arXiv:2602.07250v1 [math.NA] 6 Feb 2026

倍增算法的统一框架

Changli Liu* Tiexiang Li† Jungong Xue‡ Ren-Cang Li§
Wen-Wei Lin¶

2026年1月31日

摘要

现有的倍增算法已被证明对于若干源自实际工程应用的重要非线性矩阵方程是高效的。简而言之，这些算法迭代计算某个矩阵铅笔对应于其位于特定复区域（例如左半平面或开单位圆盘）内的特征值的特征子空间的基矩阵，且该基矩阵取两种特定形式之一；其成功与否关键取决于所关注的特征子空间确实具有取这两种特定形式之一的基矩阵。然而，这一要求通常无法得到保证。本文提出一种新的倍增算法，称为 Q-倍增算法。它将现有倍增算法作为特殊情况包含在内，并且不要求基矩阵取其中一种特定形式。研究了 Q-倍增算法在求解特征值问题中的应用，数值实验表明其相对于现有倍增算法具有更优的鲁棒性。

关键词： Q标准型， Q倍增算法，特征值，特征子空间，非线性 - 矩阵方程，Bethe-Salpeter方程
数学学科分类：15A24, 65F15, 65F30, 65H10.

*四川大学数学学院，中国成都 610065。电子邮箱：chliliu@scu.edu.cn。获国家自然科学基金委员会（NSFC）项目（编号：12571400）资助。

†东南大学数学系和丘成桐中心，中国南京 211189。电子邮箱：txli@seu.edu.cn。获国家自然科学基金委员会（NSFC）项目（编号：12371377）、江苏省应用数学科学研究中心项目（编号：BK20233002）以及上海数学与交叉科学研究院（SIMIS）项目（编号：SIMIS-ID-2024-LG）资助。

‡复旦大学数学科学学院，中国上海 200433。电子邮箱：xuej@fudan.edu.cn。获国家自然科学基金委员会（NSFC）项目（编号：12571400）资助。§德克萨斯大学阿灵顿分校数学系，美国德克萨斯州阿灵顿，邮政信箱 19408，邮编 76019。电子邮箱：rcli@uta.edu。部分获美国国家科学基金会（NSF）DMS-2407692 项目资助。

¶上海数学与交叉科学研究院，中国上海 200433。电子邮箱：wwlin@outlook.com。获上海数学与交叉科学研究院（SIMIS）资助，基金编号为 SIMIS-ID-2024-LG。

目录

1 引言 3 2 现有倍增算法概述 5 3 多功能基表示 7 4 倍增算法的新统一框架 10 4.1 初始化. 10 4.2 Q-保结构倍增算法.11 5 对偶方程 17 6 收敛性分析 21

7 Q-倍增算法与实现 23
7.1 初始化.24 7.2 倍增迭代期间.26 7.3 Q-倍增算法.30 8 特征子空间计算的应用 31 9 数值实验 32 9.1 实验 1.32 9.2 实验 2.34

10 结论
36 收敛性分析：临界情形 37

1 引言

在 [23], 中，针对两类保结构倍增算法，建立了一个通用框架和理论，用于求解与某些正则矩阵铅笔的特征子空间相关的非线性矩阵方程。本书的研究受到早期关于连续时间代数黎卡提方程（CARE）和离散时间代数黎卡提方程（DARE）的工作（ [7, 10, 11] ）以及关于M-矩阵代数黎卡提方程（MARE）的工作（ [8, 18, 39] ）的启发。简而言之，倍增算法 [23] 所针对的这类非线性矩阵方程，可以等价地转化为某些矩阵铅笔的特征值问题。例如， DARE

$$BX(I+GX)^{-1}A+H-X=0 \tag{1.1}$$

等价于

$$\underbrace{\begin{bmatrix} A & 0 \\ -H & I \end{bmatrix}}_{:=\mathscr{A}} \begin{bmatrix} I \\ X \end{bmatrix} = \underbrace{\begin{bmatrix} I & G \\ 0 & B \end{bmatrix}}_{:=\mathscr{B}} \begin{bmatrix} I \\ X \end{bmatrix} M, \tag{1.2}$$

其中 $M=(I+GX)^{-1}A$。从 (1.2) 可以直接看出两点：1）求解 DARE 等同于计算相应矩阵铅笔 $\mathscr{A}-\lambda\mathscr{B}$ 的一个特征子空间，以及 2）该特征子空间具有如下形式的基矩阵

$$Z=\begin{bmatrix} I \\ X \end{bmatrix}. \tag{1.3}$$

一般而言，存在具有这一特定形式的基矩阵并不能得到保证，必须逐案加以论证，例如，对于来自最优控制的 DARE(1.1)，这一点是可以论证的（见 [23, 第 4] 章及其中的参考文献）。另一方面，即使存在这种形式的基矩阵，有时 X 也可能具有很大的量级，从而可能引起数值不稳定性。如何有效地处理这一问题，是本文的动机之一。

由 (1.2) 可知， M 的特征值是 $\mathscr{A}-\lambda\mathscr{B}$ 的特征值的一部分。从特征值计算的角度来看，求解 DARE(1.1) 使矩阵铅笔解耦。事实上，我们将有

$$\begin{bmatrix} A & 0 \\ -H & I \end{bmatrix} \begin{bmatrix} I & 0 \\ X & I \end{bmatrix} = \begin{bmatrix} I & G \\ 0 & B \end{bmatrix} \begin{bmatrix} I & 0 \\ X & I \end{bmatrix} \begin{bmatrix} M & -(I+GX)^{-1}GM_2^{-1} \\ 0 & M_2^{-1} \end{bmatrix}, \tag{1.4}$$

其中 $M_2=B[I-X(I+GX)^{-1}G]$，因此，矩阵铅笔$\mathscr{A}-\lambda\mathscr{B}$ 的谱是多重集并 $\mathsf{eig}(M)\cup\{\mu^{-1}:\ \mu\in\mathsf{eig}(M_2)\}$，其中 $\mathsf{eig}(\cdot)$ 表示方阵的特征值集合。假设我们已知它的一个 m 维特征子空间 $\mathcal{X}$，并用某个 $(m+n)\times m$ 基矩阵 $Z\in\mathbb{C}^{(m+n)\times m}$表示，该基矩阵总可以写成

$$Z\equiv \begin{matrix} m \\ n \end{matrix} \overset{m}{\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}}. \tag{1.5}$$

如果 Z_1 也可逆，那么 $ZZ_1^{-1} = \begin{bmatrix} I \\ Z_2Z_1^{-1} \end{bmatrix} =: \begin{bmatrix} I \\ X \end{bmatrix}$ 也是一个具有形式 (1.3) 的基矩阵，并且 $\mathscr{A}\begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}\begin{bmatrix} I \\ X \end{bmatrix} M$ 对某个 $m \times m$ 矩阵 M 成立。经过详细推导，可以得到类似于 (1.4) 的方程，将 $\mathscr{A} - \lambda\mathscr{B}$ [23, 章节 3]解耦。再一次，该结果要求 Z_1 非奇异；对于最优控制、纳米研究和高速列车等实际应用中的非线性矩阵方程，这一点可以得到证明，但对于一般矩阵铅笔则未必成立。在后一种情况下，这个 m 维特征子空间没有形式为 (1.3) 的基矩阵，因此现有的倍增算法 [23] 不再适用。即使理论上 Z_1 可逆，当 Z_1 接近奇异时，也会出现计算困难。

然而，由于秩$(Z) = m$， Z 具有一个可逆的 m×m 子矩阵。这意味着 $\mathcal{X}$ 的基矩阵存在一种全天候表示，其形式为

$$Q^{\mathrm{T}}\begin{bmatrix} I \\ X \end{bmatrix} \quad \text{with a permutation matrix } Q \in \mathbb{R}^{(m+n)\times(m+n)}. \tag{1.6}$$

如果适当选取 Q ，则其量级可以保持适中，这一点我们将在本文稍后说明。遗憾的是，倍增算法 [23] 并非为处理这种形式的基矩阵而设计。另一方面， Q 通常并非[1] 先验已知。因此，任何用于计算形如(1.6)的基矩阵的迭代过程都必须能够动态更新Q ，以防止 X 的量级增长过大。本文的目标有两方面：1）建立一个统一[23], 中那些算法的倍增算法框架；2）设计一种动态更新 Q 的机制，从而使 $\|X\|$ 保持适中。为此，我们将提出 Q-标准型（SFQ），它统一了 [23]中的两种标准型SF1和SF2。

本文其余部分的组织如下。第 2 节概述现有倍增算法的框架。第 3 节解释改变基的必要性以及一种新倍增算法的理论依据。第 4 节介绍我们新的保结构倍增算法 SDASFQ，以及采用如（1.6）所示的新基矩阵表示。与现有倍增算法一样，该新倍增算法也与某种非线性矩阵方程相关，这将在第 5 节中讨论。第 6 节对所谓的正则情形进行收敛性分析；鉴于其数学复杂性，对于临界情形的分析则放在附录 A 中。第 7 节给出完整的 QDA，包括初始化以及在基矩阵表示中动态更新 Q-矩阵。第 8 节研究 QDA 在特征值问题中的应用，并在第 9 节报告两个相关的数值实验。最后，我们在第 10 节得出结论。

记号。 $\mathbb{C}^{n\times m}$ 是所有 $n \times m$ 复矩阵的集合， $\mathbb{C}^n = \mathbb{C}^{n\times 1}$，而 $\mathbb{C} = \mathbb{C}^1$.$I_n \in \mathbb{C}^{n\times n}$（若维数可由上下文确定，则简称为 I）是 $n \times n$ 单位矩阵， $\boldsymbol{e}_j$ 是 I （尺寸适当）的第 j列。 A^{T} 和 A^{H} 分别表示矩阵 A的转置和共轭转置。 $\mathcal{R}$（Z）表示列

[1]否则，我们可以处理 $\mathscr{A}Q - \lambda\mathscr{B}Q$，它将具有一个特征子空间，其基矩阵形式为$\begin{bmatrix} I \\ X \end{bmatrix}$。随后便可在 $\mathscr{A}Q - \lambda\mathscr{B}Q$上尝试 [23] 中的倍增算法。

子空间，由 Z的列张成。 $\|\cdot\|$ 表示一般的矩阵/向量范数，$\|\cdot\|_2$ 表示 ℓ_2-向量范数或矩阵的谱范数，而 $\|\cdot\|_{\mathrm{F}}$ 表示矩阵 Frobenius 范数。对于 $r>0$，定义

$$\mathbb{D}_{r-}=\{z\in\mathbb{C}:|z|<r\},\quad \mathbb{T}_r=\{z\in\mathbb{C}:|z|=r\},\quad \mathbb{D}_{r+}=\{z\in\mathbb{C}:|z|>r\},$$

$\bar{}$ $\bar{}$和 $\mathbb{D}_{r-}$ 和 $\mathbb{D}_{r+}$ 分别是 $\mathbb{D}_{r-}$ 和 $\mathbb{D}_{r+}$的闭包。默认情况下，$\mathbb{D}_-=\mathbb{D}_{1-}$、$\mathbb{T}=\mathbb{T}_1$和 $\mathbb{D}_+=\mathbb{D}_{1+}$。最后，$\mathbb{C}_-=\{z\in\mathbb{C}:\ \Re(z)<0\}$，其中 $\Re(z)$ 取 $z\in\mathbb{C}$ 的实部，类似地，$\mathbb{C}_+=\{z\in\mathbb{C}:\ \Re(z)>0\}$。

2 现有倍增算法概述

两类保结构倍增算法——SDASF1 [23, p.22] 和 SDASF2 [23, p.24] ——的一般框架如下。给定矩阵铅笔 $\mathscr{A}-\lambda\mathscr{B}$:

$$\mathscr{A}=\begin{array}{c}\\ m\\ n\end{array}\begin{array}{c}\begin{array}{cc}m & n\end{array}\\ \begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}\end{array},\quad \mathscr{B}=\begin{array}{c}\\ m\\ n\end{array}\begin{array}{c}\begin{array}{cc}m & n\end{array}\\ \begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix}\end{array},\tag{2.1}$$

我们希望求解

$$\mathscr{A}\begin{bmatrix}I\\X\end{bmatrix}\equiv\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}\begin{bmatrix}I\\X\end{bmatrix}=\begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix}\begin{bmatrix}I\\X\end{bmatrix}M\equiv\mathscr{B}\begin{bmatrix}I\\X\end{bmatrix}M\tag{2.2}$$

对于 $X\in\mathbb{C}^{n\times m}$，其中 $M\in\mathbb{C}^{m\times m}$ 可在求得 X 后确定。这是一个特征值问题，其特征值是 $\mathscr{A}-\lambda\mathscr{B}$的特征值中的一部分，位于复平面上的某些区域内，例如左半平面 $\mathbb{C}_-$ 或开单位圆盘 $\mathbb{D}_-$。为确保后续倍增算法的收敛，通常需要使 eig(M) 位于 $\mathbb{D}_-$ 内；但若不满足这一条件，则必须对 $\mathscr{A}-\lambda\mathscr{B}$ 进行某种变换，以得到另一个铅笔

$$\mathscr{A}-\lambda\mathscr{B}\ \to\ \mathscr{A}'-\lambda\mathscr{B}',\tag{2.3a}$$

例如，常用于 CARE [29] 和 MARE [18] 的 M¨obius 变换，以及用于 MARE [39] 的广义双参数 M¨obius 变换（另见 [23]）。新的铅笔 $\mathscr{A}'-\lambda\mathscr{B}'$ 应满足

$$\mathscr{A}'\begin{bmatrix}I\\X\end{bmatrix}=\mathscr{B}'\begin{bmatrix}I\\X\end{bmatrix}\mathscr{M},\ \mathrm{eig}(\mathscr{M})\subset\mathbb{D}_-\ \text{or even}\ \mathrm{eig}(\mathscr{M})\subset\bar{\mathbb{D}}_-,\tag{2.3b}$$

$\bar{}$其中 $\mathbb{D}_-$ 是 $\mathbb{D}_-$ 的闭包，即闭单位圆盘。

假设现在已执行了这样的变换 (2.3)。接下来要做的是在 $\mathscr{A}'-\lambda\mathscr{B}'$ 左乘一个非奇异矩阵 $P\in\mathbb{C}^{(m+n)\times(m+n)}$

$$\mathscr{A}_0-\lambda\mathscr{B}_0=P(\mathscr{A}'-\lambda\mathscr{B}')\tag{2.4}$$

将 $\mathscr{A}' - \lambda\mathscr{B}'$ 变换为 $\mathscr{A}_0 - \lambda\mathscr{B}_0$，使 $\mathscr{A}_0$ 和 $\mathscr{B}_0$ 取以下两种标准型之一：第一标准形 (SF1)，即

$$\mathscr{A}_0 = \begin{matrix} \\ m \\ n \end{matrix}\!\!\begin{matrix} \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix} \end{matrix}, \quad \mathscr{B}_0 = \begin{matrix} \\ m \\ n \end{matrix}\!\!\begin{matrix} \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} I & -Y_0 \\ 0 & F_0 \end{bmatrix} \end{matrix}, \tag{SF1}$$

或第二标准形（SF2），它要求 $m = n$，如下所示:

$$\mathscr{A}_0 = \begin{matrix} \\ n \\ n \end{matrix}\!\!\begin{matrix} \begin{matrix} n & n \end{matrix} \\ \begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix} \end{matrix}, \quad \mathscr{B}_0 = \begin{matrix} \\ n \\ n \end{matrix}\!\!\begin{matrix} \begin{matrix} n & n \end{matrix} \\ \begin{bmatrix} -Y_0 & I \\ F_0 & 0 \end{bmatrix} \end{matrix}. \tag{SF2}$$

由（2.3）和（2.4）可知，同样有

$$\mathscr{A}_0 \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}, \ \mathrm{eig}(\mathscr{M}) \subset \mathbb{D}_- \text{ or even } \mathrm{eig}(\mathscr{M}) \subset \bar{\mathbb{D}}_-. \tag{2.5}$$

随后构造出两个相应的倍增算法 SDASF1 [23, p.22] 和 SDASF2 [23, p.24], ，每个算法都将生成一个矩阵铅笔序列 $\mathscr{A}_i - \lambda\mathscr{B}_i$，同时保持其相应 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 中的块结构，即

$$\mathscr{A}_i = \begin{matrix} \\ m \\ n \end{matrix}\!\!\begin{matrix} \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{matrix}, \quad \mathscr{B}_i = \begin{matrix} \\ m \\ n \end{matrix}\!\!\begin{matrix} \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} I & -Y_i \\ 0 & F_i \end{bmatrix} \end{matrix} \quad \text{for SF1;} \tag{2.6a}$$

$$\mathscr{A}_i = \begin{matrix} \\ n \\ n \end{matrix}\!\!\begin{matrix} \begin{matrix} n & n \end{matrix} \\ \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{matrix}, \quad \mathscr{B}_i = \begin{matrix} \\ n \\ n \end{matrix}\!\!\begin{matrix} \begin{matrix} n & n \end{matrix} \\ \begin{bmatrix} -Y_i & I \\ F_i & 0 \end{bmatrix} \end{matrix} \quad \text{for SF2.} \tag{2.6b}$$

一般而言，要使两个倍增算法中的一个能够运行，相关矩阵铅笔 $\mathscr{A} - \lambda\mathscr{B}$ 的谱会被分成两部分，其中一部分按照（2.2）的含义对应于 X，另一部分按照

$$\text{either} \quad \mathscr{A} \begin{bmatrix} Y \\ I \end{bmatrix} N = \mathscr{B} \begin{bmatrix} Y \\ I \end{bmatrix} \text{ for SF1, or } \mathscr{A} \begin{bmatrix} I \\ Y \end{bmatrix} N = \mathscr{B} \begin{bmatrix} I \\ Y \end{bmatrix} \text{ for SF2.} \tag{2.7}$$

$\mathscr{A} - \lambda\mathscr{B}$ 的谱是 $\mathrm{eig}(M) \cup \mathrm{eig}(N)$ 的多重集并。执行从 $\mathscr{A} - \lambda\mathscr{B}$ 到 $\mathscr{A}_0 - \lambda B_0$ 的上述变换后，（2.7）相应地得到

$$\text{either} \quad \mathscr{A}_0 \begin{bmatrix} Y \\ I \end{bmatrix} \mathscr{N} = \mathscr{B}_0 \begin{bmatrix} Y \\ I \end{bmatrix} \text{ for SF1, or } \mathscr{A}_0 \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N} = \mathscr{B}_0 \begin{bmatrix} I \\ Y \end{bmatrix} \text{ for SF2,} \tag{2.8}$$

其中 $\mathrm{eig}(\mathscr{N}) \subset \mathbb{D}_-$ 甚至 $\mathrm{eig}(\mathscr{N}) \subset \mathbb{D}_{-\circ}$。$\mathscr{A}_0 - \lambda B_0$ 的谱是多重集并 $\mathrm{eig}(\mathscr{M}) \cup \{\mu^{-1} : \mu \in \mathrm{eig}(\mathscr{N})\}$。

对于 SF1 和 SF2，其各自的倍增算法生成的矩阵铅笔序列 $\mathscr{A}_i - \lambda\mathscr{B}_i$ 满足

$$\mathscr{A}_i \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_i \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}^{2^i},$$

$$\text{either} \quad \mathscr{A}_i \begin{bmatrix} Y \\ I \end{bmatrix} \mathscr{N}^{2^i} = \mathscr{B}_i \begin{bmatrix} Y \\ I \end{bmatrix} \text{ for SF1, or } \mathscr{A}_i \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}^{2^i} = \mathscr{B}_i \begin{bmatrix} I \\ Y \end{bmatrix} \text{ for SF2.}$$

显然，当 $\rho(\mathscr{M}) < 1$ 且 $\rho(\mathscr{N}) < 1$ 时，式 (2.6) 中的 X_i 收敛到解 X，并且由此还可得出 Y_i 收敛到式 (2.7) 中的解 Y。但更深入地说，若 $\rho(\mathscr{M})\rho(\mathscr{N}) < 1$，则可确保收敛；甚至若 $\rho(\mathscr{M})\rho(\mathscr{N}) = 1$，再加上对 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 的 Jordan结构施加一项附加条件（即所谓的临界情形），也可确保收敛，见第 3 节8]。为了直观地理解收敛性，设 $\rho(\mathscr{M}) < 1$ 且 $\rho(\mathscr{N}) < 1$。那么 $\mathscr{M}^{2^i} \to 0$ 和 $\mathscr{N}^{2^i} \to 0$ 都成立，因此

$$\mathscr{A}_i \begin{bmatrix} I \\ X \end{bmatrix} \to 0, \text{ and either } \mathscr{B}_i \begin{bmatrix} Y \\ I \end{bmatrix} \to 0 \text{ for SF1 or } \mathscr{B}_i \begin{bmatrix} I \\ Y \end{bmatrix} \to 0 \text{ for SF2,} \tag{2.9}$$

由于 (2.6) 中 $\mathscr{A}_i$ 和 $\mathscr{B}_i$ 具有特殊结构，从而得到 $X - X_i \to 0$ 和 $Y - Y_i \to 0$。

从理论上说，定理 2.1（下文给出的倍增变换定理）解释了倍增算法的工作原理；从计算角度看，保结构倍增算法分为两个阶段：用于获得 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 的初始设置阶段，以及用于为 $i \geq 1$生成 $\mathscr{A}_i - \lambda\mathscr{B}_i$ 的保结构倍增迭代阶段。

定理 2.1（倍增变换定理）。设 $\mathscr{A} - \lambda\mathscr{B} \in \mathbb{C}^{N\times N}$ 是正则矩阵铅笔，即 $\det(\mathscr{A} - \lambda\mathscr{B}) \not\equiv 0$，并设 $\mathscr{A}_\perp, \mathscr{B}_\perp \in \mathbb{C}^{N\times N}$ 满足

$$\operatorname{rank}([\mathscr{A}_\perp, \mathscr{B}_\perp]) = N, \quad [\mathscr{A}_\perp, \mathscr{B}_\perp] \begin{bmatrix} \mathscr{B} \\ -\mathscr{A} \end{bmatrix} = 0. \tag{2.10}$$

$\widetilde{}$ $\widetilde{}$定义 $\widetilde{\mathscr{A}} = \mathscr{A}_\perp\mathscr{A}$ 和 $\widetilde{\mathscr{B}} = \mathscr{B}_\perp\mathscr{B}$。以下结论成立。

$\widetilde{}$ $\widetilde{}$(a) $\widetilde{\mathscr{A}} - \lambda\widetilde{\mathscr{B}}$ 是正则的；

(b) 若 $\mathcal{Z} = \mathcal{R}(Z)$ 是 $\widetilde{\mathscr{A}} - \lambda\widetilde{\mathscr{B}}$的 n维特征子空间，其中 $Z \in \mathbb{C}^{N\times n}$，也就是说，存在 $M \in \mathbb{C}^{n\times n}$ 使得 $\widetilde{\mathscr{A}}Z = \widetilde{\mathscr{B}}ZM$，则

$$\widetilde{\mathscr{A}}\,Z = \widetilde{\mathscr{B}}ZM^2. \tag{2.11}$$

更一般地，如果存在 M_A、$M_B \in \mathbb{C}^{n\times n}$，且满足性质 $M_AM_B = M_BM_A$、
$\widetilde{}$ $\widetilde{}$，使得 $\widetilde{\mathscr{A}}ZM_B = \widetilde{\mathscr{B}}ZM_A$，那么 $\widetilde{\mathscr{A}}\,ZM_B^2 = \widetilde{\mathscr{B}}ZM_A^2$。

这一结果可追溯至 1980 年代的 [9, 14, 31]，以及较近期的 [1, 6, 38]。证明请参见 [23, 第 13] 页。

3 多功能基表示

两种倍增算法 SDASF1 [23, 第 22] 页和 SDASF2 [23, 第 24], 页，在求解各种非线性矩阵方程（NME）方面均取得了巨大成功，尤其是在求解 DARE 及 CARE [26, 27, 28, 32, 34] 和 MARE [18, 33, 39, 42]方面。它们是通过非线性矩阵方程与某个矩阵铅笔 $\mathscr{A} - \lambda\mathscr{B}$ 的特征值问题之间的联系实现的。

粗略地说，在考虑使用某种倍增算法之前，必须满足两个前提条件：

(1) 矩阵铅笔 $\mathscr{A} - \lambda\mathscr{B}$ 的谱被恰当地分成两个互不重叠的部分，其中一部分通常可通过对矩阵铅笔 $\mathscr{A} - \lambda\mathscr{B}$ 进行适当的变换映射到 $\mathbb{D}_-$，另一部分则映射到 $\mathbb{D}_+$；

(2) 与谱的两个不相交部分相对应的基矩阵，分别对应的特征子空间可以证明具有以下形式：对于 SF1，为 $\begin{bmatrix} I \\ X \end{bmatrix}$ 和 $\begin{bmatrix} Y \\ I \end{bmatrix}$；对于 SF2，为 $\begin{bmatrix} I \\ Y \end{bmatrix}$。

对于 DAREs、 CAREs 和 MAREs，文献中已经确立了这两个前提；此外，通过倍增迭代得到的计算近似值在某些矩阵偏序下单调收敛，例如逐项比较或矩阵半正定性。

关于谱被分成两个有重叠部分的假设可以适当放宽，从而允许变换后单位圆 $\mathbb{T}$ 上存在一些重特征值。有关更多细节，请参阅 [23, 第 3 节8] 。

