arXiv:2304.00304v1 [math.NA] 1 Apr 2023

子空间正交规范基矩阵的变化

Zhongming Teng* Ren-Cang Li†

2023年4月4日

摘要

已知子空间 $\mathcal{X}$ 的正交归一基矩阵 X 并非唯一，除非施加某种规范化要求。其中一种要求是令$X^{\mathrm{T}}D$ 为半正定矩阵，其中 D 是一个维数适当的常矩阵。在多视图子空间学习模型中，这是一种自然的要求；在此类模型中， X 作为投影矩阵，并通过 Stiefel 流形上的最大化问题确定，而该问题的目标函数包含 tr($X^{\mathrm{T}}D$)，且随其增大。本文研究在 $X^{\mathrm{T}}D$ 保持半正定的要求下，当子空间 $\mathcal{X}$ 变化时，正交归一基矩阵 X 的变化界。所得结果有助于对求解该最大化问题的 NEPv 方法（具有特征向量依赖的非线性特征值问题）进行收敛性分析。

关键词：酉不变范数、典范角、子空间、正交归一基矩阵、Stiefel 流形上的优化。

数学主题分类（2010）。15A45，65F35

1 引言

Rece在 [19], Stiefel 流形 $\mathbb{O}^{n\times k}$ 上的以下优化问题

$$\max_{X\in\mathbb{O}^{n\times k}} f(X) \quad \text{with } f(X) := \phi(X) + \psi(X)\times \mathrm{tr}(X^{\mathrm{T}}D) \tag{1.1}$$

其中 D 是常数矩阵，Stiefel 流形

$$\mathbb{O}^{n\times k} = \{X\in\mathbb{R}^{n\times k} : X^{\mathrm{T}}X = I_k \text{ (the } k\times k \text{ identity matrix)}\},$$

ϕ 和 ψ 是酉不变函数，即对于任意 $Q\in\mathbb{O}^{k\times k}$ 和 $\psi(X) > 0$， $\phi(XQ) \equiv \phi(X)$，并且 $\psi(XQ) \equiv\psi(X)$。这是对几个源于子空间学习 [26, 27, 32, 34]，的具体问题的抽象，其中目标函数包含 tr($X^{\mathrm{T}}D$)，并随 tr($X^{\mathrm{T}}D$) 的增加而增加。沿着 [27, 32, 34], 早期研究的思路，作者们

*福建农林大学计算机与信息科学学院，福州，350002，中华人民共和国。电子邮箱：zhmteng@fafu.edu.cn

†数学系，阿灵顿德克萨斯大学，阿灵顿，德克萨斯州 76019-0408，美国。电子邮箱：rcli@uta.edu。部分受美国国家科学基金会 DMS-2009689 项目资助。

[19] 首先将（1.1）的 KKT 条件转化为一个 NEPv（具有特征向量依赖的非线性特征值问题）

$$H(X)X = X\Omega, \quad X \in \mathbb{O}^{n\times k}, \tag{1.2}$$

其中 $H(X)$ 关于 X 对称。此类形式的非线性特征值问题并非新问题；事实上，在 [30, 31] 首次通过非线性特征值问题求解正交线性判别分析（OLDA）之前，它们大多源于求解密度泛函理论中的离散化 Kohn-Sham 方程 [11, 14, 23, 29]。在数值计算中，非线性特征值问题 (1.2) 通常采用所谓的自洽场（SCF）迭代求解。在 [19], 中，非线性特征值问题 (1.2) 通过自洽场类型的迭代求解：给定 $X_0 \in \mathbb{O}^{n\times k}$,

$$\text{iteratively solve } H(X_{i-1})\widehat{X}_i = \widehat{X}_i\Omega_i \text{ for } \widehat{X}_i \text{ which is postprocessed to get } X_i, \tag{1.3}$$

直至收敛；后处理得到 $X_i = \widehat{X}_i Q_i$ ，其中存在某个 $Q_i \in \mathbb{O}^{k\times k}$ ，使得$X_i^{\mathrm{T}} D \succeq 0$ （正半定）。Q_i 通常取为 $X_i^{\mathrm{T}} D$[19, 27, 32, 34], 的正交极分解因子，这是因为 $f(X)$ 关于 $\mathrm{tr}(X^{\mathrm{T}}D)$ 单调递增，并且$Q_* \in \mathbb{O}^{k\times k}$ 使得 $(XQ_*)^{\mathrm{T}}D = Q_*^{\mathrm{T}}(X^{\mathrm{T}}D)$ 为正半定，可确保 [19, 27]

$$\mathrm{tr}([XQ_*]^{\mathrm{T}}D) = \mathrm{tr}(Q_*^{\mathrm{T}}[X^{\mathrm{T}}D]) = \max_{Q\in\mathbb{O}^{k\times k}} \mathrm{tr}(Q^{\mathrm{T}}[X^{\mathrm{T}}D]) \geq \mathrm{tr}(X^{\mathrm{T}}D), \tag{1.4}$$

并且，当 $X^{\mathrm{T}}D \not\succeq 0$ 时，不等式是严格的。

SCF 型迭代（1.3）与用于求解离散 Kohn-^Sham 方程的经典自洽场的区别在于其从 X_i 到 X_i 的后处理，而对于具有酉不变性的非线性特征值问题的经典自洽场而言，这一步并不需要，即对于任意 $Q \in \mathbb{O}^{k\times k}$，均有 $H(XQ) \equiv H(X)$。在 [27, 32, 34] 中， [19],SCF 型（1.3）首次出现。在不会引起混淆的情况下，我们通常不加区分地使用 SCF 来指代经典自洽场和 SCF 型迭代。

（1.4）的一个直接推论是，对于最大化问题（1.1）的任意最大化解 X_* ，都有 $X_*^{\mathrm{T}}D \succeq 0$ 。最大化解 X_* 的另一个重要刻画是 [19,定理 3.1]

$$\mathrm{rank}(X_*^{\mathrm{T}}D) = \mathrm{rank}(D). \tag{1.5}$$

因此，对于任意 $X \in \mathbb{O}^{n\times k}$ ，只要 X 的列空间（记为 $\mathcal{R}(X)$）充分接近 $\mathcal{R}(X_*)$，就有 $\mathrm{rank}(X^{\mathrm{T}}D) = \mathrm{rank}(D)$ [19, 引理 5.1] 这意味着，当 $\mathcal{R}(X)$ 接近 $\mathcal{R}(X_*)$ [10, 15] 时， $X^{\mathrm{T}}D$ 的典范正交极分解因子具有连续性。

作为一种迭代方案，SCF 型迭代 (1.3) 的一个关键问题是，所生成的近似序列是否会收敛到预期目标。对于涉及优化问题 (1.1) 的情形，该目标是其最大化解之一。受技术限制，现有关于收敛的结果实际上讨论的是子空间收敛，即 $\mathcal{R}(X_i)$ 收敛到某个 k 维子空间 $\mathcal{X}_* := \mathcal{R}(X_*)$，当然其精确的 X_* 未知。换言之，现有结果或许能保证 $\mathcal{R}(X_i)$ 收敛到 $\mathcal{X}_*$ ，并给出收敛时子空间 $\mathcal{R}(X_i)$ 与 $\mathcal{X}_*$ 之间距离的估计，但无法给出 $\|X_i - X_*\|$ 的界，其中 $\|\cdot\|$ 是某种矩阵范数。对于 OLDA 或任何酉不变的目标函数 f ，我们能做到的最好程度就是如此，因为如果 X_* 是一个最优解，那么对于任意 $Q \in \mathbb{O}^{k\times k}$， X_*Q 也是最优解；但对于 (1.1) 中的 f ，只要 $\mathrm{rank}(X_*^{\mathrm{T}}D) = k$，就可以证明其最优解 X_* 在轨道内是唯一的

$$\mathbb{X}_* := \{X_*Q : Q \in \mathbb{O}^{k\times k}\} \tag{1.6}$$

其元素共享同一个子空间 $\mathcal{X}_* = \mathcal{R}(X_*)$，但当 $\mathrm{rank}(X_*^{\mathrm{T}} D) < k$ [27, 19] 时，X_* 仅部分唯一。请注意，对于所有 $X \in \mathbb{X}_*$，$\mathrm{rank}(X^{\mathrm{T}} D)$ 都是常数，与 $\mathcal{X}_*$ 的任何特定正交归一基矩阵无关。鉴于上述讨论，我们可以将 D 视为某种决策器，用来从轨道 (1.6) 中选出特定的 X_*。

本文旨在回答以下数学问题：

> 给定 $\mathbb{R}^n$ 中两个 k 维子空间 $\mathcal{X}$ 和 $\widetilde{\mathcal{X}}$，分别设 X、$\widetilde{X} \in \mathbb{O}^{n\times k}$ 为它们的标准正交基矩阵，使得 $X^{\mathrm{T}} D$、$\widetilde{X}^{\mathrm{T}} D \succeq 0$，并假设 $\mathrm{rank}(X^{\mathrm{T}} D) = \mathrm{rank}(\widetilde{X}^{\mathrm{T}} D)$。如何用子空间 $\mathcal{X}$ 和 $\widetilde{\mathcal{X}}$ 之间的差异来界定 X 和 $\widetilde{X}$ 之间的差异 $\widetilde{}$?

