + NN^*\, dv \end{bmatrix},$$

其中 $\mathcal{P}$ 的定义同(11)，矩阵 $M$ 和 $N$ 的定义如下：

$$M = (v^\top \otimes I_n)\mathcal{P} = \begin{bmatrix} P^{(1)}v & P^{(2)}v & \dots & P^{(p)}v \end{bmatrix} \in \mathbb{C}^{n\times p},$$
$$N = (I \otimes u^\top)\overline{\mathcal{P}} = \begin{bmatrix} (P^{(1)})^*u & (P^{(2)})^*u & \dots & (P^{(p)})^*u \end{bmatrix} \in \mathbb{C}^{n\times p}.$$

证明。 通过求导计算海森矩阵

$$d(A+\Delta)v = \Pi_{\mathcal{S}}(du\, v^*)v + \Pi_{\mathcal{S}}(u(dv)^*)v + (A+\Delta)dv,$$
$$d(A+\Delta)^*u = \Pi_{\mathcal{S}}(du\, v^*)^*u + \Pi_{\mathcal{S}}(u\,(dv)^*)^*u + (A+\Delta)^*\, du.$$

我们回顾一下在 (17) 中已经用到的一些恒等式：对于每个 $a$， $b \in \mathbb{C}^n$，都有

$$\operatorname{vec} \Pi_{\mathcal{S}}(ab^*) = \mathcal{P}\mathcal{P}^*(\operatorname{vec} ab^*) = \mathcal{P}\mathcal{P}^*(\overline{b} \otimes a) = \mathcal{P}\mathcal{P}^*(\overline{b} \otimes I_n)a = \mathcal{P}\mathcal{P}^*(I_n \otimes a)\overline{b}.$$

利用以下关系，这些恒等式可以帮助我们简化 $d(A+\Delta)v$ 和 $d(A+\Delta)^*u$ 中的若干项：

$$\Pi_{\mathcal{S}}(du\, v^*)v = (v^\top \otimes I) \operatorname{vec} \Pi_{\mathcal{S}}(du\, v^*) = (v^\top \otimes I)\mathcal{P}\mathcal{P}^*(\overline{v} \otimes I_n)du = MM^*du,$$
$$\Pi_{\mathcal{S}}(u(dv)^*)v = (v^\top \otimes I)\mathcal{P}\mathcal{P}^*(I_n \otimes u)\overline{dv} = MN^\top \overline{dv},$$
$$\Pi_{\mathcal{S}}(u\,(dv)^*)^*u = \operatorname{vec}(u^\top \overline{\Pi_{\mathcal{S}}(u(dv)^*)}) = (I_n \otimes u^\top)\overline{\mathcal{P}\mathcal{P}^*(I_n \otimes u)\overline{dv}}$$
$$= (I_n \otimes u^\top)\overline{\mathcal{P}}\mathcal{P}^\top(I_n \otimes \overline{u})dv = NN^*dv,$$
$$\Pi_{\mathcal{S}}(du\, v^*)^*u = \operatorname{vec}(u^\top \overline{\Pi_{\mathcal{S}}(du\, v^*)}) = (I_n \otimes u^\top)\overline{\mathcal{P}\mathcal{P}^*(\overline{v} \otimes I_n)du}$$
$$= (I_n \otimes u^\top)\overline{\mathcal{P}}\mathcal{P}^\top(v \otimes I_n)\overline{du} = NM^\top \overline{du}.$$

□

此外，对于投影到实稀疏结构这一特殊情况，该表达式还可以进一步简化。

**引理 8**。 设 $\mathcal{S}$ 为具有稀疏结构 $\mathcal{J}$ 的矩阵子空间，即当 $(i,j) \notin \mathcal{J}$ 时， $B_{ij}=0$。假设 $A \in \mathcal{S}$ 且 $A$、$u$、$v$ 均为实数；则算子 $H(u,v)$（作用于实向量）为

$$H = \begin{bmatrix} K_1 & A+2\Delta \\ (A+2\Delta)^* & K_2 \end{bmatrix}, \tag{30}$$

*wher*我们再次设定 $\Delta = \Pi_{\mathcal{S}}(uv^*)$，且 $K_1, K_2$ 是满足.....的两个对角矩阵 $t$

$$(K_1)_{ii} = \sum_{j \ s.t. \ (i,j)\in\mathcal{J}} v_j^2, \quad (K_2)_{jj} = \sum_{i \ s.t. \ (i,j)\in\mathcal{J}} u_i^2.$$