现在考虑一个一般的矩阵铅笔 $\mathscr{A} - \lambda\mathscr{B}$，而不是来自 DAREs、CAREs 或 MAREs 的矩阵铅笔，并假设矩阵铅笔 $\mathscr{A} - \lambda\mathscr{B}$ 的谱被很好地分成两个不相交的部分。那么，特征值问题（2.2）将与其中一个不相交部分相对应的特征子空间的基矩阵限制为形式（1.3），可能无法得到解，因为正如第 1 节所解释的， Z_1 在（1.5）中是奇异的。同样，对于谱的另一部分，（2.7）也可能无法得到解！幸运的是，对于来自控制理论的 DAREs 和 CAREs，以及来自应用概率和交通理论的 MAREs 和马尔可夫调制流体排队理论 [23]，情况并非如此。可以设想，对于一般的正则矩阵铅笔 $\mathscr{A} - \lambda\mathscr{B}$，即使谱分布满足前提条件，情况也不太可能如此。无论如何，对于（1.5）中的任何基矩阵，有一点是确定的：基矩阵总存在一个由其部分行和所有列组成的子矩阵，该子矩阵是非奇异且条件良好的，只要基矩阵本身条件良好；也就是说，基矩阵总可以写成以下形式：[2]

$$Q_1^{\mathrm{T}} \begin{bmatrix} I \\ X \end{bmatrix} \quad \text{with a permutation matrix } Q_1 \in \mathbb{R}^{(m+n)\times(m+n)} \tag{3.1a}$$

以便 X 的量级适中。事实上，如下文定理 3.1 所示，它可以很好地被界定，也就是说，我们可以借助适当的 Q_1使 $\|X\|_2 \le \sqrt{mn+1}$。

定理 3.1. 设 $U \in \mathbb{R}^{N\times m}$ 具有正交归一列，即 $U^{\mathrm{T}}U = I_m$。那么存在一个置换矩阵 $Q \in \mathbb{R}^{N\times N}$，使得在分块 $Q^{\mathrm{T}}U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$下

[2]本文其余部分的推导实际上并不要求 (3.1) 中的 Q_1 和 Q_2 是置换矩阵。我们对它们所使用的唯一性质是它们是正交矩阵。因此，下文的推导对正交的 Q_1 和 Q_2同样适用，无需任何改动。对于酉 Q_1 和 Q_2，只需将转置 Q_i^{T}替换为共轭转置 Q_i^{H}。顺便指出，经过大体上直接的修改，该推导也可适用于非奇异的 Q_1 和 Q_2，但尚不清楚这种一般性能带来什么收益，因此我们略去细节。

其中 $U_1 \in \mathbb{R}^{m\times m}$,

$$\|U_2U_1^{-1}\|_2 \le \|U_1^{-1}\|_2 \le \sqrt{(N-m)m+1}.$$

证明。众所周知， U 有一个 $m \times m$ 子矩阵 U_1 ，使得 [12, 16, 20, 41]

$$\|U_1^{-1}\|_2 \le \sqrt{(N-m)m+1}.$$

令 $Q \in \mathbb{R}^{N\times N}$ 为置换矩阵，使得 $Q^{\mathrm{T}}U$ 按引理中的方式分块。那么我们有 $\|U_2\|_2 \le \|U\|_2 = 1$ ，并且

$$\|U_2U_1^{-1}\|_2 \le \|U_1^{-1}\|_2\|U_2\|_2 \le \|U_1^{-1}\|_2 \le \sqrt{(N-m)m+1},$$

正如预期。 □

对于（2.7）中的基矩阵 $\begin{bmatrix} Y \\ I \end{bmatrix}$ 也可以作同样的处理，因此，一种替代且在数值上更优的形式为

$$Q_2^{\mathrm{T}}\begin{bmatrix} Y \\ I \end{bmatrix} \quad \text{with a permutation matrix } Q_2 \in \mathbb{R}^{(m+n)\times(m+n)} \tag{3.1b}$$

从而 Y 的量级适中。现在涉及两个置换矩阵 Q_1 和 Q_2。除少数情况外，它们不太可能是先验已知的，例如 DAREs、 CAREs或 MAREs；在这些情况下，两者都是单位矩阵。假设我们通过其他途径以某种方式已知 Q_1 和Q_2 [3] ，且它们不是单位矩阵。我们能否相应地构造一种倍增算法? 下面的方法受到近期一项工作的启发[30], ，在该工作中， $Q_1 = I$ 是已知的，而 Q_2 本质上是一个置换矩阵。

将 $\begin{bmatrix} I \\ X \end{bmatrix}$ 替换为 (3.1a) 中的表达式后，方程 (2.2) 变为

$$\mathscr{A}Q_1^{\mathrm{T}}\begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}Q_1^{\mathrm{T}}\begin{bmatrix} I \\ X \end{bmatrix} M. \tag{3.2}$$

据 此，利用新的基矩阵 (3.1b)，(2.7) 中的方程可以合并为一个方程 as

$$\mathscr{A}Q_2^{\mathrm{T}}\begin{bmatrix} Y \\ I \end{bmatrix} N = \mathscr{B}Q_2^{\mathrm{T}}\begin{bmatrix} Y \\ I \end{bmatrix}. \tag{3.3}$$

一个自然的问题是：（3.1）究竟有什么用？首先， Q_1 和 Q_2 所带来的灵活性，为求解已知问题之外的 NME铺平了道路，例如 [30], 中的问题，更重要的是，还包括未来可能出现的问题；其次，基矩阵的新表示形式可能会拓展高效倍增算法在应用型特征值问题中的新应用领域。第三，它使我们能够将两种先前的标准型 SF1 和 SF2，以及从 2018 年以前多年的研究中提炼出的两种相关倍增算法 SDASF1 和 SDASF2，统一纳入一个总体框架。

[3]我们将在第 7.2 小节提出一种自适应的构造方法。

4 倍增算法的新统一框架

回想一下我们在第 2 节中对倍增算法 SDASF1 和 SDASF2 [23]收敛性的简要讨论。对于（3.1）中的基矩阵表示，要得到 $X - X_i \to 0$ 和 $Y - Y_i \to 0$，我们自然需要采用以下形式，而不是（2.9）：

$$\mathscr{A}_i Q_1^{\mathrm{T}} \begin{bmatrix} I \\ X \end{bmatrix} \to 0 \text{ and } \mathscr{B}_i Q_2^{\mathrm{T}} \begin{bmatrix} Y \\ I \end{bmatrix} \to 0 \tag{4.1}$$

从而得到 $X - X_i \to 0$ 和 $Y - Y_i \to 0$。为此，$i = 0$时，$\mathscr{A}_i - \lambda\mathscr{B}_i$的可行取值应为 1、2、……，

$$\mathscr{A}_i = \begin{matrix} \\ m \\ n \end{matrix} \begin{matrix} \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{matrix} Q_1, \quad \mathscr{B}_i = \begin{matrix} \\ m \\ n \end{matrix} \begin{matrix} \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} I & -Y_i \\ 0 & F_i \end{bmatrix} \end{matrix} Q_2. \tag{4.2}$$

这正是我们将采用的方法，以建立倍增算法的最终统一框架。

注意，对于 $Q_1 = Q_2 = I_{m+n}$，(4.2) 变为 SF1；对于 $m = n$、$Q_1 = I_{2n}$ 和 $Q_2 = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}$，则变为 SF2。从某种意义上说，$Q_1$ 和 Q_2 衔接了从 SF1 到 SF2 的过渡。为便于今后引用，我们将 (4.2) 中的形式称为 Q-标准型（SFQ）。

4.1 初始化

现在假设我们已经完成了 (2.3)。接下来要做的是将 $\mathscr{A}' - \lambda\mathscr{B}'$ 左乘一个非奇异矩阵 P

$$\mathscr{A}_0 - \lambda\mathscr{B}_0 = P(\mathscr{A}' - \lambda\mathscr{B}') \tag{4.3}$$

将 $\mathscr{A}' - \lambda\mathscr{B}'$ 转换为 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 的 (4.2) 形式，其中 $i = 0$，Q_1 和 Q_2 是目前假定已知的两个置换矩阵。为此，我们将其划分为

$$\mathscr{A}' Q_1^{\mathrm{T}} = \begin{matrix} \\ m \\ n \end{matrix} \begin{matrix} \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} A'_{11} & A'_{12} \\ A'_{21} & A'_{22} \end{bmatrix} \end{matrix}, \quad \mathscr{B}' Q_2^{\mathrm{T}} = \begin{matrix} \\ m \\ n \end{matrix} \begin{matrix} \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} B'_{11} & B'_{12} \\ B'_{21} & B'_{22} \end{bmatrix} \end{matrix}, \tag{4.4}$$

并将 (4.3) 等价地写为 $P[\mathscr{A}', \mathscr{B}'] = [\mathscr{A}_0, \mathscr{B}_0]$，结合 (4.4) 可得

$$P \begin{bmatrix} A'_{11} & A'_{12} & B'_{11} & B'_{12} \\ A'_{21} & A'_{22} & B'_{21} & B'_{22} \end{bmatrix} = \begin{bmatrix} E_0 & 0 & I_m & -Y_0 \\ -X_0 & I_n & 0 & F_0 \end{bmatrix}. \tag{4.5}$$

等价地，

$$P \begin{bmatrix} B'_{11} & A'_{12} \\ B'_{21} & A'_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \; P \begin{bmatrix} A'_{11} & B'_{12} \\ A'_{21} & B'_{22} \end{bmatrix} = \begin{bmatrix} E_0 & -Y_0 \\ -X_0 & F_0 \end{bmatrix}, \tag{4.6}$$

从而得到

$$\begin{bmatrix} E_0 & Y_0 \\ X_0 & F_0 \end{bmatrix} = J \begin{bmatrix} E_0 & -Y_0 \\ -X_0 & F_0 \end{bmatrix} J = J \begin{bmatrix} B'_{11} & A'_{12} \\ B'_{21} & A'_{22} \end{bmatrix}^{-1} \begin{bmatrix} A'_{11} & B'_{12} \\ A'_{21} & B'_{22} \end{bmatrix} J,$$

其中

$$J = \begin{bmatrix} -I_m & 0 \\ 0 & I_n \end{bmatrix}. \tag{4.7}$$

我们将目前为止的结果总结为以下定理。以这种方式确定矩阵 E_0、 X_0、 Y_0 和 F_0 的想法受到了 Poloni [35] 的启发。

定理 4.1. 矩阵铅笔 $\mathscr{A}' - \lambda\mathscr{B}'$ 可以化为 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 的形式，如 (4.2) 所示，其中 $i = 0$:

$$\mathscr{A}_0 = \begin{matrix} \\ m \\ n \end{matrix} \begin{matrix} \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix} \end{matrix} Q_1, \quad \mathscr{B}_0 = \begin{matrix} \\ m \\ n \end{matrix} \begin{matrix} \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} I & -Y_0 \\ 0 & F_0 \end{bmatrix} \end{matrix} Q_2, \tag{4.8}$$

若且唯若 $\begin{bmatrix} B'_{11} & A'_{12} \\ B'_{21} & A'_{22} \end{bmatrix}$ 可逆，则可通过某个如 (4.3) 中的非奇异矩阵 P 化为该形式。此外，当此矩阵可逆时，矩阵 E_0、 X_0、 Y_0 和 F_0 由下式给出

$$\begin{bmatrix} E_0 & Y_0 \\ X_0 & F_0 \end{bmatrix} = - \begin{bmatrix} B'_{11} & -A'_{12} \\ B'_{21} & -A'_{22} \end{bmatrix}^{-1} \begin{bmatrix} -A'_{11} & B'_{12} \\ -A'_{21} & B'_{22} \end{bmatrix}. \tag{4.9}$$

从 $\mathscr{A} - \lambda\mathscr{B}$ 到 $\mathscr{A}' - \lambda\mathscr{B}'$，再到 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 的初始化过程保证

$$\mathscr{A}_0 Q_1^{\mathrm{T}} \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_0 Q_1^{\mathrm{T}} \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}, \tag{4.10a}$$

$$\mathscr{A}_0 Q_2^{\mathrm{T}} \begin{bmatrix} Y \\ I \end{bmatrix} \mathscr{N} = \mathscr{B}_0 Q_2^{\mathrm{T}} \begin{bmatrix} Y \\ I \end{bmatrix}, \tag{4.10b}$$

前提是原矩阵铅笔 $\mathscr{A} - \lambda\mathscr{B}$ 的特征子空间具有如 (3.1) 所示的基矩阵。

4.2 Q-保结构倍增算法

我们将构造一个相应的保结构倍增算法，它以 (4.8) 中的 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 为输入，并执行稍后推导的结构保持倍增迭代核。其基本思想是在倍增变换定理（定理 2.1）的指导下，构造一个形如 (4.2) 的矩阵铅笔序列 $\mathscr{A}_i - \lambda\mathscr{B}_i$。因此，对于 $i \geq 1$， $\mathscr{A}_i - \lambda\mathscr{B}_i$ 继承了与 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 相同的块结构形式，故称为保结构；同时，对于 $i = 0$，满足 1, . . .

$$\mathscr{A}_i Q_1^{\mathrm{T}} \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_i Q_1^{\mathrm{T}} \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}^{2^i}. \tag{4.11}$$

这一思路背后的原因如下：当谱半径 $\rho(\mathscr{M}) < 1$ 时，$\{\mathscr{M}^{2^i}\}$ 二次趋于 0；而发生这种情况时，我们有（令 $i \to \infty$）：

$$\begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} \to 0 \quad \Rightarrow \quad X - X_i \to 0,$$

假设 $\{\|\mathscr{B}_i\|\}_{i=0}^{\infty}$ 一致有界，其中 X 是方程 (4.10a) 的一个解。

根据定理 2.1，我们需要求出 $\mathscr{A}_{i\perp}$、$\mathscr{B}_{i\perp} \in \mathbb{C}^{(m+n)\times(m+n)}$，使得

$$\operatorname{rank}([\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp}]) = m+n, \quad [\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp}] \begin{bmatrix} \mathscr{B}_i \\ -\mathscr{A}_i \end{bmatrix} = 0, \tag{4.12}$$

然后令 $\mathscr{A}_{i+1} = \mathscr{A}_{i\perp}\mathscr{A}_i$ 和 $\mathscr{B}_{i+1} = \mathscr{B}_{i\perp}\mathscr{B}_i$，同时保持 (4.2) 的形式。这一最后要求意味着 $\mathscr{A}_{i\perp}$ 和 $\mathscr{B}_{i\perp}$ 具有以下形式：

$$\mathscr{A}_{i\perp} = \begin{matrix} \\ m \\ n \end{matrix} \begin{bmatrix} \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} \widehat{E}_i & 0 \\ -\widehat{X}_i & I \end{bmatrix} \end{bmatrix}, \quad \mathscr{B}_{i\perp} = \begin{matrix} \\ m \\ n \end{matrix} \begin{bmatrix} \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} I & -\widehat{Y}_i \\ 0 & \widehat{F}_i \end{bmatrix} \end{bmatrix}, \tag{4.13}$$

其中 $\widehat{E}_i$、$\widehat{F}_i$、$\widehat{X}_i$和 $\widehat{F}_i$ 需要确定，以满足 (4.12) 中的第二个方程。这是因为，暂不考虑可逆性时，我们有 $\mathscr{A}_{i\perp} = \mathscr{A}_{i+1}\mathscr{A}_i^{-1}$和 $\mathscr{B}_{i\perp} = \mathscr{B}_{i+1}\mathscr{B}_i^{-1}$，从而使 $\mathscr{A}_{i\perp}$ 和 $\mathscr{B}_{i\perp}$必须具有 (4.13) 中的形式。注意，对于 (4.13) 中的 $\mathscr{A}_{i\perp}$ 和 $\mathscr{B}_{i\perp}$，(4.12) 中的第一个方程显然成立。

为了求出 (4.13) 中的 $\mathscr{A}_{i\perp}$ 和 $\mathscr{B}_{i\perp}$，我们采用如下消元方法。令

$$Q_1 Q_2^{\mathrm{T}} = \begin{matrix} \\ m \\ n \end{matrix} \begin{bmatrix} \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \end{bmatrix}, \tag{4.14}$$

并注意到[4]

$$\begin{bmatrix} \mathscr{B}_i \\ -\mathscr{A}_i \end{bmatrix} = \begin{bmatrix} I_m & -Y_i \\ 0 & F_i \\ -E_iQ_{11} & -E_iQ_{12} \\ X_iQ_{11} - Q_{21} & X_iQ_{12} - Q_{22} \end{bmatrix} Q_2. \tag{4.15}$$

首先，进行分块高斯消元

$$\begin{bmatrix} I_m & -Y_i \\ 0 & F_i \\ -E_iQ_{11} & -E_iQ_{12} \\ X_iQ_{11} - Q_{21} & X_iQ_{12} - Q_{22} \end{bmatrix} \xrightarrow{L_1} \begin{bmatrix} I_m & -Y_i \\ 0 & F_i \\ 0 & -E_iQ_{12} - E_iQ_{11}Y_i \\ 0 & X_iQ_{12} - Q_{22} + (X_iQ_{11} - Q_{21})Y_i \end{bmatrix} =: \begin{bmatrix} I_m & -Y_i \\ 0 & F_i \\ 0 & -Z_i \\ 0 & -W_i \end{bmatrix},$$

其中 $L_1 \in \mathbb{C}^{2(m+n)\times 2(m+n)}$ 由下式给出

$$L_1 = \begin{bmatrix} I_m & 0 & 0 & 0 \\ 0 & I_n & 0 & 0 \\ E_iQ_{11} & 0 & I_m & 0 \\ -(X_iQ_{11} - Q_{21}) & 0 & 0 & I_n \end{bmatrix},$$

[4]这里用到了 Q_2 是一个正交矩阵，即 $Q_2^{\mathrm{T}}Q_2 = I_{m+n}$。如果允许 Q_2 为酉矩阵，我们就必须将（4.14）中的 Q_2^{T} 替换为 Q_2^{H}。

并且 $\overset{L_1}{\to}$ 意味着，在其左侧乘以 L_1 之前的矩阵即可得到其后的矩阵，并且

$$\begin{aligned} Z_i &:= -\big(-E_iQ_{12}-E_iQ_{11}Y_i\big) = E_i(Q_{11}Y_i+Q_{12}),\\ W_i &:= Q_{22}-X_iQ_{12}+(Q_{21}-X_iQ_{11})Y_i = [-X_i,I]Q_1Q_2^{\mathrm{T}}\begin{bmatrix}Y_i\\ I\end{bmatrix}. \end{aligned} \tag{4.16}$$

假设 W_i 是非奇异的。我们继续进行分块高斯消元 s:

$$\begin{bmatrix} I_m & -Y_i\\ 0 & F_i\\ 0 & -Z_i\\ 0 & -W_i\end{bmatrix} \overset{L_2}{\to} \begin{bmatrix} I_m & -Y_i\\ 0 & F_i\\ 0 & -Z_i\\ 0 & -I_n\end{bmatrix} \overset{L_3}{\to} \begin{bmatrix} I_m & -Y_i\\ 0 & 0\\ 0 & 0\\ 0 & -I_n\end{bmatrix} \overset{L_4}{\to} \begin{bmatrix} I_m & -Y_i\\ 0 & -I_n\\ 0 & 0\\ 0 & 0\end{bmatrix},$$

其中

$$L_2=\begin{bmatrix} I_{2m+n} & \\ & W_i^{-1}\end{bmatrix},\ L_3=\begin{bmatrix} I_m & 0 & 0 & 0\\ & I_n & 0 & F_i\\ & & I_m & -Z_i\\ & & & I_n\end{bmatrix},\ L_4=\begin{bmatrix} I_m & 0 & 0 & 0\\ 0 & 0 & 0 & I_n\\ 0 & 0 & I_m & 0\\ 0 & I_n & 0 & 0\end{bmatrix}.$$

每个 L_j 执行以下三种操作之一：(1) 将某个矩阵左乘一个块行，并将结果加到另一个块行上；(2) 将某个非奇异矩阵左乘一个块行；(3) 交换两个块行。最后，$L_4L_3L_2L_1$ 的最后两个块行给出 $[\mathscr{A}_{i\perp},\mathscr{B}_{i\perp}]$:

$$\mathscr{A}_{i\perp}=\begin{bmatrix} E_iQ_{11}+Z_iW_i^{-1}(X_iQ_{11}-Q_{21}) & 0\\ -F_iW_i^{-1}(X_iQ_{11}-Q_{21}) & I_n\end{bmatrix},\quad \mathscr{B}_{i\perp}=\begin{bmatrix} I_m & -Z_iW_i^{-1}\\ 0 & F_iW_i^{-1}\end{bmatrix} \tag{4.17}$$

其形式如 (4.13) 所示。第二种方法直接处理 (4.12) 中的第二个方程:

$$\begin{bmatrix} \widehat{E}_i & 0 & I & -\widehat{Y}_i\\ -\widehat{X}_i & I & 0 & \widehat{F}_i\end{bmatrix}\begin{bmatrix} I & -Y_i\\ 0 & F_i\\ -E_iQ_{11} & -E_iQ_{12}\\ X_iQ_{11}-Q_{21} & X_iQ_{12}-Q_{22}\end{bmatrix}=0.$$

交换第一个矩阵的第二个和最后一个块列，并相应地交换第二个矩阵的第二个和最后一个块行，得到

$$\begin{bmatrix} \widehat{E}_i & -\widehat{Y}_i & I & 0\\ -\widehat{X}_i & \widehat{F}_i & 0 & I\end{bmatrix}\begin{bmatrix} I & -Y_i\\ X_iQ_{11}-Q_{21} & X_iQ_{12}-Q_{22}\\ -E_iQ_{11} & -E_iQ_{12}\\ 0 & F_i\end{bmatrix}=0,$$

或等价地

$$\begin{bmatrix} \widehat{E}_i & -\widehat{Y}_i\\ -\widehat{X}_i & \widehat{F}_i\end{bmatrix}\begin{bmatrix} I & -Y_i\\ X_iQ_{11}-Q_{21} & X_iQ_{12}-Q_{22}\end{bmatrix}+\begin{bmatrix} -E_iQ_{11} & -E_iQ_{12}\\ 0 & F_i\end{bmatrix}=0, \tag{4.18}$$

得到

$$\begin{bmatrix} \widehat{E}_i & -\widehat{Y}_i \\ -\widehat{X}_i & \widehat{F}_i \end{bmatrix} \begin{bmatrix} I & -Y_i \\ X_iQ_{11}-Q_{21} & X_iQ_{12}-Q_{22} \end{bmatrix} = \begin{bmatrix} E_iQ_{11} & E_iQ_{12} \\ 0 & -F_i \end{bmatrix}. \tag{4.19}$$

可以看出， W_i 是（4.19）左侧第二个矩阵中（1, 1）块 I 的舒尔补，因此，当且仅当第二个矩阵可逆时， W_i 才可逆。假设 W_i 可逆，则有

$$\begin{bmatrix} I & -Y_i \\ X_iQ_{11}-Q_{21} & X_iQ_{12}-Q_{22} \end{bmatrix}^{-1} = \begin{bmatrix} I_m + Y_iW_i^{-1}(X_iQ_{11}-Q_{21}) & -Y_iW_i^{-1} \\ W_i^{-1}(X_iQ_{11}-Q_{21}) & -W_i^{-1} \end{bmatrix} \tag{4.20}$$

结合（4.19），可得

$$\widehat{E}_i = E_i\big[Q_{11} + (Q_{11}Y_i + Q_{12})W_i^{-1}(X_iQ_{11}-Q_{21})\big], \tag{4.21a}$$

$$\widehat{F}_i = F_iW_i^{-1}, \tag{4.21b}$$

$$\widehat{X}_i = F_iW_i^{-1}(X_iQ_{11}-Q_{21}), \tag{4.21c}$$

$$\widehat{Y}_i = E_i(Q_{11}Y_i + Q_{12})W_i^{-1} \tag{4.21d}$$

可以验证，这与（4.17）所推出的结果相同。

最后， $\mathscr{A}_{i+1} = \mathscr{A}_{i\perp}\mathscr{A}_i$ 和 $\mathscr{B}_{i+1} = \mathscr{B}_{i\perp}\mathscr{B}_i$ 可写成（4.2）的形式，其中

$$\begin{aligned} E_{i+1} &= \widehat{E}_iE_i \\ &= E_i\big[Q_{11} + (Q_{11}Y_i + Q_{12})W_i^{-1}(X_iQ_{11}-Q_{21})\big]E_i, \end{aligned} \tag{4.22a}$$

$$\begin{aligned} F_{i+1} &= \widehat{F}_iF_i \\ &= F_iW_i^{-1}F_i, \end{aligned} \tag{4.22b}$$

$$\begin{aligned} X_{i+1} &= X_i + \widehat{X}_iE_i \\ &= X_i + F_iW_i^{-1}(X_iQ_{11}-Q_{21})E_i, \end{aligned} \tag{4.22c}$$

$$\begin{aligned} Y_{i+1} &= Y_i + \widehat{Y}_iF_i \\ &= Y_i + E_i(Q_{11}Y_i + Q_{12})W_i^{-1}F_i, \end{aligned} \tag{4.22d}$$

其中， W_i 由（4.16）定义。

倍增变换定理（定理 2.1）保证：若 $i=0$ 成立，则 $i \geq 1$ 的 (4.11) 也成立；此外

$$\mathscr{A}_iQ_2^{\mathrm{T}} \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}^{2^i} = \mathscr{B}_iQ_2^{\mathrm{T}} \begin{bmatrix} I \\ Y \end{bmatrix} \tag{4.23}$$

若 $i=0$ 成立，则 $i \geq 1$ 也成立。

寻找 $\mathscr{A}_{i\perp}$ 和 $\mathscr{B}_{i\perp}$ 的第三种方法基于这样一个事实：[5] 也

$$\begin{bmatrix} \mathscr{B}_i \\ -\mathscr{A}_i \end{bmatrix} = \begin{bmatrix} Q_{11}^{\mathrm{T}} - Y_iQ_{12}^{\mathrm{T}} & Q_{21}^{\mathrm{T}} - Y_iQ_{22}^{\mathrm{T}} \\ F_iQ_{12}^{\mathrm{T}} & F_iQ_{22}^{\mathrm{T}} \\ -E_i & 0 \\ X_i & -I_n \end{bmatrix} Q_1. \tag{4.24}$$