(1.7)

关于刚才提出的自洽场中 X_i 的收敛问题，我们的主要结果可通过子空间 $\mathcal{R}(X_i)$ 与 $\mathcal{X}_*$ 之间的距离来界定 $\|X_i - X_*\|$，其中令 $X = X_*$ 和 $\widetilde{X} = X_i$。第一个条件 $X_i^{\mathrm{T}} D \succeq 0$ 按设计成立，而 $X_*^{\mathrm{T}} D \succeq 0$ 是最大化解的必要条件；收敛附近 $\mathrm{rank}(X_i^{\mathrm{T}} D) = \mathrm{rank}(X_*^{\mathrm{T}} D)$ 则源于 [19, 引理 5 的 (1.5)。1]

记号。$\mathbb{R}^{m\times n}$ 表示 $m \times n$ 实矩阵的集合，$\mathbb{R}^n = \mathbb{R}^{n\times 1}$ 和 $\mathbb{R} = \mathbb{R}^1$。$I_n \in \mathbb{R}^{n\times n}$ 表示单位矩阵。对于 $B \in \mathbb{R}^{m\times n}$，$\mathcal{R}(B)$表示由其列张成的列空间，其奇异值记为 $\sigma_i(B)$，其中 $i = 1$, ..., $\min(m, n)$按非增顺序排列，并且

$$\|B\|_2 = \sigma_1(B),\ \|B\|_{\mathrm{F}} = \sqrt{\sum_{i=1}^{\mathrm{rank}(B)} [\sigma_i(B)]^2},\ \|B\|_{\mathrm{tr}} = \sum_{i=1}^{\mathrm{rank}(B)} \sigma_i(B)$$

分别是 B的谱范数、Frobenius 范数和迹范数。B^{T} 表示 B的转置。迹范数也称为核范数。对于对称 $A \in \mathbb{R}^{n\times n}$，$A \succ 0$ ($\succeq 0$)表示 A 是正定（半正定）的。使用类似 MATLAB 的记号访问矩阵或向量的元素：$X_{(i:j,k:l)}$ 表示矩阵 X的子矩阵，由第 i 至 j 行与第 k 至 l列的交集组成；当 i: j 被替换为 : 时，表示所有行。

2 预备知识

本节汇集了一些已知结果，以供后续推导使用。

2.1 子空间之间的典范角

给定 $\mathbb{R}^n$ 中两个 k 维子空间 $\mathcal{X}$ 和 $\widetilde{\mathcal{X}}$，设它们的标准正交基矩阵分别为 $X \in \mathbb{O}^{n\times k}$ 和 $\widetilde{X} \in \mathbb{O}^{n\times k}$，即，

$$X^{\mathrm{T}} X = I_k,\ \mathcal{R}(X) = \mathcal{X} \quad \text{and} \quad \widetilde{X}^{\mathrm{T}} \widetilde{X} = I_k,\ \mathcal{R}(\widetilde{X}) = \widetilde{\mathcal{X}}.$$

记 ω_i $1 \le i \le k$ 为 $X^{\mathrm{T}} \widetilde{X}$ 的奇异值，并按降序排列，即 $\omega_1 \ge \cdots \ge \omega_k$。

~ ~ k 与 $\mathcal{X}$和 $\mathcal{X}$之间的典范角 $\theta_i(\mathcal{X}, \mathcal{X})$ 定义为

$$0 \le \theta_i(\mathcal{X}, \widetilde{\mathcal{X}}) := \arccos(\omega_{k-i+1}) \le \frac{\pi}{2}, \quad \text{for } 1 \le i \le k.$$

Set

$$\Theta(\mathcal{X},\widetilde{\mathcal{X}})=\operatorname{diag}\left(\theta_1(\mathcal{X},\widetilde{\mathcal{X}}),\ldots,\theta_k(\mathcal{X},\widetilde{\mathcal{X}})\right)\in\mathbb{R}^{k\times k}. \tag{2.1}$$

~可以看出，(2.1) 中的角度矩阵 $\Theta(\mathcal{X},\mathcal{X})$ 与为 $\mathcal{X}$ 和 $\mathcal{X}$选取正交规范~基矩阵的方式无关。

本文中，我们所指的任何酉不变范数 $\|\cdot\|_{\mathrm{ui}}$ [24] 均假定为与维数无关，即可一致地应用于任意大小的矩阵，例如矩阵谱范数和弗罗贝尼乌斯范数。稍微放宽这一要求，我们也可以将酉不变范数限制为由对称规范函数 Φ 在 $\mathbb{R}^N$ [24, 第二节4]上诱导的范数，其中 N 足够大，使所有感兴趣的矩阵 B 的行数和列数均不超过 N，然后令 [24, p。79]

$$\|B\|_{\mathrm{ui}}=\Phi(\sigma_1(B),\ldots,\sigma_r(B),0,\ldots,0),$$

即在 B 的奇异值集合中补入 0，使其包含 N 个奇异值。已知对于大小相容的矩阵 A、B 和 C，有

$$\|ABC\|_{\mathrm{ui}}\le\|A\|_2\|B\|_{\mathrm{ui}}\|C\|_2. \tag{2.2}$$

Sun [25, 第 95] 页证明，对于任意酉不变范数 $\|\cdot\|_{\mathrm{ui}}$，$\|\sin\Theta(\mathcal{X},\mathcal{Y})\|_{\mathrm{ui}}$在由所有 k维子~空间构成的格拉斯曼流形上定义了一种酉不变度量。处理 sin (,) 的一种便捷方法如下。令 ，使得 且 ，分别。则

$$\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{ui}}=\|X_\perp^{\mathrm{T}}\widetilde{X}\|_{\mathrm{ui}}=\|\widetilde{X}_\perp^{\mathrm{T}}X\|_{\mathrm{ui}}. \tag{2.3}$$

引理 2.1（[33, 引理 4.1]）。存在一个正交矩阵 $Q\in\mathbb{O}^{k\times k}$，使得

$$\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{ui}}\le\|X-\widetilde{X}Q\|_{\mathrm{ui}}\le\sqrt{2}\,\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{ui}}. \tag{2.4}$$

2.2 奇异值分解扰动

对于任意大小合适的矩阵 B，我们用 $B_{\text{best-}r}$ 表示其最佳秩-r 近似，即在其奇异值分解中将除前 r 个最大奇异值之外的所有奇异值置零后得到的矩阵。利用 Fan 定理 [24, 第 86], 页可以证明，对于任意酉不变范数 $\|\cdot\|_{\mathrm{ui}}$，$\|(\cdot)_{\text{best-}r}\|_{\mathrm{ui}}$对于 $r\ge1$ 也同样是酉不变范数。(2.2) 中的一致性不等式可以稍作加强:

$$\|ABC\|_{\mathrm{ui}}\le\|A\|_2\|B_{\text{best-}r}\|_{\mathrm{ui}}\|C\|_2\quad\text{if }\min\{\operatorname{rank}(A),\operatorname{rank}(B),\operatorname{rank}(C)\}\le r. \tag{2.5}$$

The 下一个引理是由经典的 Wedin 的结果 [28, (3.1)] 和 (2.5) 推出的推论 .

~ ~引理 2.2. 设 B、$\widetilde{B}=B+F\in\mathbb{R}^{m\times n}$ 满足 rank(B) = rank($\widetilde{B}$) $=r$，且它们的奇异值分解为

$$B=U\Sigma V^{\mathrm{T}}\quad and\quad\widetilde{B}=\widetilde{U}\widetilde{\Sigma}\widetilde{V}^{\mathrm{T}}, \tag{2.6}$$

~其中 $\Sigma_{(1:r,1:r)} \succ 0$ 且 $\Sigma_{(1:r,1:r)} \succ 0$。则有

$$\max\left\{\|\sin\Theta(\mathcal{U},\widetilde{\mathcal{U}})\|_{\mathrm{ui}},\ \|\sin\Theta(\mathcal{V},\widetilde{\mathcal{V}})\|_{\mathrm{ui}}\right\} \le \frac{\|F_{\text{best-}r}\|_{\mathrm{ui}}}{\max\{\sigma_r(B),\sigma_r(\widetilde{B})\}} \tag{2.7a}$$

$$\le \frac{\|F\|_{\mathrm{ui}}}{\max\{\sigma_r(B),\sigma_r(\widetilde{B})\}}. \tag{2.7b}$$

~ ~ ~其中 $\mathcal{U}=\mathcal{R}(U_{(:,1:r)})$、$\mathcal{U}=\mathcal{R}(U_{(:,1:r)})$、$\mathcal{V}=\mathcal{R}(V_{(:,1:r)})$ 和 $\mathcal{V}=\mathcal{R}(V_{(:,1:r)})$。

证明。令

$$R := B\widetilde{V}_{(:,1:r)} - \widetilde{U}_{(:,1:r)}\widetilde{\Sigma}_{(1:r,1:r)} = (B-\widetilde{B})\widetilde{V}_{(:,1:r)},$$
$$S := B^{\mathrm{T}}\widetilde{U}_{(:,1:r)} - \widetilde{V}_{(:,1:r)}\widetilde{\Sigma}_{(1:r,1:r)} = (B-\widetilde{B})^{\mathrm{T}}\widetilde{U}_{(:,1:r)}.$$

根据（2.5），我们得到 $\|R\|_{\mathrm{ui}} \le \|F_{\text{best-}r}\|_{\mathrm{ui}}$ 和 $\|S\|_{\mathrm{ui}} \le \|F_{\text{best-}r}\|_{\mathrm{ui}}$。因此，借助经典的 Wedin 的结果 [28, (3.1)]（对于该情形，另见 [17, 事实 4，第 21-7]页），我们有

$$\max\left\{\|\sin\Theta(\mathcal{U},\widetilde{\mathcal{U}})\|_{\mathrm{ui}},\ \|\sin\Theta(\mathcal{V},\widetilde{\mathcal{V}})\|_{\mathrm{ui}}\right\} \le \frac{\max\{\|R\|_{\mathrm{ui}},\|S\|_{\mathrm{ui}}\}}{\sigma_r(\widetilde{B})}$$

$$\le \frac{\|F_{\text{best-}r}\|_{\mathrm{ui}}}{\sigma_r(\widetilde{B})}. \tag{2.8}$$

~在（2.8）中交换 B 和 B 的角色，我们得到

$$\max\left\{\|\sin\Theta(\mathcal{U},\widetilde{\mathcal{U}})\|_{\mathrm{ui}},\ \|\sin\Theta(\mathcal{V},\widetilde{\mathcal{V}})\|_{\mathrm{ui}}\right\} \le \frac{\|F_{\text{best-}r}\|_{\mathrm{ui}}}{\sigma_r(B)}. \tag{2.9}$$

（2.7）中的不等式是（2.8）和（2.9）的推论。 □

2.3 极分解

任意的 $B \in \mathbb{R}^{n\times m}$ $(n \ge m)$ 都可以分解为 $B = QH$，称为极分解 [12,p.449]，其中 $Q \in \mathbb{O}^{n\times m}$，且 $H = (B^{\mathrm{T}}B)^{1/2} \succeq 0$ 是 $B^{\mathrm{T}}B$ 的唯一半正定平方根。已知，当且仅当 $\mathrm{rank}(B) = m$ [17] 时，正交因子 Q 唯一。当 $\mathrm{rank}(B) < m$ 时，存在所谓的规范极分解 $B = QH$，其中 $Q \in \mathbb{R}^{n\times m}$ 是部分等距算子，并满足 $\mathcal{R}(Q^{\mathrm{T}}) = \mathcal{R}(H)$，且再次有 $H = (B^{\mathrm{T}}B)^{1/2}$。在规范极分解中，$Q$ 是唯一的（参见 [2, p.220], [10, 第 8], [15] 章）。