证明。 我们可以取 $(e_i e_j^\top : (i,j) \in \mathcal{J})$ 作为 $\mathcal{S}$ 的标准正交基。于是， $M$ 的列为 $e_i e_j^* v = e_i v_j$，而 $N$ 的列为 $e_j e_i^* u = e_j u_i$。由此可知，在引理 7 中

$$\begin{aligned} MM^* = MM^\top &= \sum_{(i,j)\in\mathcal{J}} e_i v_j v_j e_i^\top = K_1, \\ NN^* = NN^\top &= \sum_{(i,j)\in\mathcal{J}} e_j u_i u_i e_j^\top = K_2, \\ MN^\top &= \sum_{(i,j)\in\mathcal{J}} e_i v_j u_i e_j^\top = \Delta. \end{aligned}$$

□

**例 9**。 设 $A = \mathrm{diag}(\sigma_1, \sigma_2)$，其中 $\sigma_1 > \sigma_2 > 0$，而 $\mathcal{S} = \mathbb{R}^{2\times 2}$ 是每个 $2\times 2$ 矩阵都满足的平凡稀疏结构（$\mathcal{J} = \{(1,1)(1,2)(2,1)(2,2)\}, \Pi_{\mathcal{S}} = I$）。那么，根据 *Eckhart-Young* 定理，距离 $A$ 最近的（结构化）奇异矩阵是 $\mathrm{diag}(\sigma_1, 0)$，对应的秩一扰动为 $\Delta = uv^*$，其中 $u = -e_2\sigma_2$, $v = e_2$。我们可以验证 $(A+\Delta)v = (A+\Delta)^* u = 0$ 成立，即满足(18)。

此外，在点 $(u,v) = (-\sigma e_2, e_2)$ 处，(30) 给出

$$H = \nabla^2_{[u,v]} F = \begin{bmatrix} 1 & 0 & \sigma_1 & 0 \\ 0 & 1 & 0 & -\sigma_2 \\ \sigma_1 & 0 & \sigma_2^2 & 0 \\ 0 & -\sigma_2 & 0 & \sigma_2^2 \end{bmatrix}.$$

$H$ 的特征值是以下各项的并集：

$$H_1 = \begin{bmatrix} 1 & \sigma_1 \\ \sigma_1 & \sigma_2^2 \end{bmatrix}, \quad H_2 = \begin{bmatrix} 1 & -\sigma_2 \\ -\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

矩阵 $H_1$ 是不定的，因为 $\sigma_1 > \sigma_2$，因此它有一个正特征值和一个负特征值。矩阵 $H_2$ 是半定的，其特征值为0 和 $1+\sigma_2^2$。这个零特征值反映了解是过度参数化的事实： $F(u,v) = F(\alpha u, \frac{1}{\alpha} v)$，因此 $\frac{d}{d\alpha} F(\alpha u, \frac{1}{\alpha} v) = 0$。

由于海森矩阵 $H$ 是不定的，奇异性距离问题的解是 $F(u,v)$的鞍点，而不是局部极小值或极大值。

# 4 基于牛顿方法的算法

可以用牛顿法求解非线性方程组（18）。这里有几个问题需要讨论。

## 4.1 固定归一化

回想一下，(18) 的解如 (28) 所示，仅在一个归一化因子的意义下确定；因此，微分 $H$ 是奇异的。为避免出现多余的自由度，我们只希望计算满足 $\|v\| = 1$ 的解。为了制定强制实现这一归一化的策略，我们先考虑 $G(u, v)$ 是梯度系统的情形。此时，我们可以选择 $\beta > 0$ 并定义

$$F_\beta(u, v) = \frac{1}{2}\|A + \Pi_{\mathcal{S}}(uv^*)\|_F^2 + \frac{\beta}{4}\left(\|v\|^2 - 1\right)^2.$$

$F_\beta(u, v)$ 的梯度为

$$G_\beta(u, v) = \begin{bmatrix} (A + \Delta)v \\ (A + \Delta)^* u + \beta(\|v\|^2 - 1)v \end{bmatrix}, \quad \Delta = \Pi_{\mathcal{S}}(uv^*).$$

显然，如果 $u, v$ 是 $F(u, v)$ 的驻点，且 $\|v\| = 1$，那么它们也是 $F_\beta(u, v)$ 的驻点，因为两个加项的梯度都为零。

即使 $A \notin \mathcal{S}$ 和 $G(u, v)$ 不是梯度系统，我们仍然可以构造修正方程 $G_\beta(u, v) = 0$：事实上，我们可以直接验证，$G(u, v) = 0$ 的任意解 $(u, v)$，只要其中 $\|v\| = 1$，也是 $G_\beta(u, v)$ 的解。