[5]这里我们需要 Q_1 是一个正交矩阵，即 $Q_1^{\mathrm{T}}Q_1 = I_{m+n}$。如果允许 Q_1 为酉矩阵，我们就必须将Q_1^{T}替换为Q_1^{H}，并将（4.24）中的每个 Q_{ij}^{T} 替换为 Q_{ij}^{H}。

将（4.24）代入（4.12）中的第二个方程，并将 $\mathscr{A}_{i\perp}$ 和 $\mathscr{B}_{i\perp}$替换为它们在（4.13）中的表达式后，我们得到

$$\begin{bmatrix} \widehat{E}_i & -\widehat{Y}_i \\ -\widehat{X}_i & \widehat{F}_i \end{bmatrix} \begin{bmatrix} Q_{11}^{\mathrm{T}} - Y_i Q_{12}^{\mathrm{T}} & Q_{21}^{\mathrm{T}} - Y_i Q_{22}^{\mathrm{T}} \\ X_i & -I_n \end{bmatrix} + \begin{bmatrix} -E_i & 0 \\ F_i Q_{12}^{\mathrm{T}} & F_i Q_{22}^{\mathrm{T}} \end{bmatrix} = 0, \tag{4.25}$$

与 (4.18) 类似。比较 (4.15) 和 (4.24)，得到

$$\underbrace{\begin{bmatrix} I & -Y_i \\ X_i Q_{11} - Q_{21} & X_i Q_{12} - Q_{22} \end{bmatrix}}_{=:G_1} Q_2 = \underbrace{\begin{bmatrix} Q_{11}^{\mathrm{T}} - Y_i Q_{12}^{\mathrm{T}} & Q_{21}^{\mathrm{T}} - Y_i Q_{22}^{\mathrm{T}} \\ X_i & -I_n \end{bmatrix}}_{=:G_2} Q_1. \tag{4.26}$$

因此，当且仅当 G_2 可逆时， G_1 可逆。此前，我们假设 W_i 可逆，并且还证明了当且仅当 G_1 可逆时， W_i 可逆。另一方面，我们还看到，当且仅当以下条件成立时， G_2 可逆

$$\widetilde{W}_i = Q_{11}^{\mathrm{T}} - Y_i Q_{12}^{\mathrm{T}} + (Q_{21}^{\mathrm{T}} - Y_i Q_{22}^{\mathrm{T}}) X_i \tag{4.27}$$

由于 W_i 是 G_2 中 (2,2) 块 $-I_n$ 的舒尔补，因此 $\widetilde{\ }$ 可逆。总之，如果 G_1、 G_2、 W_i 和 W_i 中任意一个是非奇异的，那么其余三个也都是非奇异的。

假设 W_i 可逆。由 (4.25) 可得

$$\begin{bmatrix} \widehat{E}_i & \widehat{Y}_i \\ \widehat{X}_i & \widehat{F}_i \end{bmatrix} = J \begin{bmatrix} E_i & 0 \\ -F_i Q_{12}^{\mathrm{T}} & -F_i Q_{22}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} Q_{11}^{\mathrm{T}} - Y_i Q_{12}^{\mathrm{T}} & Q_{21}^{\mathrm{T}} - Y_i Q_{22}^{\mathrm{T}} \\ X_i & -I_n \end{bmatrix}^{-1} J, \tag{4.28}$$

其中 J 的定义见 (4.7)。注意

$$\begin{bmatrix} Q_{11}^{\mathrm{T}} - Y_i Q_{12}^{\mathrm{T}} & Q_{21}^{\mathrm{T}} - Y_i Q_{22}^{\mathrm{T}} \\ X_i & -I_n \end{bmatrix}^{-1} = \begin{bmatrix} \widetilde{W}_i^{-1} & \widetilde{W}_i^{-1}(Q_{21}^{\mathrm{T}} - Y_i Q_{22}^{\mathrm{T}}) \\ X_i \widetilde{W}_i^{-1} & -I_n + X_i \widetilde{W}_i^{-1}(Q_{21}^{\mathrm{T}} - Y_i Q_{22}^{\mathrm{T}}) \end{bmatrix}$$

结合 (4.28)，可得

$$\widehat{E}_i = E_i \widetilde{W}_i^{-1}, \tag{4.29a}$$
$$\widehat{F}_i = F_i \big[Q_{22}^{\mathrm{T}} + (Q_{22}^{\mathrm{T}} X_i + Q_{12}^{\mathrm{T}}) \widetilde{W}_i^{-1} (Y_i Q_{22}^{\mathrm{T}} - Q_{21}^{\mathrm{T}}) \big], \tag{4.29b}$$
$$\widehat{X}_i = F_i (Q_{22}^{\mathrm{T}} X_i + Q_{12}^{\mathrm{T}}) \widetilde{W}_i^{-1}, \tag{4.29c}$$
$$\widehat{Y}_i = E_i \widetilde{W}_i^{-1} (Y_i Q_{22}^{\mathrm{T}} - Q_{21}^{\mathrm{T}}). \tag{4.29d}$$

最后，我们得到 $\mathscr{A}_{i+1} = \mathscr{A}_{i\perp}\mathscr{A}_i$ 和 $\mathscr{B}_{i+1} = \mathscr{B}_{i\perp}\mathscr{B}_i$ ，其形式为 (4.2)，其中

$$\begin{aligned} E_{i+1} &= \widehat{E}_i E_i \\ &= E_i \widetilde{W}_i^{-1} E_i, \\ F_{i+1} &= \widehat{F}_i F_i \end{aligned} \tag{4.30a}$$

算法 4.1 SDASFQ：SFQ 的结构化加倍算法

输入： $X_0 \in \mathbb{C}^{n\times m}$、 $Y_0 \in \mathbb{C}^{m\times n}$、 $E_0 \in \mathbb{C}^{m\times m}$、 $F_0 \in \mathbb{C}^{n\times n}$和 $Q_1Q_2^{\mathrm{T}}$，按如下方式分块 (4.14);

输出：若 X_i 收敛，则输出其极限。

1: 对于 $i = 0,1,\ldots,$ 直至收敛

2: $W_i = Q_{22} - X_iQ_{12} - (X_iQ_{11} - Q_{21})Y_i$

（或者，也可以，$\widetilde{W}_i = Q_{11}^{\mathrm{T}} - Y_iQ_{12}^{\mathrm{T}} + (Q_{21}^{\mathrm{T}} - Y_iQ_{22}^{\mathrm{T}})X_i$ 如果 $n > m$);

3: 根据 (4.22) 计算 E_{i+1}、 F_{i+1}、 X_{i+1}、 Y_{i+1}

（或者，如果 $n > m$，也可以用 (4.30)）；

4: 结束循环

5: 返回 X_i 收敛时作为计算所得的解。

$$= F_i\big[Q_{22}^{\mathrm{T}} + (Q_{22}^{\mathrm{T}}X_i + Q_{12}^{\mathrm{T}})\widetilde{W}_i^{-1}(Y_iQ_{22}^{\mathrm{T}} - Q_{21}^{\mathrm{T}})\big]F_i, \tag{4.30b}$$

$$\begin{aligned} X_{i+1} &= X_i + \widehat{X}_iE_i \\ &= X_i + F_i(Q_{22}^{\mathrm{T}}X_i + Q_{12}^{\mathrm{T}})\widetilde{W}_i^{-1}E_i, \end{aligned} \tag{4.30c}$$

$$\begin{aligned} Y_{i+1} &= Y_i + \widehat{Y}_iF_i \\ &= Y_i + E_i\widetilde{W}_i^{-1}(Y_iQ_{22}^{\mathrm{T}} - Q_{21}^{\mathrm{T}})F_i, \end{aligned} \tag{4.30d}$$

$\widetilde{\ }$其中 W_i 由 (4.27) 定义。

在 (4.22) 和 (4.30) 中，我们为同一目的建立了两组倍增迭代公式，但每组公式都只涉及 W_i^{-1}或 $\widetilde{W}_i^{-1}$。这为实现提供了灵活性$\widetilde{\ }$，具体取决于 m 和 n。注意到 W_i 是 $n\times n$，而 W_i 是 $m\times m$，因此当 $n < m$ 时应使用 (4.22)，否则使用 (4.30)。算法 4.1 概述了适用于 SFQ 的保结构倍增算法。我们将在第 4.2 小节稍后讨论应在线 1 使用何种停止准则来终止倍增迭代。由于该算法在每个迭代步骤中都涉及矩阵逆，因此一般而言可能发生失效，即某些逆矩阵不存在。在诸如 CARE、 DARE、 MARE 和 H^*ARE 等各种重要应用中，可以证明不会发生失效。

特殊情况

当 $Q_1 = Q_2$ 时，(4.22)/(4.30)中的迭代公式可以大大简化，此时 $Q_1Q_2^{\mathrm{T}} = I$，从而意味着对 $i \neq j$ 有 $Q_{11} = I$、 $Q_{22} = I$和 $Q_{ij} = 0$ ，因此$W_i = I - X_iY_i$ 以及

$$\begin{aligned} E_{i+1} &= E_i\big[I + Y_iW_i^{-1}X_i\big]E_i \\ &= E_i\big(I - Y_iX_i\big)^{-1}E_i \end{aligned} \tag{4.31a}$$

$$\begin{aligned} F_{i+1} &= F_iW_i^{-1}F_i \\ &= F_i\big(I - X_iY_i\big)^{-1}F_i, \end{aligned} \tag{4.31b}$$

$$X_{i+1} = X_i + F_iW_i^{-1}(X_iQ_{11} - Q_{21})E_i$$

$$= X_i + F_i\big(I - X_iY_i\big)^{-1}X_iE_i, \tag{4.31c}$$

$$\begin{aligned} Y_{i+1} &= Y_i + E_i(Q_{11}Y_i + Q_{12})W_i^{-1}F_i \\ &= Y_i + E_iY_i\big(I - X_iY_i\big)^{-1}F_i, \end{aligned} \tag{4.31d}$$

which与SDASF1 [23, 第21], 页的SF1倍增算法完全相同 .

另一种情况也能使迭代公式比（4.22）大幅简化，即

$$m = n, \quad Q_1Q_2^{\mathrm{T}} = \begin{bmatrix} & I_n \\ I_n & \end{bmatrix}, \tag{4.32}$$

这意味着 $Q_{11} = 0$、 $Q_{22} = 0$ 和 $Q_{ij} = I_n$ 对于 $i \neq j$，因此 $W_i = Y_i - X_i$ 和

$$E_{i+1} = E_i\big(X_i - Y_i\big)^{-1}E_i \tag{4.33a}$$

$$F_{i+1} = F_i\big(Y_i - X_i\big)^{-1}F_i, \tag{4.33b}$$

$$X_{i+1} = X_i + F_i\big(X_i - Y_i\big)^{-1}E_i, \tag{4.33c}$$

$$Y_{i+1} = Y_i + E_i\big(Y_i - X_i\big)^{-1}F_i, \tag{4.33d}$$

whi它们与 SDASF2 [23, p.24], SF2 倍增算法完全相同。

5 对偶方程

回顾（4.8）和（4.10a）。将该方程按块展开为两个方程，然后消去 $\mathscr{M}$，即可得到一个关于 X 的方程，其中没有明确提及 $\mathscr{M}$。我们称所得方程为原始方程，保结构倍增算法 SDASFQ（算法 4.1）正是为利用 X 序列 $\{X_i\}_{i=0}^{\infty}$ 求解该方程而设计的。自然会问， Y 序列 $\{Y_i\}_{i=0}^{\infty}$ 会收敛到什么（如果它确实收敛的话）。事实上，若不了解 Y 序列 $\{Y_i\}_{i=0}^{\infty}$ 的变化，就无法充分理解倍增算法，尤其是其收敛行为。乍看之下，Y 序列只是辅助性的，其存在的唯一目的就是使 X 序列的计算成为可能。但这只是表象。事实上， Y 序列也会趋近某个值，而这个值最终被证明是另一个非线性矩阵方程的解，我们称之为对偶方程。

将 $\mathscr{A}_0$ 和 $\mathscr{B}_0$ 代入 (4.10a) 中的 (4.8)，并注意到 (4.14)，可得，

$$\begin{aligned} \begin{bmatrix} E_0 & 0 \\ -X_0 & I_n \end{bmatrix}\begin{bmatrix} I_m \\ X \end{bmatrix} &= \begin{bmatrix} I_m & -Y_0 \\ 0 & F_0 \end{bmatrix} Q_2Q_1^{\mathrm{T}}\begin{bmatrix} I_m \\ X \end{bmatrix}\mathscr{M} \\ &= \begin{bmatrix} I_m & -Y_0 \\ 0 & F_0 \end{bmatrix}\begin{bmatrix} Q_{11}^{\mathrm{T}} & Q_{21}^{\mathrm{T}} \\ Q_{12}^{\mathrm{T}} & Q_{22}^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix}\mathscr{M} \\ &= \begin{bmatrix} Q_{11}^{\mathrm{T}} - Y_0Q_{12}^{\mathrm{T}} + (Q_{21}^{\mathrm{T}} - Y_0Q_{22}^{\mathrm{T}})X \\ F_0(Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}X) \end{bmatrix}\mathscr{M}, \end{aligned} \tag{5.1}$$

得到原始方程

$$X = X_0 + F_0(Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}X)\big[Q_{11}^{\mathrm{T}} - Y_0Q_{12}^{\mathrm{T}} + (Q_{21}^{\mathrm{T}} - Y_0Q_{22}^{\mathrm{T}})X\big]^{-1}E_0. \tag{5.2}$$

接下来，将 $\mathscr{A}_0$ 和 $\mathscr{B}_0$ 代入（4.8）中的（4.10b），得到

$$\begin{aligned}\begin{bmatrix} I_m & -Y_0 \\ 0 & F_0 \end{bmatrix}\begin{bmatrix} Y \\ I_n \end{bmatrix} &= \begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix} Q_1 Q_2^{\mathrm{T}} \begin{bmatrix} Y \\ I \end{bmatrix} \mathscr{N} \\ &= \begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix}\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}\begin{bmatrix} Y \\ I \end{bmatrix} \mathscr{N} \\ &= \begin{bmatrix} E_0(Q_{12}+Q_{11}Y) \\ Q_{22}-X_0Q_{12}+(Q_{21}-X_0Q_{11})Y \end{bmatrix} \mathscr{N},\end{aligned} \tag{5.3}$$

从而得到对偶方程

$$Y = Y_0 + E_0(Q_{12}+Q_{11}Y)\big[Q_{22}-X_0Q_{12}+(Q_{21}-X_0Q_{11})Y\big]^{-1}F_0. \tag{5.4}$$

Let

$$\mathit{\Pi}_{m,n} = \begin{bmatrix} 0 & I_m \\ I_n & 0 \end{bmatrix}. \tag{5.5}$$

可以看出

$$\mathit{\Pi}_{m,n}^{\mathrm{T}}\begin{bmatrix} I_m & -Y_0 \\ 0 & F_0 \end{bmatrix}\begin{bmatrix} Y \\ I_n \end{bmatrix} = \mathit{\Pi}_{m,n}^{\mathrm{T}}\begin{bmatrix} I_m & -Y_0 \\ 0 & F_0 \end{bmatrix}\mathit{\Pi}_{m,n}\mathit{\Pi}_{m,n}^{\mathrm{T}}\begin{bmatrix} Y \\ I_n \end{bmatrix} = \begin{bmatrix} F_0 & 0 \\ -Y_0 & I_m \end{bmatrix}\begin{bmatrix} I_n \\ Y \end{bmatrix}$$

同理,

$$\mathit{\Pi}_{m,n}^{\mathrm{T}}\begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix}\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}\begin{bmatrix} Y \\ I \end{bmatrix} = \begin{bmatrix} I_n & -X_0 \\ 0 & E_0 \end{bmatrix}\begin{bmatrix} Q_{22} & Q_{21} \\ Q_{12} & Q_{11} \end{bmatrix}\begin{bmatrix} I_n \\ Y \end{bmatrix}.$$

我们可以将（5.3）等价地改写为

$$\begin{bmatrix} F_0 & 0 \\ -Y_0 & I_m \end{bmatrix}\begin{bmatrix} I_n \\ Y \end{bmatrix} = \begin{bmatrix} I_n & -X_0 \\ 0 & E_0 \end{bmatrix}\begin{bmatrix} Q_{22} & Q_{21} \\ Q_{12} & Q_{11} \end{bmatrix}\begin{bmatrix} I_n \\ Y \end{bmatrix}\mathscr{N}, \tag{5.6}$$

其结构与（5.1）相同。这将成为我们定义对偶矩阵束 $\mathscr{A}_{(\mathrm{d})\,0} \quad -\lambda\mathscr{B}_0^{(\mathrm{d})}$ 以满足下式的基础： $\mathscr{A}_0 - \lambda\mathscr{B}_0$

$$\mathscr{A}_0^{(\mathrm{d})}(Q_1^{(\mathrm{d})})^{\mathrm{T}}\begin{bmatrix} I \\ Y \end{bmatrix} = \mathscr{B}_0^{(\mathrm{d})}(Q_1^{(\mathrm{d})})^{\mathrm{T}}\begin{bmatrix} I \\ Y \end{bmatrix}\mathscr{N}, \tag{5.7}$$

从（5.6）出发，我们定义对偶矩阵束 $\mathscr{A}_{(\mathrm{d})}{}_0 \quad -\lambda\mathscr{B}_0^{(\mathrm{d})}$ 如下：

$$\mathscr{A}_0^{(\mathrm{d})} := \mathit{\Pi}_{m,n}^{\mathrm{T}}\mathscr{B}_0\mathit{\Pi}_{m,n},\ \mathscr{B}_0^{(\mathrm{d})} := \mathit{\Pi}_{m,n}^{\mathrm{T}}\mathscr{A}_0\mathit{\Pi}_{m,n} \tag{5.8a}$$

以给出

$$\mathscr{A}_0^{(\mathrm{d})} = \begin{matrix} \\ n \\ m \end{matrix}\begin{matrix} \begin{matrix} n & m \end{matrix} \\ \begin{bmatrix} F_0 & 0 \\ -Y_0 & I \end{bmatrix} \end{matrix}\underbrace{\mathit{\Pi}_{m,n}^{\mathrm{T}}Q_2\mathit{\Pi}_{m,n}}_{=:Q_1^{(\mathrm{d})}},\ \mathscr{B}_0^{(\mathrm{d})} = \begin{matrix} \\ n \\ m \end{matrix}\begin{matrix} \begin{matrix} n & m \end{matrix} \\ \begin{bmatrix} I & -X_0 \\ 0 & E_0 \end{bmatrix} \end{matrix}\underbrace{\mathit{\Pi}_{m,n}^{\mathrm{T}}Q_1\mathit{\Pi}_{m,n}}_{=:Q_2^{(\mathrm{d})}}, \tag{5.8b}$$

其中 $Q_1^{(\mathrm{d})}$ 和 $Q_2^{(\mathrm{d})}$ 也有定义。有趣的是，只需交换 X_0 和 Y_0，并在 $\mathscr{A}_0$ 和 $\mathscr{B}_0$ 中交换 E_0 和 F_0，再配以适当的 Q-矩阵，就能直接推导出它们。注意，$\mathscr{A}_0^{(\mathrm{d})} - \lambda\mathscr{B}_0^{(\mathrm{d})}$ 仍为 SFQ 的形式，并通过 (5.7) 得到 (4.10b)。可以验证

$$\Pi_{m,n}\mathscr{A}_0^{(\mathrm{d})} = \mathscr{B}_0\Pi_{m,n}, \quad \Pi_{m,n}\mathscr{B}_0^{(\mathrm{d})} = \mathscr{A}_0\Pi_{m,n}. \tag{5.9}$$

定理 5.1。对于 (4.8) 中的 $\mathscr{A}_0 - \lambda\mathscr{B}_0$，以及由 (5.8) 定义的 $\mathscr{A}^{(\mathrm{d})}{}_0 \quad - \lambda\mathscr{B}_0^{(\mathrm{d})}$，若 $(\lambda, \boldsymbol{z})$ 是 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 的特征对，则 $(1/\lambda, \Pi_{m,\ n}^{\mathrm{T}}\boldsymbol{z})$ 是 $\mathscr{A}^{(\mathrm{d})}{}_0 \quad - \lambda\mathscr{B}_0^{(\mathrm{d})}$ 的特征对，反之亦然。

*Pro*的。设 $(\lambda, \boldsymbol{z})$ 是 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 的特征对，即 $\mathscr{A}_0\boldsymbol{z} = \lambda\mathscr{B}_0\boldsymbol{z}$。由 (5.8a) 可得

$$\Pi_{m,n}\mathscr{B}_0^{(\mathrm{d})}\Pi_{m,n}^{\mathrm{T}}\boldsymbol{z} = \lambda\Pi_{m,n}\mathscr{A}_0^{(\mathrm{d})}\Pi_{m,n}^{\mathrm{T}}\boldsymbol{z},$$

这意味着 $(1/\lambda, \Pi_{m,\ n}^{\mathrm{T}}\boldsymbol{z})$ 是 $\mathscr{A}^{(\mathrm{d})}{}_0 \quad - \lambda\mathscr{B}_0^{(\mathrm{d})}$ 的特征对。同样，我们可以证明其逆命题。

□

备注 5.1。该定理为理解 SDASFQ（算法 4.1）的收敛提供了重要启示：具体而言，若 $\rho(\mathscr{M}) < 1$ 且 $\rho(\mathscr{N}) < 1$，则 X-序列和 Y-序列的极限分别与 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 在其特征值属于 $\mathbb{D}_-$ 和 $\mathbb{D}_+$时对应的特征子空间相关。

接下来，我们分别比较以输入 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 和输入 $\mathscr{A}_0^{(\mathrm{d})} - \lambda\mathscr{B}_0^{(\mathrm{d})}$ 运行 SDASFQ（算法 4.1）所得的结果。忽略任何可能发生的崩溃，我们将算法对输入 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 生成的矩阵记为 X_i, Y_i, E_i, F_i，将算法对输入 $\mathscr{A}^{(\mathrm{d})}{}_0 \quad - \lambda\mathscr{B}_0^{(\mathrm{d})}$ 生成的矩阵记为 $X_i^{(\mathrm{d})}, Y^{(\mathrm{d})}{}_i, E_i^{(\mathrm{d})}, F_i^{(\mathrm{d})}$。我们有以下定理。

定理5.2。对于所有 $i \geq 0$，我们有 $X_i^{(\mathrm{d})} = Y_i, Y^{(\mathrm{d})}{}_i \quad = X_i, E_i^{(\mathrm{d})} = F_i$ 和 $F_i^{(\mathrm{d})} = E_i$。特别地

$$\mathscr{A}_i^{(\mathrm{d})} \equiv \begin{matrix} n \\ m \end{matrix} \overset{\begin{matrix} n & m \end{matrix}}{\begin{bmatrix} E_i^{(\mathrm{d})} & 0 \\ -X_i^{(\mathrm{d})} & I \end{bmatrix}} Q_1^{(\mathrm{d})} = \Pi_{m,n}^{\mathrm{T}}\mathscr{B}_i\Pi_{m,n} \equiv \begin{matrix} n \\ m \end{matrix} \overset{\begin{matrix} n & m \end{matrix}}{\begin{bmatrix} F_i & 0 \\ -Y_i & I \end{bmatrix}} (\Pi_{m,n}^{\mathrm{T}}Q_2\Pi_{m,n}),$$

$$\mathscr{B}_i^{(\mathrm{d})} \equiv \begin{matrix} n \\ m \end{matrix} \overset{\begin{matrix} n & m \end{matrix}}{\begin{bmatrix} I & -Y_i^{(\mathrm{d})} \\ 0 & F_i^{(\mathrm{d})} \end{bmatrix}} Q_2^{(\mathrm{d})} = \Pi_{m,n}^{\mathrm{T}}\mathscr{A}_i\Pi_{m,n} \equiv \begin{matrix} n \\ m \end{matrix} \overset{\begin{matrix} n & m \end{matrix}}{\begin{bmatrix} I & -X_i \\ 0 & E_i \end{bmatrix}} (\Pi_{m,n}^{\mathrm{T}}Q_1\Pi_{m,n}).$$

$\widehat{\ }\quad\widehat{\ }\quad\widehat{\ }\quad\widehat{\ }$证明。需要指出的是，$(E_{i+1}, F_{i+1}, X_{i+1}, Y_{i+1})$ 是通过 (E_i, F_i, X_i, Y_i) 从 (E_i, F_i, X_i, Y_i) 定义得到的，而后者由以下内容唯一确定：

$$\begin{bmatrix} \widehat{E}_i & 0 & I & -\widehat{Y}_i \\ -\widehat{X}_i & I & 0 & \widehat{F}_i \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_m & -Y_i \\ 0 & F_i \end{bmatrix} Q_2 \\ \begin{bmatrix} E_i & 0 \\ -X_i & I_n \end{bmatrix} Q_1 \end{bmatrix} = 0. \tag{5.10}$$