引理 2.3（[16, 定理 1], [18, 定理 3.4]）。设 $B \in \mathbb{R}^{n\times m}$ $(n \ge m)$ 被 B 扰动为 ~ ~，使得 $\mathrm{rank}(B) = \mathrm{rank}(B) = r$。设 B 的奇异值分解为

$$B = \begin{array}{c} \overset{r}{} \quad \overset{n-r}{} \\ \begin{bmatrix} U_1 & U_2 \end{bmatrix} \end{array} \times \begin{array}{c} \overset{r}{} \quad \overset{m-r}{} \\ \begin{bmatrix} \Sigma_1 & \\ & 0 \end{bmatrix} \end{array} \times \begin{array}{c} \\ \begin{bmatrix} V_1^{\mathrm{T}} \\ V_2^{\mathrm{T}} \end{bmatrix} \end{array}, \tag{2.10}$$

~其中 r =$\mathrm{rank}(B)$，类似地，B 的奇异值分解形式与式 (2.10) 相同，只是其中每个符号上都加了波浪号。于是

$$Q = U_1V_1^{\mathrm{T}} \quad and \quad \widetilde{Q} = \widetilde{U}_1\widetilde{V}_1^{\mathrm{T}} \tag{2.11}$$

$\widetilde{\ }$分别是 B 和 B的典范极分解中的唯一部分等距算子因子，并且

$$\|Q-\widetilde{Q}\|_{\mathrm{ui}} \le \begin{cases} \dfrac{2}{\sigma_n(B)+\sigma_n(\widetilde{B})}\|\widetilde{B}-B\|_{\mathrm{ui}}, & \textit{if } r=n=m; \\ \left(\dfrac{2}{\sigma_r(B)+\sigma_r(\widetilde{B})}+\dfrac{2}{\max\{\sigma_r(B),\sigma_r(\widetilde{B})\}}\right)\|\widetilde{B}-B\|_{\mathrm{ui}}, & \textit{otherwise.} \end{cases} \tag{2.12}$$

当$n>m$ 或 $r<n$时，矩阵谱范数和弗罗贝尼乌斯范数的这个不等式可以得到改进：

$$\|Q-\widetilde{Q}\|_{\mathrm{F}} \le \frac{2}{\sigma_r(B)+\sigma_r(\widetilde{B})}\|\widetilde{B}-B\|_{\mathrm{F}}, \tag{2.13a}$$

$$\|Q-\widetilde{Q}\|_2 \le \sqrt{\frac{4}{\left[\sigma_r(B)+\sigma_r(\widetilde{B})\right]^2}+\frac{2}{\left[\max\{\sigma_r(B),\sigma_r(\widetilde{B})\}\right]^2}}\|\widetilde{B}-B\|_2. \tag{2.13b}$$

下一个引理刻画了 $X\in\mathbb{O}^{n\times k}$，使得 $X^{\mathrm{T}}D\succeq 0$ 可分解为两项之和，其中一项仅取决于 $\mathcal{R}(X)$。

引理 2.4（[27, 定理 3.2]）。给定 $\mathbb{R}^n$中的一个 k 维子空间 $\mathcal{X}$，设 $X_\diamond\in\mathbb{O}^{n\times k}$满足 $\mathcal{R}(X_\diamond)=\mathcal{X}$，且令 $r=\mathrm{rank}(X_\diamond^{\mathrm{T}}D)$，其中 $D\in\mathbb{R}^{n\times k}$。设 $X_\diamond^{\mathrm{T}}D$ 的奇异值分解为

$$X_\diamond^{\mathrm{T}}D = \begin{array}{c} \begin{array}{cc} r & k-r \end{array} \\ \begin{bmatrix} U_1 & U_2 \end{bmatrix} \end{array} \times \begin{array}{c} r \\ k-r \end{array} \begin{array}{c} \begin{array}{cc} r & k-r \end{array} \\ \begin{bmatrix} \Sigma_1 & \\ & 0 \end{bmatrix} \end{array} \times \begin{array}{c} r \\ k-r \end{array} \begin{bmatrix} V_1^{\mathrm{T}} \\ V_2^{\mathrm{T}} \end{bmatrix}. \tag{2.14}$$

任何满足 $\mathcal{R}(X)=\mathcal{X}$ 且使得 $X^{\mathrm{T}}D\succeq 0$ 的 $X\in\mathbb{O}^{n\times k}$ 都具有如下形式

$$X = X_\diamond U_1 V_1^{\mathrm{T}} + X_\diamond U_2 W V_2^{\mathrm{T}},$$

其中，第一项虽然由 $X_\diamond$构造而成，却仅取决于 $\mathcal{X}$；而第二项在 $W\in\mathbb{O}^{(k-r)\times(k-r)}$中具有自由度。特别地，若 $r=k$，则给定 $\mathcal{X}$后， $X=X_\diamond UV\mathrm{T}$ 是唯一的。

3 主要结果

在本节中，我们将给出主要结果，回答（1.7）中的问题。给定k维子空间 $\mathcal{X}$ of $\mathbb{R}^n$，考虑所有满足以下条件的 $X\in\mathbb{O}^{n\times k}$

$$\mathcal{R}(X)=\mathcal{X}, \quad X^{\mathrm{T}}D\succeq 0. \tag{3.1}$$

第一个条件 $\mathcal{R}(X)=\mathcal{X}$ 仅仅表示 X 是 $\mathcal{X}$ 的一个正交归一基矩阵，并且它并不唯一；事实上，它具有如下自由度： $k^2-\frac{1}{2}k(k+1)=\frac{1}{2}k(k-1)$。正是第二个刻画 $X^{\mathrm{T}}D\succeq 0$ 将决定正交归一

基矩阵中的哪一个或哪些应当是 $\mathcal{X}$。设 $r = \mathrm{rank}(X^{\mathrm{T}}D)$ 为 $r' = k - r$，并设 $X^{\mathrm{T}}D$的奇异值分解为

$$X^{\mathrm{T}}D = V\Sigma V^{\mathrm{T}} \equiv \begin{array}{c} \begin{matrix} r & r' \end{matrix} \\ \left[\begin{array}{cc} V_1 & V_2 \end{array}\right] \end{array} \times \begin{array}{c} \begin{matrix} r \\ r' \end{matrix} \\ \left[\begin{array}{cc} \Sigma_1 & \\ & 0 \end{array}\right] \end{array} \times \begin{array}{c} \begin{matrix} r \\ r' \end{matrix} \\ \left[\begin{array}{c} V_1^{\mathrm{T}} \\ V_2^{\mathrm{T}} \end{array}\right] \end{array}. \tag{3.2}$$

根据引理 2.4，我们知道 $Y \in \mathbb{O}^{n\times k}$ 使得 $\mathcal{R}(Y) = \mathcal{X}$且 $Y_{\mathrm{T}}D \succeq 0$ 当且仅当 if

$$Y \in \mathbb{X} := \left\{ XV \begin{bmatrix} I_r & \\ & W \end{bmatrix} V = XV_1V_1^{\mathrm{T}} + XV_2WV_2^{\mathrm{T}} \,:\, W \in \mathbb{O}^{r'\times r'} \right\}. \tag{3.3}$$

当 $\mathrm{rank}(X^{\mathrm{T}}D) = k$ 时，第二项 $XV_2WV_{\mathrm{T}2}$ 消失，且 $\mathbb{X} = \{X\}$，这意味着 X 是唯一的。但当 $\mathrm{rank}(X^{\mathrm{T}}D) < k$ 时， $\mathbb{X}$ 由矩阵变量 $W \in \mathbb{O}^{r'\times r'}$ 参数化，其自由度为$\frac{1}{2}r'(r'-1)$。由于 $\mathbb{X}$ 中的任意元素都可以作为 X，因此 X 并非由 $X^{\mathrm{T}}D \succeq 0$ 唯一确定。值得强调的是，式 (3.3) 中的 $XV_1V_1^{\mathrm{T}}$ 仅取决于 $\mathcal{X}$，尽管它是在 $\mathcal{X}$ 的特定正交归一基矩阵 X 的帮助下构造的。

对于满足以下条件的 $\widetilde{X} \in \mathbb{O}^{n\times k}$，也可以作同样的论述：

$$\mathcal{R}(\widetilde{X}) = \widetilde{\mathcal{X}}, \quad \widetilde{X}^{\mathrm{T}}D \succeq 0. \tag{3.4}$$

鉴于此，只有在 $\mathrm{rank}(X^{\mathrm{T}}D) = \mathrm{rank}(\widetilde{X}^{\mathrm{T}}D) = k$ 时，对 $\|X - \widetilde{X}\|$ 进行界定才有意义；而当 $\mathrm{rank}(X^{\mathrm{T}}D) = \mathrm{rank}(\widetilde{X}^{\mathrm{T}}D) < k$ 时，则应在满足 $Y \in \mathbb{X}$ 的条件下对 $\min \|\widetilde{X} - Y\|$ 进行界定。

令 $\widetilde{X}^{\mathrm{T}}D$ 相同的方式写出其奇异值分解，如式 (3.2) 所示，但将其中的符号 V_i 和 Σ_1 都加上波浪号。记

$$\Sigma_1 = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r), \quad \sigma_1 \geq \cdots \geq \sigma_r > 0. \tag{3.5}$$