$G_\beta(u, v)$ 的微分为

$$H_\beta(u, v) = H(u, v) + \begin{bmatrix} 0 & 0 \\ 0 & 2\beta vv^* + \beta(\|v\|^2 - 1)I \end{bmatrix},$$

其中，$H(u, v)$ 在引理 7 中定义。利用该微分，我们写出用于求解非线性方程组 $G_\beta(u, v) = 0$ 的多元牛顿法，其解为 $(u, v)$：

$$\begin{bmatrix} u_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} u_k \\ v_k \end{bmatrix} + \begin{bmatrix} \delta_{u_k} \\ \delta_{v_k} \end{bmatrix} \tag{31}$$

其中

$$\begin{bmatrix} \delta_{u_k} \\ \delta_{v_k} \end{bmatrix} = -H_\beta^{-1}(u, v)G_\beta(u, v). \tag{32}$$

## 4.2 线搜索

我们修改更新公式 (32)，得到一种带线搜索的算法，并将其表述为算法 1。

换言之，如果牛顿步 (32) 会产生一个新的迭代点 $(u_{k+1}, v_{k+1}) = (u_k + \delta_{u_k}, v_k + \delta_{v_k})$，且该迭代点不能减小 $\|G_\beta(u_k, v_k)\|$ 的值，那么我们将步长减半，并检验一个新的候选点 $u_{k+1} = u_k + \frac{1}{2}\delta_{u_k}$，$v_{k+1} = v_k + \frac{1}{2}\delta_{v_k}$；我们以这种方式继续将步长减半，直到范数减小。如果 $H_\beta(u, v)$ 是正定的，牛顿方向始终是下降方向，这一策略应能确保最终出现下降。

**算法 1：** 采用回溯线搜索的牛顿法

**数据：** 非奇异矩阵 $A$，结构 $\mathcal{S}$

**结果：** $G(u,v)=0, \|v\|=1$ 的一个解，以及满足 $A+\Delta$ 为 $=0$ 的相关 $\Delta \in \mathcal{S}$ 奇异（希望是一个全局最小值点）

选择 $\beta>0$ （例如 $\beta=\|A\|_F$）；

$u, v \leftarrow$ 的最小左右奇异值 $A$;

**当**收敛 **执行**

  $(\delta_u, \delta_v) \leftarrow$ (32) 中的牛顿增量;

  $\alpha \leftarrow 1$; // 候选步长

  **当** $\|G_\beta(u+\alpha\delta_u, v+\alpha\delta_v)\| \geq \|G_\beta(u,v)\|$ **执行**

    $\alpha \leftarrow \alpha/2$;

  $(u,v) \leftarrow (u+\alpha\delta_u, v+\alpha\delta_v)$; // 确保 $\|G_\beta(u,\ v)\|$ 减小

$\Delta \leftarrow \Pi_{\mathcal{S}}(uv^*)$;

这种回溯策略是牛顿法全局化的一种非常简单的形式，用以确保我们产生一个递减序列 $\|G_\beta(u_k, v_k)\|$。我们建议读者参阅 [19, 第 3] 章，以更广泛地了解优化中的线搜索方法以及更精细的策略。

在实践中，我们观察到，在大多数情况下，选择 $\alpha = 1$（不带线搜索的经典牛顿法）就足以使范数减小，无需执行回溯步骤；但这一保护措施只需极少的额外开销，便能提高算法的鲁棒性。

## 4.3 初始值

由于牛顿法的表现取决于初始点的选择，我们介绍一种有效策略，用于为算法 1 中的向量 $u,\ v$ 选择初始值。

回顾一下，在非结构化情形下，(1) 的最小值由 $\Delta = -\sigma_n u_n v_n^*$ 给出，其中$u_n,\ v_n$ 是与矩阵 $A$ 的最小奇异值 $\sigma_n$ 相对应的左右奇异向量。我们可以将这个非结构化极小化点用作结构化问题的初始点，令 $u_{\text{initial}} = -\sigma_n u_n,\ v_{\text{initial}} = v_n$，但结果表明，对于缩放系数而言，通常存在比$\sigma_n$ 更好的选择。接下来，我们将介绍一种方法，用来为如下形式的初始值选择 $\sigma$

$$u_{\text{initial}} = -\sigma u_n, \quad v_{\text{initial}} = v_n, \tag{33}$$

其灵感来源于第 2.1 小节中的思路。

根据第 2.1 小节，我们考虑以下泛函，并关于 $E$ （具有单位弗罗贝尼乌斯范数）对其进行极小化：

$$\widetilde{F}_\varepsilon(E) = \sigma_n(A+\varepsilon E), \qquad \text{where} \qquad \sigma_n(A+\varepsilon E) = \sigma_{\min}(A+\varepsilon E).$$