$\widehat{\ }\ \widehat{\ }\ \widehat{\ }\ \widehat{\ }$该方程导出 (4.19)，由于 (5.10) 中最左侧矩阵具有特殊的$\widehat{\ }\ \widehat{\ }\ \widehat{\ }\ \widehat{\ }$结构形式，(4.19) 唯一确定 (E_i, F_i, X_i, Y_i)。换言之，(5.10) 唯一确定 (E_i, F_i, X_i, Y_i)。出于同样的原因，$(E_{i+1}^{(\mathrm{d})}, F_{i+1}^{(\mathrm{d})}, X_{i+1}^{(\mathrm{d})}, Y_{(\mathrm{d})\,i+1})$ 是通过 $(\widehat{E}_i^{(\mathrm{d})}, \widehat{F}_{(\mathrm{d})\,i}, \widehat{X}_i^{(\mathrm{d})}, \widehat{Y}_{(\mathrm{d})\,i})$ 从 $(E_i^{(\mathrm{d})}, F_i^{(\mathrm{d})}, X_i^{(\mathrm{d})}, Y_{(\mathrm{d})\,i})$ 定义得到的，而后者由以下内容唯一确定：

$$\begin{bmatrix} \widehat{E}_i^{(\mathrm{d})} & 0 & I & -\widehat{Y}_i^{(\mathrm{d})} \\ -\widehat{X}_i^{(\mathrm{d})} & I & 0 & \widehat{F}_i^{(\mathrm{d})} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_n & -Y_i^{(\mathrm{d})} \\ 0 & F_i^{(\mathrm{d})} \end{bmatrix} Q_2^{(\mathrm{d})} \\ \begin{bmatrix} E_i^{(\mathrm{d})} & 0 \\ -X_i^{(\mathrm{d})} & I_m \end{bmatrix} Q_1^{(\mathrm{d})} \end{bmatrix} = 0 \tag{5.11}$$

因为最左侧的矩阵恰好被限制为特殊结构形式。回顾式 (5.8b)，我们得到

$$\begin{aligned} \begin{bmatrix} \begin{bmatrix} I_n & -Y_i^{(\mathrm{d})} \\ 0 & F_i^{(\mathrm{d})} \end{bmatrix} Q_2^{(\mathrm{d})} \\ \begin{bmatrix} E_i^{(\mathrm{d})} & 0 \\ -X_i^{(\mathrm{d})} & I_m \end{bmatrix} Q_1^{(\mathrm{d})} \end{bmatrix} &= \begin{bmatrix} \begin{bmatrix} I_n & -Y_i^{(\mathrm{d})} \\ 0 & F_i^{(\mathrm{d})} \end{bmatrix} \Pi_{m,n}^{\mathrm{T}} Q_1 \Pi_{m,n} \\ \begin{bmatrix} E_i^{(\mathrm{d})} & 0 \\ -X_i^{(\mathrm{d})} & I_m \end{bmatrix} \Pi_{m,n}^{\mathrm{T}} Q_2 \Pi_{m,n} \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} -Y_i^{(\mathrm{d})} & I_n \\ F_i^{(\mathrm{d})} & 0 \end{bmatrix} Q_1 \\ \begin{bmatrix} 0 & E_i^{(\mathrm{d})} \\ I_m & -X_i^{(\mathrm{d})} \end{bmatrix} Q_2 \end{bmatrix} \Pi_{m,n} \\ &= \begin{bmatrix} & & & I_m \\ & & I_n & \\ & I_m & & \\ I_n & & & \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_m & -X_i^{(\mathrm{d})} \\ 0 & E_i^{(\mathrm{d})} \end{bmatrix} Q_2 \\ \begin{bmatrix} F_i^{(\mathrm{d})} & 0 \\ -Y_i^{(\mathrm{d})} & I_n \end{bmatrix} Q_1 \end{bmatrix} \Pi_{m,n}. \end{aligned}$$

因此，根据式 (5.11)，我们得出结论：$(\widehat{E}_i^{(\mathrm{d})}, \widehat{F}_i^{(\mathrm{d})}, \widehat{X}_i^{(\mathrm{d})}, \widehat{Y}_{(\mathrm{d})\,i})$ 满足

$$\begin{bmatrix} \widehat{E}_i^{(\mathrm{d})} & 0 & I & -\widehat{Y}_i^{(\mathrm{d})} \\ -\widehat{X}_i^{(\mathrm{d})} & I & 0 & \widehat{F}_i^{(\mathrm{d})} \end{bmatrix} \begin{bmatrix} & & & I_m \\ & & I_n & \\ & I_m & & \\ I_n & & & \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_m & -X_i^{(\mathrm{d})} \\ 0 & E_i^{(\mathrm{d})} \end{bmatrix} Q_2 \\ \begin{bmatrix} F_i^{(\mathrm{d})} & 0 \\ -Y_i^{(\mathrm{d})} & I_n \end{bmatrix} Q_1 \end{bmatrix} = 0,$$

或者等价地，

$$\begin{bmatrix} -\widehat{Y}_i^{(\mathrm{d})} & I & 0 & \widehat{E}_i^{(\mathrm{d})} \\ \widehat{F}_i^{(\mathrm{d})} & 0 & I & -\widehat{X}_i^{(\mathrm{d})} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_m & -X_i^{(\mathrm{d})} \\ 0 & E_i^{(\mathrm{d})} \end{bmatrix} Q_2 \\ \begin{bmatrix} F_i^{(\mathrm{d})} & 0 \\ -Y_i^{(\mathrm{d})} & I_n \end{bmatrix} Q_1 \end{bmatrix} = 0,$$

或者等价地,

$$\begin{bmatrix} \widehat{F}_i^{(\mathrm{d})} & 0 & I & -\widehat{X}_i^{(\mathrm{d})} \\ -\widehat{Y}_i^{(\mathrm{d})} & I & 0 & \widehat{E}_i^{(\mathrm{d})} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_m & -X_i^{(\mathrm{d})} \\ 0 & E_i^{(\mathrm{d})} \end{bmatrix} Q_2 \\ \begin{bmatrix} F_i^{(\mathrm{d})} & 0 \\ -Y_i^{(\mathrm{d})} & I_n \end{bmatrix} Q_1 \end{bmatrix} = 0. \tag{5.12}$$

现在我们可以通过归纳证明定理的结论。比较 (4.8) 与 (5.8b)，可得 $X_i^{(\mathrm{d})} = Y_i$, $Y_{(\mathrm{d})\,i} = X_i$, $E_i^{(\mathrm{d})} = F_i$，且 $F_i^{(\mathrm{d})} = E_i$ 对 $i = 0$成立。假设它们对 i 成立，现在考虑 $i+1$。由于 $X_i^{(\mathrm{d})} = Y_i$, $Y_i^{(\mathrm{d})} = X_i$,$E_i^{(\mathrm{d})} = F_i$ 和 $F_i^{(\mathrm{d})} = E_i$，(5.10) 与 (5.12) 中的第二个大型矩阵相同。根据每个方程中特殊结构形式下首个矩阵所具有的唯一性，我们得出 $\widehat{X}_i^{(\mathrm{d})} = \widehat{Y}_i$, $\widehat{Y}_{(\mathrm{d})\,i} = \widehat{X}_i$, $\widehat{E}_i^{(\mathrm{d})} = \widehat{F}_i$，且 $\widehat{F}_i^{(\mathrm{d})} = \widehat{E}_i$。最后，根据 (4.22)，有 $E_{i+1}^{(\mathrm{d})} = \widehat{E}_i^{(\mathrm{d})} E_i^{(\mathrm{d})} = \widehat{F}_i \widehat{F}_i = F_{i+1}$, $F_{(\mathrm{d})\,i+1} = \widehat{F}_i^{(\mathrm{d})} F_i^{(\mathrm{d})} = E_{i+1}$,

$X_{i+1}^{(\mathrm{d})} = X_i^{(\mathrm{d})} + \widehat{X}_i^{(\mathrm{d})} E_i^{(\mathrm{d})} = Y_{i+1}$，且 $Y_{(\mathrm{d})\,i+1} = Y_{(\mathrm{d})\,i} + \widehat{Y}^{(\mathrm{d})\,i} F_i^{(\mathrm{d})} = X_{i+1}$，归纳证明至此完成。 □

6 收敛性分析

当保结构倍增算法 SDASFQ 应用于 $\mathscr{A}_0 - \lambda\mathscr{B}_0$和 $\mathscr{A}_{(\mathrm{d})\,0} - \lambda\mathscr{B}_0^{(\mathrm{d})}$ 时，在不发生崩溃的假设下，会生成两列矩阵束 $\{\mathscr{A}_i - \lambda\mathscr{B}_i\}_{i=0}^{\infty}$ 和 $\{\mathscr{A}_{(\mathrm{d})\,i} - \lambda\mathscr{B}_i^{(\mathrm{d})}\}_{i=0}^{\infty}$。根据定理5.2，这两列序列在“相同”的意义下是一致的。此外,

$$\mathscr{A}_i Q_1^{\mathrm{T}} \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_i Q_1^{\mathrm{T}} \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}^{2^i} \quad \text{for } i = 0, 1, \ldots, \tag{6.1a}$$

$$\mathscr{A}_i Q_2^{\mathrm{T}} \begin{bmatrix} Y \\ I \end{bmatrix} \mathscr{N}^{2^i} = \mathscr{B}_i Q_2^{\mathrm{T}} \begin{bmatrix} Y \\ I \end{bmatrix} \quad \text{for } i = 0, 1, \ldots. \tag{6.1b}$$

出于实际考虑，在将$\mathscr{A} - \lambda\mathscr{B}$ 转换为 $\mathscr{A}' - \lambda\mathscr{B}'$的预设置过程中，我们应注意，应对 $\rho(\mathscr{M})$ 和 $\rho(\mathscr{N})$ 施加某些条件，以确保收敛性。还可以验证，(6.1b) 与下式相同:

$$\mathscr{A}_i^{(\mathrm{d})} (Q_1^{(\mathrm{d})})^{\mathrm{T}} \begin{bmatrix} I \\ Y \end{bmatrix} = \mathscr{B}_i^{(\mathrm{d})} (Q_1^{(\mathrm{d})})^{\mathrm{T}} \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}^{2^i} \quad \text{for } i = 0, 1, \ldots$$

根据定理5.2。

假设我们要寻找原始方程 (5.2) 的一个特殊解 $X = \Phi$，以及可能与之相关的对偶方程 (5.4) 的解 $Y = \Psi$，或者等价地，寻找方程 (4.10) 的解。对于 $X = \Phi$ 和 $Y = \Psi$，方程 (6.1) 也分别成立。将 (4.2) 中 $\mathscr{A}_i$ 和 $\mathscr{B}_i$ 的表达式代入方程 (6.1)，得到

$$E_i = \big[Q_{11}^{\mathrm{T}} + Q_{21}^{\mathrm{T}}\Phi - Y_i(Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}\Phi)\big]\mathscr{M}^{2^i}, \quad \Phi - X_i = F_i(Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}\Phi)\mathscr{M}^{2^i}, \tag{6.2a}$$

$$F_i = \big[Q_{22} + Q_{21}\Psi - X_i(Q_{12} + Q_{11}\Psi)\big]\mathscr{N}^{2^i}, \quad \Psi - Y_i = E_i(Q_{12} + Q_{11}\Psi)\mathscr{N}^{2^i}. \tag{6.2b}$$

因此，

$$\Phi - X_i = \big[Q_{22} + Q_{21}\Psi - X_i(Q_{12} + Q_{11}\Psi)\big]\mathscr{N}^{2^i}(Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}\Phi)\mathscr{M}^{2^i}, \tag{6.3a}$$

$$\Psi - Y_i = \big[Q_{11}^{\mathrm{T}} + Q_{21}^{\mathrm{T}}\Phi - Y_i(Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}\Phi)\big]\mathscr{M}^{2^i}(Q_{12} + Q_{11}\Psi)\mathscr{N}^{2^i}. \tag{6.3b}$$

定理 6.1。假设方程 (5.2) 和 (5.4) 分别存在解 $X = \Phi$ 和 $Y = \Psi$，使得

$$\rho(\mathscr{M}) \cdot \rho(\mathscr{N}) < 1, \tag{6.4}$$

并且假设将 SDASFQ（算法 4.1）应用于 SFQ (4.8) 中的 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 时，执行过程中不会发生崩溃，即在倍增迭代过程中所有逆矩阵均存在。那么，X_i 和 Y_i 分别以二次速度收敛到 Φ 和 Ψ，此外，

$$\limsup_{i\to\infty} \|X_i - \Phi\|^{1/2^i}, \quad \limsup_{i\to\infty} \|Y_i - \Psi\|^{1/2^i} \le \rho(\mathscr{M}) \cdot \rho(\mathscr{N}), \tag{6.5a}$$

$$\limsup_{i\to\infty} \|E_i\|^{1/2^i} \le \rho(\mathscr{M}), \quad \limsup_{i\to\infty} \|F_i\|^{1/2^i} \le \rho(\mathscr{N}). \tag{6.5b}$$

该定理表明，在适当条件下，X_i 和 Y_i 都至少以 $[\rho(\mathscr{M}) \cdot \rho(\mathscr{N})]^{2^i}$ 的速度收敛到 Φ 和 Ψ。

证明。我们只证明 (6.5a) 中的第一个不等式和 (6.5b) 中的第二个不等式，其余两个不等式可用类似方法证明。令 $Z_i = \Phi - X_i$。由 (6.3a) 可得

$$\begin{aligned} Z_i &= \big[Q_{22} + Q_{21}\Psi - (\Phi - Z_i)(Q_{12} + Q_{11}\Psi)\big]\mathscr{N}^{2^i}(Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}\Phi)\mathscr{M}^{2^i} \\ \Rightarrow \|Z_i\| &\le \Big[\|Q_{22} + Q_{21}\Psi - \Phi(Q_{12} + Q_{11}\Psi)\| + \|Z_i\|\,\|Q_{12} + Q_{11}\Psi\|\Big] \\ &\quad \times \|\mathscr{N}^{2^i}\|\,\|Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}\Phi\|\,\|\mathscr{M}^{2^i}\|. \end{aligned}$$

令 $\epsilon_i = \|\mathscr{N}^{2^i}\|\,\|\mathscr{M}^{2^i}\|$。众所周知，

$$\lim_{i\to\infty} (\epsilon_i)^{1/2^i} = \lim_{i\to\infty} \|\mathscr{N}^{2^i}\|^{1/2^i}\|\mathscr{M}^{2^i}\|^{1/2^i} = \rho(\mathscr{M}) \cdot \rho(\mathscr{N}).$$

特别地，当 $i \to \infty$ 时，$\epsilon_i \to 0$；若 $\rho(\mathscr{M}) \cdot \rho(\mathscr{N}) < 1$，这正是我们的假设。因此，对于充分大的 i，$1 - \|Q_{12} + Q_{11}\Psi\|\,\|Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}\Phi\|\,\epsilon_i > 0$，从而

$$\|Z_i\| \le \frac{\|Q_{22} + Q_{21}\Psi - \Phi(Q_{12} + Q_{11}\Psi)\|\,\|Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}\Phi\|}{1 - \|Q_{12} + Q_{11}\Psi\|\,\|Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}\Phi\|\,\epsilon_i}\epsilon_i \to 0 \quad \text{as } i \to \infty.$$

此外，由于 $i \to \infty$

$$\|Z_i\|^{1/2^i} \le \left[\frac{\|Q_{22} + Q_{21}\Psi - \Phi(Q_{12} + Q_{11}\Psi)\|\,\|Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}\Phi\|}{1 - \|Q_{12} + Q_{11}\Psi\|\,\|Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}\Phi\|\,\epsilon_i}\right]^{1/2^i} (\epsilon_i)^{1/2^i} \to \rho(\mathscr{M}) \cdot \rho(\mathscr{N}).$$

这证明了 (6.5a) 中的第一个不等式，也意味着 $\|X_i\|$ 一致有界，即存在一个常数 ω，使得对于所有 i，都有 $\|X_i\| \le \omega$。由 (6.2b) 中的第一个方程可知

$$\begin{aligned} & \|F_i\| \le \big(\|Q_{22}+Q_{21}\Psi\|+\omega\|Q_{12}+Q_{11}\Psi\|\big)\|\mathscr{N}^{2^i}\| \\ \Rightarrow\quad & \|F_i\|^{1/2^i} \le \big(\|Q_{22}+Q_{21}\Psi\|+\omega\|Q_{12}+Q_{11}\Psi\|\big)^{1/2^i}\|\mathscr{N}^{2^i}\|^{1/2^i} \to \rho(\mathscr{N}) \end{aligned}$$

如 $i\to\infty$。这就证明了 (6.5b) 中的第二个不等式。 □

注 6.1。在定理 6.1 中，假设方程 (5.2) 和 (5.4) 分别存在解 $X=\Phi$ 和 $Y=\Psi$，使得 $\rho(\mathscr{M})\cdot\rho(\mathscr{N})<1$，这对 $\mathscr{A}_0-\lambda\mathscr{B}_0$ 的特征值分布有重要影响：它有 m 个特征值位于 $\mathbb{D}_{\varrho-}$，其余 n 个特征值位于 $\mathbb{D}_{\varrho+}$，其中 ϱ 是任意满足 $\rho(\mathscr{M})<\varrho<1/\rho(\mathscr{N})$ 的数。此外，$\mathcal{R}(Q_1^{\mathrm{T}}\begin{bmatrix}I_m\\ \Phi\end{bmatrix})$ 是与位于 $\mathbb{D}_{\varrho-}$ 中的 m 个特征值对应的特征子空间，而 $\mathcal{R}(Q_2^{\mathrm{T}}\begin{bmatrix}\Psi\\ I_n\end{bmatrix})$ 是与位于 $\mathbb{D}_{\varrho+}$ 中的 n 个特征值对应的特征子空间。为说明这些结论，我们注意到 $\rho(\mathscr{N})<1/\varrho$，且 $\mathrm{eig}(\mathscr{N})\subset\mathrm{eig}(\mathscr{A}(\mathrm{d})\ 0\ ,\mathscr{B}_0^{(\mathrm{d})})$，因此由定理 5.1 可知，$\mathscr{A}_0-\lambda\mathscr{B}_0$ 在 $\mathbb{D}_{\varrho+}$ 中有 n 个特征值，它们是 $\mathrm{eig}(\mathscr{N})$ 中特征值的倒数。在本节之前，我们的讨论和动机一直聚焦于这样一种情形：$\mathscr{A}_0-\lambda B_0$ 的谱可分为两部分，分别位于 $\mathbb{D}_-$ 或 $\mathbb{D}_+$。这样做只是为了便于阐述和理解。

定理 6.1 的一个关键条件是 (6.4)，我们称之为正则情形。正如我们所见，其分析相当简单。值得注意的是，在正则情形下，只要 $\rho(\mathscr{M})$ 和 $\rho(\mathscr{N})$ 的乘积小于 1，其中一个就可能大于 1。SDASFQ 的 (6.3) 中的误差方程强烈表明，如果 $\rho(\mathscr{M})\cdot\rho(\mathscr{N})>1$，倍增算法可能会发散。然而，即使在

$$\rho(\mathscr{M})\cdot\rho(\mathscr{N})=1, \tag{6.6}$$

我们将其称为临界情形。临界情形下的收敛性分析相当复杂。为避免打断文章的主要脉络，我们将分析推迟至附录 A。

7 Q-倍增算法及实现

在第3节中，我们论证了引入置换矩阵 Q_1 和 Q_2 的必要性，以便如(3.1)那样稳健地表示所关注特征子空间的基矩阵，并且我们还指出，除了DARE、 CARE或 MARE等少数情形（此时两者均为单位矩阵）外，它们不太可能事先已知。在本节中，我们将描述最初构造 Q_1 和 Q_2 以及在倍增迭代过程中自适应更新它们的方法。我们将假定当前处于阶段 (2.3b)。

为便于表述，我们将采用MATLAB式约定来访问向量和矩阵的元素。 $i\colon\ j$ 是从 i 到 j （含端点）的整数集合。对于向量 u 和矩阵 X， $u_{(j)}$ 是 u的第 j个元素， $X_{(i,j)}$ 是 X 的第(i,j)个元素； X的子矩阵 $X_{(k:\ell,i:j)}$、$X_{(k:\ell,:)}$和 $X_{(:,i:j)}$ 分别由第 k 行至第 ℓ行与第 i 列至第 j列的交集、第 k 行至第 ℓ行，以及第 i 列至第 j列构成。

7.1 初始化

回顾一下，在初始化结束时，我们需要得到

$$P\left[\mathscr{A}'Q_1^{\mathrm{T}},\mathscr{B}'Q_2^{\mathrm{T}}\right]=\begin{bmatrix}E_0 & 0 & I_m & -Y_0\\ -X_0 & I_n & 0 & F_0\end{bmatrix}, \tag{7.1}$$

其中 Q_1 和 Q_2 是两个置换矩阵，而 P 是某个非奇异矩阵，实际上完全不需要它本身，而只需其存在性。由于我们不知道 Q_1 和Q_2，因此无法使用定理 4.1 进行初始化。在本小节中，我们说明三种思路，它们可能通过适当的置换 Q_1 和 Q_2 得到 (7.1)，从而使 $\|X_0\|$ 和 $\|Y_0\|$ 有望不会太大。每种思路都会产生两种约化版本，分别从先处理 $\mathscr{A}'$ 或先处理 $\mathscr{B}'$ 开始。从数学上看，这两种版本是相同的，但对于给定的 $\mathscr{A}'-\lambda B'$，数值表现可能不同。

思路 1

由该思路得到的第一个版本包含三个步骤:

（1）我们对 $\mathscr{A}'$ 在最后 n 行内进行带完全选主元的反向高斯消元，从位置 $(m+n,$ $m+n)$ 向后进行到位置 $(m+1,m+1)$，共 n 步。若成功，我们将由行选主元引起的所有列置换收集到 Q_1^{T}中；此时，更新后的矩阵 $\mathscr{A}'$ 的右下 $n\times n$ 子矩阵是一个下三角矩阵，为了继续进行，必须假定它是非奇异的[6]。现在，我们将 n 次高斯消元步骤中的所有矩阵，以及将最后 n 行乘以下三角矩阵的逆矩阵所得的那个矩阵，一并收集到 P_1中。至此，我们有

$$P_1\mathscr{A}'Q_1^{\mathrm{T}}=\begin{bmatrix}\widetilde{E}_0 & 0\\ -\widetilde{X}_0 & I_n\end{bmatrix}. \tag{7.2}$$

在反向高斯过程中得到 (7.2) 时， P_1 由一系列从左侧进行的简单矩阵乘法和行置换组成。这些操作也同时作用于 $\mathscr{B}'$ ，即当得到 (7.2) 时， $P_1\mathscr{B}'$也随之得到。

(2) 接下来，我们在 $P_1\mathscr{B}'$ 上执行常规的高斯消元，在前 m 行内采用完全选主元，从位置 (1,1) 开始，向位置 (m,m) 推进，共进行 m 步。如果成功，我们将行选主元产生的所有列置换收集到

[6]否则，我们可以尝试这一思路的第二种版本，或下文思路 2 和思路 3 中的任何其他约化。

Q_2^{T} 中；此时，更新后的矩阵 $P_1\mathscr{B}'$的 $m\times m$ 左上子矩阵是一个上三角矩阵，，为继续后续步骤，假定它是非奇异的。现在，我们将 m 步高斯消元产生的所有矩阵，以及将前 m 行乘以上三角矩阵的逆矩阵所得的矩阵，收集到 P_2中。此时，

$$P_2P_1\mathscr{B}'Q_2^{\mathrm{T}} = \begin{bmatrix} I_m & -Y_0 \\ 0 & F_0 \end{bmatrix}.$$

(3) 最后，令 $P = P_2P_1$。注意，$P\mathscr{A}'Q_1^{\mathrm{T}} = P_2(P_1\mathscr{A}'Q_1^{\mathrm{T}})$ 保持了 (7.2) 中的块$\widetilde{\quad}$ $\widetilde{\quad}$结构，但其中的 E_0 和 X_0 已更新为 (7.1) 中的 E_0 和 X_0 。因此，我们将得到 (7.1)。

在上文的两个不同位置，都必须假设一个下三角矩阵和一个上三角矩阵是非奇异的。然而，如果其中任意一个是奇异的，那么思路 1 将无法得到 (7.1)。

与刚才所述类似，第二个版本与之相同，只是交换 $\mathscr{A}'$ 或与 $\mathscr{B}'$ 的角色；也就是说，先在 $\mathscr{B}'$ 的前 m 行内对 $\mathscr{B}'$ 进行完全选主元的常规高斯消元，同时更新 $\mathscr{A}'$，然后在最后 n 行内对更新后的 $\mathscr{A}'$ 进行完全选主元的反向高斯消元。细节从略。

思路 2

这一思路的第一个版本也由三个步骤组成:
(1) 从位置（$m+n$，$m+n$）开始，向后朝位置（$m+1$，$m+1$）进行 n 步，在整个矩阵中采用完全选主元，对 $\mathscr{A}'$ 进行反向高斯消元。同样，如果成功，我们就得到 (7.2) 和 $P_1\mathscr{B}'$。这一步与思路 1 的第一步不同之处在于，它是在整个矩阵中选择主元。

步骤 (2) 和 (3) 与思路 1 中的相同。这是因为一旦建立了 (7.2)，我们就不希望在处理 $P_1\mathscr{B}'$ 时破坏其右侧的结构。与思路 1 一样，思路 2 也会遇到需要求下三角矩阵和上三角矩阵逆的情况。如果其中任一个矩阵是奇异的，那么思路 2 就无法得到 (7.1)。但预计思路 2 成功的机会更大，因为步骤 (1) 在整个矩阵中进行了完全选主元；出于同样的原因，它很可能会产生一个 $\|X_0\|$和 $\|Y_0\|$较小的初始化。

另一种做法是先处理 $\mathscr{B}'$，再处理 updated$\mathscr{A}'$。这是第二个版本。细节从略。

思路 3

在思路 1 和思路 2 中，我们都提出先只处理 $\mathscr{A}'$ 或 $\mathscr{B}'$，同时对另一个执行相同的操作，然后切换到处理另一个矩阵，并在高斯消元后保留第一个矩阵的结构。

