arXiv:2602.22084v1 [math.NA] 25 Feb 2026

# 等谱矩阵切空间中的矩阵扰动理论 *

FRANCESCO HROBAT† 和YUJI NAKATSUKASA†

**摘要**。特征值与特征向量扰动理论是数值线性代数、量子物理及相关领域等多个学科中的基本课题。核心问题是理解矩阵 $A \in \mathbb{C}^{n\times n}$ 的特征值与特征向量在加入扰动矩阵 $E \in \mathbb{C}^{n\times n}$后如何变化。现有文献大多关注结构化扰动。例如，在 [C.-K. Li 和 R.-C. Li, Linear Algebra Appl. 2005], 中，矩阵 $A$被假定为厄米且分块对角，而扰动 $E$ 为厄米且分块非对角。在本工作中，我们研究一种不同的结构化设定，其中扰动具有交换子形式 $E = AB - BA$ ，对应于某个矩阵 $B$；我们证明，这是对 Li 和 Li 所考虑的分块对角结构的推广。首先，我们推广其主要结果，证明第 $i$个特征值（属于 $A$）的扰动（记为 $\lambda_i$）的阶为 $\|E\|^2/\eta_i$，其中$\eta_i = \min_{j\neq i}|\lambda_i - \lambda_j|$ 是与 $\lambda_i$相关联的谱间隙。其次，我们详细分析矩阵 $B$ 在特征向量扰动中所起的作用。该分析进一步推广到具有重特征值的分块对角矩阵情形，以及对奇异值的扰动和若尔当块特征值的扰动。

**关键词**。扰动，切空间，等谱流，高阶扰动界，特征值，特征向量，奇异值

**美国数学学会主题分类号**。15A22,15A42, 65F15

**1. 引言**。矩阵扰动理论是众多领域中的一个经典且核心课题，包括数值线性代数 [13], 量子物理[9] 以及动力系统 [1]。一个基本问题涉及谱量在扰动下的稳定性：给定矩阵 $A \in \mathbb{C}^{n\times n}$ 和一个扰动$E \in \mathbb{C}^{n\times n}$， $A + E$ 的特征值和特征向量如何偏离$A$的相应量?

对于特征值，已有一套广泛的理论。诸如 Weyl 定理 [30],、Bauer-Fike 定理 [2], 及相关界 [28, §IV] [3, §III] 等经典结果表明，对于正规矩阵 $A$,

$$|\lambda_i(A + E) - \lambda_i(A)| \leq \|E\|,$$

其中 $\|\cdot\|$ 表示谱范数。该界在一般情况下是尖锐的，若不附加额外假设则无法改进，这一点在标量情形中已显而易见。

这类界很简单，只取决于扰动的范数。使用这些界的主要缺点是，它们没有利用 $E$ 可能具有的结构性质。因此，许多研究分析了结构化$E$ 的情形，并得出了更好的界。Li 和 Li [20], 的研究是一个显著例子，他们考虑了如下形式的厄米矩阵

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \qquad E = \begin{bmatrix} 0 & \widetilde{E} \\ \widetilde{E}^* & 0 \end{bmatrix}.$$

在此设定下，特征值扰动被证明关于 $\|E\|$是二阶的,

*版本日期：2026年2月26日。**资助：** YN 获得英国工程与自然科学研究委员会（EPSRC）资助，项目编号为 EP/Y010086/1 和 EP/Y030990/1。FH 获得数学研究所奖学金资助。

†牛津大学数学研究所，OX2 6GG，英国（francesco.hrobat@maths.ox.ac.uk, yuji.nakatsukasa@maths.ox.ac.uk）。

即

$$|\lambda_i(A+E)-\lambda_i(A)| \le \frac{\|E\|^2}{\eta},$$

其中

$$\eta = \min_{\lambda\in\Lambda(A_1),\ \mu\in\Lambda(A_2)} |\lambda-\mu|$$

表示两个对角块之间的谱间隙。该结果凸显了谱分离的关键作用，并表明结构化扰动可能显著改善特征值稳定性。这一原理也已在后续研究中得到应用 [13, 19]。

相比之下，关于特征向量的类似结果要微妙得多 [10,27]。特征向量本质上较不稳定，尤其是在谱间隙与扰动幅度相当时。例如，考虑对角矩阵

$$\begin{bmatrix} 1 & 0 \\ 0 & 1+\varepsilon \end{bmatrix}$$

其中 $\varepsilon$ 被认为很小，以及其扰动后的对应矩阵

$$\begin{bmatrix} 1 & \varepsilon \\ \varepsilon & 1+\varepsilon \end{bmatrix}.$$

尽管扰动的范数为 $\mathcal{O}(\varepsilon)$，且特征值以相同的阶数发生偏离，但相应的特征向量却旋转了一个 $\mathcal{O}(1)$ 角。尽管存在这一困难，特征向量信息通常比特征值估计更有力。事实上，左右特征向量的近似可通过瑞利商得到精确的特征值估计 [3]。

在本文中，我们对矩阵 $E$ 具有特定结构时的特征值和特征向量扰动进行了深入分析

$$E = AB - BA$$

其中 $B \in \mathbb{C}^{n\times n}$。换言之， $E$ 具有交换子的形式

$$E = [A, B] := AB - BA$$

涉及矩阵 $A$ 本身以及另一个矩阵 $B$。通过考察这一结构，我们有两个目标：首先，将 Li 和 Li 的结果 [20] 推广到这类扰动；与分块非对角结构相比，这是一种更一般的结构（见第 5 节）。

其次，我们说明矩阵 $B$ 在特征值，尤其是特征向量的扰动理论中发挥的重要作用，并揭示其与经典结果之间的紧密联系。我们还说明，如果用矩阵 $B$来表示，扰动理论中的一些已有量会大大简化。考察渐近展开（另见第3.1小节）即可看出这一点：令 $x_i$， $y_i$ 表示矩阵 $A$ 对应于特征值 $\lambda_i(A)$ 的右特征向量和左特征向量，我们证明扰动特征值$\lambda_i(A+E)$ 的渐近展开为

$$\lambda_i(A+E) = \lambda_i(A) + y_i^* BEx_i - y_i^* BEBx_i + \mathcal{O}(\|E\|^4)$$