~ ~回顾一下，$\Delta = \Pi_{\mathcal{S}}(uv^*) = \varepsilon E$，其中 $\|E\|_F = 1$。令 $G$ 为 $E \mapsto F_\varepsilon(E)$在$\sim\varepsilon = 0$处的梯度，即 $G = \Pi_{\mathcal{S}}(u_n v_n^*)$，其中 $u_n$ 和 $v_n$ 是与最小奇异值 $\sigma_n(A)$ 相对应的左右奇异向量。我们假设该奇异值是单重且非零的。

我们希望近似求出以下方程的最小解

$$\varphi(\varepsilon) := \widetilde{F}_\varepsilon(E(\varepsilon)) = 0,$$

其中， $E(\varepsilon)$表示梯度系统 (6) 的一个驻点，因此也是在单位弗罗贝尼乌斯范数矩阵集合上，针对给定的 $\varepsilon \geq 0$ ，使 $\sim F_\varepsilon(E)$ 取得极小值的一个（局部）极小化点。显然，在 $\varepsilon = 0$ 处， $E$ 不起任何作用，且 $\varphi(0) = \sigma_n(A)$。为了计算一个有效的初始值 $\varepsilon = \varepsilon_0$，我们从值 $\varepsilon = \varepsilon_{-1} = 0$ 出发，形式上对方程 (4) 应用牛顿步。

对于 $k = -1$ 且 $\varepsilon_{-1} = 0$，牛顿迭代给出

$$\varepsilon_{\text{initial}} := \varepsilon_0 = 0 - \frac{\varphi(0)}{\varphi'(0)} = -\frac{\sigma_n(A)}{\varphi'(0)}. \tag{34}$$

假设 $E(\varepsilon)$ 关于 $\varepsilon$ 是光滑的，我们使用 [10, 定理 IV.1.4]以显式形式表示$\varphi(\varepsilon)$ 的导数（对于小于结构化奇异距离的 $\varepsilon$ ，即满足 $\varphi(\varepsilon) > 0$），即

$$\varphi'(\varepsilon) = -\|\widetilde{G}(\varepsilon)\|_F = -\|\Pi_{\mathcal{S}}\left(\hat{u}(\varepsilon)v(\varepsilon)^*\right)\|_F,$$

w这里我们回顾一下，$\hat{}u = u/\|u\|$。因此，在 $\varepsilon = 0$ 处，其中 $\hat{}u = u_n$ 且 $v = v_n$，我们有

$$\varphi'(0) = -\|\Pi_{\mathcal{S}}\left(u_n v_n^*\right)\|_F.$$

因此，由（34）可得

$$\varepsilon_0 = \frac{\sigma_n(A)}{\|\Pi_{\mathcal{S}}\left(u_n v_n^*\right)\|_F}; \tag{35}$$

我们将此 $\varepsilon_0$ 值用作初始扰动 $\Delta$ 的范数。因此，我们在（33）中选择 $\sigma$ ，使得 $\|\Delta\| = \varepsilon_0$:

$$\sigma = \frac{\varepsilon_0}{\|\Pi_{\mathcal{S}}\left(u_n v_n^*\right)\|_F} = \frac{\sigma_n}{\|\Pi_{\mathcal{S}}\left(u_n v_n^*\right)\|_F^2}. \tag{36}$$

注释 3. （36）中的起始值启发式 $\sigma$ 也可以通过另一种论证得到：对于形如（33）的初始值，我们选择使 $A + \Delta$ 与 $u_n v_n^*$ 正交（在弗罗贝尼乌斯内积下）的 $\sigma$ 值。事实上，如果将 $\Delta = \Pi_{\mathcal{S}}(u_{\text{initial}} v_{\text{initial}}^*) = -\sigma\Pi_{\mathcal{S}}(u_n v_n^*)$ 代入$\langle A + \Delta,\ u_n v_n^* \rangle = 0$ 并求解 $\sigma$，我们得到

$$\sigma = \frac{u_n^* A v_n}{\langle \Pi_{\mathcal{S}}(u_n v_n^*), u_n v_n^* \rangle} = \frac{\sigma_n}{\|\Pi_{\mathcal{S}}(u_n v_n^*)\|_F^2},$$