在思路 3 中，我们提出交替处理 $\mathscr{A}'$ 和 $\mathscr{B}'$。第一个版本如下：

(1) 对 $\mathscr{A}'$ 执行反向高斯消元的第一步，在整个 $\mathscr{A}'$ 范围内进行完全选主元；随后对更新后的 $\mathscr{B}'$ 执行高斯消元的第一步，在前 $m+n-1$ 行内进行完全选主元；

(2) 对更新后的 $\mathscr{A}'$ 执行反向高斯消元的第二步，在更新后的 $\mathscr{A}'_{(2:m+n-1,1:m+n-1)}$ 范围内进行完全选主元；随后对更新后的 $\mathscr{B}'$ 执行一步高斯消元，在更新后的 $\mathscr{B}'_{(2:m+n-2,2:m+n)}$ 范围内进行完全选主元；

(3) 重复这一过程。

先处理 $\mathscr{B}'$，再以相同的交替方式处理更新后的 $\mathscr{A}'$，即可得到第二个版本。

7.2 倍增迭代

在倍增迭代过程中，我们必须监控 $\|X_i\|$ 和 $\|Y_i\|$。如果其中一个或两者被认为过大，我们就必须更新 Q_1 和/或 Q_2，以使 $\|X_i\|$ 和 $\|Y_i\|$ 降下来。但多大才算太大？前面的定理 3.1 为我们提供了指导，下面将对此进行说明。引理 7.1 是一个众所周知的引理，之后会派上用场。

引理 7.1. 设 $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^k$。若 $1-\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u} \neq 0$，则 $I_k - \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ 可逆，且

$$(I_k - \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})^{-1} = I_k + \frac{\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}}{1-\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u}}.$$

回到倍增迭代。在收敛时，将 X_i 和 Y_i 的极限代入 (3.1)，得到对应特征子空间的两个基矩阵；它们也可以由某些标准正交基矩阵导出。因此，理想情况下，通过适当的置换矩阵 Q_1和 Q_2，可以将 $\|X\|_2$ 和 $\|Y\|_2$ 都控制在不超过 $\sqrt{nm+1}$ 的范围内。这表明，在倍增迭代过程中，将 X_i 和 Y_i 的元素控制在某个 $\sqrt{nm+1}$倍数以内是合理的。目前，在我们的数值算例中，我们会检查 X_i 或 Y_i 是否存在大于

$$\tau = \max\{10^3, 10\sqrt{nm+1}\}. \tag{7.3}$$

如果存在这样的元素，就需要将其减小。最简单的做法是调用第7.1小节中的某个初始化过程，以得到

$$P[(\mathscr{A}_i Q_1^{\mathrm{T}})S_1^{\mathrm{T}}, (\mathscr{B}_i Q_2^{\mathrm{T}})S_2^{\mathrm{T}}] = \begin{bmatrix} \widetilde{E}_i & 0 & I_m & -\widetilde{Y}_i \\ -\widetilde{X}_i & I_n & 0 & \widetilde{F}_i \end{bmatrix} \tag{7.4}$$

~ ~并且希望 X_i 和 Y_i 的 2-范数都远小于预先选定的 τ。现在，~ ~ ~ ~将 (X_i, Y_i, E_i, F_i) 替换为 (X_i, Y_i, E_i, F_i)，将 Q_1 更新为 S_1Q_1，并将 Q_2 更新为 S_2Q_2，然后从相应更新后的 $\mathscr{A}_i$ 和 $\mathscr{B}_i$继续倍增过程。

我们刚才讨论的是在第7.1小节中重用初始化过程。这样做时，我们完全忽略了当前 $\mathscr{A}_i Q_1^{\mathrm{T}}$ 和 $\mathscr{B}_i Q_2^{\mathrm{T}}$ 中已有的结构。可以设想，利用这些结构能够提高数值效率。回想一下，

$$\left[\mathscr{A}_i Q_1^{\mathrm{T}}, \mathscr{B}_i Q_2^{\mathrm{T}}\right] = \begin{bmatrix} E_i & 0 & I_m & -Y_i \\ -X_i & I_n & 0 & F_i \end{bmatrix}.$$

为便于表述，我们省略 i 到 X_i、Y_i、E_i、F_i 的下标。我们的思路是反复执行以下两个操作，以控制 X 和 Y。

操作（i）假设 $X_{(j,\ell)}$ 是 X 中幅值最大的元素，即

$$(j,\ell) = \arg\max_{(i,k)} \{|X_{(i,k)}| : 1 \le i \le n,\, 1 \le k \le m\}, \tag{7.5}$$

同时也有 $|X_{(j,\ell)}| > \tau$，即最大幅值容差，例如 (7.3) 中所示。记为 $\boldsymbol{x} = X_{(:,\ell)}$ 和 $\boldsymbol{h} = E_{(:,\ell)}$。令 S^{T} 为满足以下条件的置换矩阵

$$\begin{bmatrix} E & 0 \\ -X & I_n \end{bmatrix} S^{\mathrm{T}} = \begin{bmatrix} E - \boldsymbol{h}\boldsymbol{e}_\ell^{\mathrm{T}} & \boldsymbol{h}\boldsymbol{e}_j^{\mathrm{T}} \\ -X + (\boldsymbol{x}+\boldsymbol{e}_j)\boldsymbol{e}_\ell^{\mathrm{T}} & I_n - (\boldsymbol{x}+\boldsymbol{e}_j)\boldsymbol{e}_j^{\mathrm{T}} \end{bmatrix},$$

即交换最左侧矩阵的第 ℓ 列和第 $(m+j)$ 列。根据引理 7.1，我们得到

$$\left[I_n - (\boldsymbol{x}+\boldsymbol{e}_j)\boldsymbol{e}_j^{\mathrm{T}}\right]^{-1} = I_n + \frac{(\boldsymbol{x}+\boldsymbol{e}_j)\boldsymbol{e}_j^{\mathrm{T}}}{1 - \boldsymbol{e}_j^{\mathrm{T}}(\boldsymbol{x}+\boldsymbol{e}_j)} = I_n - \frac{\boldsymbol{x}+\boldsymbol{e}_j}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}}\,\boldsymbol{e}_j^{\mathrm{T}}.$$

注意，$\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x} = X_{(j,\ell)}$ 是 X 中幅值最大的元素。令

$$P_1 = \begin{bmatrix} I_m & \\ & \left[I_n - (\boldsymbol{x}+\boldsymbol{e}_j)\boldsymbol{e}_j^{\mathrm{T}}\right]^{-1} \end{bmatrix}, \quad P_2 = \begin{bmatrix} I_m & -\boldsymbol{h}\boldsymbol{e}_j^{\mathrm{T}} \\ 0 & I_n \end{bmatrix}.$$

我们有

$$P_1\left[(\mathscr{A}_i Q_1^{\mathrm{T}})S^{\mathrm{T}}, \mathscr{B}_i Q_2^{\mathrm{T}}\right] = \begin{bmatrix} E - \boldsymbol{h}\boldsymbol{e}_\ell^{\mathrm{T}} & \boldsymbol{h}\boldsymbol{e}_j^{\mathrm{T}} & I_m & -Y \\ -\widetilde{X} & I_n & 0 & \widetilde{F} \end{bmatrix},$$

其中

$$\begin{aligned} \widetilde{X} &= \left[I_n - \frac{(\boldsymbol{x}+\boldsymbol{e}_j)\boldsymbol{e}_j^{\mathrm{T}}}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}}\right]\left[X - (\boldsymbol{x}+\boldsymbol{e}_j)\boldsymbol{e}_\ell^{\mathrm{T}}\right] \\ &= X + \frac{(\boldsymbol{x}+\boldsymbol{e}_j)\boldsymbol{e}_\ell^{\mathrm{T}}}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}} - \frac{(\boldsymbol{x}+\boldsymbol{e}_j)}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}}\boldsymbol{e}_j^{\mathrm{T}}X \\ &= X + \frac{\boldsymbol{x}+\boldsymbol{e}_j}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}}\left[\boldsymbol{e}_\ell^{\mathrm{T}} - \boldsymbol{e}_j^{\mathrm{T}}X\right], \end{aligned} \tag{7.6a}$$

$$\widetilde{F} = \left[I_n - \frac{(\boldsymbol{x}+\boldsymbol{e}_j)\boldsymbol{e}_j^{\mathrm{T}}}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}}\right]F$$

$$= F - \frac{\boldsymbol{x}+\boldsymbol{e}_j}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}}\left(\boldsymbol{e}_j^{\mathrm{T}}F\right), \tag{7.6b}$$

and

$$P_2P_1\left[(\mathscr{A}_iQ_1^{\mathrm{T}})S^{\mathrm{T}}, \mathscr{B}_iQ_2^{\mathrm{T}}\right] = \begin{bmatrix} \widetilde{E} & 0 & I_m & -\widetilde{Y} \\ -\widetilde{X} & I_n & 0 & \widetilde{F} \end{bmatrix},$$

其形式符合要求，其中

$$\widetilde{E} = E - \boldsymbol{h}(\boldsymbol{e}_\ell^{\mathrm{T}} - \boldsymbol{e}_j^{\mathrm{T}}\widetilde{X}) = E + \frac{\boldsymbol{h}}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}}\left[\boldsymbol{e}_\ell^{\mathrm{T}} - e_j^{\mathrm{T}}X\right], \tag{7.6c}$$

$$\widetilde{Y} = Y + \boldsymbol{h}(\boldsymbol{e}_j^{\mathrm{T}}\widetilde{F}) = Y - \frac{\boldsymbol{h}}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}}\left(\boldsymbol{e}_j^{\mathrm{T}}F\right). \tag{7.6d}$$

除 $Q_1 \leftarrow SQ_1$ 外，公式 (7.6) 还展示了当 X 中的某个元素被认为过大时，如何更新 Q-标准型。

操作（ii）假设 $Y_{(j,\ell)}$ 是 Y 中幅值最大的元素，同时也超过了 $|Y_{(j,\ell)}| > \tau$，即最大幅值容差，例如 (7.3) 中所示。记为 $\boldsymbol{y} = Y_{(:,\ell)}$ 和 $\boldsymbol{h} = F_{(:,\ell)}$。令 S^{T} 为置换矩阵，使得

$$\begin{bmatrix} I_m & -Y \\ 0 & F \end{bmatrix} S^{\mathrm{T}} = \begin{bmatrix} I_m - (\boldsymbol{y}+\boldsymbol{e}_j)\boldsymbol{e}_j^{\mathrm{T}} & -Y + (\boldsymbol{y}+\boldsymbol{e}_j)\boldsymbol{e}_\ell^{\mathrm{T}} \\ \boldsymbol{h}\boldsymbol{e}_j^{\mathrm{T}} & F - \boldsymbol{h}\boldsymbol{e}_\ell^{\mathrm{T}} \end{bmatrix},$$

也就是说，交换最左侧矩阵的第 j 列与第 $(m+\ell)$ 列。根据引理 7.1，可得

$$\left[I_m - (\boldsymbol{y}+\boldsymbol{e}_j)\boldsymbol{e}_j^{\mathrm{T}}\right]^{-1} = I_m + \frac{(\boldsymbol{y}+\boldsymbol{e}_j)\boldsymbol{e}_j^{\mathrm{T}}}{1 - \boldsymbol{e}_j^{\mathrm{T}}(\boldsymbol{y}+\boldsymbol{e}_j)} = I_m - \frac{(\boldsymbol{y}+\boldsymbol{e}_j)\boldsymbol{e}_j^{\mathrm{T}}}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{y}}.$$

Let

$$P_1 = \begin{bmatrix} \left[I_m - (\boldsymbol{y}+\boldsymbol{e}_j)\boldsymbol{e}_j^{\mathrm{T}}\right]^{-1} & \\ & I_n \end{bmatrix}, \quad P_2 = \begin{bmatrix} I_m & 0 \\ -\boldsymbol{h}\boldsymbol{e}_j^{\mathrm{T}} & I_n \end{bmatrix}.$$

我们有

$$P_1\left[\mathscr{A}_iQ_1^{\mathrm{T}}, (\mathscr{B}_iQ_2^{\mathrm{T}})S^{\mathrm{T}}\right] = \begin{bmatrix} \widetilde{E} & 0 & I_m & -\widetilde{Y} \\ -X & I_n & \boldsymbol{h}\boldsymbol{e}_j^{\mathrm{T}} & F - \boldsymbol{h}\boldsymbol{e}_\ell^{\mathrm{T}} \end{bmatrix},$$

其中

$$\begin{aligned} \widetilde{E} &= \left[I_m - \frac{(\boldsymbol{y}+\boldsymbol{e}_j)\boldsymbol{e}_j^{\mathrm{T}}}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{y}}\right]E \\ &= E - \frac{\boldsymbol{y}+\boldsymbol{e}_j}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{y}}\left(\boldsymbol{e}_j^{\mathrm{T}}E\right), \end{aligned} \tag{7.7a}$$

$$\widetilde{Y} = \left[I_n - \frac{(\boldsymbol{y}+\boldsymbol{e}_j)\boldsymbol{e}_j^{\mathrm{T}}}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{y}}\right]\left[Y - (\boldsymbol{y}+\boldsymbol{e}_j)\boldsymbol{e}_\ell^{\mathrm{T}}\right]$$

$$= Y + \frac{(\boldsymbol{y}+\boldsymbol{e}_j)\boldsymbol{e}_\ell^{\mathrm{T}}}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{y}} - \frac{(\boldsymbol{y}+\boldsymbol{e}_j)}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{y}}\boldsymbol{e}_j^{\mathrm{T}}Y$$
$$= Y + \frac{\boldsymbol{y}+\boldsymbol{e}_j}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{y}}\left[\boldsymbol{e}_\ell^{\mathrm{T}} - \boldsymbol{e}_j^{\mathrm{T}}Y\right], \tag{7.7b}$$

and

$$P_2P_1\left[\mathscr{A}_iQ_1^{\mathrm{T}}, (\mathscr{B}_iQ_2^{\mathrm{T}})S^{\mathrm{T}}\right] = \begin{bmatrix} \widetilde{E} & 0 & I_m & -\widetilde{Y} \\ -\widetilde{X} & I_n & 0 & \widetilde{F} \end{bmatrix},$$

其形式符合要求，其中

$$\widetilde{X} = X + \boldsymbol{h}(\boldsymbol{e}_j^{\mathrm{T}}\widetilde{E}) = X - \frac{\boldsymbol{h}}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{y}}\left(\boldsymbol{e}_j^{\mathrm{T}}E\right), \tag{7.7c}$$

$$\widetilde{F} = F - \boldsymbol{h}(\boldsymbol{e}_\ell^{\mathrm{T}} - \boldsymbol{e}_j^{\mathrm{T}}\widetilde{Y}) = F + \frac{\boldsymbol{h}}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{y}}\left[\boldsymbol{e}_\ell^{\mathrm{T}} - e_j^{\mathrm{T}}Y\right]. \tag{7.7d}$$

结合 $Q_2 \leftarrow SQ_2$，公式（7.7）展示了当 Y 的某个元素被认为过大时，如何更新 Q-标准型。

~ ~ ~ ~在下一个定理中，我们将通过上述两种操作来界定所得 X、Y、E和 F的元素。

定理7.1。在上述操作（i）之后，我们有

$$|\widetilde{X}_{(i,k)}| \le \begin{cases} \tau^{-1}, & \textit{for } i = j,\ k = \ell, \\ 1, & \textit{for } i \ne j,\ k = \ell,\ \textit{or } i = j,\ k \ne \ell, \\ 2|X_{(i,k)}|, & \textit{for } i \ne j,\ k \ne \ell, \end{cases} \tag{7.8a}$$

$$|\widetilde{Y}_{(i,k)}| \le |Y_{(i,k)}| + |E_{(i,\ell)}|\,|F_{(j,k)}|/\tau. \tag{7.8b}$$

证明。回顾操作（i）中的（7.5）。由（7.6a）可知

$$\widetilde{X}_{(:,\ell)} = \widetilde{X}\boldsymbol{e}_\ell = \boldsymbol{x} + \frac{\boldsymbol{x}+\boldsymbol{e}_j}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}}\left[1 - \boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}\right] = \frac{\boldsymbol{x}+\boldsymbol{e}_j}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}} - \boldsymbol{e}_j,$$

$$\widetilde{X}_{(i,\ell)} = \boldsymbol{e}_i^{\mathrm{T}}\widetilde{X}\boldsymbol{e}_\ell = \begin{cases} 1/(\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}), & \text{for } i = j, \\ (\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{x})/(\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}), & \text{for } i \ne j, \end{cases}$$

得到 ng $|\widetilde{X}_{(i,\ell)}| \le \tau^{-1}$ 对于 $i = j$ 为 $i \ne j$，对于 k列中的第 $k \ne \ell$ 个，我们有 e

$$\widetilde{X}_{(:,k)} = \widetilde{X}\boldsymbol{e}_k = X\boldsymbol{e}_k + \frac{\boldsymbol{x}+\boldsymbol{e}_j}{\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}}\left[-\boldsymbol{e}_j^{\mathrm{T}}X\boldsymbol{e}_k\right],$$

$$\widetilde{X}_{(i,k)} = \boldsymbol{e}_i^{\mathrm{T}}\widetilde{X}\boldsymbol{e}_k = \begin{cases} -(\boldsymbol{e}_j^{\mathrm{T}}X\boldsymbol{e}_k)/(\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x}), & \text{for } i = j, \\ \left[1 - (\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{x})/(\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{x})\right](\boldsymbol{e}_i^{\mathrm{T}}X\boldsymbol{e}_k), & \text{for } i \ne j, \end{cases}$$

~从而，对于 $k \neq \ell$， $|X_{(i,k)}| \leq 1$ 为 $i = j$，而 $2|X_{(i,k)}|$ 为 $i \neq j$。现在我们考察 Y 的~个元素。由 (7.6d) 可得

$$|\widetilde{Y}_{(i,k)}| \leq |Y_{(i,k)}| + |E_{(i,\ell)}|\,|F_{(j,k)}|/\tau,$$

正如预期。 □

从 (7.8a) 可以看出，包含绝对值最大的元素 $\sim X_{(j,\ell)}$ 的整行和整列都显著减小，但对于 (i,k) 位置的元素 X，当 $i \neq j$ 或 $k \neq \ell$ 时，其绝对值可能增至原来的两倍。虽然这可能令人担忧，但我们认为这种极端情况很少见，因为它只有在以下条件同时满足时才会发生：1）$X_{(i,\ell)}$ 与 $X_{(j,\ell)}$ 符号相反，且二者绝对值几乎相等；2）$|X_{(i,k)}|\tilde{}$ 约为 τ 或更大。从 (7.8b) 可以看出，对于 (i,k) 位置的元素 Y，其绝对值即使增大，增幅也很小，因为 τ 通常设得相当大，例如 (7.3) 中的情形；并且随着倍增迭代的推进，$|E_{(i,\ell)}|\,|F_{(j,k)}|$ 趋于 0。

类似地，对于操作（ii），我们有以下估计；由于证明类似，故略去。

定理 7.2。执行上述操作（ii）后，我们有

$$|\widetilde{X}_{(i,k)}| \leq |X_{(i,k)}| + |F_{(i,\ell)}|\,|E_{(j,k)}|/\tau, \tag{7.9a}$$

$$|\widetilde{Y}_{(i,k)}| \leq \begin{cases} \tau^{-1}, & \text{for } i = j,\ k = \ell, \\ 1, & \text{for } i \neq j,\ k = \ell,\ \text{or } i = j,\ k \neq \ell, \\ 2|Y_{(i,k)}|, & \text{for } i \neq j,\ k \neq \ell. \end{cases} \tag{7.9b}$$

7.3 Q-倍增算法

综合考虑所有因素后，我们给出完整的倍增算法。我们将其称为 Q-倍增算法（QDA）。它以 $\mathscr{A}' - \lambda\mathscr{B}' \in \mathbb{C}^{(m+n)\times(m+n)}$ 为起点，其中 $\mathbb{D}_-$ 中有 m 个特征值，$\mathbb{D}_+$ 中有 n 个特征值。此条件可以放宽，以包含附录 A 中的临界情形。为简单起见，我们只讨论正则情形。

（对于任何迭代方法）一种常用的停止准则是

$$\|X_i - X_{i-1}\|_{\mathrm{F}} \leq \texttt{rtol} \times \|X_i\|_{\mathrm{F}}, \tag{7.10}$$

其中 rtol 是给定的相对容差，但这里有一个危险的陷阱：当迭代经历一段极其缓慢的变化、因而满足（7.10）时，可能出现假收敛，而此时 X_i 仍远未达到目标。在这种情况下，我们可以检查 $Q_1^{\mathrm{T}}\begin{bmatrix} I_m \\ X_i \end{bmatrix}$ 的归一化残差，将其视为 $\mathscr{A} - \lambda\mathscr{B}$ 的近似特征基矩阵，而 $\mathscr{A}' - \lambda\mathscr{B}'$ 是通过变换从该矩阵得到的。

或者，我们可以使用 Kahan 准则

$$\frac{\|X_i - X_{i-1}\|_{\mathrm{F}}^2}{\|X_{i-1} - X_{i-2}\|_{\mathrm{F}} - \|X_i - X_{i-1}\|_{\mathrm{F}}} \leq \texttt{rtol} \times \|X_i\|_{\mathrm{F}}. \tag{7.11}$$

算法 7.1 QDA：Q-倍增算法 输入：$\mathscr{A}'-\lambda\mathscr{B}'\in\mathbb{C}^{(m+n)\times(m+n)}$，其在 $\mathbb{D}_-$ 中有 m 个特征值，在 $\mathbb{D}_+$ 中有 n 个特征值；输出：Q_1、X、Q_2 和 Y，其中 $Q_1^{\mathrm{T}}\begin{bmatrix}I_m\\X\end{bmatrix}$ 和 $Q_2^{\mathrm{T}}\begin{bmatrix}Y\\I_n\end{bmatrix}$ 分别是与 $\mathscr{A}'-\lambda\mathscr{B}'$ 在 $\mathbb{D}_-$ 中的 m 个特征值以及在 $\mathbb{D}_+$ 中的 n 个特征值相关联的基矩阵。1：根据第7.1小节中三种方法之一，将 (4.8) 中的 $\mathscr{A}'-\lambda\mathscr{B}'$ 初始化为 $\mathscr{A}_0-\lambda\mathscr{B}_0$；2：对于 $i=0$，1，.....，直至收敛，执行；3：按照 (4.14) 对 $Q_1Q_2^{\mathrm{T}}$ 进行分块；4：$W_i=Q_{22}-X_iQ_{12}-(X_iQ_{11}-Q_{21})Y_i$（或者，如果 $n>m$，则使用 $\widetilde{W}_i=Q_{11}^{\mathrm{T}}-Y_iQ_{12}^{\mathrm{T}}+(Q_{21}^{\mathrm{T}}-Y_iQ_{22}^{\mathrm{T}})X_i$）；5：按照 (4.22) 计算 E_{i+1}、F_{i+1}、X_{i+1}、Y_{i+1}（或者，如果 $n>m$，则按照 (4.30) 计算）；6：如果 X_{i+1} 和/或 Y_{i+1} 的某些元素过大，则按照第7.2小节更新 Q_1、Q_2，以及E_{i+1}、F_{i+1}、X_{i+1}、Y_{i+1}；7：循环结束；8：返回收敛时最后的 Q_1、X_i、Q_2 和 Y_i。

这一停止准则也存在我们刚才讨论过的缺陷，因此应与保护措施配合使用。若迭代至少以超线性速度收敛，则对于相同的相对容差，Kahan停止准则 (7.11) 通常比 (7.10) 更快地终止迭代。关于 Kahan 准则的更多讨论，请参阅 [23,第 3.9] 节。

8 特征子空间计算的应用

Q-倍增算法 (QDA) 的一种可能应用是计算矩阵铅笔的不变子空间

$$\mathscr{A}-\lambda\mathscr{B}\in\mathbb{R}^{(m+n)\times(m+n)},$$

假设 $\mathscr{A}-\lambda\mathscr{B}$ 在 $\mathbb{C}_-$ 中有 m 个特征值，在 $\mathbb{C}_+$ 中有 n 个特征值（或者 m 在 $\mathbb{D}_-$ 中有 n 个特征值，在 $\mathbb{D}_+$中有 n 个特征值）。用 eig($\mathscr{A},\mathscr{B}$) 表示 $\mathscr{A}-\lambda\mathscr{B}$ 的谱。

与 $\mathbb{C}_-$ 中的 m个特征值和 $\mathbb{C}_+$ 中的 n 个特征值（或者 $\mathbb{D}_-$ 中的 m个特征值和 $\mathbb{D}_+$ 中的 n 个特征值）对应的 $\mathscr{A}-\lambda\mathscr{B}$ 的不变子空间的基矩阵表示为

$$Q_1^{\mathrm{T}}\times\begin{matrix}m\\n\end{matrix}\begin{bmatrix}\overset{m}{I}\\X\end{bmatrix},\quad Q_2^{\mathrm{T}}\times\begin{matrix}m\\n\end{matrix}\begin{bmatrix}\overset{n}{Y}\\I\end{bmatrix},\tag{8.1}$$

分别地，其中未知置换矩阵 Q_1、$Q_2\in\mathbb{C}^{(m+n)\times(m+n)}$ 以及未知矩阵 $X\in\mathbb{C}^{n\times m}$、$Y\in\mathbb{C}^{m\times n}$ 将由 QDA 计算，QDA 会自适应地更新 Q_1 和 Q_2。

当 $\mathscr{A}-\lambda\mathscr{B}$ 在 $\mathbb{C}_-$ 中有 m 个特征值、在 $\mathbb{C}_+$ 中有 n 个特征值时，我们首先选择一个参数 $\gamma<0$。倍增迭代的渐近收敛速率为

$$\varrho(\gamma)=\max\left\{\varrho_-(\lambda):=\max_{\lambda\in\mathrm{eig}(\mathscr{A},\mathscr{B})\cap\mathbb{C}_-}\left|\frac{\gamma-\lambda}{\gamma+\lambda}\right|,\varrho_+(\lambda):=\max_{\lambda\in\mathrm{eig}(\mathscr{A},\mathscr{B})\cap\mathbb{C}_+}\left|\frac{\gamma+\lambda}{\gamma-\lambda}\right|\right\}<1$$