$\widetilde{\ }$，类似地，记 Σ_1。本文的主要结果陈述如下定理。

定理 3.1。给定 $D \in \mathbb{R}^{n\times k}$ 维和 k 维子空间 $\mathcal{X}$ 和 $\widetilde{\mathcal{X}}$，它们属于 $\mathbb{R}^n$。令 X、$\widetilde{X} \in \mathbb{O}^{n\times k}$，使得 (3.1) 和 (3.4) 均成立。假设秩$(X^{\mathrm{T}}D) =$ 秩$(\widetilde{X}^{\mathrm{T}}D) =: r$。那么，对于任意酉不变范数 $\|\cdot\|_{\mathrm{ui}}$

$$\min_{Y\in\mathbb{X}} \|\widetilde{X} - Y\|_{\mathrm{ui}} \leq \eta \,\|\sin\Theta(\mathcal{X}, \widetilde{\mathcal{X}})\|_{\mathrm{ui}}, \tag{3.6}$$

其中 $\mathbb{X}$ 如式 (3.3) 所定义，并且

$$\eta = \begin{cases} \sqrt{2}\left(1 + \dfrac{2\,\|D\|_2}{\sigma_k + \tilde{\sigma}_k}\right), & \text{if } r = k, \\ \sqrt{2}\left(1 + \dfrac{2\,\|D\|_2}{\sigma_r + \tilde{\sigma}_r}\right) + \dfrac{(2\sqrt{2}+4)\,\|D\|_2}{\max\{\sigma_r, \tilde{\sigma}_r\}}, & \text{if } r < k. \end{cases} \tag{3.7}$$

对于矩阵谱范数和弗罗贝尼乌斯范数，当$r < k$时，不等式 (3.6) 可以改进，其中较小的 η 为

$$\eta = \begin{cases} \sqrt{2}\left(1 + \dfrac{2\,\|D\|_2}{\sigma_r + \tilde{\sigma}_r}\right) + \dfrac{4\,\|D\|_2}{\max\{\sigma_r, \tilde{\sigma}_r\}}, & \text{for } \|\cdot\|_{\mathrm{ui}} = \|\cdot\|_{\mathrm{F}}, \\ \sqrt{2} + \sqrt{\dfrac{8\,\|D\|_2^2}{(\sigma_r + \tilde{\sigma}_r)^2} + \dfrac{4\,\|D\|_2^2}{[\max\{\sigma_r, \tilde{\sigma}_r\}]^2}} + \dfrac{4\,\|D\|_2}{\max\{\sigma_r, \tilde{\sigma}_r\}}, & \text{for } \|\cdot\|_{\mathrm{ui}} = \|\cdot\|_2. \end{cases} \tag{3.8}$$

备注 3.1。以下依次给出几点评论。

(a) 对于 $r = k$ 的情形，(3.6) 左端实际上是 $\|X - X\|_{\mathrm{ui}}$，因此得到

$$\|\widetilde{X} - X\|_{\mathrm{ui}} \le \eta \,\|\sin\Theta(\mathcal{X}, \widetilde{\mathcal{X}})\|_{\mathrm{ui}}, \tag{3.9}$$

因为此时 $\mathbb{X} = \{X\}$，正如我们之前所解释的。

(b) 对于任意酉不变范数以及两种特定范数——矩阵谱范数和弗罗贝尼乌斯范数，系数 η 在 $r = k$ 时最小。对于 $r < k$ 的情形，(3.8) 中 η 的两个值均小于 (3.7) 中的对应值。第一个值显然如此；对于第二个值，我们可以对所有 a 使用 $\sqrt{a^2 + b^2} \le a + b$， $b \ge 0$，从而看出这一点。

(c) 类似于集合 $\mathbb{X}$ 的定义，我们可以定义与 X 相关的

$$\widetilde{\mathbb{X}} = \left\{ \widetilde{X}\widetilde{V} \begin{bmatrix} I_r & \\ & W \end{bmatrix} \widetilde{V} = \widetilde{X}\widetilde{V}_1\widetilde{V}_1^{\mathrm{T}} + \widetilde{X}\widetilde{V}_2 W \widetilde{V}_2^{\mathrm{T}} \,:\, W \in \mathbb{O}^{r' \times r'} \right\}. \tag{3.10}$$

正如我们之前所解释的，项 $\widetilde{X}\widetilde{V}_1\widetilde{V}_{\mathrm{T}1}$ 只取决于 $\widetilde{\mathcal{X}}$，而 $\widetilde{X}$ 只是 $\mathbb{X}$ 中的个元素之一；就定理 3.1 的结论而言，它可以是该集合中的任意一个。因此，(3.6) 给出了 $\mathbb{X}$ 与 $\mathbb{X}$ 之间 Hausdorff距离的一个上界 Section 11.1 .

$$\mathrm{dist}(\mathbb{X}, \widetilde{\mathbb{X}}) = \max_{\widetilde{Y} \in \widetilde{\mathbb{X}}} \min_{Y \in \mathbb{X}} \|\widetilde{Y} - Y\|_{\mathrm{ui}} \le \eta \,\|\sin\Theta(\mathcal{X}, \widetilde{\mathcal{X}})\|_{\mathrm{ui}}, \tag{3.11}$$

这可以解释为：对于 $\mathbb{X}$ 中的任意一点 Y，在 $\mathbb{X}$ 中都存在一点 Y，使得其与 Y 的距离不超过 $\eta\|\sin\Theta(\mathcal{X}, \mathcal{X})\|_{\mathrm{ui}}$。

本节余下部分将用于证明定理 3.1。为此，我们注意到，根据引理 2.1，存在一个正交矩阵 $Q \in \mathbb{O}^{k \times k}$，使得 $\widehat{X} = \widetilde{X}Q^{\mathrm{T}} \in \mathbb{O}^{n \times k}$ 满足

$$\|X - \widehat{X}\|_{\mathrm{ui}} \le \sqrt{2}\,\|\sin\Theta(\mathcal{X}, \widetilde{\mathcal{X}})\|_{\mathrm{ui}}. \tag{3.12}$$

注意，$\mathcal{R}(X) = \mathcal{X}$，并且 $X = XQ$。利用 (3.12)，可得

$$\|X^{\mathrm{T}}D - \widehat{X}^{\mathrm{T}}D\|_{\mathrm{ui}} \le \|X^{\mathrm{T}} - \widehat{X}^{\mathrm{T}}\|_{\mathrm{ui}}\|D\|_2 \le \sqrt{2}\,\|D\|_2\,\|\sin\Theta(\mathcal{X}, \widetilde{\mathcal{X}})\|_{\mathrm{ui}}. \tag{3.13}$$

3.1 情况 秩$(X^{\mathrm{T}}D) =$ 秩$(X^{\mathrm{T}}D) = k$

在这种情况下，$X^{\mathrm{T}}D \succ 0$ 和 $\widetilde{X}^{\mathrm{T}}D \succ 0$，且 X 和 $\widetilde{X}$ 都是唯一的。因此，我们可以界定 $\|X - X\|_{\mathrm{ui}}$。

注意，因为 $Q \in \mathbb{O}^{k \times k}$ 和 $\widetilde{X}^{\mathrm{T}}D \succ 0$，这是 $X^{\mathrm{T}}D$ 的极分解，$(X^{\mathrm{T}}D)$ 是 $X^{\mathrm{T}}D$ 的极分解。因此，根据引理 2.3，我们有

$$\|I_k - Q\|_{\mathrm{ui}} \le \frac{2}{\sigma_k + \tilde{\sigma}_k}\|X^{\mathrm{T}}D - \widehat{X}^{\mathrm{T}}D\|_{\mathrm{ui}} \le \frac{2\,\|D\|_2}{\sigma_k + \tilde{\sigma}_k}\|X^{\mathrm{T}} - \widehat{X}^{\mathrm{T}}\|_{\mathrm{ui}}.$$

最后，利用 (3.12)，我们得到

$$\begin{aligned}\|X-\widetilde{X}\|_{\mathrm{ui}} &= \|X-\widehat{X}Q\|_{\mathrm{ui}}\\ &= \|X-\widehat{X}+\widehat{X}-\widehat{X}Q\|_{\mathrm{ui}}\\ &\le \|X-\widehat{X}\|_{\mathrm{ui}}+\|\widehat{X}\|_2\|I_k-Q\|_{\mathrm{ui}}\\ &\le \|X-\widehat{X}\|_{\mathrm{ui}}+\frac{2\|D\|_2}{\sigma_k+\tilde{\sigma}_k}\|X^{\mathrm{T}}-\widehat{X}^{\mathrm{T}}\|_{\mathrm{ui}}\\ &\le \left(1+\frac{2\|D\|_2}{\sigma_k+\tilde{\sigma}_k}\right)\sqrt{2}\,\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{ui}},\end{aligned} \tag{3.14}$$

从而得到该情形下的（3.6）。

3.2 秩$(X^{\mathrm{T}}D)$ $=$ 秩$(\widetilde{X}^{\mathrm{T}}D)<k$

在当前情形下， X 和 $\widetilde{X}$ 都不能由 $X^{\mathrm{T}}D\succeq 0$ 和 $\widetilde{X}^{\mathrm{T}}D\succeq 0$唯一确定。因此，仅对满足 $Y\in\mathbb{X}$的 $\|X-Y\|_{\mathrm{ui}}$ 进行界定才有意义。

回顾 $X^{\mathrm{T}}D)\widehat{X}=$秩$(\widetilde{X}^{\mathrm{T}}D)$ 用于满足 $X^{\mathrm{T}}D\succeq 0$ 最小秩 $\widehat{X}$，显然，秩$(\widehat{X}^{\mathrm{T}}D)=$秩$(\widetilde{X}^{\mathrm{T}}D)$。令 Q 为 $\widehat{X}^{\mathrm{T}}D$ 的奇异值分解（$\widetilde{X}^{\mathrm{T}}D$）可表示为

$$\widehat{X}^{\mathrm{T}}D=(Q\widetilde{U})\widetilde{\Sigma}\widetilde{V}^{\mathrm{T}}\equiv \begin{array}{c}\begin{matrix}\ \ r & r'\end{matrix}\\ \left[\begin{array}{cc} Q\widetilde{U}_1 & Q\widetilde{U}_2\end{array}\right]\end{array}\times\begin{array}{c}\begin{matrix}r & r'\end{matrix}\\ \left[\begin{array}{cc}\widetilde{\Sigma}_1 & \\ & 0\end{array}\right]\end{array}\times\begin{array}{c}\begin{matrix}r\\ r'\end{matrix}\\ \left[\begin{array}{c}\widetilde{V}_1^{\mathrm{T}}\\ \widetilde{V}_2^{\mathrm{T}}\end{array}\right]\end{array}, \tag{3.15}$$