其中

$$\eta_i = \min_{i \neq j} |\lambda_i(A) - \lambda_j(A)|$$

是相对于第 $i$ 个特征值的谱间隙。更简单地说，扰动特征向量 $x_i(A+E)$ 和 $y_i(A+E)$（分别对应于对 $x_i$ 和 $y_i$ 的扰动）的渐近展开为

$$x_i(A+E) = x_i - Bx_i + \mathcal{O}(\|E\|^2) \quad \text{and} \quad y_i(A+E) = y_i + B^* y_i + \mathcal{O}(\|E\|^2).$$

请注意，这些渐近展开不需要任何关于不同于 $y_i$、$x_i$ 或 $\lambda_i(A)$ 的特征值或特征向量的信息，因为这些信息已内在地编码在矩阵 $B$ 中。除了渐近展开之外，我们还给出了非渐近界；这类界在数值线性代数等领域往往更有用。

并非所有矩阵都能写成交换子的形式：事实上，给定矩阵 $E$，矩阵 $B$ 存在，当且仅当 $E$ 属于等谱流形的切空间，而 $A$ 属于该流形。这一思想见于先前有关等谱流的研究。其中一项奠基性研究是 Lax [18] 在如下形式的矩阵常微分方程背景下所做的工作

$$\frac{d}{dt}A(t) = [A(t), B(t)]. \tag{1.1}$$

在这一设定中，$B(t)$扮演我们的矩阵 $B$的角色。Lax论文的主要结果之一是，对于所有 $t$，$A(t)$与 $A(0)$具有相同的谱。如果 $B(t) \equiv B$是常数，则这一点可以清楚地看出：

$$A(t) = e^{-tB} A(0) e^{tB}.$$

等谱流被应用于不同的场景：$QR$算法的连续版本，即 $QR$ 流 [7]，其形式为

$$\dot{A} = [A, B(A)]$$

其中

$$B(A) = A_- - A_-^*$$

其中，$A_-$ 是 $A$的严格下三角部分。Brockett-Wegner流 [7, §7]

$$\dot{A} = [A, [A, N]]$$

其中 $N$ 是对角的。此外，Lax流还出现在等谱流形 [8]上的黎曼优化中。除此之外，等谱流也出现在物理学中，例如海森堡运动方程 [26, §2.2.3]。

形如方程（1.1）其中 $A$（$t$）为厄米矩阵的方程通常通过考虑李群分裂 [29]

$$\begin{aligned} \mathcal{M} &\iff U(n) \\ A(t) &\iff Q(t) \end{aligned}$$

其中，$\mathcal{M}$ 是等谱流形，$A(t) = Q(t) D_0 Q(t)^*$ 是特征值分解。在此情况下，$Q(t)$满足

$$\dot{Q}(t) = -B(t)Q(t).$$

这表明矩阵 $B$ 与扰动特征向量矩阵之间存在密切联系。

本文其余部分安排如下：第2节介绍等谱矩阵的切空间。第3节给出主要结果和非渐近界。为全面了解这一问题，第4节还将分析 $E$ 不在切空间中的情形，方法是将 $E$ 分解为一个属于切空间的矩阵与另一个和切空间正交的矩阵之和。本节将说明如何通过考察扰动在切空间上的投影来构造矩阵 $B$ ，以及 $B$ 在这种情况下也发挥着重要作用。在第5节中，我们讨论块对角情形，其中明确包含 Li 与 Li 设定。当特征值不是单重的，或存在特征值簇时，都需要考虑这种情形。在第6节中，我们在厄米矩阵的假设下改进前述结果。在第7节中，我们将相同的结果推广到奇异值分解。在这一设置下，流形由具有相同奇异值的矩阵构成。该流形没有广为采用的特定名称。切空间的刻画如下： $E$ 当且仅当 $E$ 可以写成

$$E = S_1 A + A S_2$$

其中 $S_1$， $S_2$ 是维数适当的反对称矩阵（$A$ 可能是矩形矩阵）。最后，第 8 节 将说明如何把这一设定推广到若尔当块。

**2. 等谱矩阵切空间的描述。** 给定一个可对角化矩阵 $A \in \mathbb{C}^{n\times n}$，我们希望研究在等谱流形的切空间中施加加性扰动 $E \in \mathbb{C}^{n\times n}$ 时特征值的扰动。首先，我们给出一个非正式的论证，说明 $E$属于切空间当且仅当

$$E = AB - BA.$$

李群一般情形的严格证明可见 [14]。

除非另有说明，我们始终假设 $A$ 的所有特征值都是单的。考虑特征值分解 $A = XDX^{-1}$。等谱流形可参数化为

$$\begin{aligned} GL(n,\mathbb{C})/D^{\times} &\longrightarrow \mathbb{C}^{n\times n} \\ W &\longrightarrow WDW^{-1} \end{aligned} \tag{2.1}$$

其中 $GL(n,\ \mathbb{C})/D^{\times}$ 是可逆矩阵空间关于以下等价关系所得的商空间

$$W_1 \sim W_2 \quad \text{if and only if} \quad W_1 = W_2 D$$

其中 $D$ 是一个可逆对角矩阵。这是因为特征向量矩阵在特征向量的缩放意义下是唯一的。如果我们只考虑厄米矩阵，则采用如下参数化方式

$$\begin{aligned} U(n)/T^n &\longrightarrow \mathbb{C}^{n\times n} \\ Q &\longrightarrow QDQ^{*} \end{aligned}$$

其中 $U(n)/T^n$ 是酉矩阵对酉对角矩阵的商 .

为了确定切空间，我们对 $W$施加一个小的扰动 $P$

$$W + P \longrightarrow (W + P)D(W + P)^{-1}.$$

假设 $P$足够小，我们可以将 $(W + P)D(W + P)^{-1}$ 展开到一阶

$$\begin{aligned}(W + P)D(W + P)^{-1} &= W(I + W^{-1}P)D(I + W^{-1}P)^{-1}W^{-1} \\ &\approx W(I + W^{-1}P)D(I - W^{-1}P)W^{-1} = A + A(-PW^{-1}) - (-PW^{-1})A.\end{aligned}$$