也就是说，随着 $[\varrho(\gamma)]^{2^i}$ 趋于 0，倍增迭代以相同的速度收敛，其中 i 是迭代指标 [23]。因此，理想情况下，能够带来最佳渐近收敛速度的最优 γ 为

$$\arg\min_{\gamma<0}\varrho(\gamma).$$

在实践中，对这一最优 γ 进行粗略估计也同样有效。如果已知 $\mathrm{eig}(\mathscr{A},\mathscr{B})$ 的分布情况，则可以估算出一个较好的 γ [24]。接下来，我们将矩阵铅笔 $\mathscr{A}'-\lambda\mathscr{B}'\in\mathbb{R}^{(m+n)\times(m+n)}$ 定义为

$$\begin{bmatrix}\mathscr{A}'\\ \mathscr{B}'\end{bmatrix}=\begin{bmatrix}I_{m+n} & -\gamma I_{m+n}\\ I_{m+n} & \gamma I_{m+n}\end{bmatrix}\begin{bmatrix}\mathscr{A}\\ \mathscr{B}\end{bmatrix}. \tag{8.2}$$

矩阵铅笔 $\mathscr{A}'-\lambda\mathscr{B}'$ 的特征值为 $\frac{\lambda-\gamma}{\lambda+\gamma}$，其中 $\lambda\in\mathrm{eig}(\mathscr{A},\mathscr{B})$。可以看出，$_\mathscr{A}'-\lambda\mathscr{B}'$ 在 $\mathbb{D}_{[\varrho_-(\gamma)]-}$ 中有 m 个特征值，它们来自 $_\mathbb{C}_$中 $\mathscr{A}-\lambda\mathscr{B}$ 的特征值；而在 $\mathbb{D}_{[1/\varrho_+(\gamma)]+}$ 中有 n 个特征值，它们来自 $\mathbb{C}_+$中 A 的特征值。回顾式 (8.1) 中的基矩阵，我们有

$$\mathscr{A}'Q_1^{\mathrm{T}}\begin{bmatrix}I\\ X\end{bmatrix}=\mathscr{B}'Q_1^{\mathrm{T}}\begin{bmatrix}I\\ X\end{bmatrix}\mathscr{M},\quad \mathscr{A}'Q_2^{\mathrm{T}}\begin{bmatrix}Y\\ I\end{bmatrix}\mathscr{N}=\mathscr{B}'Q_2^{\mathrm{T}}\begin{bmatrix}Y\\ I\end{bmatrix},$$

其中 $_\mathscr{M}\in\mathbb{R}^{m\times m}$ 且 $\mathscr{N}\in\mathbb{R}^{n\times n}$ 使得 $\mathrm{eig}(\mathscr{M})\subset\mathbb{D}_{[\varrho_-(\gamma)]-}$，$\mathrm{eig}(\mathscr{N})\subset\mathbb{D}_{[1/\varrho_+(\gamma)]+}$。

若 $\mathscr{A}-\lambda\mathscr{B}$ 已经在 $\mathbb{D}_-$ 中有 m 个特征值，且在 $\mathbb{D}_+$中有 n 个特征值，我们只需令 $\mathscr{A}'-\lambda\mathscr{B}'\equiv\mathscr{A}-\lambda\mathscr{B}$。最后，我们使用QDA（算法7.1）计算(8.1)中的 Q_1、Q_2、 X和 Y。

9 数值实验

在本节中，我们将设计并开展两项实验，以展示QDA相较于现有倍增算法SDASF1和SDASF2 [23]的鲁棒性。 $\mathscr{A}_0-\lambda\mathscr{B}_0$ 的初始化按照第7.1小节中的思路 3进行。

9.1 实验1

我们在三个离散化的 Bethe-Salpeter 方程（BSE）上测试 QDA（算法 7.1），分别对应[7] 萘（$C_{10}H_8$）、砷化镓（GaAs）和氮化硼（BN）。这些方程此前曾用于 [19] 中的实验，基本上采用了 SDASF1，但经过了重新表述，以便利用 BSE 的特殊结构

[7] Th作者感谢 Z.-C. Guo [19] 和 M. Shao [37] 提供这三个示例。

利用 BSE 中的特殊结构:

$$H = \begin{bmatrix} A & B \\ -\bar{B} & -\bar{A} \end{bmatrix},$$

其中 $A \in \mathbb{C}^{n\times n}$ 是厄米矩阵， $B \in \mathbb{C}^{n\times n}$ 是对称矩阵，即 $B^{\mathrm{T}} = B$。具有相同矩阵结构的特征值问题还包括线性响应特征值问题[2, 3, 4, 5, 36]。已知，如果虚轴上没有特征值，则 H 在 $\mathbb{C}_-$ 上有 n 个特征值，在 $\mathbb{C}_+$上有 n 个特征值，因为 H 是哈密顿矩阵 [23,第 2 节9]。如 [19], 所报告，SDASF1 能够计算所有三个示例中 H 特征子空间的基矩阵。下面我们将展示，与 SDASF1 相比，QDA 具有更强的鲁棒性，具体表现为收敛行为平滑得多、计算得到的 X幅值小得多，以及归一化残差小得多。为方便起见，我们可以将归一化残差定义为

$$\mathrm{NRes}_1 := \frac{\|HZ - ZM\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}(\|H\|_2 + \|M\|_2)}, \tag{9.1}$$

其中 $M = (Z^{\mathrm{H}}Z)^{-1}Z^{\mathrm{H}}HZ$，并且

$$\text{either } Z = \begin{bmatrix} I_n \\ X \end{bmatrix} \text{ for SDASF1, or } Z = Q_1^{\mathrm{T}} \begin{bmatrix} I_n \\ X \end{bmatrix} \text{ for QDA.} \tag{9.2}$$

为数值便利起见，我们用$\sqrt{\|H\|_1\|H\|_\infty}$ 估计 $\|H\|_2$ ，并类似地估计 $\|M\|_2$。但 NRes_1 可能并不合适，尤其是在 $\|X\|_{\mathrm{F}}$ 过大时；此时 Z 表示基矩阵的条件较差，更合适的归一化残差为

$$\mathrm{NRes}_2 = \frac{\|HU - U(U^{\mathrm{H}}HU)\|_{\mathrm{F}}}{\sqrt{n}(\|H\|_2 + \|U^{\mathrm{H}}HU\|_2)} \tag{9.3}$$

其中， U 是通过对 (9.2) 中的 Z 的列进行正交化得到的，例如采用薄 QR 分解： $Z = UR$，其中 $U \in \mathbb{C}^{2n\times n}$ 且 $U_{\mathrm{H}}U = I_{n\circ}$

在应用 SDASF1 和 QDA 之前，我们需要选取 $\gamma < 0$ ，并将 (8.2) 中的变换从 $\mathscr{A} - \lambda\mathscr{B} := H - \lambda I_{2n}$ 转换为 $\mathscr{A}' - \lambda\mathscr{B}'$。为简单起见，我们仅使用 $\gamma = -1$ 进行说明。表 1 展示了本文提出的 SDASF1 [19, 23] 和 QDA 在三个 BSE 上的性能统计。我们的观察总结如下：1）两者所需的倍增迭代次数相同；2）SDASF1 的运行速度远快于 QDA，这符合预期，因为 QDA 涉及的计算要复杂得多，且其当前实现中的高斯消元远未优化；3）QDA 的鲁棒性强得多，在收敛时，其归一化残差和范数 $\|X\|_{\mathrm{F}}$ 均小得多。特别需要指出的是，对于 BN，$\mathrm{NRes}_1 = 1.7 \cdot 10^{-10}$，而 $\mathrm{NRes}_2 = 6.0 \cdot 10^{-4}$，这表明 SDASF1 计算得到的 Z 对目标特征子空间的表示效果较差。在所有示例中，QDA 得到的 NRes_1 和 NRes_2 均约为 10^{-16}，这表明 QDA 计算得到的 Z 能够准确表示该特征子空间。

图 1 绘出了 X-更新的历史，其中第 i 次绝对和相对 X-更新分别定义为

$$\|\Delta_i\|_{\mathrm{F}}, \quad \frac{\|\Delta_i\|_{\mathrm{F}}}{\|X_{i+1}\|_{\mathrm{F}}}$$

表 1：BSE 的性能统计

		$C_{10}H_8$	GaAs	BN
n		32	128	2304
$\|X\|_F$	SDASF1	$1.6 \cdot 10^{5}$	$4.2 \cdot 10^{4}$	$1.7 \cdot 10^{9}$
	QDA	$5.3 \cdot 10^{-1}$	$2.6 \cdot 10^{-1}$	$3.1 \cdot 10^{-1}$
CPU	SDASF1	$9.7 \cdot 10^{-3}$	$2.3 \cdot 10^{-2}$	$4.6 \cdot 10^{1}$
	QDA	$4.7 \cdot 10^{-2}$	$1.9 \cdot 10^{-1}$	$1.8 \cdot 10^{3}$
$NRes_1$	SDASF1	$1.9 \cdot 10^{-13}$	$1.8 \cdot 10^{-13}$	$1.7 \cdot 10^{-10}$
	QDA	$7.8 \cdot 10^{-17}$	$6.3 \cdot 10^{-17}$	$1.0 \cdot 10^{-16}$
$NRes_2$	SDASF1	$6.9 \cdot 10^{-10}$	$4.2 \cdot 10^{-11}$	$6.0 \cdot 10^{-4}$
	QDA	$1.3 \cdot 10^{-16}$	$1.7 \cdot 10^{-16}$	$2.9 \cdot 10^{-16}$
#it'n	SDASF1	7	9	7
	QDA	7	9	7

在 $X_{i+1} = X_i + \Delta_i$ 的倍增迭代过程中。再次展现了 QDA 出色的鲁棒性。我们观察到：1）SDASF1 的绝对和相对 X-更新之间存在相当大的差异，因为末期较大的 $\|X\|_F$ 必然会导致较大的更新，以实现快速收敛；2）SDASF1 的收敛过程非常粗糙，事实上，相对 X-更新直到最后 2 到 3 次迭代才显示出收敛趋势，极端情况下达到 $C_{10}H_8$，此时 SDASF1 直到最后一次倍增迭代才收敛！3）对于全部三个 BSE，QDA 都表现出更平稳的收敛行为。

9.2 实验 2

我们首先构造一些随机矩阵，使其在 $\mathbb{C}_-$ 中有 m 个特征值，在$\mathbb{C}_+$ 中有 n 个特征值，并满足：与 $\mathbb{C}_-$ 中 m 个特征值相关的、如 (1.5) 所示的任意基矩阵 Z，其 Z_1 几乎奇异。假设 $m \le n$，我们用类似 MATLAB 的伪代码表示如下

$$U = \texttt{randn}(N) + \texttt{1i}*\texttt{randn}(N), \tag{9.4a}$$

$$T = \texttt{triu}(\texttt{randn}(N) + \texttt{1i}*\texttt{randn}(N), 1) + \texttt{diag}([(2\texttt{rand}(m,1) - \alpha), (2\texttt{rand}(n,1) + \alpha)]) + \texttt{1i}*\operatorname{diag}(\texttt{randn}(N,1)), \tag{9.4b}$$

其中 $N = m + n$、$\alpha > 2$ 是一个参数，然后令其随 η变化，

$$U_{(1:m,1:m)} = \eta U_{(1:m,1:m)}, \quad \mathscr{A} = UTU^{-1}, \quad \mathscr{B} = I_{m+n}, \tag{9.4c}$$

其中 η 是另一个可调参数。根据设计和理论，$\mathscr{A} - \lambda\mathscr{B}$ 在 $\mathbb{C}_-$ 中有 m个特征值，在 $\mathbb{C}_+$中有 n 个特征值，这对倍增算法的收敛至关重要；但已有观察发现，在式 (9.4c) 中构造 $\mathscr{A} - \lambda\mathscr{B}$ 时产生的舍入误差可能会破坏这一理论性质，除非将式 (9.4b) 中的参数 α 设得足够大。

接下来选取 $\gamma < 0$ 并对 $\mathscr{A} - \lambda\mathscr{B}$ 到 $\mathscr{A}' - \lambda\mathscr{B}'$执行变换 (8.2)。可以验证，$\mathscr{A}'Z = \mathscr{B}'ZM'$，其中 $M' = (M - \gamma I_m)^{-1}(M + \gamma I_m)$，其特征值位于 $\mathbb{D}_-$。从此处开始，我们可以选择两种方法来计算特征子空间 $\mathcal{R}(Z)$，其中 $Z = U_{(:,1:m)}$：1) 第 [23, 章 3] 中的现有倍增算法 SDASF1，它将 $\mathscr{A}' - \lambda\mathscr{B}'$ 变换为标准型 SF1，然后调用 SDASF1；2)

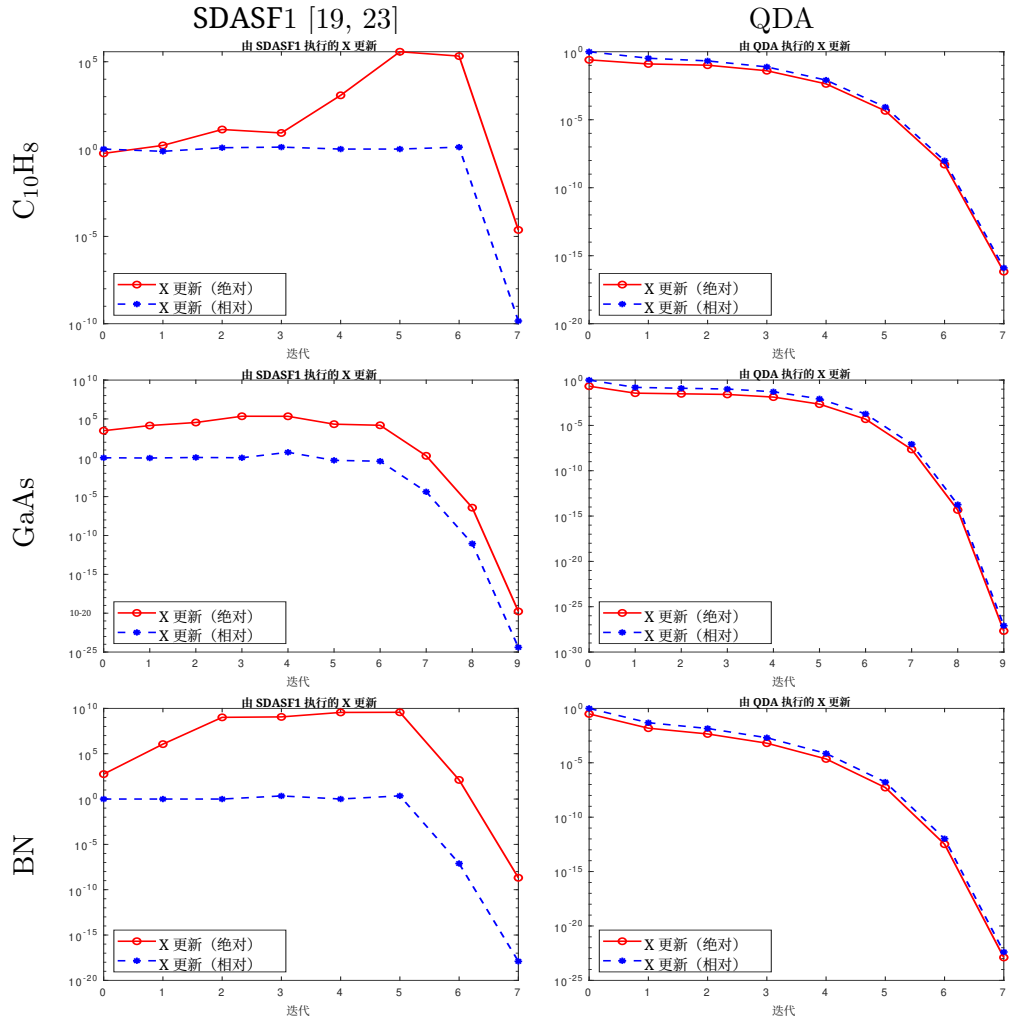

图 1：SDASF1 和 QDA 在 BSE 上进行倍增迭代期间的 X 更新历史。

新的 QDA。当 η 趋于 0 时，$Z_1 := Z_{(1:m,:)}$ 趋于 0，因此 $Z_2 Z_1^{-1}$趋于无穷大，其中 $Z_2 = Z_{(m+1:m+n,:)}$。预计随着 η 变得非常小，现有倍增算法 SDASF1 在计算 $Z_2 Z_1^{-1}$时将遇到越来越大的困难。

虽然我们按照（9.4）生成了许多随机的 $\mathscr{A} - \lambda\mathscr{B}$ 并进行了实验，但下面将报告一个典型示例：首先，使用 $m = 200$、$n = 250$ 和 $\alpha = 8$，通过（9.4a）和（9.4b）生成并保存 U 和 T，然后改变η，通过（9.4c）生成 $\mathscr{A} - \lambda\mathscr{B}$。表2展示了当 η 从 10^{-4} 变化到 10^{-7} 时的性能统计，其中 NRes_1 和 NRes_2 的定义方式与（9.1）和（9.3）相同（将其中的 H 替换为这里的 $\mathscr{A}$，注意 $\mathscr{B} = I$）。可以看出，SDASF1 在能够正常运行时比 QDA 快，但结果精度较低。事实上，当 $\eta = 10^{-7}$ 时，SDASF1 会崩溃，即产生非数值；当 $\eta = 10^{-6}$ 时，则会产生错误结果。

表2：随机（9.4）的性能统计

η		10^{-4}	10^{-5}	10^{-6}	10^{-7}
$\|X\|_{\mathrm{F}}$	SDASF1	$2.2 \cdot 10^{6}$	$2.9 \cdot 10^{7}$	–	–
	QDA	$7.8 \cdot 10^{1}$	$3.2 \cdot 10^{1}$	$3.2 \cdot 10^{1}$	$3.3 \cdot 10^{1}$
CPU	SDASF1	$1.1 \cdot 10^{-1}$	$1.0 \cdot 10^{-1}$	–	–
	QDA	$1.2 \cdot 10^{0}$	$2.3 \cdot 10^{0}$	$2.3 \cdot 10^{0}$	$2.3 \cdot 10^{0}$
NRes_1	SDASF1	$5.3 \cdot 10^{-9}$	$1.5 \cdot 10^{-8}$	–	–
	QDA	$5.2 \cdot 10^{-11}$	$8.0 \cdot 10^{-11}$	$2.4 \cdot 10^{-10}$	$1.0 \cdot 10^{-9}$
NRes_2	SDASF1	$2.8 \cdot 10^{-8}$	$1.0 \cdot 10^{-6}$	–	–
	QDA	$5.6 \cdot 10^{-11}$	$8.0 \cdot 10^{-11}$	$2.5 \cdot 10^{-10}$	$8.9 \cdot 10^{-10}$
#it'n	SDASF1	9	9	–	–
	QDA	9	8	8	8

10 结论

现有的倍增算法 [23] 用于求解源自现实世界工程应用的各种非线性矩阵方程，其实质是根据某个矩阵铅笔的特征子空间的基矩阵来计算该特征子空间。已知 m 维子空间在 $\mathbb{C}^{m+n}$ 中的基矩阵并不唯一。设 $Z = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$ 是该子空间的一个基矩阵，其中 Z_1 是一个 $m \times m$ 矩阵。当 Z_1 非奇异时，$ZZ_1^{-1} = \begin{bmatrix} I \\ Z_2 Z_1^{-1} \end{bmatrix} =: \begin{bmatrix} I \\ X \end{bmatrix}$ 也是一个基矩阵。需要指出的是，Z_1 的非奇异性取决于子空间，而非某个并非唯一的基矩阵的选取。类似地，如果 Z 的由其最后 m 行组成的子矩阵非奇异，那么它还可以表示为另一种基矩阵形式 $\begin{bmatrix} Y \\ I \end{bmatrix}$。

根据基矩阵的两种特定形式，现有的倍增算法可分为两类 [23]。本文提出了一种统一的倍增算法，通过采用灵活的基矩阵表示 $Q^{\mathrm{T}} \begin{bmatrix} I \\ X \end{bmatrix}$，弥合现有两类倍增算法之间的差距，其中 Q 是某个置换矩阵。值得注意的是，我们的表述无需任何改动即可适用于正交 Q；若将矩阵转置替换为共轭转置，也可适用于酉 Q。此外，我们还提出了一种新的标准型，称为 Q 标准型（SFQ），它统一了与现有两种倍增算法相关的第一标准型和第二标准型 SF1 与 SF2 [23]。我们将新算法命名为 Q 倍增算法（QDA），该算法会动态更新基矩阵表示中的 Q 矩阵。我们还对 QDA 进行了收敛性分析。此外，我们展示了两组实验，一组针对离散化的 Bethe-Salpeter 方程（BSE），另一组针对随机矩阵。实验表明，与现有倍增算法相比，QDA 的鲁棒性和精度更高。遗憾的是，这需要付出代价：QDA 的运行速度比现有倍增算法慢，这是预料之中的。

收敛性分析：临界情形

在第 6 节中，我们曾对正则情形下的 SDASFQ（算法 4.1）进行了收敛性分析，即 ρ（$\mathscr{M}$）$\cdot\,\rho$（$\mathscr{N}$）<1。在本附录中，我们将对临界情形进行收敛性分析，即

$$\rho(\mathscr{M})\cdot\rho(\mathscr{N})=1. \tag{A.1a}$$

我们主要采用 [23, section 3.8] 中使用的技巧。

等价地，我们可以将 (A.1a) 表示为

$$0<\varrho:=\rho(\mathscr{M})=1/\rho(\mathscr{N}). \tag{A.1b}$$

在本节中，我们始终采用这一假设。引入记号 $J_p(\omega)$ 表示对应于标量 ω 的 $p\times p$若尔当块：

$$J_p(\omega)=\begin{bmatrix}\omega & 1 & 0 & \cdots & 0\\ 0 & \omega & 1 & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & 0\\ \vdots & & \ddots & \ddots & 1\\ 0 & \cdots & \cdots & 0 & \omega\end{bmatrix}\in\mathbb{C}^{p\times p}. \tag{A.2}$$

它是上三角矩阵。众所周知，它的任意次幂仍为上三角矩阵。为了后续分析，我们需要求正整数 i 次幂，即 2^i 次幂 $J_p(\omega)$。

引理 A.1（[15,第 557] 页）。设 $J_p(\omega)$ 由 (A.2) 给出，并记

$$\big[J_p(\omega)\big]^{2^i}=\begin{bmatrix}\gamma_{1,i} & \gamma_{2,i} & \cdots & \gamma_{p,i}\\ 0 & \gamma_{1,i} & \ddots & \vdots\\ \vdots & \ddots & \ddots & \gamma_{2,i}\\ 0 & \cdots & 0 & \gamma_{1,i}\end{bmatrix}. \tag{A.3}$$

然后 $\gamma_{1,\ i}=\omega^{2^i}$ 以及

$$\gamma_{j,i}=\frac{2^i(2^i-1)\cdots(2^i-j+2)}{(j-1)!}\omega^{2^i-j+1}\quad for\ j=2,\dots,p.$$

鉴于引理 A.1，我们设定（对于 $p=2k$）

$$\mathit{\Gamma}_{i,k}:=\big\{\big[J_{2k}(\omega)\big]^{2^i}\big\}_{(1:k,k+1:2k)}=\begin{bmatrix}\gamma_{k+1,i} & \gamma_{k+2,i} & \cdots & \gamma_{2k-1,i} & \gamma_{2k,i}\\ \gamma_{k,i} & \ddots & \ddots & & \gamma_{2k-1,i}\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ \gamma_{3,i} & & \ddots & \ddots & \gamma_{k+2,i}\\ \gamma_{2,i} & \gamma_{3,i} & \cdots & \gamma_{k,i} & \gamma_{k+1,i}\end{bmatrix}. \tag{A.4}$$

特别地，$\Gamma_{0,k}=\boldsymbol{e}_k\boldsymbol{e}_1^{\mathrm{T}}$。在此处以及下文的引理 A.1 和引理 A.2 中，为使记号清晰，我们省略 $\Gamma_{i,k}$ 和 $\gamma_{i,j}$ 对 ω 的依赖。下一个引理是我们收敛性分析的核心。

引理 A.2（[22]）。(A.4) 中的托普利茨矩阵 $\Gamma_{i,\ k}$ 当i充分大时是非奇异的，并且满足

$$\left\|[J_k(\omega)]^{2^i}\Gamma_{i,k}^{-1}\right\|=\left\|\Gamma_{i,k}^{-1}[J_k(\omega)]^{2^i}\right\|=O(2^{-i}),\ \left\|[J_k(\omega)]^{2^i}\Gamma_{i,k}^{-1}[J_k(\omega)]^{2^i}\right\|=O(2^{-i}).$$

在开始分析 SDASFQ 的收敛性之前，我们先对 $\mathscr{A}_0-\lambda\mathscr{B}_0$ 作出一些通用假设。回顾 (A.1)。根据特征值位于 $\mathbb{D}_{\varrho^-}$ 内、位于 $\mathbb{T}_\varrho$ 上或位于 $\mathbb{D}_{\varrho^-}$ 之外（即位于 $\mathbb{D}_{\varrho^+}$ 内），我们将 $\mathscr{A}_0-\lambda\mathscr{B}_0$ 的特征值分为三个子集¯。下文的 (A.5) 给出了关于 $\mathscr{A}_0-\lambda\mathscr{B}_0$ 的魏尔斯特拉斯标准形的主要假设，其中，某个特征值的部分重数是与该特征值相关联的所有若尔当块的阶数。