并令 $\widehat{U}=Q\widetilde{U}$、 $\widehat{U}_1=Q\widetilde{U}_1$ 和 $\widehat{U}_2=Q\widetilde{U}_2$。根据引理 2.4，存在一个 $\widetilde{W}\in\mathbb{O}^{r'\times r'}$ 使得

$$\widetilde{X}=\widehat{X}\widehat{U}\begin{bmatrix}I_r & \\ & \widetilde{W}\end{bmatrix}\widetilde{V}^{\mathrm{T}}=\widehat{X}\widehat{U}_1\widetilde{V}_1^{\mathrm{T}}+\widehat{X}\widehat{U}_2\widetilde{W}\widetilde{V}_2^{\mathrm{T}}. \tag{3.16}$$

由 (3.3) 和 (3.16)，我们有

$$\begin{aligned}\min_{Y\in\mathbb{X}}\|\widetilde{X}-Y\|_{\mathrm{ui}} &= \min_{W\in\mathbb{O}^{r'\times r'}}\left\|\widehat{X}\widehat{U}\begin{bmatrix}I_r & \\ & \widetilde{W}\end{bmatrix}\widetilde{V}^{\mathrm{T}}-XV\begin{bmatrix}I_r & \\ & W\end{bmatrix}V^{\mathrm{T}}\right\|_{\mathrm{ui}}\\ &\le \min_{W\in\mathbb{O}^{r'\times r'}}\left(\left\|\widehat{X}\widehat{U}\begin{bmatrix}I_r & \\ & \widetilde{W}\end{bmatrix}\widetilde{V}^{\mathrm{T}}-\widehat{X}V\begin{bmatrix}I_r & \\ & W\end{bmatrix}V^{\mathrm{T}}\right\|_{\mathrm{ui}}\right.\\ &\qquad\left.+\left\|\widehat{X}V\begin{bmatrix}I_r & \\ & W\end{bmatrix}V^{\mathrm{T}}-XV\begin{bmatrix}I_r & \\ & W\end{bmatrix}V^{\mathrm{T}}\right\|_{\mathrm{ui}}\right)\\ &= \|\widehat{X}-X\|_{\mathrm{ui}}+\min_{W\in\mathbb{O}^{r'\times r'}}\left\|\widehat{X}\widehat{U}\begin{bmatrix}I_r & \\ & \widetilde{W}\end{bmatrix}\widetilde{V}^{\mathrm{T}}-\widehat{X}V\begin{bmatrix}I_r & \\ & W\end{bmatrix}V^{\mathrm{T}}\right\|_{\mathrm{ui}}\\ &\le \|\widehat{X}-X\|_{\mathrm{ui}}+\min_{W\in\mathbb{O}^{r'\times r'}}\|\widehat{U}_1\widetilde{V}_1^{\mathrm{T}}-V_1V_1^{\mathrm{T}}+\widehat{U}_2\widetilde{W}\widetilde{V}_2^{\mathrm{T}}-V_2WV_2^{\mathrm{T}}\|_{\mathrm{ui}}\\ &\le \|\widehat{X}-X\|_{\mathrm{ui}}+\|\widehat{U}_1\widetilde{V}_1^{\mathrm{T}}-V_1V_1^{\mathrm{T}}\|_{\mathrm{ui}}+\min_{W\in\mathbb{O}^{r'\times r'}}\|\widehat{U}_2\widetilde{W}\widetilde{V}_2^{\mathrm{T}}-V_2WV_2^{\mathrm{T}}\|_{\mathrm{ui}}.\end{aligned} \tag{3.17}$$

~现在，我们将用 $\|\sin\Theta(\mathcal{X},\mathcal{X})\|_{\mathrm{ui}}$ 来界定 (3.17) 右端的三项。首先， $\|\widehat{X}-X\|_{\mathrm{ui}}$ 已由 (3.12) 处理。接下来可以看出， $\widehat{U}_1\widetilde{V}_{\mathrm{T}1}$ 和 $V_1V_1^{\mathrm{T}}$ 分别是 $\widehat{X}^{\mathrm{T}}D$ 和 $X^{\mathrm{T}}D$ 的典范等距极因子。由引理 2.3 和 (3.13)，我们有

$$\begin{aligned}\|\widehat{U}_1\widetilde{V}_1^{\mathrm{T}}-V_1V_1^{\mathrm{T}}\|_{\mathrm{ui}} &\le \left(\frac{2}{\sigma_r+\tilde{\sigma}_r}+\frac{2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\right)\|X^{\mathrm{T}}D-\widehat{X}^{\mathrm{T}}D\|_{\mathrm{ui}}\\ &\le \left(\frac{2}{\sigma_r+\tilde{\sigma}_r}+\frac{2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\right)\sqrt{2}\,\|D\|_2\,\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{ui}}. \end{aligned}\tag{3.18}$$

^ ^最后，令 $\mathcal{U}_i=\mathcal{R}(U_i)$ 和 $\mathcal{V}_i=\mathcal{R}(V_i)$，其中 $i=1,2$。由 (2.3) 和引理 2.2 可得

$$\begin{aligned}\|\sin\Theta(\mathcal{V}_2,\widehat{\mathcal{U}}_2)\|_{\mathrm{ui}} &= \|V_1^{\mathrm{T}}\widehat{U}_2\|_{\mathrm{ui}}=\|\widehat{U}_2^{\mathrm{T}}V_1\|_{\mathrm{ui}}\\ &= \|\sin\Theta(\mathcal{V}_1,\widehat{\mathcal{U}}_1)\|_{\mathrm{ui}}\le\frac{\|X^{\mathrm{T}}D-\widehat{X}^{\mathrm{T}}D\|_{\mathrm{ui}}}{\max\{\sigma_r,\tilde{\sigma}_r\}}\\ &\le\frac{\sqrt{2}\,\|D\|_2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{ui}}, \end{aligned}\tag{3.19}$$

其中，最后一个不等式成立是因为 (3.13)。注意， $\widehat{U}_2\widetilde{W}, V_2\in\mathbb{O}^{k\times r'}$ 满足^ ~ ^$\mathcal{R}$($U_2W)=\mathcal{U}_2$ 且 $\mathcal{R}(V_2)=\mathcal{V}_2$。因此，根据引理 2.1，存在一个正交矩阵 $W_1\in\mathbb{O}^{r'\times r'}$ 使得

$$\|V_2W_1-\widehat{U}_2\widetilde{W}\|_{\mathrm{ui}}\le\sqrt{2}\,\|\sin\Theta(\mathcal{V}_2,\widehat{\mathcal{U}}_2)\|_{\mathrm{ui}}\le\frac{2\,\|D\|_2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{ui}}.\tag{3.20}$$

类似地，存在 $W_2\in\mathbb{O}^{r'\times r'}$ 满足

$$\|V_2W_2-\widetilde{V}_2\|_{\mathrm{ui}}\le\sqrt{2}\,\|\sin\Theta(\mathcal{V}_2,\widetilde{\mathcal{V}}_2)\|_{\mathrm{ui}}\le\frac{2\,\|D\|_2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{ui}}.\tag{3.21}$$

记住 (3.20) 和 (3.21)，我们有

$$\begin{aligned}&\min_{W\in\mathbb{O}^{r'\times r'}}\|\widehat{U}_2\widetilde{W}\widetilde{V}_2^{\mathrm{T}}-V_2WV_2^{\mathrm{T}}\|_{\mathrm{ui}}\\ &\quad\le\|\widehat{U}_2\widetilde{W}\widetilde{V}_2^{\mathrm{T}}-V_2W_1W_2^{\mathrm{T}}V_2^{\mathrm{T}}\|_{\mathrm{ui}}\\ &\quad=\|\widehat{U}_2\widetilde{W}\widetilde{V}_2^{\mathrm{T}}-V_2W_1\widetilde{V}_2^{\mathrm{T}}+V_2W_1\widetilde{V}_2^{\mathrm{T}}-V_2W_1W_2^{\mathrm{T}}V_2^{\mathrm{T}}\|_{\mathrm{ui}}\\ &\quad\le\|\widehat{U}_2\widetilde{W}\widetilde{V}_2^{\mathrm{T}}-V_2W_1\widetilde{V}_2^{\mathrm{T}}\|_{\mathrm{ui}}+\|V_2W_1\widetilde{V}_2^{\mathrm{T}}-V_2W_1W_2^{\mathrm{T}}V_2^{\mathrm{T}}\|_{\mathrm{ui}}\\ &\quad\le\|\widehat{U}_2\widetilde{W}-V_2W_1\|_{\mathrm{ui}}+\|\widetilde{V}_2^{\mathrm{T}}-W_2^{\mathrm{T}}V_2^{\mathrm{T}}\|_{\mathrm{ui}}\\ &\quad\le\frac{4\,\|D\|_2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{ui}}. \end{aligned}\tag{3.22}$$

结合 (3.12)、(3.17)、(3.18) 和 (3.22)，我们有

$$\begin{aligned}\min_{Y\in\mathbb{X}}\|\widetilde{X}-Y\|_{\mathrm{ui}} &\le \left[\sqrt{2}+\left(\frac{2}{\sigma_r+\tilde{\sigma}_r}+\frac{2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\right)\sqrt{2}\,\|D\|_2\right.\\ &\qquad\left.+\frac{4\,\|D\|_2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\right]\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{ui}}\\ &\le \left(\sqrt{2}+\frac{2\sqrt{2}\,\|D\|_2}{\sigma_r+\tilde{\sigma}_r}+\frac{(2\sqrt{2}+4)\,\|D\|_2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\right)\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{ui}}. \end{aligned}\tag{3.23}$$

In等式 (3.14) 和 (3.23) 推得，对于任意一般的酉不变范数，均有 (3.6)　　.