由于 $P$是一般的，因此切空间可参数化为

$$\mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{n\times n}$$

$$B \longrightarrow AB - BA = [A, B].$$

此外，我们已经看到，这意味着特征向量矩阵的扰动等于

$$P = -BW.$$

注意，矩阵 $WDW^{-1}$在 公式（2.1） 由 $n^2 - n$个参数确定，因为 $W$ 的列可以重新缩放（即， $GL(n, \mathbb{C})/D^{\times}$ 的自由度为$n^2 - n$）。这意味着矩阵 $AB - BA$也由 $n^2 - n$个参数确定。定义 $x_i$ 为 $i$的第 $X$ 列，并令 $y_i$ 为 $i$的第 $X^{-*}$列。则有

$$y_i^*(AB - BA)x_i = e_i^*DXBx_i - y_i^*BX^{-1}De_i = \lambda_i(e_i^*XBX^{-1}e_i - e_i^*XBX^{-1}e_i) = 0.$$

正如预期的那样，我们有 $n$ 个约束。此外，切空间还具有如下刻画

Theorem 2.1（切空间的另一种刻画）。

设 $A \in \mathbb{C}^{n\times n}$ 为可对角化矩阵。设 $\lambda_i$ 为 $A$ 的第 $i$ 个特征值，且 $x_i$、$y_i \in \mathbb{C}^n$ 为对应于 $\lambda_i$ 的右特征向量和左特征向量。那么，当且仅当满足以下条件时， $E \in \mathbb{C}^{n\times n}$ 属于等谱流形的切空间

$$y_i^*Ex_i = 0, \qquad i = 1, \ldots, n.$$

注意到若 $B = p(A)$ 是 $A$ 的多项式，则也可以推导出这些约束：

$$Ap(A) - p(A)A = 0.$$

这些约束并不令人意外：事实上，根据经典微扰理论， $i$ 阶特征值的一阶扰动等于 [28, 12]，

$$\lambda_i(A + E) - \lambda_i(A) \approx y_i^*Ex_i.$$

现在我们考虑矩阵 $A$ 为分块对角矩阵的特殊情形

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}.$$

任意块非对角扰动

$$E = \begin{bmatrix} 0 & E_1 \\ E_2 & 0 \end{bmatrix}$$

都满足 $n$ 约束，这意味着我们可以找到一个矩阵 $B$，使得

$$E = AB - BA.$$

事实上，如果我们考虑与第一分块的某个特征值对应的右特征向量和左特征向量，则它们具有如下形式

$$x_i = \begin{bmatrix} \tilde{x}_i \\ 0 \end{bmatrix} \qquad y_i = \begin{bmatrix} \tilde{y}_i \\ 0 \end{bmatrix}$$

其中，0 向量的长度与 $A_2$ 的维数相同。这意味着

$$y_i^* E x_i = 0.$$

更明确地说，矩阵 $B$ 等于（相差一个关于 $A$ 的多项式）

$$B = \begin{bmatrix} 0 & B_1 \\ B_2 & 0 \end{bmatrix}$$

其中，$B_1$ 和 $B_2$ 满足

$$A_1 B_1 - B_1 A_2 = E_1 \qquad A_2 B_2 - B_2 A_1 = E_2.$$

因此，我们在此研究的设置是先前研究内容的推广 [20, 23]。注意，在分块情形下，即使 $B_1$ 和$B_2$ 的特征值并不简单，我们仍然可以计算 $A_1$ 和 $A_2$。我们所要求的只是 $A_1$ 的谱与 $A_2$的谱之间存在分离。

不失一般性（除可对角化这一假设外），我们考虑矩阵 $A$ 为对角矩阵的情形（因此记作 $D$）。事实上，若 $A = XDX^{-1}$，矩阵 $A + E$ 可以改写为

$$A + E = XDX^{-1} + E = X(D + X^{-1}EX)X^{-1}.$$

若 $E$ 等于 $AB - BA$，则

$$A + E = X(D + X^{-1}ABX - X^{-1}BAX)X^{-1} = X(D + DX^{-1}BX - X^{-1}BXD)X^{-1}$$

于是我们归结为对角情形，此时矩阵 $B$ 为 $X^{-1}BX$。如果作如下替换，以下各节中的所有结果都适用于一般的可对角化矩阵

$$B \leftarrow X^{-1}BX, \quad E \leftarrow X^{-1}EX$$

以及后续的估计

$$\|E\| \leftarrow \kappa_2(X)\|E\|, \quad \|B\| \leftarrow \kappa_2(X)\|B\|.$$

**3. 特征向量和特征值的界**。我们先列出一些稍后会用到的结果。

**3.1. 初步计算与渐近展开**。本节中，我们假设所考虑矩阵的特征值均为单特征值。我们将在第 5 节讨论分块情形。此外，我们用 $D$ 表示对角矩阵，用 $A$ 表示一般矩阵。

LEMMA 3.1. 考虑一个对角矩阵 $D$，其对角线上有互不相同的特征值 $\{\lambda_i\}_1^n$。扰动 $E$ 属于 $D$ 的等谱流形，当且仅当

$$diag(E) = 0.$$

证明。设 $e_i$ 为 $i$标准基的第Theorem2.1，矩阵 $E$ 属于等谱流形，当且仅当

$$e_i^* E e_i = 0$$

对所有 $i$，因为 $e_i$ 是与 $\lambda_i$对应的左右特征向量。 □

给定切空间中的矩阵 $E$，我们可以利用以下引理计算满足$E = AB - BA$ 的矩阵 $B$。

LEMMA 3.2. 考虑一个对角矩阵 $D$，其对角线上有互不相同的特征值 $\{\lambda_i\}_1^n$，以及一个满足 $diag(E) = 0$ 的扰动 $E$。那么存在一个矩阵 $B$，使得

$$E = DB - BD$$

并且元素 $B(i,j)$ $i \neq j$ 等于

$$B(i,j) = \frac{E_{i,j}}{\lambda_i - \lambda_j} \quad i \neq j.$$

对角线上的元素是自由的（因为我们有 $n$ 个自由度）。为简便起见，我们将 $B$ 的对角线设为 0。等价地，

$$B(:,i) = (D - \lambda_i I)^\dagger E(:,i) \tag{3.1}$$

在一般情况下，当 $A$不是对角矩阵时，我们可以将 引理 3.2 改写如下。

LEMMA 3.3. 考虑矩阵 $A = XDX^{-1} \in \mathbb{C}^{n\times n}$，其中 $D$ 是具有互异特征值的对角矩阵。设 $E$ 是切空间中的扰动，则

$$E = AB - BA$$

*and*

$$Bx_i = (A - \lambda_i I)^\# E x_i$$

其中$(\cdot)^\#$ 是群逆 *[24, 12], [28, pp. 240-241],*，而 $x_i$ 表示 $X$的第 $i$ 列。

引理 3.3 将我们的框架与特征向量微扰的经典框架联系起来[24]。我们首先给出特征值和特征向量的渐近展开。这些展开已为人所知，尤其是在物理学微扰理论领域；Landau 等人已经推导出了这些渐近展开。 [17, §VI]。为使结果更加清晰，我们引入以下定义。