利用 $\Pi_{\mathcal{S}}$ 是正交投影这一事实。假设 $A \in \mathcal{S}$， $\sigma$ 的选择也可以从几何角度解释如下：我们选择 $\sigma$ ，使得 $A + \Delta$ 与结构化最速下降方向（结构化负梯度） $\Pi_S(u_n v_n^*)$ 正交，即，

$$\langle A + \Delta,\ \Pi_{\mathcal{S}}(u_n v_n^*) \rangle = 0.$$

由于 $A \in \mathcal{S}$ 和 $\Pi_{\mathcal{S}}$ 是到 $\mathcal{S}$ 上的正交投影算子，我们有 $\langle A,\ \Pi_{\mathcal{S}}(u_n v_n^*) \rangle = \langle A,\ u_n v_n^* \rangle = \sigma_n$.因此，初始扰动是通过消去 $A$ 在梯度方向 $\Pi_{\mathcal{S}}(u_n v_n^*)$ 上的分量而得到的。

### 多个初始值

虽然当仅使用单个初始点时，上述为 $u$和 $v$ 选择的初始值是合理的，但在某些问题中，可能有必要对若干初始值进行测试运行，以降低

陷入局部极小值的风险。事实上，可能存在这样的例子：非奇异矩阵$A$ 具有一组量级较小且彼此相当的奇异值。在此类情形下，加入结构可能会使优化问题(1)出现多个量级相当的局部极小值；而(33)–(36)中的初始选择并不能保证收敛到全局最优解。因此，我们建议针对初始向量 $u_{\text{initial}}$和 $v_{\text{initial}}$的若干选择，运行算法 1中的方法。具体而言，我们选取与第 $n-k+1$小的奇异值 $\sigma_{n-K+1}$、$k=1,\dots,K$相关联的 $K$ 对左右奇异向量 $u_{n-k+1}$ 和 $v_{n-k+1}$ ，并将该方法运行 $K$次，每次选择（对于每个 $k$）

$$u_{\text{initial}} = \widehat{\sigma}_{n-k+1} u_{n-k+1}, \quad v_{\text{initial}} = v_{n-k+1},$$

with

$$\widehat{\sigma}_{n-k+1} = \frac{\sigma_{n-k+1}}{\left\| \Pi_{\mathcal{S}} \left( u_{n-k+1} v_{n-k+1}^* \right) \right\|_F^2} \tag{37}$$

计算得到的最终扰动 $\Delta_* \in \mathcal{S}$（用于确定 $A+\Delta_*$ 的奇异性）随后被选为 $K$ 个计算结果中弗罗贝尼乌斯范数最小的一个。

一个解得益于多重初始化的示例将在随后的第5.2节中给出。

作为上述方案的一种计算代价更低的替代，可以选择

$$u_{\text{initial}} = \widehat{\sigma}_{n-\ell+1} u_{n-\ell+1}, \quad v_{\text{initial}} = v_{n-\ell+1}, \qquad \text{with} \qquad \ell = \operatorname*{arg\,min}_{1 \le k \le K} \widehat{\sigma}_{n-k+1},$$

从而避免多次应用所提出的方法。

# 5 数值实验

在本节中，我们报告了若干数值实验，这些实验使用了可在 https://github.com/fph/NearestSingularAsSyst em/ 获取的本方法 Matlab 实现。所报告的计时结果对应于一台配备英特尔酷睿 i5-1135G7 并运行 Matlab R2025b 的笔记本电脑。

## 5.1 大规模矩阵

为说明新方法的数值效率，我们采用了 [6, 11] 中的一个实验：在保持相同稀疏模式结构的情况下，寻找与 SuiteSparse 矩阵集合中的矩阵 orani678 最近的奇异矩阵。这是一个真实的非对称 $2529 \times 2529$稀疏矩阵，包含 90158 个非零元素；在黎曼预言机方法中，这是一个对于 $M(v)$ 极小化点出现在秩下降点的情形，因此需要进行正则化。我们针对该问题运行算法 1，并利用其稀疏性：使用 svds 计算与最小奇异值对应的左右奇异向量，将其作为初始点；然后使用具有非常宽松容差 $10^{-2}$ 的 minres 求解系统（32）。