(1) $\mathrm{eig}(\mathscr{A}_0,\mathscr{B}_0)\cap\mathbb{T}_\varrho\neq\emptyset$；(2) $\mathbb{T}_\varrho$ 上每个特征值的部分重数均为偶数；(3) 计入代数重数后，$\mathbb{T}_\varrho$ 上的特征值数量 $2n_0$ 不超过 $2\times\min\{m,n\}$，即 $n_0\le\min\{m,n\}$；以及 (4) $\mathbb{D}_{\varrho^-}$内、$\mathbb{T}_\varrho$上和 $\mathbb{D}_{\varrho^+}$ 内的特征值数量分别为 $m-n_0$、$2n_0$和 $n-n_0$。 (A.5)

(A.5) 中的第一个假设是 (A.1b) 的推论，但第二个假设相当复杂，且无法通过数值方法验证，因为魏尔斯特拉斯标准形并不连续：对 $\mathscr{A}_0$ 和 $\mathscr{B}_0$的元素进行任意微小的扰动，都可能使其发生巨大变化。因此，它仅具有理论价值。要使我们能够将 $2n_0$ 特征值分成相等的两半，一半与 $\mathbb{D}_{\varrho^-}$ 中的特征值合并，另一半与 $\mathbb{D}_{\varrho^+}$中的特征值合并（具体方式将在后续¯证明中详述），就需要这个假设。这样我们便可以得出如下结论：$\mathscr{A}_0-\lambda\mathscr{B}_0$ 在 $\mathbb{D}_{\varrho^-}$¯中有 m 个特征值，而另外 n 个特征值在 $\mathbb{D}_{\varrho^+}$中。为便于后续使用，我们引入

定义 A.1. 若矩阵铅笔 $\mathscr{A}_0-\lambda\mathscr{B}_0$ 在 $\mathbb{T}_\varrho$ 上至少有一个特征值，且 $\mathbb{T}_\varrho$ 上每个特征值的部分重数都是偶数，则称其具有性质 $\mathrm{P}^{\mathbb{T}_\varrho}$。默认情况下，当 $\varrho=1$时，我们写作性质 $\mathrm{P}_{\mathbb{T}}$。

鉴于 (A.5) 中的这些假设，我们令 $J_{2m_j}(\omega_j)$（ $j=1$、2、.....、r）表示与 $\mathbb{T}_\varrho$ 上的特征值相关联的所有若尔当块，并记为

$$J_{2m_j}(\omega_j)=\begin{bmatrix}J_{m_j}(\omega_j) & \Gamma_{0,m_j}\\ 0_{m_j} & J_{m_j}(\omega_j)\end{bmatrix},\tag{A.6}$$

其中 $\Gamma_{0,\ m_j}=\boldsymbol{e}_{m_j}\boldsymbol{e}_1^{\mathrm{T}}$。于是 $n_0=\sum_{j=1}^r m_j$。令

$$m'=m-n_0,\quad n'=n-n_0.\tag{A.7}$$

于是，根据 (A.5)，$m' \geq 0, n' \geq 0$，并且 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 在 $\mathbb{D}_{\varrho^-}$ 中有 m' 个特征值，在$\mathbb{D}_{\varrho^+}$ 中有 n' 个特征值。注意，对于 $1 \leq j \leq r$， $\omega_j \in \mathbb{T}_\varrho$ 不一定互不相同。做好这些准备后，我们发现， $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 的魏尔斯特拉斯标准形 [13, p.28] 可写为

$$J_{\mathscr{A}_0} - \lambda J_{\mathscr{B}_0} = \begin{array}{c} \\ m' \\ n_0 \\ n' \\ n_0 \end{array} \begin{array}{c} \begin{array}{cccc} m' & n_0 & n' & n_0 \end{array} \\ \left[\begin{array}{cc|cc} \mathscr{M}_{\mathrm{s}} & 0 & 0 & 0 \\ 0 & J_1 & 0 & \varGamma_0 \\ \hline 0 & 0 & I & 0 \\ 0 & 0 & 0 & J_1 \end{array}\right] \end{array} - \lambda \begin{array}{c} \\ m' \\ n_0 \\ n' \\ n_0 \end{array} \begin{array}{c} \begin{array}{cccc} m' & n_0 & n' & n_0 \end{array} \\ \left[\begin{array}{cc|cc} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ \hline 0 & 0 & \mathscr{N}_{\mathrm{s}} & 0 \\ 0 & 0 & 0 & I \end{array}\right] \end{array}, \tag{A.8a}$$

其中 $\mathscr{M}_{\mathrm{s}}$ 和 $\mathscr{N}_{\mathrm{s}}$ 满足 $\rho(\mathscr{M}_{\mathrm{s}}) < \varrho$ 以及 $\rho(\mathscr{N}_{\mathrm{s}}) < \varrho^{-1}$，并且

$$J_1 = \oplus_{j=1}^r J_{m_j}(\omega_j), \quad \varGamma_0 = \oplus_{j=1}^r \varGamma_{0,m_j}, \tag{A.8b}$$

并且存在 $(m+n)\times(m+n)$ 非奇异矩阵 $\mathscr{P}$ 和 $\mathscr{U}$，使得

$$\mathscr{P}\mathscr{A}_0\mathscr{U} = J_{\mathscr{A}_0}, \quad \mathscr{P}\mathscr{B}_0\mathscr{U} = J_{\mathscr{B}_0}. \tag{A.8c}$$

另一方面，如果交换 $\mathscr{A}_0$ 和 $\mathscr{B}_0$ 的作用，我们发现 $\mathscr{B}_0 - \lambda\mathscr{A}_0$ 的魏尔斯特拉斯标准形可以写成

$$\widehat{J}_{\mathscr{B}_0} - \lambda\widehat{J}_{\mathscr{A}_0} = \begin{array}{c} \\ m' \\ n_0 \\ n' \\ n_0 \end{array} \begin{array}{c} \begin{array}{cccc} m' & n_0 & n' & n_0 \end{array} \\ \left[\begin{array}{cc|cc} I & 0 & 0 & 0 \\ 0 & \widehat{J}_1^{\mathrm{T}} & 0 & 0 \\ \hline 0 & 0 & \mathscr{N}_{\mathrm{s}} & 0 \\ 0 & \varGamma_0^{\mathrm{T}} & 0 & \widehat{J}_1^{\mathrm{T}} \end{array}\right] \end{array} - \lambda \begin{array}{c} \\ m' \\ n_0 \\ n' \\ n_0 \end{array} \begin{array}{c} \begin{array}{cccc} m' & n_0 & n' & n_0 \end{array} \\ \left[\begin{array}{cc|cc} \mathscr{M}_{\mathrm{s}} & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ \hline 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{array}\right] \end{array}, \tag{A.9a}$$

其中 $\mathscr{M}_{\mathrm{s}}$、 $\mathscr{N}_{\mathrm{s}}$ 和 $\varGamma_0$ 与 (A.8) 中的相同，并且

$$\widehat{J}_1 = \oplus_{j=1}^r J_{m_j}(1/\omega_j), \tag{A.9b}$$

并且存在非奇异矩阵 $\mathscr{Q}$ 和 $\mathscr{V}$，使得

$$\mathscr{Q}\mathscr{A}_0\mathscr{V} = \widehat{J}_{\mathscr{A}_0}, \quad \mathscr{Q}\mathscr{B}_0\mathscr{V} = \widehat{J}_{\mathscr{B}_0}. \tag{A.9c}$$

引理 A.3. 设 $\{\mathscr{A}_i - \lambda\mathscr{B}_i\}_{i=1}^{\infty}$ 为由 SDASFQ（算法 4.1）生成的序列，其中矩阵 $\mathscr{A}_i$ 和 $\mathscr{B}_i$ 如 (4.2) 所示。则对于 $i \geq 0$

$$\mathscr{A}_i\mathscr{U}\,J_{\mathscr{B}_0}^{2^i} = \mathscr{B}_i\mathscr{U}\,J_{\mathscr{A}_0}^{2^i}, \quad \mathscr{A}_i\mathscr{V}\,\widehat{J}_{\mathscr{B}_0}^{2^i} = \mathscr{B}_i\mathscr{V}\,\widehat{J}_{\mathscr{A}_0}^{2^i}. \tag{A.10}$$

证明。由于 $J_{\mathscr{A}_0}$与 $J_{\mathscr{B}_0}$在 (A.8a) 中可交换，即 $J_{\mathscr{A}_0}J_{\mathscr{B}_0} = J_{\mathscr{B}_0}J_{\mathscr{A}_0}$，由 (A.8c)

$$\mathscr{P}\mathscr{A}_0\mathscr{U}\,J_{\mathscr{B}_0} = J_{\mathscr{A}_0}J_{\mathscr{B}_0} = J_{\mathscr{B}_0}J_{\mathscr{A}_0} = \mathscr{P}\mathscr{B}_0\mathscr{U}\,J_{\mathscr{A}_0}$$

在两边左乘 $\mathscr{P}^{-1}$ 后，即可得到关于 $i = 0$ 的 (A.10) 中的第一个方程。利用定理 2.1 进行归纳，可得该结论对所有 i 均成立。(A.10) 中的第二个方程也可用同样的方法证明。 □

将 (A.8) 和 (A.9) 中的 $\mathscr{U}$ 和 $\mathscr{V}$ 分别划分为

$$\mathscr{U} = \begin{array}{c} \\ m \\ n \end{array} \begin{array}{c} \begin{array}{cc} m & n \end{array} \\ \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \end{array}, \quad \mathscr{V} = \begin{array}{c} \\ m \\ n \end{array} \begin{array}{c} \begin{array}{cc} m & n \end{array} \\ \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \end{array}. \tag{A.11}$$

引理 A.4. 设

$$\begin{array}{c} \\ m \\ n \end{array} \begin{array}{c} \begin{array}{c} m \end{array} \\ \begin{bmatrix} Z_{11} \\ Z_{21} \end{bmatrix} \end{array} := Q_1 \begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix}, \quad \begin{array}{c} \\ m \\ n \end{array} \begin{array}{c} \begin{array}{c} n \end{array} \\ \begin{bmatrix} \widetilde{Z}_{12} \\ \widetilde{Z}_{22} \end{bmatrix} \end{array} =: Q_2 \begin{bmatrix} V_{12} \\ V_{22} \end{bmatrix}. \tag{A.12}$$

若 Z_{11} 和 $Q_{11}^{\mathrm{T}} - Y_0 Q_{12}^{\mathrm{T}} + (Q_{21}^{\mathrm{T}} - Y_0 Q_{22}^{\mathrm{T}}) Z_{21} Z_{11}^{-1}$ 均为非奇异，则 $\Phi := Z_{21} Z_{11}^{-1}$ 是 (5.2) 的解。若 $\widetilde{Z}_{22}$ 和 $Q_{22} - X_0 Q_{12} + (Q_{21} - X_0 Q_{11})\, \widetilde{Z}_{12} \widetilde{Z}_{22}^{-1}$ 均为非奇异，则 $\Psi := \widetilde{Z}_{12} \widetilde{Z}_{22}^{-1}$ 是 (5.4) 的解。此外

$$\mathrm{eig}([Q_{11}^{\mathrm{T}} - Y_0 Q_{12}^{\mathrm{T}} + (Q_{21}^{\mathrm{T}} - Y_0 Q_{22}^{\mathrm{T}})\Phi]^{-1} E_0) \subset \bar{\mathbb{D}}_{\varrho -} \tag{A.13a}$$

and

$$\mathrm{eig}([Q_{22} - X_0 Q_{12} + (Q_{21} - X_0 Q_{11})\Psi]^{-1} F_0) \subset \bar{\mathbb{D}}_{\varrho^{-1} -}. \tag{A.13b}$$

证明. 由 (A.10) 可验证，取 $i = 0$，有

$$\mathscr{A}_0 \begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix} \begin{bmatrix} \mathscr{M}_{\mathrm{s}} & \\ & J_1 \end{bmatrix}, \quad \mathscr{A}_0 \begin{bmatrix} V_{12} \\ V_{22} \end{bmatrix} \begin{bmatrix} \mathscr{N}_{\mathrm{s}} & \\ & \hat{J}_1^{\mathrm{T}} \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} V_{12} \\ V_{22} \end{bmatrix}. \tag{A.14}$$

将 (4.8) 和 (A.12) 代入 (A.14)，再消去 $\mathscr{M}_{\mathrm{s}} \oplus J_1$ 和 $\mathscr{N}_{\mathrm{s}} \oplus \hat{} J_1^{\mathrm{T}}$，可得 $\Phi = Z_{21} Z_{11}^{-1}$ 和 $\Psi = \widetilde{Z}_{12} \widetilde{Z}_{22}^{-1}$ 分别是 (5.2) 和 (5.4) 的解。此外，

$$[Q_{11}^{\mathrm{T}} - Y_0 Q_{12}^{\mathrm{T}} + (Q_{21}^{\mathrm{T}} - Y_0 Q_{22}^{\mathrm{T}})\Phi]^{-1} E_0 = Z_{11} (\mathscr{M}_{\mathrm{s}} \oplus J_1) Z_{11}^{-1}$$

and

$$[Q_{22} - X_0 Q_{12} + (Q_{21} - X_0 Q_{11})\Psi]^{-1} F_0 = \widetilde{Z}_{22} (\mathscr{N}_{\mathrm{s}} \oplus \hat{J}_1^{\mathrm{T}}) \widetilde{Z}_{22}^{-1}$$

从而得到 (A.13)。 □

定义 A.2. 设 Φ 是 (5.2) 的一个解，Ψ 是 (5.4) 的一个解。若 (A.13) 成立，则称 (Φ, Ψ) 为原问题 (5.2) 及其对偶问题 (5.4) 的广义弱稳定解对。在 $\varrho = 1$ 的情况下，我们简称其为弱稳定解对。

弱稳定解这一概念的灵感来自最优控制理论中的类似概念。接下来，我们的主要目标是为原问题 (5.2) 及其对偶问题 (5.4) 存在广义弱稳定解对建立充分条件，并证明由 SDASFQ（算法 4.1）生成的 X 序列和 Y 序列分别线性收敛到弱稳定解，线性收敛速率为 $1/2$。

定理 A.1. 设 $\mathscr{A}_0 - \lambda \mathscr{B}_0$ 属于 (4.8) 中的 SFQ，并假设 (A.5) 成立。设 Φ 和 Ψ 为引理 A.4 中的对象，且 $\rho_{\max} = \max\{\rho(\mathscr{M}_{\mathrm{s}})/\varrho, \varrho\, \rho(\mathscr{N}_{\mathrm{s}})\} < 1$。如果由 SDASFQ（算法 4.1）生成的序列 $\{(E_i, X_i, Y_i, F_i)\}_{i=1}^{\infty}$ 定义良好，即不存在任何崩溃，则

(a) $\max\{\varrho^{-2^i}\|E_i\|, \varrho^{2^i}\|F_i\|\} \le O(2^{-i}) + O(\rho^{2^i}$最大值$)$ 当 $i \to \infty$,

(b) $\|\Phi - X_i\| \le O(2^{-i}) + O(\rho^{2^i}$ 最大值$)$ 作为 $i \to \infty$,

(c) $\|\Psi - Y_i\| \le O(2^{-i}) + O(\rho^{2^i}$ 最大值$)$ 作为 $i \to \infty$, 并且

(d) $\widetilde{W}_i$ 和 W_i 分别趋近于奇异矩阵 $\widetilde{W} = Q_{11}^{\mathrm{T}} - \Psi Q_{12}^{\mathrm{T}} + (Q_{21}^{\mathrm{T}} - \Psi Q_{22}^{\mathrm{T}})\Phi$ 和 $W = Q_{22} - \Phi Q_{12} + (Q_{21} - \Phi Q_{11})\Psi$。此外,

$$\widetilde{W}(Q_1 \boldsymbol{u}_j)_{(1:m)} = 0, \quad W(Q_2 \boldsymbol{u}_j)_{(m+1:m+n)} = 0,$$

其中 $\boldsymbol{u}_j \in \mathbb{C}^{m+n}$ 是与 ω_j相关联的 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 的特征向量，即，对于 $1 \le j \le r$, $(\mathscr{A}_0 - \omega_j \mathscr{B}_0)\boldsymbol{u}_j = 0$ 。

*Pro*的。不失一般性，我们可以假设 $\varrho = 1$；否则我们可以考虑 er

$$\mathscr{A}(\varrho) - \lambda\mathscr{B}(\varrho) := \begin{bmatrix} \frac{1}{\varrho}E_0 & 0 \\ -X_0 & I \end{bmatrix} Q_1 - \lambda \begin{bmatrix} I & -Y_0 \\ 0 & \varrho F_0 \end{bmatrix} Q_2,$$

改为，因为

(i) SDASFQ（算法 4.1）应用于 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 时，生成的近似值 X_i 和 Y_i 与应用于 $\mathscr{A}(\varrho) - \lambda\mathscr{B}(\varrho)$ 时相同，但 E_i 和 F_i 分别缩放为 $\varrho^{-2^i}E_i$ 和 $\varrho^{2^i}F_i$，并且

(ii) 若 $\mathscr{A}_0$ 和 $\mathscr{B}_0$ 满足 (A.5)，则 $\mathscr{A}(\varrho) - \lambda\mathscr{B}(\varrho)$ 满足 (A.5)，其中 ϱ 设为 1。

逐步检查倍增迭代内核 (4.22)，即可直接验证第 (i) 项。对于第 (ii) 项，我们注意到

$$\begin{aligned} \mathscr{A}_0 Z = \mathscr{B}_0 Z M \quad &\Leftrightarrow \quad \mathscr{A}(\varrho)Z = \mathscr{B}(\varrho)Z(\frac{1}{\varrho}M), \\ \mathscr{A}_0 Z N = \mathscr{B}_0 Z \quad &\Leftrightarrow \quad \mathscr{A}(\varrho)Z(\varrho N) = \mathscr{B}(\varrho)Z, \end{aligned}$$

这意味着 $\mathscr{A}_0 - \lambda\mathscr{B}_0$ 和 $\mathscr{A}(\varrho) - \lambda\mathscr{B}(\varrho)$ 具有相同的魏尔斯特拉斯标准形，只是所有特征值 λ 都缩放为 λ/ϱ。

从现在起，假设 $\varrho = 1$。因此，在 (A.9a) 中, $\widehat{J}_1^{\mathrm{T}} = J^{\mathrm{H1}}$。

代入 $J_{\mathscr{A}_0}$, $J_{\mathscr{B}_0}$, $\widehat{J}_{\mathscr{A}_0}$, 以及 $\widehat{J}_{\mathscr{B}_0}$ 到 (A.10) 中，得到

$$\mathscr{A}_i \mathscr{U} \begin{bmatrix} I_m & 0_{m\times n} \\ 0_{n\times m} & \mathscr{N}_{\mathrm{s}}^{2^i} \oplus I_{n_0} \end{bmatrix} = \mathscr{B}_i \mathscr{U} \begin{bmatrix} \mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i} & 0_{m'\times n'} \oplus \varGamma_i \\ 0_{n\times m} & I_{n'} \oplus J_1^{2^i} \end{bmatrix}, \tag{A.15a}$$

$$\mathscr{A}_i \mathscr{V} \begin{bmatrix} I_{m'} \oplus (J_1^{\mathrm{H}})^{2^i} & 0_{m\times n} \\ 0_{n'\times m'} \oplus \varGamma_i^{\mathrm{T}} & \mathscr{N}_{\mathrm{s}}^{2^i} \oplus (J_1^{\mathrm{H}})^{2^i} \end{bmatrix} = \mathscr{B}_i \mathscr{V} \begin{bmatrix} \mathscr{M}_{\mathrm{s}}^{2^i} \oplus I_{n_0} & 0_{m\times n} \\ 0_{n\times m} & I_n \end{bmatrix}, \tag{A.15b}$$

其中 $\varGamma_i = \varGamma_{i,m_1} \oplus \cdots \oplus \varGamma_{i,m_r}$, $\varGamma_{i,m_j}$的定义见 (A.4)。

Let

$$\mathscr{Z} = Q_1 \mathscr{U} = \begin{array}{c} \\ m \\ n \end{array} \begin{array}{c} \begin{array}{cc} m & n \end{array} \\ \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \end{array} \tag{A.16}$$

and

$$\mathscr{G} = Q_2 Q_1^{\mathrm{T}} \mathscr{Z} =: \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} = \begin{bmatrix} Q_{11}^{\mathrm{T}} Z_{11} + Q_{21}^{\mathrm{T}} Z_{21} & Q_{11}^{\mathrm{T}} Z_{12} + Q_{21}^{\mathrm{T}} Z_{22} \\ Q_{12}^{\mathrm{T}} Z_{11} + Q_{22}^{\mathrm{T}} Z_{21} & Q_{12}^{\mathrm{T}} Z_{12} + Q_{22}^{\mathrm{T}} Z_{22} \end{bmatrix}. \tag{A.17}$$

将 (4.2) 中的 $\mathscr{A}_i$ 和 $\mathscr{B}_i$、(A.16) 中的 $\mathscr{Z}$ 以及 (A.17) 中的 $\mathscr{G}$ 代入 (A.15a)，得到

$$E_i Z_{11} = (G_{11} - Y_i G_{21})(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i}), \tag{A.18a}$$

$$\begin{aligned} E_i Z_{12}(\mathscr{N}_{\mathrm{s}}^{2^i} \oplus I_{n_0}) = {} & (G_{11} - Y_i G_{21})(0_{m' \times n'} \oplus \Gamma_i) \\ & + (G_{12} - Y_i G_{22})(I_{n'} \oplus J_1^{2^i}), \end{aligned} \tag{A.18b}$$

$$-X_i Z_{11} + Z_{21} = F_i G_{21}(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i}), \tag{A.18c}$$

$$(-X_i Z_{12} + Z_{22})(\mathscr{N}_{\mathrm{s}}^{2^i} \oplus I_{n_0}) = F_i G_{21}(0_{m' \times n'} \oplus \Gamma_i) + F_i G_{22}(I_{n'} \oplus J_1^{2^i}). \tag{A.18d}$$

后乘 以 Z_{11}^{-1} 乘 (A.18a)，以 $(0_{n' \times m'} \oplus \Gamma_i^{-1} J_1^{2^i}) Z_{11}^{-1}$ 乘 (A.18b)，分别得到 y,

$$E_i = (G_{11} - Y_i G_{21})(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i}) Z_{11}^{-1}, \tag{A.18a'}$$

$$\begin{aligned} E_i Z_{12}(0_{n' \times m'} \oplus \Gamma_i^{-1} J_1^{2^i}) Z_{11}^{-1} = {} & (G_{11} - Y_i G_{21})(0_{m' \times m'} \oplus J_1^{2^i}) Z_{11}^{-1} \\ & + (G_{12} - Y_i G_{22})(0_{n' \times m'} \oplus J_1^{2^i} \Gamma_i^{-1} J_1^{2^i}) Z_{11}^{-1}. \end{aligned} \tag{A.18b'}$$

用 (A.18a') 减去 (A.18b')，得到

$$\begin{aligned} & E_i \left[I - Z_{12}(0_{n' \times m'} \oplus \Gamma_i^{-1} J_1^{2^i}) Z_{11}^{-1} \right] \\ & = (G_{11} - Y_i G_{21})(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus 0_{n_0 \times n_0}) Z_{11}^{-1} - (G_{12} - Y_i G_{22})(0_{n' \times m'} \oplus J_1^{2^i} \Gamma_i^{-1} J_1^{2^i}) Z_{11}^{-1}. \end{aligned} \tag{A.19}$$

后乘 以 Z_{11}^{-1} 乘 (A.18c)，以 $(0_{n' \times m'} \oplus \Gamma_i^{-1} J_1^{2^i}) Z_{11}^{-1}$ 乘 (A.18d)，分别得到 y,

$$-X_i + \Phi = F_i G_{21}(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i}) Z_{11}^{-1}, \tag{A.18c'}$$

$$\begin{aligned} (-X_i Z_{12} + Z_{22})(0_{n' \times m'} \oplus \Gamma_i^{-1} J_1^{2^i}) Z_{11}^{-1} = {} & F_i G_{21}(0_{m' \times m'} \oplus J_1^{2^i}) Z_{11}^{-1} \\ & + F_i G_{22}(0_{n' \times m'} \oplus J_1^{2^i} \Gamma_i^{-1} J_1^{2^i}) Z_{11}^{-1}. \end{aligned} \tag{A.18d'}$$

用 (A.18c') 减去 (A.18d')，得到

$$\begin{aligned} & -X_i \left[I - Z_{12}(0_{n' \times m'} \oplus \Gamma_i^{-1} J_1^{2^i}) Z_{11}^{-1} \right] + \Phi = Z_{22}(0_{n' \times m'} \oplus \Gamma_i^{-1} J_1^{2^i}) Z_{11}^{-1} \\ & \qquad - F_i G_{22}(0_{n' \times m'} \oplus J_1^{2^i} \Gamma_i^{-1} J_1^{2^i}) Z_{11}^{-1} + F_i G_{21}(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus 0_{n_0 \times n_0}) Z_{11}^{-1}. \end{aligned} \tag{A.20}$$

由引理 A.2，(A.20) 可写为

$$X_i[I + O_{m \times m}(2^{-i})] = \Phi + O_{n \times m}(2^{-i}) + F_i\big[O_{n \times m}(2^{-i}) + O_{n \times m}(\rho(\mathscr{M}_{\mathrm{s}})^{2^i})\big], \tag{A.21}$$