DEFINITION 3.4. 设 $t$ 为 $[0,1]$中的一个参数。定义 $x_i(t)$ 为 $D+tE$对应于特征值 $\lambda_i(t)$的右特征向量，该特征值是 $D$的第 $i$个特征值的扰动，记为 $\lambda_i := \lambda_i(0)$。类似地，用 $y_i(t)$ 表示与同一特征值 $\lambda_i(t)$相关联的左特征向量。

LEMMA 3.5. 根据定义3.4，$x_i(t)$ 的渐近展开为

$$x_i(t) = e_i - tB(:,i) + t^2(D - \lambda_i I)^\dagger E B(:,i) + \mathcal{O}((t\|E\|)^3)$$

假设 $B$ 的对角线为零。

证明。 回想一下，简单特征值及其特征向量是矩阵元素的解析函数 [16], 通过考察以下内容即可得出该结果

$$x_i(t) = e_i + tx_i^{(1)} + t^2x_i^{(2)} + \dots$$

$$\lambda_i(t) = \lambda_i + t\lambda_i^{(1)} + t^2\lambda_i^{(2)} + \dots$$

其中 $t$ 取值介于 0 和 1 之间，然后在等式中匹配 $t$ 的相同幂次项

$$(D+tE)(e_i + tx_i^{(1)} + t^2x_i^{(2)} + \dots) = (\lambda_i + t\lambda_i^{(1)} + t^2\lambda_i^{(2)} + \dots)(e_i + tx_i^{(1)} + t^2x_i^{(2)} + \dots). \qquad \blacksquare$$

对于左特征向量 $y_i(t)$，也可以推导出类似的结果，因为它是 $D^* + tE^*$ 的右特征向量，且 $-B^*$ 满足

$$D^*(-B^*) - (-B^*)D^* = E^*.$$

特别地，我们有

$$y_i^*(t) = e_i + tB(i,:) - t^2B(i,:)E(D-\lambda_i I)^\dagger + \mathcal{O}((t\|E\|)^3).$$

引理 3.5 意味着整个特征向量矩阵 $X(t)$ 的 $D+tE$ 可由下式很好地近似

$$I - tB.$$

逆矩阵 $X(t)^{-1}$ 也同样如此，它可由下式很好地近似

$$I + tB.$$

这并不令人意外，因为我们利用单位矩阵推导出了切空间。

$$(I+tB)D(I+tB)^{-1} \approx D + t(DB - BD).$$

值得注意的是，根据特征向量的渐近展开，我们能够得到阶数更高的特征值渐近展开。

LEMMA 3.6. $\lambda_i(t)$ 的渐近展开为

$$\lambda_i(t) = \lambda_i + t^2B(i,:)E(:,i) - t^3B(i,:)EB(:,i) + \mathcal{O}((t\|E\|)^4).$$

证明。 这一结果由瑞利-里茨商恒等式推出

$$\lambda_i(t) = \frac{y_i(t)^*(D+tE)x_i(t)}{y_i(t)^*x_i(t)}$$

并利用 引理 3.5中的特征向量渐近展开即可得到。 $\square$

这些结果已经令人惊讶。事实上，若已知 $B$ ，渐近展开仅取决于 $B$、 $E$、 $x_i$ 和 $y_i$ ，而不需要了解其他特征值或特征向量。此外，我们有

$$\sum_i B(i,:)E(:,i) = 0 \qquad \sum_i B(i,:)EB(:,i) = 0$$

因此，这些展开保留了特征值之和。最后，我们回顾特征值导数的经典表达式

LEMMA 3.7. 设 $A \in \mathbb{C}^{n\times n}$，并令 $E \in \mathbb{C}^{n\times n}$ 为一个扰动。令 $A(t) = A + tE$。$A + tE$ 的第 $i$ 个特征值的导数，记作 $\lambda_i(t)$，等于

$$\frac{d\lambda_i(t)}{dt} = \frac{y_i(t)^* E x_i(t)}{y_i(t)^* x_i(t)}$$

其中，$y_i$（$t$）、$x_i$（$t$）分别是 $\lambda_i$（$t$）的左右特征向量。

在下一节中，我们将给出非渐近结果。以下量在接下来的各节中将至关重要:

$$\rho_i = \frac{\|E\|}{\eta_i}$$

其中

$$\eta_i = \min_{\lambda_j,\ j\neq i} |\lambda_i - \lambda_j|.$$

称为相对于 $i$ 阶特征值的谱间隙。

3.2. 扰动特征向量的非渐近界。本节将给出特征向量的非渐近界。文献中已有多项工作研究扰动的不变子空间（包括特征向量） [15] [28,§V] [6, §2 §4] [24]。再次考虑矩阵 $D+tE$，其中 $t$ 的取值范围为 0 到 1。令 $x_i(t)$ 表示第 $i$ 个右特征向量，令 $y_i(t)$ 表示第 $i$ 个特征向量，并将二者归一化，使其第 $i$ 个分量等于 1（只要扰动的范数足够小，我们就可以这样做；例如参见 [28, §V]）。此外，令 $\eta_i = \min_{i\neq j} |\lambda_i - \lambda_j|$ 为第 $i$ 个谱间隙。为简便起见，我们仅给出关于 $x_i(t)$ 的界（$y_i(t)$ 的情形类似）。回顾引理 3.5 中关于特征向量的渐近展开

$$x_i(t) = e_i - tB(:,i) + t^2(D-\lambda_i I)^\dagger E B(:,i) + \mathcal{O}((t\|E\|)^3)$$

定义

(3.2)
$$x_i^{(0)} := e_i \quad x_i^{(1)} := -B(:,i) \quad x_i^{(2)} := (D-\lambda_i I)^\dagger E B(:,i)$$
$$\Delta_i(t) := \lambda_i(t) - \lambda_i \quad \Pi_i[z] := (I - e_i e_i^*)z$$
$$r_i^{(0)}(t) := x_i(t) - x_i^{(0)} \quad r_i^{(1)}(t) := x_i(t) - x_i^{(0)} - t x_i^{(1)} \quad r_i^{(2)}(t) := x_i(t) - x_i^{(0)} - t x_i^{(1)} - t^2 x_i^{(2)}.$$