该方法经过 5 次迭代（在 minres 中每次迭代平均进行 2083.6 次矩阵-向量乘积）后，收敛到矩阵 $A+\Delta$，其 $\sigma_{\min}(A+\Delta) \approx 1.5481 \times 10^{-13}$和 $\sigma_{\max}(A+\Delta) \approx 3.20 \times 10^{1}$。该算法耗时不到 3 秒：与 [6] 和 [11], 中的方法相比，速度优势显著；后两者在同一测试机器上求解该问题耗时超过 30 秒。计算得到的极小值至少在 7 位有效数字上相符。

## 5.2 需要多个初始值的情形

我们在矩阵 $C$ 上测试该算法，该矩阵可从 https://github.com/fph/NearestSingularAsSystem/blob/main/example_starting_value.mat 获取。这是一个非对称 $50 \times 50$ 矩阵，非零元素密度为 50%；它由随机生成的正交矩阵经稀疏化得到。选择这个特定矩阵，是因为最小奇异值及其向量 $\sigma_n$，$u_n$，$v_n$ 并非产生式 (35) 中最小 $\varepsilon_0$ 的那些。表1列出了 $C$的 5 个最小奇异值，以及以它们为初始值所得局部极小值 $\|\Delta_*\|_F$ 的大小。我们看到，最小范数是用第二小的

| $k$ | $\sigma_k$ | $\|\Delta_*\|_F$ |
|---|---|---|
| 46 | 0.1552 | 0.2321 |
| 47 | 0.1015 | 0.1489 |
| 48 | 0.0574 | 0.0793 |
| 49 | 0.0401 | **0.0571** |
| 50 | 0.0389 | 0.0639 |

表1：从最小的 5 个奇异值及其对应向量出发，按照第4.3节的方法构造初始值后，得到的 $\|\Delta_*\|_F$ 关于 $G_\beta(u,\ v) = 0$ 的解的取值，其中 $C$。

奇异值，而非最小奇异值。此示例表明，使用多个初始值可以得到更好的解。

由于初始矩阵较小，因此使用直接 $O(n^3)$ 方法计算奇异值分解，并求解式 (32) 中的线性方程组。将算法 (1) 使用 5 个不同的初始值重复运行 5 次，运行时间不到 0.02 秒。

## 5.3 多项式 ε-最大公因式计算示例

我们还用另一个出现在 $[6, 24]$ 中的示例测试了新方法：给定多项式

$$p(x) = \gamma_p \prod_{j=1}^{10} (x - \alpha_j), \quad q(x) = \gamma_q \prod_{j=1}^{10} (x - \alpha_j + 10^{-j}), \quad \alpha_j = (-1)^j \tfrac{j}{2}, \tag{38}$$

我们寻找最小扰动 $\tilde{p}$，$\tilde{q}$ 使多项式对的最大公因式次数达到指定次数 $d$。(38) 中出现的归一化系数 $\gamma_p$ 和 $\gamma_q$ 的选取方式是，使 $p$和 $q$ 的系数向量的欧几里得范数为 1。

该问题可以表述为西尔维斯特矩阵$A_S$ 的奇异性距离问题，定义为

$$A_S = \begin{bmatrix} \frac{1}{\sqrt{\deg(q)-d+1}} \mathcal{T}_p & \frac{1}{\sqrt{\deg(p)-d+1}} \mathcal{T}_q \end{bmatrix}, \tag{39}$$

| 式 (38) 的近似最大公因式 | | | | | |
|---|---|---|---|---|---|
| $d$ | 算法 1 | Riemann-Oracle | $VP_S$ [24] | $VP_g$, $VP_h$ [24] | UVGCD |
| 9 | $3.9964 \cdot 10^{-3}$ | $3.9964 \cdot 10^{-3}$ | $4.00 \cdot 10^{-3}$ | $4.00 \cdot 10^{-3}$ | $3.996 \cdot 10^{-3}$ |
| 8 | $1.7288 \cdot 10^{-4}$ | $1.7288 \cdot 10^{-4}$ | $1.73 \cdot 10^{-4}$ | $1.73 \cdot 10^{-4}$ | $1.729 \cdot 10^{-4}$ |
| 7 | $7.0890 \cdot 10^{-6}$ | $7.0890 \cdot 10^{-6}$ | $\underline{3.93 \cdot 10^{-4}}$ | $7.09 \cdot 10^{-6}$ | $7.089 \cdot 10^{-6}$ |
| 6 | $1.8293 \cdot 10^{-7}$ | $1.8293 \cdot 10^{-7}$ | $\underline{2.7 \cdot 10^{-5}}$ | $1.83 \cdot 10^{-7}$ | $1.829 \cdot 10^{-7}$ |
| 5 | $\underline{2.0201} \cdot 10^{-9}$ **(*)** | $4.4\underline{913} \cdot 10^{-9}$ | $\underline{1.72 \cdot 10^{-5}}$ | $4.49 \cdot 10^{-9}$ | $4.487 \cdot 10^{-9}$ |
| 4 | $\underline{2.7776} \cdot 10^{-11}$ **(*)** | $\underline{1.8293 \cdot 10^{-7}}$ | $\underline{1.55 \cdot 10^{-14}}$ | $8.40 \cdot 10^{-11}$ | $8.40 \cdot 10^{-11}$ |