对于两种特定的酉不变范数——矩阵谱范数和弗罗贝尼乌斯范数，不等式 (3.6) 可以得到改进。在本例中，借助引理 2.3，对于这两种特定范数，我们可以利用比 (3.18) 更优的界，从而实现改进。

根据引理 2.3，对于弗罗贝尼乌斯范数，不等式 (3.18) 可改进为

$$\begin{aligned}\|\widehat{U}_1\widetilde{V}_1^{\mathrm{T}}-V_1V_1^{\mathrm{T}}\|_{\mathrm{F}} &\le \frac{2}{\sigma_r+\tilde{\sigma}_r}\|X^{\mathrm{T}}D-\widehat{X}^{\mathrm{T}}D\|_{\mathrm{F}}\\ &\le \frac{2\sqrt{2}\,\|D\|_2}{\sigma_r+\tilde{\sigma}_r}\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{F}}. \end{aligned}\tag{3.24}$$

因此，由 (3.17)、(3.22) 和 (3.24) 可得

$$\min_{Y\in\mathbb{X}}\|\widetilde{X}-Y\|_{\mathrm{F}} \le \left(\sqrt{2}+\frac{2\sqrt{2}\,\|D\|_2}{\sigma_r+\tilde{\sigma}_r}+\frac{4\,\|D\|_2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\right)\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{F}}.\tag{3.25}$$

类似地，当 $\|\cdot\|_{\mathrm{ui}}=\|\cdot\|_2$ 时，根据 (2.13b) 有

$$\min_{Y\in\mathbb{X}}\|\widetilde{X}-Y\|_2 \le \left(\sqrt{2}+\sqrt{\frac{8\,\|D\|_2^2}{(\sigma_r+\tilde{\sigma}_r)^2}+\frac{4\,\|D\|_2^2}{\max\{\sigma_r^2,\tilde{\sigma}_r^2\}}}+\frac{4\,\|D\|_2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\right)\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_2.\tag{3.26}$$

不等式 (3.25) 和 (3.26) 推得 (3.6)，其中改进后的 η 如 (3.8) 所示。

注 3.2. 也可以从另一个角度对任意一般的酉不变范数下的 (3.6) 作些许改进。假设 $r=$ 秩$(X^{\mathrm{T}}D)$ $=$ 秩$(\widetilde{X}^{\mathrm{T}}D)<k$，则可得

$$\max\left\{\|\sin\Theta(\mathcal{V}_1,\widehat{\mathcal{U}}_1)\|_{\mathrm{ui}},\ \|\sin\Theta(\mathcal{V}_1,\widetilde{\mathcal{V}}_1)\|_{\mathrm{ui}}\right\} \le \frac{\left\|\left(X^{\mathrm{T}}D-\widehat{X}^{\mathrm{T}}D\right)_{\text{best-}r}\right\|_{\mathrm{ui}}}{\max\{\sigma_r,\tilde{\sigma}_r\}}$$

由 (2.7a)，并将 (3.13) 改进为

$$\begin{aligned}\left\|\left(X^{\mathrm{T}}D-\widehat{X}^{\mathrm{T}}D\right)_{\text{best-}r}\right\|_{\mathrm{ui}} &\le \|D\|_2\left\|\left(X^{\mathrm{T}}-\widehat{X}^{\mathrm{T}}\right)_{\text{best-}r}\right\|_{\mathrm{ui}}\\ &\le \sqrt{2}\,\|D\|_2\left\|\left[\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\right]_{\text{best-}r}\right\|_{\mathrm{ui}}. \end{aligned}\tag{3.27}$$

注意，在 (3.19) 中取 $\left\|\left[\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\right]_{\text{best-}r}\right\|_{\mathrm{ui}}=\left\|\left(X_\perp^{\mathrm{T}}\widehat{X}\right)_{\text{best-}r}\right\|_{\mathrm{ui}}$，可得

$$\|\sin\Theta(\mathcal{V}_2,\widehat{\mathcal{U}}_2)\|_{\mathrm{ui}} \le \frac{\sqrt{2}\,\|D\|_2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\left\|\left[\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\right]_{\text{best-}r}\right\|_{\mathrm{ui}}.$$

同样，对 (3.20) 和 (3.21) 的改进可按如下方式得到：

$$\|V_2W_1-\widehat{U}_2\widetilde{W}\|_{\mathrm{ui}}\le\sqrt{2}\,\|\sin\Theta(\mathcal{V}_2,\widehat{\mathcal{U}}_2)\|_{\mathrm{ui}}\le\frac{2\,\|D\|_2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\big\|\big[\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\big]_{\text{best-}r}\big\|_{\mathrm{ui}},$$
$$\|V_2W_2-\widetilde{V}_2\|_{\mathrm{ui}}\le\sqrt{2}\,\|\sin\Theta(\mathcal{V}_2,\widetilde{\mathcal{V}}_2)\|_{\mathrm{ui}}\le\frac{2\,\|D\|_2}{\max\{\sigma_r,\tilde{\sigma}_r\}}\big\|\big[\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\big]_{\text{best-}r}\big\|_{\mathrm{ui}},$$

因此，通过将式（3.6）中的 $\sin\Theta(\mathcal{X},\mathcal{X})$ 替换为 $\big[\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\big]_{\text{最佳}r}$，可以得到一个略微更紧的 $\min_{Y\in\mathbb{X}}\|X-Y\|_{\mathrm{ui}}$ 界。

4 数值示例

本节通过数值实验展示本文主要结果的有效性。

令 $M=\frac{1}{\sqrt{n}}\mathrm{hadamard}(n)$，其中 hadamard 是 MATLAB 函数，用于生成正交的 Hadamard矩阵。令

$$X_\diamond=M_{(:,1:k)},\quad \widetilde{X}_\diamond=\sqrt{1-\delta^2}\,M_{(:,1:k)}Q_1+\delta M_{(:,k+1:2k)}Q_2,$$

其中 δ 是用于控制 $\mathcal{X}=\mathcal{R}(X_\diamond)$ 与 $\mathcal{X}=\mathcal{R}(X_\diamond)$ 之间距离的参数，$Q_1,Q_2\in\mathbb{O}^{k\times k}$ 由 MATLAB 内置函数 orth 和 randn 按 orth(randn(k)) 生成。可以计算得出 $X_\diamond^{\mathrm{T}}\widetilde{X}_\diamond=\sqrt{1-\delta^2}Q_1$，其奇异值为 $\sqrt{1-\delta^2}$，重数为 k，因此，$\mathcal{X}$ 与 $\widetilde{\mathcal{X}}$ 之间的 k 典范角 θ_i 均相同，且 $\cos\theta_i=\sqrt{1-\delta^2}$，从而得到

$$\|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_2=\delta,\ \|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{F}}=\sqrt{k}\,\delta,\ \|\sin\Theta(\mathcal{X},\widetilde{\mathcal{X}})\|_{\mathrm{tr}}=k\,\delta.\tag{4.1}$$

当 δ 趋于 0 时，它们也都趋于 0；但作为 $\mathcal{X}$ 和 $\mathcal{X}$的标准正交基矩阵，$X_\diamond$和 $X_\diamond$ 却相差甚远。

本文的主要结果表明，任意 $D\in\mathbb{R}^{n\times k}$ 只要满足 $X^{\mathrm{T}}D\succ 0,\widetilde{X}^{\mathrm{T}}D\succ 0$，就能确定 $\mathcal{X}$ 的特定标准正交基矩阵 X 以及 $\mathcal{X}$的特定标准正交基矩阵 X，并且确保 $X-X=O(\delta)$。接下来，我们将首先通过数值计算展示矩阵范数 $\|\cdot\|_{\mathrm{ui}}=\|\cdot\|_2$、$\|\cdot\|_{\mathrm{F}}$和 $\|\cdot\|_{\mathrm{tr}}$的这个上界 $\delta\to 0$是尖锐的。具体而言，令 $X=X_\diamond UV^{\mathrm{T}}$和 $\widetilde{X}=\widetilde{X}_\diamond\widetilde{U}\widetilde{V}^{\mathrm{T}}$，其中 UV^{T}和 $\widetilde{U}\widetilde{V}^{\mathrm{T}}$分别是 $X_\diamond^{\mathrm{T}}D$ 和 $\widetilde{X}_\diamond^{\mathrm{T}}D$的正交极分解因子。考虑 $n=96$、$k=5$，以及

$$D=\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\\ \frac{6}{8n}&\frac{6}{8n+1}&\frac{6}{8n+2}&\frac{6}{8n+3}&\frac{6}{8n+4}\\ \vdots&\vdots&\vdots&\vdots&\vdots\\ \frac{n}{8n}&\frac{n}{8n+1}&\frac{n}{8n+2}&\frac{n}{8n+3}&\frac{n}{8n+4}\end{bmatrix}.\tag{4.2}$$

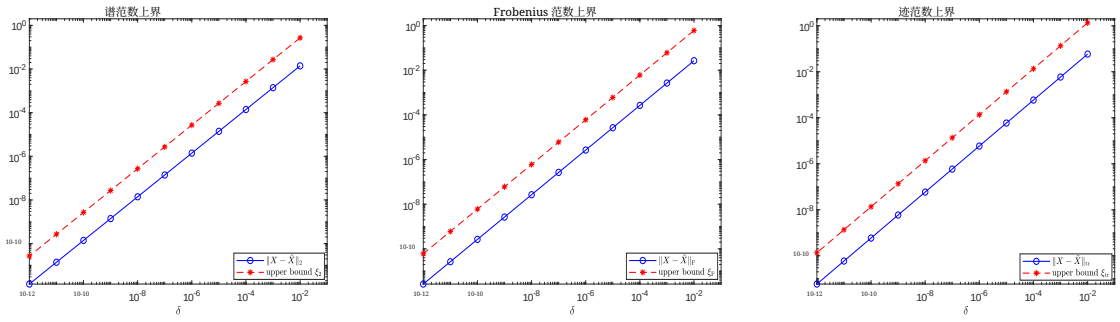

图4.1：满秩情形： $r = k$ 其中 D 如（4.2）所示。式（4.3）中的上界 ξ_{ui} 以及精确的$\|X - \widetilde{X}\|_{\mathrm{ui}}$，随着 δ 从 10^{-12} 变化到 10^{-2}而变化。左：谱范数；中：弗罗贝尼乌斯范数；右：迹范数。

由此，根据定理3.6，我们有：

$$\|X - \widetilde{X}\|_{\mathrm{ui}} \le \xi_{\mathrm{ui}} := \left(\sqrt{2} + \frac{2\sqrt{2}\,\|D\|_2}{\sigma_k + \tilde{\sigma}_k}\right) \times \|\sin\Theta(\mathcal{X}, \widetilde{\mathcal{X}})\|_{\mathrm{ui}}. \tag{4.3}$$