换言之，$x_i^{(k)}$ 是 $x_i(t)$渐近展开中 $t^k$ 的系数，而$r_i^{(k)}$ $(t)$是相对于 $k$阶渐近展开的残差。我们将 $y_i^{(\cdot)}$ 和 $\ell_i^{(\cdot)}$ $(t)$定义为 $y_i(t)$中 $x_i^{(\cdot)}$ 和 $r_i^{(\cdot)}$ $(t)$的类似物。本节的主要结果如下。

THEOREM 3.8（特征向量的非渐近界）。设 $D$ 为一个对角矩阵，其对角线上的特征值为 $\{\lambda_i\}_{i=1}^n$。设 $E$、$B$、$x_i(t)$ 如前所定义。回顾其定义

$$\rho_i = \frac{\|E\|}{\eta_i}.$$

则 $r_i^{(k)}(t)$ 如方程(3.2)所示，满足

$$\begin{aligned}
\|r_i^{(0)}(t)\| &\le \frac{t\|B(:,i)\|}{1-\frac{|\Delta(t)|+t\|E\|}{\eta_i}} \le \frac{t\rho_i}{1-2t\rho_i} \\
\|r_i^{(1)}(t)\| &\le \frac{t(|\Delta(t)|+t\|E\|)\|B(:,i)\|}{\eta_i(1-\frac{\Delta(t)+t\|E\|}{\eta_i})} \le \frac{2(t\rho_i)^2}{1-2t\rho_i} \\
\|r_i^{(2)}(t)\| &\le \left(|\Delta(t)|\frac{t\|B(:,i)\|}{\eta_i} + \frac{t^2(|\Delta(t)|\|E\|+t\|E\|^2)\|B(:,i)\|}{\eta_i^2}\right)\frac{1}{1-\frac{2t\|E\|}{\eta_i}} \\
&\le \left(t\frac{|\Delta(t)|}{\eta_i}\rho_i + 2(t\rho_i)^3\right)\frac{1}{1-2t\rho_i}.
\end{aligned}$$

证明。 根据特征向量的定义

$$(D+tE)x_i(t) = \lambda_i(t)x_i(t)$$

我们可以写出以下方程

$$\begin{aligned}
(D-\lambda_i I)r_i^{(0)}(t) &= \Delta(t)r_i^{(0)}(t) - t\Pi_i[Er_i^{(0)}(t)] - tEe_i \\
(D-\lambda_i I)r_i^{(1)}(t) &= \Delta(t)r_i^{(0)}(t) - t\Pi_i[Er_i^{(0)}(t)] \\
(D-\lambda_i I)r_i^{(2)}(t) &= \Delta(t)r_i^{(0)}(t) - t\Pi_i[Er_i^{(1)}(t)].
\end{aligned} \tag{3.3}$$

线性界。 由公式（3.3）中的第一个方程可得

$$\|r_i^{(0)}(t)\| \le \frac{|\Delta(t)|}{\eta_i}\|r_i^{(0)}(t)\| + t\frac{\|E\|}{\eta_i}\|r_i^{(0)}(t)\| + t\|B(:,i)\|$$

其中我们利用了 $(D-\lambda_i I)^\dagger Ee_i = Be_i$ 正如 引理 3.2所述。因此,

$$\|r_i^{(0)}(t)\| \le \frac{t\|B(:,i)\|}{1-\frac{|\Delta(t)|+t\|E\|}{\eta_i}} \tag{3.4}$$

二次界。 由 公式（3.3）中的第二个方程以及 公式（3.4）可得

$$\|r_i^{(1)}(t)\| = \frac{t(|\Delta(t)|+t\|E\|)\|B(:,i)\|}{\eta_i\left(1-\frac{|\Delta(t)|+t\|E\|}{\eta_i}\right)} \tag{3.5}$$

三次[1] 界。 由公式（3.3）、公式（3.4）和公式（3.5）可得

$$\|r_i^{(2)}(t)\| \le \left(|\Delta(t)|\frac{t\|E\|}{\eta_i^2} + \frac{t^2(|\Delta(t)|\|E\|^2+t\|E\|^3)}{\eta_i^3}\right)\frac{1}{1-\frac{2t\|E\|}{\eta_i}} \tag{3.6}$$

其中我们使用了 $\|B(:\ \ i)\| \le \|E\|/\eta_i$ 由公式 (3.1) 和 $|\Delta(t)| \le \|E\|$ 根据 Bauer-Fike 得到。 □