表 2：多项式对 $p$，$q$ 到具有指定次数 $d$的最大公因式 $g$ 的多项式对的计算距离 $\|(p - gu, q - gw)\|$，以及与 [6,24] 和来自 [18] 的 UVGCD 结果的比较。并非所有来源都报告相同位数的有效数字。第一列中以星号 **(*)** 标注的结果，是所得 $\Delta$ 无法用于显式构造三元多项式组 $g$，$u$，$w$ 的情形，因而表明出现数值失败。带下划线的数字是与 UVGCD 结果不同的数字，我们认为 UVGCD 的结果最为可靠。

由两个具有托普利兹结构的块组成

$$\mathcal{T}_p = \underbrace{\begin{bmatrix} p_0 & & & \\ p_1 & p_0 & & \\ \vdots & p_1 & \ddots & \\ p_{10} & \vdots & \ddots & p_0 \\ & p_{10} & \ddots & p_1 \\ & & \ddots & \vdots \\ & & & p_{10} \end{bmatrix}}_{\deg(q) \ -\ d+1 \text{ 列}}, \quad \mathcal{T}_q = \underbrace{\begin{bmatrix} q_0 & & & \\ q_1 & q_0 & & \\ \vdots & q_1 & \ddots & \\ q_{10} & \vdots & \ddots & q_0 \\ & q_{10} & \ddots & q_1 \\ & & \ddots & \vdots \\ & & & q_{10} \end{bmatrix}}_{\deg(p) \ -\ d+1 \text{ 列}}.$$

更多细节请参见 [6,第 7]节。该问题在数值上颇具挑战性，因为对于 $d \leq 5$，目标函数 $\|\Delta\|_F^2$ 小于机器精度 $\approx 10^{-16}$。

我们报告了新方法得到的结果，并在 [6, 24], 表 2 中进行了比较。结果表明，该方法能够获得可靠的结果，直到 $\|\Delta\|_F^2$ 小于机器精度。与 [24, 表 1] 中的结果进行比较很有意思：对于多项式最大公因式问题的相同矩阵形式 $A_S$（[24, 表 1]中的第 $VP_S$ 列），新方法的表现优于 SLRA；然而，当 SLRA 使用该问题的另一种矩阵形式（第 $VP_g$、$VP_h$列）时，其结果更好，而这种形式不属于第 2.2 节所述的完全相同的框架。[24] 中 $VP_g$、$VP_h$ 的结果与最先进的专用算法的结果一致（第 $UVGCD$列；[18]报告了更多有效数字）。总的来说，结果表明，新求解算法能够可靠地工作，直到 $\|\Delta\|_F^2$ 小于机器精度：在相同矩阵 $A_S$上运行时，其精度与 [6] 中的方法相当，且略高于 [24], 中的方法。

# 6 结论

我们介绍并研究了一种用于数值近似结构化奇异距离的新方法。在对两种现有方法进行理论比较后，我们提出了一种独立于现有方法的新颖方法。该技术通过对一组非线性方程使用牛顿法来求解近邻问题，从而避免了 [6] 和 [11]中出现的嵌套优化。我们将该技术应用于大型稀疏矩阵以及具有类托普利兹结构的较小问题。未来的研究方向包括扩展对所提方法的理论分析，并将其推广到稳定性距离问题。

**致谢** 所有作者均隶属于意大利 INdAM-GNCS（国家科学计算组）。FP 感谢国家高性能计算、大数据与量子计算中心—HPC 提供的支持，项目编号为 CUP B83C22002940006，资金来自欧盟—下一代欧盟；同时感谢意大利大学与研究部通过 2022年 PRIN 项目“MOLE：流形约束优化与学习”提供的支持，项目编号为 CUP B53C24006410006。SS 感谢欧盟的支持（ERC 协同资助项目，eLinoR，编号 101085607）。NG 感谢意大利大学与研究部（Ministero dell’Universit`a e della Ricerca）通过 2022年 PRIN 项目“用于时间依赖参数偏微分方程及其应用的先进数值方法”以及 PRIN-PNRR 项目“FIN4GEO”提供的研究经费。他还感谢意大利大学与研究部（MUR）授予格兰萨索科学研究所（GSSI）的 2023—2027 年卓越学科项目提供的资助。