在图4.1中，我们给出了三幅图，每幅图分别包含三种范数在式（4.3）中的上界 ξ_{ui} 和 ~精确的 $\|X - X\|_{\mathrm{ui}}$。可以观察到，上界 ~很紧，能够反映真实差异 $\|X - X\|_{\mathrm{ui}}$。

接下来，我们考虑秩亏情形：$\mathrm{rank}(X^{\mathrm{T}}D) = \mathrm{rank}(\widetilde{X}^{\mathrm{T}}D) =: r < k$。根据定理 3.1，我们有

$$\min_{Y \in \mathbb{X}} \|\widetilde{X} - Y\|_{\mathrm{ui}} \le \xi_{\mathrm{ui}}, \tag{4.4a}$$

其中，对于矩阵范数 $\|\cdot\|_{\mathrm{ui}} = \|\cdot\|_2$、 $\|\cdot\|_{\mathrm{F}}$ 和 $\|\cdot\|_{\mathrm{tr}}$

$$\xi_2 = \left(\sqrt{2} + \sqrt{\frac{8\,\|D\|_2^2}{(\sigma_r + \tilde{\sigma}_r)^2} + \frac{4\,\|D\|_2^2}{\max\{\sigma_r^2, \tilde{\sigma}_r^2\}}} + \frac{4\,\|D\|_2}{\max\{\sigma_r, \tilde{\sigma}_r\}}\right) \|\sin\Theta(\mathcal{X}, \widetilde{\mathcal{X}})\|_2, \tag{4.4b}$$

$$\xi_{\mathrm{F}} = \left(\sqrt{2} + \frac{2\sqrt{2}\,\|D\|_2}{\sigma_r + \tilde{\sigma}_r} + \frac{4\,\|D\|_2}{\max\{\sigma_r, \tilde{\sigma}_r\}}\right) \times \|\sin\Theta(\mathcal{X}, \widetilde{\mathcal{X}})\|_{\mathrm{F}}, \tag{4.4c}$$

$$\xi_{\mathrm{tr}} = \left(\sqrt{2} + \frac{2\sqrt{2}\,\|D\|_2}{\sigma_r + \tilde{\sigma}_r} + \frac{(2\sqrt{2}+4)\,\|D\|_2}{\max\{\sigma_r, \tilde{\sigma}_r\}}\right) \times \|\sin\Theta(\mathcal{X}, \widetilde{\mathcal{X}})\|_{\mathrm{tr}}. \tag{4.4d}$$

对于 $r = k - 1$，我们只需取 (4.2) 中相同的 D，但将其最后一列重置为 0。根据引理 2.4，只有两个 X 满足 $X^{\mathrm{T}}D \succeq 0$ 和 $\mathcal{R}(X) = \mathcal{X} := \mathcal{R}(X_\diamond)$:

$$X = X_\diamond U_{(:,1:k-1)} V_{(:,1:k-1)}^{\mathrm{T}} + X_\diamond U_{(:,k)} V_{(:,k)}^{\mathrm{T}}, \quad X_- = X_\diamond U_{(:,1:k-1)} V_{(:,1:k-1)}^{\mathrm{T}} - X_\diamond U_{(:,k)} V_{(:,k)}^{\mathrm{T}}.$$

相同的对于满足 $\widetilde{X}^{\mathrm{T}}D \succeq 0$ 和 $\mathcal{R}(\widetilde{X}) = \widetilde{\mathcal{X}} := \mathcal{R}(\widetilde{X}_\diamond)$ 的 $\widetilde{X}$，也可以这样说。因此

$$\min_{Y \in \mathbb{X}} \|\widetilde{X} - Y\|_{\mathrm{ui}} = \min\{\|\widetilde{X} - X\|_{\mathrm{ui}}, \|\widetilde{X} - X_-\|_{\mathrm{ui}}\}. \tag{4.5}$$

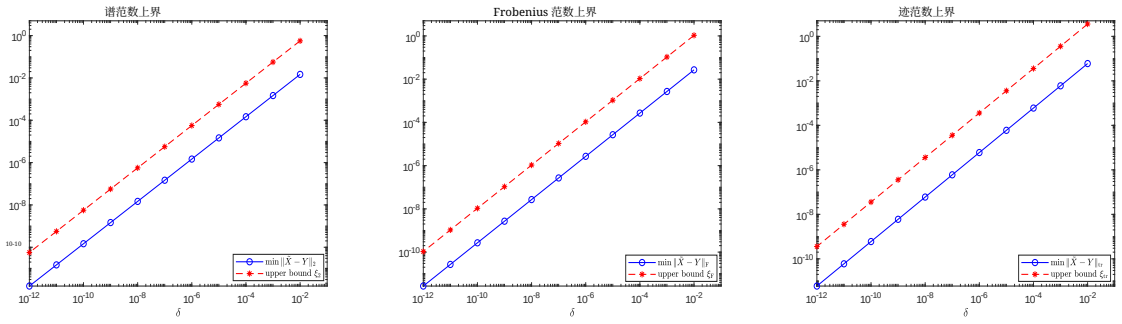

图 4.2：秩亏情形：在 (4.2) 中将 D 的最后一列重置为 0：$r = k-1$。当 δ 从 10^{-12} 变化到 10^{-2} 时，(4.4) 中的上界 ξ_{ui} 和 (4.5) 中的精确最小值 $\|\widetilde{X} - Y\|_{\mathrm{ui}}$。左：谱范数；中：弗罗贝尼乌斯范数；右：迹范数。

在图4.2中，我们再次给出了三幅图，每幅图分别包含三种范数对应的上界 ξ_{ui} 和式（4.5）中的精确量。我们观察到，对于满秩情形，其行为与图4.1中的情况类似。

当 $r < k-1$ 时，集合 $\mathbb{X} = \left\{X_\diamond U_1 V_{\mathrm{T}1} + X_\diamond U_2 W V_{\mathrm{T}2} : \ W \in \mathbb{O}^{(k-r)\times(k-r)}\right\}$ 包含无穷多个元素。幸运的是，对于弗罗贝尼乌斯范数，我们有

$$\begin{aligned}
\min_{Y\in\mathbb{X}} \|\widetilde{X} - Y\|_{\mathrm{F}}^2 &= \min_{W\in\mathbb{O}^{(k-r)\times(k-r)}} \left\| \widetilde{X} - X_\diamond U \begin{bmatrix} I_r & \\ & W \end{bmatrix} V^{\mathrm{T}} \right\|_{\mathrm{F}}^2 \\
&= \min_{W\in\mathbb{O}^{(k-r)\times(k-r)}} \left\| \widetilde{X}V - X_\diamond U \begin{bmatrix} I_r & \\ & W \end{bmatrix} \right\|_{\mathrm{F}}^2 \\
&= \min_{W\in\mathbb{O}^{(k-r)\times(k-r)}} \left\| \left[\widetilde{X}V_1, \quad \widetilde{X}V_2\right] - \left[X_\diamond U_1, \quad X_\diamond U_2 W\right] \right\|_{\mathrm{F}}^2 \\
&= \|\widetilde{X}V_1 - X_\diamond U_1\|_{\mathrm{F}}^2 + \min_{W\in\mathbb{O}^{(k-r)\times(k-r)}} \|\widetilde{X}V_2 - X_\diamond U_2 W\|_{\mathrm{F}}^2,
\end{aligned}$$

其中最后一项可表示为

$$\min_{W\in\mathbb{O}^{(k-r)\times(k-r)}} \|\widetilde{X}V_2 - X_\diamond U_2 W\|_{\mathrm{F}}^2 = 2(k-r)^2 - \max_{W\in\mathbb{O}^{(k-r)\times(k-r)}} \mathrm{tr}(W^{\mathrm{T}} U_2^{\mathrm{T}} X_\diamond^{\mathrm{T}} \widetilde{X} V_2).$$

根据冯·诺伊曼迹不等式 [24, section II.3.1], , W_{opt} 的最优解为

$$\max_{W\in\mathbb{O}^{(k-r)\times(k-r)}} \mathrm{tr}(W^{\mathrm{T}} U_2^{\mathrm{T}} X_\diamond^{\mathrm{T}} \widetilde{X} V_2)$$

是 $U_2^{\mathrm{T}} X_\diamond^{\mathrm{T}} \widetilde{X} V_2$ 的正交极分解因子，最后

$$\min_{Y\in\mathbb{X}} \|\widetilde{X} - Y\|_{\mathrm{F}} = \|\widetilde{X} - Y_{\mathrm{opt}}\|_{\mathrm{F}}, \tag{4.6}$$

其中 $Y_{\mathrm{opt}} = X_\diamond U_1 V_{\mathrm{T}1} + X_\diamond U_2 W_{\mathrm{opt}} V_{\mathrm{T}2}$。对于 $\|\cdot\|_{\mathrm{F}}$ 以外的范数，计算满足 $Y \in \mathbb{X}$ 的 $\widetilde{X}$ 的 $\min \|X - Y\|_{\mathrm{ui}}$ 并不容易。对于这些范数，我们可以借助以下方法对精确最小值进行紧确界定 $_Y \|X -$

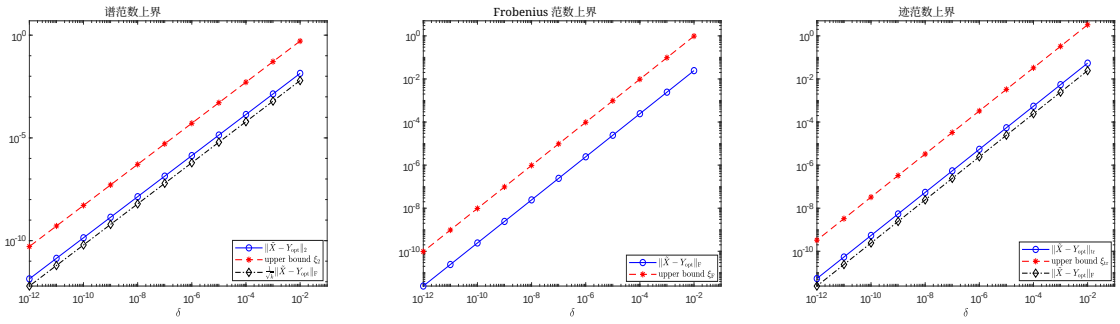

图 4.3：秩亏情形：$r = k - 2$ 将 (4.2) 中 D 的最后两列重置为 0。~，(4.4) 中的 $\|X - Y_{\text{opt}}\|_{\text{ui}}$ ξ_{ui} 以及由 (4.7) 给出的界。左：谱范数；中：弗罗贝尼乌斯范数；右：迹范数。