[1]完整证明请见下一节。

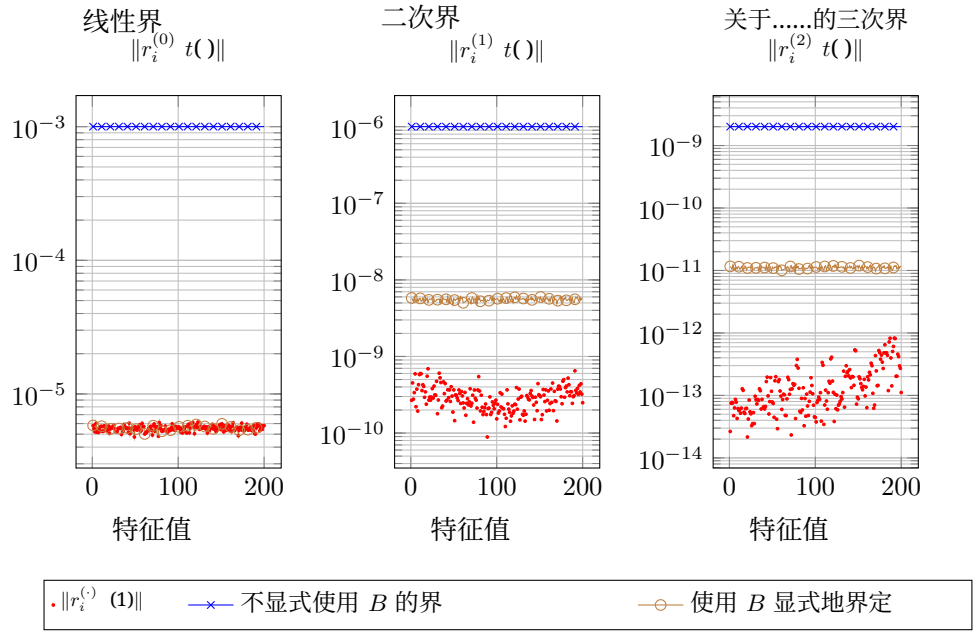


FIG. 1。 *Plot of bounds in Theorem 3。 8. The red dots are the actual norm of the residuals, the blue line is the bound when we have used* $\|B \quad i\| \leq \|E\|/\eta_i$ *and the brown circles are the bound*(:)(:)()。 *where* $\|B \quad i\|$ *is kept as it is.* $|\Delta\ t|$ *in is bounded by Theorem 3 9. The matrix considered is a diagonal matrix with eigenvalues equispaced from* 1 *to* 200 *and* $B$ *is a random off-diagonal matrix such that* $\|E\| = 10^{-3}$. *Note that* $\rho_i$ *is constant and equal to* $10^{-3}$. *The blue line in the linear bound is* $\approx \|E\|/\eta_i$, *which is the classical bound presented for perturbed eigenvectors.*

我们对定理 3.8作几点说明。首先，界

$$\|B(:,i)\| \leq \frac{\|E\|}{\eta_i} = \rho_i$$

如果 $E$ 具有特定结构，则可能较为宽松，例如如果 $E$ 具有较大的元素 $E(j,\ i)$，其中 $|\lambda_j - \lambda_i|$ 也较大。这是因为

$$\|B(:,i)\|^2 = \sum_j \frac{E_{j,i}^2}{(\lambda_j - \lambda_i)^2}.$$

其次，该项

$$|\Delta(t)|\frac{t\|E\|}{\eta_i^2}$$

在定理 3.8 的最后一个不等式中，必须是关于 $\|E\|$ 的三次项，该界才有意义，因为我们要界定一个阶为 $\|E\|^2$的近似。换言之，我们需要证明 $|\Delta(t)|$ 的阶为 $\|E\|^2$，这一结果比简单的 Bauer-Fike 更强，与 Li 和 Li 界类似 [20]。这一结果的证明将是下一节的核心内容。在 图 1、图 2和图 3中，我们绘出了若干示例矩阵对应的定理 3.8 中的界。从这些图中可以看出，对 $\|B(:\quad i)\|$ 进行界定时，若取 $\|E\|/\eta_i$ ，会使该界增大几个数量级。

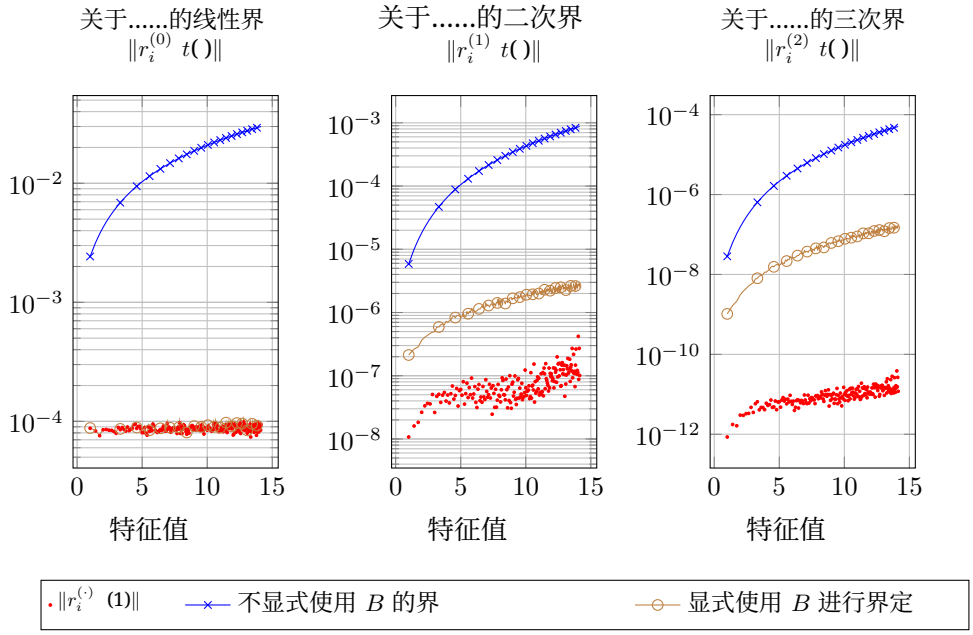


FIG. 2。 *Plot of bounds in Theorem 3。8. The red dots are the actual norm of the residual, the blue line is the bound when we have used* $\|B\ \ i\ \| \le \|E\|/\eta_i$ *and the brown circles are the bound* $(:)(:)()$。 *where* $\|B\ \ i\ \|$ *is the eigenvalues* $|\Delta|$ *in the bounded box of the equispaced points from 1 to 200 and* $B$ *is a random off-diagonal matrix such that* $\|E\| = 10^{-3}$. *Note that for the largest eigenvalues* , , $\rho_i \approx 10^{-1}$. *The blue line in the linear bound is* $\approx \|E\|/\eta_i$ , *which is the classical bound presented for perturbed eigenvectors.*

**3.3. 利用特征向量渐近展开对扰动特征值给出非渐近界**。本节中，我们利用定理 3.8 为特征值给出非渐近界。主要定理如下。

THEOREM 3.9（特征向量的非渐近界）。 设 $D$ 为一个对角矩阵，其对角线上的特征值为 $\{\lambda_i\}_{i=1}^n$ 。设 $E$, $B$, $x_i(t)$ 如前所定义。回顾定义

$$\rho_i = \frac{\|E\|}{\eta_i}.$$

我们有如下结论；

$$|\lambda_i(t) - \lambda_i| \le \frac{t^2\|E\|\frac{\|B(i,:)\|+\|B(:,i)\|}{2} + \frac{t^3\|E\|\|B(i,:)\|\|B(:,i)\|}{3\left(1-\frac{2\|E\|}{\eta_i}\right)}}{1 - \frac{2\|E\|}{\eta_i} - \frac{\|B(i,:)\|\|B(:,i)\|}{\left(1-\frac{2\|E\|}{\eta_i}\right)}} \le \frac{t^2\|E\|(\rho_i + \frac{t}{3}\frac{\rho_i^2}{1-\rho_i})}{1 - 2\rho_i - \frac{\rho_i^2}{1-2\rho_i}}.$$