其中，每个 $O_{n \times m}(\epsilon)$ 都表示一个 $n \times m$ 矩阵，其元素按 ϵ 的顺序排列。

Let

$$\widetilde{\mathscr{Z}} = Q_2 \mathscr{V} = \begin{array}{c} \\ m \\ n \end{array} \begin{array}{c} \begin{array}{cc} m & n \end{array} \\ \begin{bmatrix} \widetilde{Z}_{11} & \widetilde{Z}_{12} \\ \widetilde{Z}_{21} & \widetilde{Z}_{22} \end{bmatrix} \end{array} \tag{A.22}$$

and

$$\widetilde{\mathscr{G}} = Q_1 Q_2^{\mathrm{T}} \widetilde{\mathscr{Z}} =: \begin{bmatrix} \widetilde{G}_{11} & \widetilde{G}_{12} \\ \widetilde{G}_{21} & \widetilde{G}_{22} \end{bmatrix} = \begin{bmatrix} Q_{11}\widetilde{Z}_{11} + Q_{12}\widetilde{Z}_{21} & Q_{11}\widetilde{Z}_{12} + Q_{12}\widetilde{Z}_{22} \\ Q_{21}\widetilde{Z}_{11} + Q_{22}\widetilde{Z}_{21} & Q_{21}\widetilde{Z}_{12} + Q_{22}\widetilde{Z}_{22} \end{bmatrix}. \tag{A.23}$$

接下来，将 (4.2) 中的 $\mathscr{A}_i$ 和 $\mathscr{B}_i$、(A.22) 中的 $\mathscr{Z}$ 以及 (A.23) 中的 $\mathscr{G}$ 代入 (A.15b)，得到

$$F_i \widetilde{Z}_{22} = (-X_i \widetilde{G}_{12} + \widetilde{G}_{22}) \big[\mathscr{N}_{\mathrm{s}}^{2^i} \oplus (J_1^{\mathrm{H}})^{2^i} \big], \tag{A.24a}$$

$$\begin{aligned} F_i \widetilde{Z}_{21} (\mathscr{M}_{\mathrm{s}}^{2^i} \oplus I_{n_0}) &= (-X_i \widetilde{G}_{12} + \widetilde{G}_{22})(0_{n' \times m'} \oplus \Gamma_i^{\mathrm{T}}) \\ &\quad + (-X_i \widetilde{G}_{11} + \widetilde{G}_{21}) \big[I_{m'} \oplus (J_1^{\mathrm{H}})^{2^i} \big], \end{aligned} \tag{A.24b}$$

$$\begin{aligned} \widetilde{Z}_{12} - Y_i \widetilde{Z}_{22} &= E_i \widetilde{G}_{12} \big[\mathscr{N}^{2^i}{}_{\mathrm{s}} \oplus (J_1^{\mathrm{H}})^{2^i} \big] \\ &= E_i \widetilde{G}_{12} (\mathscr{N}_{\mathrm{s}}^{2^i} \oplus 0_{n_0 \times n_0}) + E_i \widetilde{G}_{12} (0_{n' \times n'} \oplus (J_1^{\mathrm{H}})^{2^i}), \end{aligned} \tag{A.24c}$$

$$(\widetilde{Z}_{11} - Y_i \widetilde{Z}_{21})(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus I_{n_0}) = E_i \widetilde{G}_{11} \big[I_{m'} \oplus (J_1^{\mathrm{H}})^{2^i} \big] + E_i \widetilde{G}_{12} (0_{n' \times m'} \oplus \Gamma_i^{\mathrm{T}}). \tag{A.24d}$$

将 (A.24a) 右乘 $\widetilde{Z}_{22}^{-1}$，并将 (A.24b) 右乘 $\big[0_{m' \times n'} \oplus \Gamma_i^{-\mathrm{T}} (J_1^{2^i})^{\mathrm{H}}\big] \widetilde{Z}_{22}^{-1}$，分别得到：

$$F_i = (-X_i \widetilde{G}_{12} + \widetilde{G}_{22}) \big[\mathscr{N}_{\mathrm{s}}^{2^i} \oplus (J_1^{\mathrm{H}})^{2^i} \big] \widetilde{Z}_{22}^{-1}, \tag{A.24a'}$$

$$\begin{aligned} F_i \widetilde{Z}_{21} \big[0_{m' \times n'} \oplus \Gamma_i^{-\mathrm{T}} (J_1^{2^i})^{\mathrm{H}}\big] \widetilde{Z}_{22}^{-1} &= (-X_i \widetilde{G}_{12} + \widetilde{G}_{22}) \big[0_{n' \times n'} \oplus (J_1^{2^i})^{\mathrm{H}}\big] \widetilde{Z}_{22}^{-1} \\ &\quad + (-X_i \widetilde{G}_{11} + \widetilde{G}_{21}) \big[0_{m' \times n'} \oplus (J_1^{\mathrm{H}})^{2^i} \Gamma^{-\mathrm{T}i} (J_1^{2^i})^{\mathrm{H}}\big] \widetilde{Z}_{22}^{-1}. \end{aligned} \tag{A.24b'}$$

将 (A.24b') 从 (A.24a') 中相减，得到

$$\begin{aligned} F_i \Big\{ I - \widetilde{Z}_{21} \big[0_{m' \times n'} \oplus \Gamma_i^{-\mathrm{T}} (J_1^{2^i})^{\mathrm{H}}\big] \widetilde{Z}_{22}^{-1} \Big\} &= (-X_i \widetilde{G}_{12} + \widetilde{G}_{22})(\mathscr{N}_{\mathrm{s}}^{2^i} \oplus 0_{n_0 \times n_0}) \widetilde{Z}_{22}^{-1} \\ &\quad - (-X_i \widetilde{G}_{11} + \widetilde{G}_{21}) \big[0_{m' \times n'} \oplus (J_1^{\mathrm{H}})^{2^i} \Gamma_i^{-\mathrm{T}} (J_1^{2^i})^{\mathrm{H}}\big] \widetilde{Z}_{22}^{-1}. \end{aligned} \tag{A.25}$$

将 (A.24c) 右乘 $\widetilde{Z}_{22}^{-1}$，将 (A.24d) 右乘 $\big[0_{m' \times n'} \oplus \Gamma_i^{-\mathrm{T}} (J_1^{\mathrm{H}})^{2^i}\big] \widetilde{Z}_{22}^{-1}$，分别得到，

$$\Psi - Y_i = E_i \widetilde{G}_{12} (\mathscr{N}_{\mathrm{s}}^{2^i} \oplus 0_{n_0 \times n_0}) \widetilde{Z}_{22}^{-1} + E_i \widetilde{G}_{12} \big[0_{n' \times n'} \oplus (J_1^{\mathrm{H}})^{2^i}\big] \widetilde{Z}_{22}^{-1}, \tag{A.24c'}$$

$$\begin{aligned} (\widetilde{Z}_{11} - Y_i \widetilde{Z}_{21}) \big[0_{m' \times n'} \oplus \Gamma_i^{-\mathrm{T}} (J_1^{2^i})^{\mathrm{H}}\big] \widetilde{Z}_{22}^{-1} &= E_i G_{11} \big[0_{m' \times n'} \oplus (J_1^{\mathrm{H}})^{2^i} \Gamma^{-\mathrm{T}i} (J_1^{\mathrm{H}})^{2^i}\big] \widetilde{Z}_{22}^{-1} \\ &\quad + E_i \widetilde{G}_{12} \big[0_{n' \times n'} \oplus (J_1^{\mathrm{H}})^{2^i}\big] \widetilde{Z}_{22}^{-1}. \end{aligned} \tag{A.24d'}$$

将 (A.24d') 从 (A.24c') 中相减，我们得到

$$\Psi - Y_i \big\{ I - \widetilde{Z}_{21} \big[0_{m' \times n'} \oplus \Gamma_i^{-\mathrm{T}} (J_1^{\mathrm{H}})^{2^i}\big] \widetilde{Z}_{22}^{-1} \big\} = \widetilde{Z}_{11} \big[0_{m' \times n'} \oplus \Gamma_i^{-\mathrm{T}} (J_1^{\mathrm{H}})^{2^i}\big] \widetilde{Z}_{22}^{-1}$$

$$- E_i\widetilde{G}_{11}\big[0_{m'\times n'} \oplus (J_1^{\mathrm{H}})^{2^i}\varGamma_i^{-\,\mathrm{T}}(J_1^{\mathrm{H}})^{2^i}\big]\widetilde{Z}_{22}^{-1} + E_i\widetilde{G}_{12}(\mathscr{N}_{\mathrm{s}}^{2^i} \oplus 0_{n_0\times n_0})\widetilde{Z}_{22}^{-1}. \quad \text{(A.26)}$$

根据引理 A.2，(A.26) 可写为

$$Y_i[I + O_{n\times n}(2^{-i})] = \varPsi + O_{m\times n}(2^{-i}) + E_i\big[O_{m\times n}(2^{-i}) + O_{m\times n}(\rho(\mathscr{N}_{\mathrm{s}})^{2^i})\big]. \quad \text{(A.27)}$$

将 Y_i 代入 (A.27)，再代入 (A.19)，并由引理 A.2 得

$$\|E_i\| \le O(2^{-i}) + O(\rho_{\max}^{2^i}) \to 0, \text{ as } i \to \infty. \quad \text{(A.28)}$$

将 X_i 代入 (A.21)，再代入 (A.25)，并由引理 A.2 得

$$\|F_i\| \le O(2^{-i}) + O(\rho_{\max}^{2^i}) \to 0, \text{ as } i \to \infty. \quad \text{(A.29)}$$

结论 (a) 可由 (A.28) 和 (A.29) 推出。

由 (A.29) 和 (A.21) 可得

$$\|\varPhi - X_i\| \le O(2^{-i}) + O(\rho_{\max}^{2^i}) \to 0.$$

这证明了结论 (b)。由 (A.28) 和 (A.27) 可得

$$\|\varPsi - Y_i\| \le O(2^{-i}) + O(\rho_{\max}^{2^i}) \to 0.$$

这证明了结论 (c)。

对于 (d) 项，将 G_{11} 和 G_{21} 的公式代入 (A.17) 至 (A.18a)，可得

$$\begin{aligned} E_iZ_{11} &= \big[Q_{11}^{\mathrm{T}}Z_{11} + Q_{21}^{\mathrm{T}}Z_{21} - Y_i(Q_{12}^{\mathrm{T}}Z_{11} + Q_{22}^{\mathrm{T}}Z_{21})\big](\mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i}) \\ &= (Q_{11}^{\mathrm{T}} - Y_iQ_{12}^{\mathrm{T}})Z_{11}(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i}) + (Q_{21}^{\mathrm{T}} - Y_iQ_{22}^{\mathrm{T}})\varPhi Z_{11}(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i}). \end{aligned} \quad \text{(A.30)}$$

将 (A.18c) 改写为

$$\varPhi Z_{11} = X_iZ_{22} + F_i(Q_{22}^{\mathrm{T}}\varPhi + Q_{12}^{\mathrm{T}}Z_{11}(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i}), \quad \text{(A.31)}$$

然后将 (A.31) 右乘 $\mathscr{M}^{2^i\mathrm{s}} \oplus J_1^{2^i}$，得

$$\varPhi Z_{11}(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i}) = X_iZ_{22}(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i}) + F_i(Q_{22}^{\mathrm{T}}\varPhi + Q_{12}^{\mathrm{T}}Z_{11}\big(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i}\big)^2. \quad \text{(A.32)}$$

将 (A.32) 代入 (A.30)，得到：

$$E_iZ_{11} = \widetilde{W}_iZ_{11}(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i}) + (Q_{21}^{\mathrm{T}} - Y_iQ_{22}^{\mathrm{T}})F_i(Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}\varPhi)Z_{11}(\mathscr{M}_{\mathrm{s}}^{2^i} \oplus J_1^{2^i})^2. \quad \text{(A.33)}$$

用 $(0_{m'\times m'} \oplus J_1^{-2^i})\begin{bmatrix}0_{m'\times 1}\\ \boldsymbol{e}_1\end{bmatrix}$ 右乘 (A.33)，可得

$$\widetilde{W}_iZ_{11}\begin{bmatrix}0_{m'\times 1}\\ \boldsymbol{e}_1\end{bmatrix} = -E_iZ_{11}\begin{bmatrix}0_{m'\times 1}\\ \omega_1^{-2^i}\boldsymbol{e}_1\end{bmatrix} + (Q_{21}^{\mathrm{T}} - Y_iQ_{22}^{\mathrm{T}})F_i(Q_{12}^{\mathrm{T}} + Q_{22}^{\mathrm{T}}\varPhi)Z_{11}\begin{bmatrix}0_{m'\times 1}\\ \omega_1^{2^i}\boldsymbol{e}_1\end{bmatrix}, \quad \text{(A.34)}$$

$\widetilde{}$利用第 (a) 项和第 (c) 项，当 $i \to \infty$ 时收敛到 0。因此， W_i 收敛到奇异矩阵 $\widetilde{W} = Q_{11}^{\mathrm{T}} - \Psi Q_{12}^{\mathrm{T}} + (Q_{21}^{\mathrm{T}} - \Psi Q_{22}^{\mathrm{T}})\Phi$，其中

$$\widetilde{W} Z_{11} \begin{bmatrix} 0_{m' \times 1} \\ \boldsymbol{e}_1 \end{bmatrix} = 0. \tag{A.35}$$

注意到，由 (A.8a) 可知，(A.34) 中的 $\boldsymbol{e}_1$ 可替换为 $\boldsymbol{e}_{m_1+\cdots+m_j+1}$，其中$1 \le j \le r-1$。

$\widetilde{}$类似地，我们利用 (A.24a) 和 (A.24c) 证明 $W_i Z_{22} \boldsymbol{e}_{n'+1} \to 0$，其中 $i \to \infty$。$\widetilde{}$ $\widetilde{}$ 将 (A.23) 中的 G_{12} 和 G_{22} 公式代入 (A.24a)，得到

$$\begin{aligned} F_i \widetilde{Z}_{22} &= \big[Q_{22} - X_i Q_{12} + (Q_{21} - X_i Q_{11})\Psi\big] \widetilde{Z}_{22} \big[\mathscr{N}_{\mathrm{s}}^{2^i} \oplus (J_1^{\mathrm{H}})^{2^i}\big] \\ &= (Q_{22} - X_i Q_{12}) \widetilde{Z}_{22} \big[\mathscr{N}_{\mathrm{s}}^{2^i} \oplus (J_1^{\mathrm{H}})^{2^i}\big] + (Q_{21} - X_i Q_{11}) \Psi \widetilde{Z}_{22} \big[\mathscr{N}_{\mathrm{s}}^{2^i} \oplus (J_1^{\mathrm{H}})^{2^i}\big]. \end{aligned} \tag{A.36}$$

将 (A.24c) 改写为

$$\Psi \widetilde{Z}_{22} = Y_i \widetilde{Z}_{22} + E_i (Q_{11}\Psi + Q_{12}) \widetilde{Z}_{22} \big[\mathscr{N}_{\mathrm{s}}^{2^i} \oplus (J_1^{\mathrm{H}})^{2^i}\big], \tag{A.37}$$

然后将 (A.37) 右乘 $\mathscr{N}^{2^i}{}_{\mathrm{s}} \oplus (J_1^{\mathrm{H}})^{2^i}$，得到

$$\Psi \widetilde{Z}_{22} \big[\mathscr{N}_{\mathrm{s}}^{2^i} \oplus (J_1^{\mathrm{H}})^{2^i}\big] = Y_i \widetilde{Z}_{22} \big[\mathscr{N}_{\mathrm{s}}^{2^i} \oplus (J_1^{\mathrm{H}})^{2^i}\big] + E_i (Q_{11}\Psi + Q_{12}) \widetilde{Z}_{22} \big[\mathscr{N}_{\mathrm{s}}^{2^i} \oplus (J_1^{\mathrm{H}})^{2^i}\big]^2. \tag{A.38}$$

将 (A.38) 代入 (A.36)，得到:

$$F_i \widetilde{Z}_{22} = W_i \big[\mathscr{N}_{\mathrm{s}}^{2^i} \oplus (J_1^{\mathrm{H}})^{2^i}\big] + (Q_{21} - X_i Q_{11}) E_i (Q_{11}\Psi + Q_{12}) \widetilde{Z}_{22} \big[\mathscr{N}_{\mathrm{s}}^{2^i} \oplus (J_1^{\mathrm{H}})^{2^i}\big]^2. \tag{A.39}$$

将 (A.39) 右乘 $(0_{n' \times n'} \oplus (J_1^{\mathrm{H}})^{-2^i}) \begin{bmatrix} 0_{n' \times 1} \\ \boldsymbol{e}_{m_1} \end{bmatrix}$，得到

$$W_i \widetilde{Z}_{22} \begin{bmatrix} 0_{n' \times 1} \\ \boldsymbol{e}_{m_1} \end{bmatrix} = -F_i \widetilde{Z}_{22} \begin{bmatrix} 0_{n' \times 1} \\ (\overline{\omega}_1)^{-2^i} \boldsymbol{e}_{m_1} \end{bmatrix} + (Q_{21} - X_i Q_{11}) E_i (Q_{11}\Psi + Q_{12}) \widetilde{Z}_{22} \begin{bmatrix} 0_{n' \times 1} \\ \overline{\omega}_1^{2^i} \boldsymbol{e}_{m_1} \end{bmatrix}, \tag{A.40}$$

利用 (a) 和 (b) 项，可知其随 $i \to \infty$ 趋于 0。因此， W_i 收敛于奇异矩阵 $W = Q_{22} - \Phi Q_{12} + (Q_{21} - \Phi Q_{11})\Psi$，且

$$W \widetilde{Z}_{22} \begin{bmatrix} 0_{n' \times 1} \\ \boldsymbol{e}_{m_1} \end{bmatrix} = 0. \tag{A.41}$$

注意，由 (A.9a) 可知，(A.40) 中的 $\boldsymbol{e}_{m_1}$ 可替换为 $\boldsymbol{e}_{m_1+m_2+\cdots+m_j}$，其中 $2 \le j \le r$。

□

注 A.1. 需要认识到，定理中的广义弱稳定解对（Φ，Ψ）是在引理 A.4 中专门构造的，并且正是由 SDASFQ（算法 4.1）计算得到的那个解对。在定义 A.2 的意义下，可能还存在其他广义弱稳定解对，但这些解对并非由 SDASFQ 计算得到。

注 A.2. 线性收敛速度 $1/2$ 在实际应用中通常已经相当好，但仍慢于二次收敛。对于特定的非线性矩阵方程，仍有可能借助特殊技巧保持二次收敛。其中一个例子是临界情形中的 MARE，为此已作出努力 [17, 21, 25, 40] 以实现二次收敛。

参考文献

[1] Zhaojun Bai、James Demmel 和 Ming Gu。用于非对称特征值问题的无逆并行谱分治算法。《数值数学》，76：279–308，1997年。[2] Zhaojun Bai 和 Ren-Cang Li。线性响应特征值问题的极小化原理（一）：理论。SIAM J. Matrix Anal. Appl.，33(4)：1075–1100，2012年。[3] Zhaojun Bai 和 Ren-Cang Li。线性响应特征值问题的极小化原理（二）：计算。SIAM J. Matrix Anal. Appl.，34(2)：392–416，2013年。[4] Zhaojun Bai 和 Ren-Cang Li。广义线性响应特征值问题的极小化原理与计算。BIT Numer. Math.，54(1)：31–54，2014年。[5] Zhaojun Bai、Ren-Cang Li 和 Wen-Wei Lin。采用扩展局部最优预条件共轭梯度法求解线性响应特征值问题。SCIENCE CHINA Math.，59(8)：1443–1460，2016年。[6] P. Benner。代数黎卡提方程及相关特征值问题的数值解研究。博士学位论文，Fakult¨ät f¨ur Mathematik，开姆尼茨-茨维考工业大学，1997年。[7] Dario A. Bini、Bruno Iannazzo 和 Beatrice Meini。代数黎卡提方程的数值解。SIAM，费城，2012年。[8] Dario A. Bini、Beatrice Meini 和 Federico Poloni。将代数黎卡提方程变换为单边二次矩阵方程。《数值数学》，116：553–578，2010年。[9] A. Ya. Bulgakov 和 S. K. Godunov。矩阵谱的圆形二分。Siberian Math. J.，29：734–744，1988年。[10] E. K.-W. Chu、H.-Y. Fan 和 W.-W. Lin。连续时间代数黎卡提方程的保结构倍增算法。Linear Algebra Appl.，396：55–80，2005年。[11] E. K. W. Chu、H.-Y. Fan、W. W. Lin 和 C. S. Wang。周期离散时间代数黎卡提方程的保结构算法。Int. J. Control，77(8)：767–788，2004年。[12] F. R. de Hoog 和 R. M. M. Mattheij。矩阵的子集选取。Linear Algebra Appl.，422：349–359，2007年。[13] F. R. Gantmacher。矩阵理论，第I、II卷。Chelsea Publishing Company，纽约，1959年。[14] S. K. Godunov。矩阵谱二分问题。Siberian Math. J.，27：649–660，1986年。[15] G. H. Golub 和 C. F. Van Loan。矩阵计算。约翰·霍普金斯大学出版社，巴尔的摩，马里兰州，第三版，1996年。[16] S. A. Goreinov、E. E. Tyrtyshnikov 和 N. L. Zamarashkin。伪骨架近似理论。Linear Algebra Appl.，261(1-3)：1–21，1997年。

[17] Chun-Hua Guo、Bruno Iannazzo 和 Beatrice Meini。关于（移位）非对称代数黎卡提方程的倍增算法。SIAM J. Matrix Anal. Appl.，29(4)：1083–1100，2007年。[18] X. Guo、W. Lin 和 S. Xu。非对称代数黎卡提方程的保结构倍增算法。《数值数学》，103：393–412，2006年。[19] Zhen-Chen Guo、Eric King-Wah Chu 和 Wen-Wei Lin。离散化 Bethe-Salpeter特征值问题的倍增算法。《数学计算》，88(319)：2325–2350，2019年。[20] Y. P. Hong 和 C.-T. Pan。秩揭示QR分解与奇异值分解。《数学计算》，58(197)：213–232，1992年。[21] T.-M. Huang、W.-Q. Huang、R.-C. Li 和 W.-W. Lin。用于临界奇异 M-矩阵代数黎卡提方程的新型两阶段保结构倍增算法。Numer. Linear Algebra Appl.，23：291–313，2016年。[22] T.-M. Huang 和 W.-W. Lin。用于代数黎卡提方程弱稳定厄米解的结构化倍增算法。Linear Algebra Appl.，430：1452–1478，2009年。[23] Tsung-Ming Huang、Ren-Cang Li 和 Wen-Wei Lin。非线性矩阵方程的保结构倍增算法。算法基础，第14卷。SIAM，费城，2018年9月。[24] Tsung-Ming Huang、Ren-Cang Li、Wen-Wei Lin 和 Linzhang Lu。倍增算法的最优参数。J. Math. Study，50(4)：339–357，2017年。[25] Bruno Iannazzo 和 Federico Poloni。用于与 M-矩阵相关的非对称代数黎卡提方程的子空间移位技术。Numer. Linear Algebra Appl.，20(3)：440–452，2013年。[26] Peter Lancaster 和 Leiba Rodman。代数黎卡提方程。牛津大学出版社，纽约，美国，1995年。[27] A. Laub。求解代数黎卡提方程的舒尔方法。IEEE《自动控制汇刊》，24(6)：913–921，1979年。[28] A. J. Laub。用于数值求解黎卡提方程的不变子空间方法。载于 S. Bittanti、A. J. Laub 和 J. C. Willems 编，《黎卡提方程》，第163–196页。施普林格出版社，柏林，1991年。[29] W.-W. Lin 和 S.-F. Xu。黎卡提型矩阵方程保结构倍增算法的收敛性分析。SIAM J. Matrix Anal. Appl.，28(1)：26–39，2006年。[30] Changli Liu 和 Jungong Xue。Markov调制布朗运动 Wiener-Hopf 分解的高效计算。《数学计算》，94(355)：2367–2408，2025年。[31] A. N. Malyshev。正则线性矩阵束不变子空间的计算。Siberian Math. J.，30：559–567，1989年。[32] V. Mehrmann。自治线性二次控制问题：理论与数值解。施普林格出版社，1991年。[33] G. T. Nguyen 和 F. Poloni。采用倍增算法进行逐分量精确的流体队列计算。《数值数学》，130(4)：763–792，2015年。[34] T. Pappas、A. J. Laub 和 N. R. Sandell。离散时间代数黎卡提方程的数值求解。IEEE《自动控制汇刊》，25：631–641，1980年。

[35] F. Poloni。《二次矩阵与向量方程算法》。博士学位论文，比萨高等师范学院，意大利比萨，2010年。[36] 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。《化学物理学杂志》，136:034111，2012年。[37] M. Shao、F. H. da Jornada、C. Yang、J. Deslippe 和 S. G. Louie。用于求解 Bethe-Salpeter特征值问题的保持结构的并行算法。《Linear Algebra Appl.》，488:148–167，2016年。[38] Xiaobai Sun 和 Enrique S. Quintana-Ort´ı。块广义Schur分解的谱分割方法。《数学计算》，73:1827–1847，2004年。[39] Wei-Guo Wang、Wei-Chao Wang 和 Ren-Cang Li。用于 M-矩阵代数黎卡提方程的交替方向倍增算法。《SIAM J. Matrix Anal. Appl.》，33(1):170–194，2012年。[40] Wei-Guo Wang、Wei-Chao Wang 和 Ren-Cang Li。消去不可约奇异 M-矩阵代数黎卡提方程。《Numer. Alg., Contr. Optim.》，3:491–518，2013年。[41] Hongteng Xu、Hongyuan Zha、Ren-Cang Li 和 Mark A. Davenport。通过 Gershgorin圆引导的样本选择进行主动流形学习。载于《AAAI人工智能会议》，2015年。网址：www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9479。[42] Jungong Xue 和 Ren-Cang Li。用于 M-矩阵代数黎卡提方程的高精度加倍算法。《Numer. Math.》，135(3):733–767，2017年。