$Y\|_{\text{ui}}$ 借助此 Y_{opt}，我们可以对其进行紧确界定，如下所示：

$$\frac{1}{\sqrt{k}}\|\widetilde{X} - Y_{\text{opt}}\|_{\text{F}} \le \min_{Y\in\mathbb{X}} \|\widetilde{X} - Y\|_2 \le \|\widetilde{X} - Y_{\text{opt}}\|_2, \tag{4.7a}$$

$$\|\widetilde{X} - Y_{\text{opt}}\|_{\text{F}} \le \min_{Y\in\mathbb{X}} \|\widetilde{X} - Y\|_{\text{tr}} \le \|\widetilde{X} - Y_{\text{opt}}\|_{\text{tr}}, \tag{4.7b}$$

因为 $Y_{\text{opt}} \in \mathbb{X}$ 且对于任意 $Y \in \mathbb{X}$,

$$\frac{1}{\sqrt{k}}\|\widetilde{X} - Y\|_{\text{F}} \le \|\widetilde{X} - Y\|_2, \quad \|\widetilde{X} - Y\|_{\text{F}} \le \|\widetilde{X} - Y\|_{\text{tr}}.$$

~我们在图 4.3 中绘出了三种范数的 $\min_Y \|X - Y\|_{\text{ui}}$ 上界 ξ_{ui}，以及式 (4.7) 中的下界和上界。值得注意的是，我们的上界 ξ_{ui} 依然非常精确，并且以与真实值相同的速率趋于 0。

5 结语

设 $\mathcal{X}$ 是 $\mathbb{R}^n$ 的 k 维子空间，且 $D \in \mathbb{R}^{n\times k}$，并设 X 是 $\mathcal{X}$ 的标准正交基矩阵。若 $X^{\mathrm{T}}D \succ 0$，则在 $\mathcal{X}$ 的所有标准正交基矩阵中，X 是唯一的。然而，如果 $X^{\mathrm{T}}D \succeq 0$ 且 $r = \text{rank}(X^{\mathrm{T}}D)$ $< k$，则 X 可以是集合 $\mathbb{X}$ 中的任意一个，即（3.3）中 $\mathcal{X}$ 的标准正交基矩阵。这些结果最近在 [19, 27] 中得到。然而，本文研究的是在这两种情况下，随着 $\mathcal{X}$ 的变化，X 如何定量变化。

具体来说，假设 $\widetilde{\mathcal{X}}$ 与其中 ~ $\widetilde{\mathcal{X}}$，并通过 $\sin\Theta(\mathcal{X},$ ~ 是它们的典范角的正弦；设 X 是在 $X^{\mathrm{T}}D \succ 0$、$\widetilde{X}^{\mathrm{T}}D \succ 0$ 均成立的情况下，我们针对谱范数、Frobenius 范数以及更一般的任意酉不变范数，$\|X - \widetilde{X}\|$ 的上界，该上界以 $\sim\|\sin\Theta(\mathcal{X},\ \mathcal{X})\|$ 表示；而在 $X^{\mathrm{T}}D \succeq 0$、$\widetilde{X}^{\mathrm{T}}D \succeq 0$ 均成立且 $\text{rank}(X^{\mathrm{T}}D) =$ 的情况下，~，我们的界本质上是关于两个集合 $\mathbb{X}$ 和 $\mathbb{X}$ 的 Hausdorff距离的上界（见（3 测试展示了我们所给上界的精确性。

本文的结果可用于理解采用 NEPv方法求解 Stiefel 流形上最大化问题时的自洽场迭代收敛性，其中目标

函数包含 $\mathrm{tr}(X^{\mathrm{T}}D)$ 并随之增加，甚至可用于评估自洽场迭代期间的近似精度。

尽管我们迄今为止的分析主要关注实数域，但所有推导都可以直接扩展到复数域。为此，只需进行一些细微的修改，即将所有 $\mathbb{R}$ 替换为 $\mathbb{C}$，将所有矩阵/向量转置替换为共轭转置，并将 $\mathrm{tr}(\cdot)$ 替换为 $\mathrm{Re}(\mathrm{tr}(\cdot))$，其中 $\mathrm{Re}(\cdot)$ 表示取复数的实部。

参考文献

[1] Z. Bai、R.-C. Li 和 D. Lu。求解特征向量依赖的非线性特征值问题时自洽场迭代收敛速率的精确估计。SIAM 矩阵分析与应用期刊，43(1):301–327，2022。[2] A. Ben-Israel 和 T. N. E. Greville。广义逆：理论与应用。Springer，纽约，第 2 版，2003。[3] Y. Cai、L.-H. Zhang、Z. Bai 和 R.-C. Li。关于一个特征向量依赖的非线性特征值问题。SIAM 矩阵分析与应用期刊，39(3):1360–1382，2018。[4] M. T. Chu 和 N. T. Trendafilov。重新审视正交约束回归。计算与图形统计学期刊，10(4):746–771，2001。[5] J. P. Cunningham 和 Z. Ghahramani。线性降维：综述、见解与推广。机器学习研究期刊，16:2859–2900，2015。[6] A. Edelman、T. A. Arias 和 S. T. Smith。带正交约束的算法的几何。SIAM 矩阵分析与应用期刊，20(2):303–353，1999。[7] L. Eld´en 和 H. Park。Stiefel 流形上的一个普罗克鲁斯忒斯问题。Numer. Math.，82:599–619，1999。[8] G. H. Golub 和 C. F. Van Loan。矩阵计算。约翰斯·霍普金斯大学出版社，巴尔的摩，马里兰州，第 4 版，2013。[9] J. C. Gower 和 G. B. Dijksterhuis。普罗克鲁斯忒斯问题。牛津大学出版社，纽约，2004。[10] N. J. Higham。矩阵函数：理论与计算。工业与应用数学学会，费城，宾夕法尼亚州，美国，2008。[11] P. Hohenberg 和 W. Kohn。非均匀电子气。物理评论，136:B864–B871，1964。[12] R. A. Horn 和 C. R. Johnson。矩阵分析。剑桥大学出版社，纽约，第 2 版，2013。[13] J. R. Hurley 和 R. B. Cattell。普罗克鲁斯忒斯程序：通过直接旋转检验假设的因子结构。计算行为科学，7:258–262，1962。[14] W. Kohn 和 L. J. Sham。包含交换与关联效应的自洽方程。物理评论，140:A1133–A1138，1965。[15] R.-C. Li。广义极分解的一个扰动界。BIT，33:304–308，1993。[16] R.-C. Li。酉极因子的新的扰动界。SIAM 矩阵分析与应用期刊，16:327–332，1995。

[17] R.-C. Li。矩阵扰动理论。载于 L. Hogben、R. Brualdi 和 G. W. Stewart 编，线性代数手册，第 21 章。CRC Press，Boca Raton，佛罗里达州，第 2 版，2014。[18] W. Li 和 W. Sun。酉极分解因子的新的扰动界。SIAM 矩阵分析与应用期刊，25(2):362–372，2003。[19] D. Lu 和 R.-C. Li。局部酉不变化非线性特征值问题的自洽场迭代收敛性，2022。DOI: 10.48550/ARXIV.2212.14098。[20] T. Ngo、M. Bellalij 和 Y. Saad。用于降维的迹比优化问题。SIAM 矩阵分析与应用期刊，31(5):2950–2971，2010。[21] B. N. Parlett。对称特征值问题。SIAM，费城，1998。[22] P. Petersen。黎曼几何。Springer，2006。[23] D. Rocca、Z. Bai、R.-C. Li 和 G. Galli。用于非厄米随机相位近似矩阵迭代对角化的块变分过程。J. Chem. Phys.，136:034111，2012。[24] G. W. Stewart 和 J.-G. Sun。矩阵扰动理论。Academic Press，波士顿，1990。[25] J.-G. Sun。矩阵扰动分析。Academic Press，北京，1987。中文。[26] L. Wang、L.-H. Zhang 和 R.-C. Li。耦合迹之和的最大化及其应用。Numer. Math.，152:587–629，2022。doi.org/10.1007/s00211-022-01322-y。[27] L. Wang、L.-H. Zhang 和 R.-C. Li。迹比优化及其在多视图学习中的应用。Math. Program.，2022。doi.org/10.1007/s10107-022-01900-w。[28] P.-˚A. Wedin。奇异值分解相关的扰动界。BIT，12:99–111，1972。[29] C. Yang、J. C. Meza、B. Lee 和 L.-W. Wang。KSSOLV——用于求解 Kohn-Sham 方程的 MATLAB 工具箱。ACM Trans. Math. Software，36(2):1–35，2009。[30] L.-H. Zhang、L.-Z. Liao 和 M. K. Ng。广义 Foley-Sammon 判别分析的快速算法。SIAM 矩阵分析与应用期刊，31(4):1584–1605，2010。[31] L.-H. Zhang、L.-Z. Liao 和 M. K. Ng。广义 Foley-Sammon 判别分析的一般算法的超线性收敛。J. Optim. Theory Appl.，157(3):853–865，2013。[32] L.-H. Zhang、W. H. Yang、C. Shen 和 J. Ying。求解非平衡 Procrustes 问题的一种基于特征值的方法。SIAM 矩阵分析与应用期刊，41(3):957–983，2020。[33] L.-H. Zhang 和 R.-C. Li。Stiefel 流形上迹比之和的最大化，II：计算。SCIENCE CHINA Math.，58(7):1549–1566，2015。[34] L.-H. Zhang、L. Wang、Z. Bai 和 R.-C. Li。用于正交典型相关分析的自洽场迭代。IEEE Trans. Pattern Anal. Mach. Intell.，44(2):890–904，2022。[35] Z. Zhang 和 K. Du。求解非平衡 Procrustes 问题的逐次投影法。SCIENCE CHINA Math.，49(7):971–986，2006。