证明。 证明依赖于特征值导数的公式引理 3.7 $d\lambda_i(t)\ dt = {}^{y_i}{}_{(t)^*Ex_i(t)\ y_i(t)^*x_i(t)}$ . 首先考虑这些分解

$$x_i(t) = x_i^{(0)} + r_i^{(0)}(t)$$
$$y_i(t) = y_i^{(0)} + \ell_i^{(0)}(t).$$

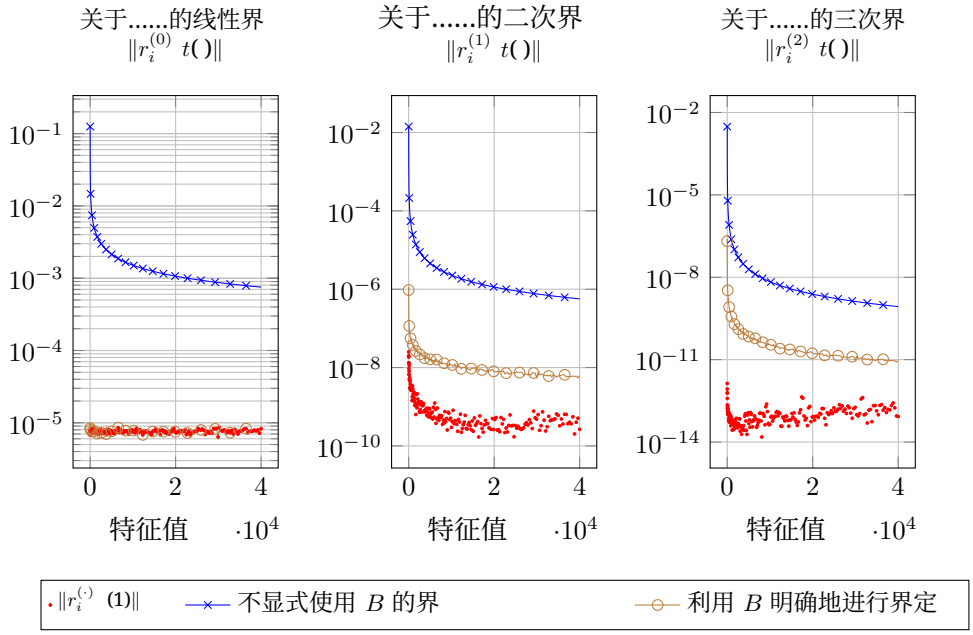


FIG. 3。 *Plot of bounds in Theorem 3。8. The red dots are the actual norm of the residual, the blue line is the bound when we have used* $\|B\ \ i\| \le \|E\|/\eta_i$ *and the brown circles are the bound* (:)(:)()。 *diagonal matrix with eigenvalues* $|\Delta t|$ *in the square of the equispaced points from 1 to 200 and* $B$ *is a random off-diagonal matrix such that* $\|E\| = 0.3$. *Note that for the largest eigenvalues* $\rho_i \approx 10^{-3}$. , ,

*The blue line in the linear bound is* $\approx \|E\|/\eta_i$ *which is the classical bound presented for perturbed , eigenvectors.*

我们可以将导数改写为

$$\frac{d\lambda_i(t)}{dt} = \frac{y_i(t)^* E x_i(t)}{y_i(t)^* x_i(t)} = \frac{(y_i^{(0)} + \ell_i^{(0)}(t))^* E (x_i^{(0)} + r_i^{(0)}(t))}{(y_i^{(0)} + \ell_i^{(0)}(t))^* (x_i^{(0)} + r_i^{(0)}(t))} =$$

$$= \frac{\ell_i^{(0)}(t)^* E x_i^{(0)} + y_i^{(0)} E r_i^{(0)}(t) + \ell_i^{(0)}(t)^* E r_i^{(0)}(t)}{1 + \ell_i^{(0)}(t)^* r_i^{(0)}(t)}$$

其中我们利用了 $x_i^{(0)}$， $y_i^{(0)} \perp r_i^{(0)}(t)\ \ \ell_i^{(0)}(t)$. 根据 定理 3.8 我们有

$$\|r_i^{(0)}(t)\| \le \frac{t\|B(:,i)\|}{1 - \frac{2\|E\|}{\eta_i}}$$

$$\|\ell_i^{(0)}(t)\| \le \frac{t\|B(i,:)\|}{1 - \frac{2\|E\|}{\eta_i}}$$

这意味着

$$\left|\frac{d\lambda_i(t)}{dt}\right| \le \frac{t\|E\|(\|B(i,:)\| + \|B(:,i)\|) + \frac{t^2\|E\|\|B(i,:)\|\|B(:,i)\|}{\left(1 - \frac{2\|E\|}{\eta_i}\right)}}{1 - \frac{2\|E\|}{\eta_i} - \frac{\|B(i,:)\|\|B(:,i)\|}{\left(1 - \frac{2\|E\|}{\eta_i}\right)}}.$$

通过积分，我们得到该界。 □

该定理表明，特征值的扰动关于 $\|E\|$是二次的，并特别地证明了定理 3.8中的最后一个不等式为 $\mathcal{O}\left(\|E\|^3\right)$ 。

**3.4. 特征值界的进一步改进。** 鉴于我们掌握了特征向量的非渐近界，且精度达到二阶修正，原则上可以将其代入特征值导数的表达式，从而得到高阶特征值的非渐近界。遗憾的是，引理 3.7分母中的项数随特征向量界的阶数呈二次增长，计算也十分繁琐。幸运的是，如果我们对 $\rho_i$作进一步假设，例如

$$\rho_i \leq \frac{1}{4}$$

原则上，我们可以证明另外两个定理。

THEOREM 3.10. 设 $D \in \mathbb{C}^{n\times n}$ 为对角矩阵，且 $E \in \mathbb{C}^{n\times n}$ 满足 $diag(E) = 0$ ，矩阵 $B \in \mathbb{C}^{n\times n}$ 满足 $E = AB - BA$ 和 $diag(B) = 0$。设 $\lambda_i$ 为 $D$ 的第$i$个特征值， $\lambda_i(1)$ 为 $D + E$的第 $i$ 个特征值。那么，若

$$\rho_i = \frac{\|E\|}{\eta_i} \leq \frac{1}{4}$$

成立：

$$|\lambda_i(1) - \lambda_i - B(i,:)E(:,i)| \leq C_1 \frac{\|E\|^3}{\eta_i^2} = C_1\|E\|\rho_i^2$$