# 参考文献

[1] E. Andreotti、D. Edelmann、N. Guglielmi 和 C. Lubich。谱聚类稳定性的度量。《线性代数及其应用》，610:673–697，2021年。[2] F. H. Clarke。广义梯度及其应用。《美国数学会汇刊》，205:247–262，1975年。[3] A. De Marinis、N. Guglielmi、S. Sicilia 和 F. Tudisco。以最小权重扰动提高神经常微分方程的鲁棒性。*arXiv* 预印本 *arXiv:2501.10740*，2025年。[4] C. Eckart 和 G. Young。用低秩矩阵逼近另一矩阵。《心理测量学》，1(3):211–218，1936年。[5] M. Gnazzo 和 N. Guglielmi。关于矩阵值函数奇异距离的数值逼近。*SIAM* 矩阵分析与应用杂志，46(2):1484–1517，2025年。[6] M. Gnazzo、V. Noferini、L. Nyman 和 F. Poloni。Riemann-Oracle：一种用于求解矩阵理论中邻近问题的通用黎曼优化器。《计算数学基础》，2025年。[7] G. H. Golub 和 V. Pereyra。伪逆的微分以及变量可分离的非线性最小二乘问题。*SIAM* 数值分析杂志，10(2):413–432，1973年。

[8] N. Guglielmi、D. Kressner 和 C. Lubich。辛伪谱极值点的计算与辛矩阵逼近问题的求解。*SIAM* 矩阵分析与应用杂志，35(4):1407–1428，2014年。[9] N. Guglielmi、M. L´opez-Fern´andez 和 M. Manucci。对流扩散方程的伪谱游走轮廓积分方法。科学计算杂志，89(1)，2021年。[10]N. Guglielmi 和 C. Lubich。矩阵逼近问题与特征值优化。arXiv:2503.14750，2025年。[11] N. Guglielmi、C. Lubich 和 S. Sicilia。用于结构化特征值优化的秩一矩阵微分方程。*SIAM* 数值分析杂志，61(4):1737–1762，2023年。[12] N. Guglielmi 和 I. Markovsky。一种基于常微分方程的方法，用于计算互素多项式到公因子整除的距离。*SIAM* 数值分析杂志，55(3):1456–1482，2017年。[13] N. Guglielmi、M.-U. Rehman 和 D. Kressner。一种逼近结构奇异值的新型迭代方法。*SIAM* 矩阵分析与应用杂志，38(2):361–386，2017年。[14] N. Guglielmi 和 S. Sicilia。通过低秩自适应常微分方程实现矩阵稳定化。《*BIT* 数值数学》，64(4):38，2024年。[15] N. Guglielmi 和 S. Sicilia。用于谱聚类稳定性的低秩常微分方程。《线性代数及其应用》，721:250–276，2025年。[16] N. J. Higham。矩阵逼近问题及其应用。矩阵理论的应用，会议论文集，布拉德福德/英国 1988年，数学及其应用研究所会议丛书，新辑 22，1-27（1989年）。，1989年。[17] I. Markovsky 和 K. Usevich。加权结构化低秩近似软件。计算与应用数学杂志，256:278–292，2014年。[18] K. Nagasaka。迈向单变量多项式近似最大公因式的最佳算法。符号计算杂志，105:4–27，2021年。[19] J. Nocedal 和 S. J. Wright。数值优化。施普林格，纽约，第2版，2006年。[20] V. Noferini、L. Nyman 和 F. Poloni。通过黎曼优化求具有重特征值的最近矩阵。预印本，arXiv:2509.26344 [math.NA](2025)，2025年。[21] E. 施密特。关于线性和非线性积分方程理论。数学年刊，63(4):433–476，1907年。[22] S. Sicilia。结构矩阵邻近问题中的低秩性质。博士论文，格兰萨索科学研究所，2025年1月。可于 https://iris.gssi.it/handle/20.500.12571 /33684?mode=simple 获取。[23] K. Usevich 和 I. Markovsky。加权二范数下仿射结构化低秩逼近的变量投影。计算与应用数学杂志，272:430–448，2014年。[24] K. Usevich 和 I. Markovsky。用于近似（最大）公因式计算的变量投影方法。《理论计算机科学》，681:176–198，2017年。