其中， $C_1$ 是一个与所有涉及的量无关的常数。

THEOREM 3.11. 给定一个对角矩阵 $D \in \mathbb{C}^{n\times n}$、一个满足 $diag(E) = 0$ 的一般矩阵 $E \in \mathbb{C}^{n\times n}$，以及矩阵 $B \in \mathbb{C}^{n\times n}$ ，使得 $E = AB-BA$ 且 $diag(B) = 0$。令 $\lambda_i$ 为 $D$ 的第 $i$ 个特征值， $\lambda_i(1)$ 为 $D + E$的第 $i$个特征向量。那么，若

$$\frac{\|E\|}{\eta_i} \leq \frac{1}{4}$$

成立：

$$|\lambda_i(1) - \lambda_i - B(i,:)E(:,i) + B(i,:)EB(:,i)| \leq C_2 \frac{\|E\|^4}{\eta_i^3}$$

其中， $C_2$ 是一个与所有涉及的量无关的常数。

**4. 一般扰动 $E$.** 本节将先前的结果推广到扰动 $E$ 不属于切空间的情形。与之前一样，我们将问题归结为对角情形，并考虑对角矩阵 $D$。定义 $B$ 为

$$DB - BD = E - diag(E)$$

其中， $diag(E)$ 是一个对角矩阵，其对角线元素与 $E$ 的对角线元素相同。注意，

$$E = (E - diag(E)) + diag(E)$$

是一种加性分解，在

$$\mathcal{T}_D\mathcal{M} \times \mathcal{T}_D\mathcal{M}^{\perp}$$

其中 $\mathcal{M}$ 是等谱流形，而 $\mathcal{T}_D\mathcal{M}$ 是 $D$中的切空间。事实上，如果我们使用自然度量，

$$\langle (E - diag(E), diag(E)\rangle = \mathrm{tr}((E - diag(E))^* diag(E)) = 0$$

由于零对角矩阵与对角矩阵的乘积具有零对角线，并且根据 引理 3.1 的特征，第一矩阵属于切空间。再次考虑 公式（3.3）。如果 $E$ 具有非零对角线，则新方程为

$$\begin{aligned}
(D - \lambda_i I) r_i^{(0)}(t) &= \Delta(t) r_i^{(0)}(t) + tE r_i^{(0)}(t) + t(E - E(i,i)I)e_i \\
(D - \lambda_i I) r_i^{(1)}(t) &= \Delta(t) r_i^{(0)}(t) - t\Pi_i[E r_i^{(0)}(t)] \\
(D - \lambda_i I) r_i^{(2)}(t) &= (\Delta(t) - tE(i,i)) r_i^{(0)}(t) - t\Pi_i[(E - E(i,i)I) r_i^{(1)}(t)].
\end{aligned}$$

利用这些方程，我们可以将 定理 3.8 推广到这一情形，并给出特征值的界；需要注意的是，此时特征向量的渐近展开为

$$x_i(t) \approx e_i - tB(:,i) + t^2 (D - \lambda_i I)^{-1} E B(:,i) + t^2 E(i,i)(D - \lambda_i I)^{-1} B(:,i)$$

而特征值的展开为

$$\lambda_i(t) \approx \lambda_i + tE(i,i) + t^2 B(i,:)E(:,i).$$

**5.** $2\times 2$ **分块对角矩阵**。 现在假设 $D$ 为对角矩阵，但分为两个分块

$$D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}.$$

在每个分块内，特征值可能是重特征值，也可能彼此接近。特别地，暂时假设 $D_1 = \lambda_1 I$。假设 $\lambda_1$ 与 $D_2$的谱之间的距离为 $gap$ ，即

$$gap = \min_{\lambda \in \Lambda(D_2)} |\lambda_1 - \lambda|.$$

考虑扰动

$$E = \begin{bmatrix} 0 & E_1 \\ E_2 & 0 \end{bmatrix}$$

其中， $E_1$ 和 $E_2$ 的维数与 $D$的分块相匹配。在这种情况下，即使存在重特征值，矩阵 $B$ 仍有明确定义。特别地， $B$ 等于（可以选为）

$$B = \begin{bmatrix} 0 & B_1 \\ B_2 & 0 \end{bmatrix}$$

and

$$E_1 = D_1 B_1 - B_1 D_2, \qquad E_2 = D_2 B_2 - B_2 D_1$$

请注意，与一般情况不同， $B$ 中的元素只涉及属于不同块的特征值之差（在本例中，只涉及 $\lambda_1$与 $D_2$中的一个特征值之差）。在这种情况下，我们希望像之前那样，针对第一块的特征向量给出特征子空间的一阶近似，正如我们

之前所做的那样。请注意，由于第一个特征值是重特征值，我们无法对单个特征向量给出任何结论，只能整体分析特征子空间。与单个特征向量的情况类似，我们希望证明，该特征子空间可以很好地由

$$span\left(\begin{bmatrix} I \\ -tB_2 \end{bmatrix}\right)$$

这与我们针对$x_i(t)$得到的 $e_i - tB(:,i)$一阶展开的对应分块相同。

设 $x_i(t)$为对应于 $\lambda_i(t)$的归一化右特征向量。考虑 $x_i(t)$的分解：

$$x_i(t) = \begin{bmatrix} I \\ -tB_2 \end{bmatrix} x_i^{(top)}(t) + \begin{bmatrix} 0 \\ r_i(t) \end{bmatrix} \tag{5.1}$$

其中， $x_i^{(top)}(t)$是一个向量，其长度等于 $D_1$ 的宽度，且其各个分量是 $x_i(t)$的前几个分量。我们要证明， $r_i(t)$由 $\|E\|$的二次项控制。我们再次考虑特征向量方程：

$$(D+tE)x_i(t) = \lambda_i(t)x_i(t)$$

$$(D+tE)(\begin{bmatrix} I \\ -tB_2 \end{bmatrix} x_i^{(top)}(t) + \begin{bmatrix} 0 \\ r_i(t) \end{bmatrix}) = \lambda_i(t)(\begin{bmatrix} I \\ -tB_2 \end{bmatrix} x_i^{(top)}(t) + \begin{bmatrix} 0 \\ r_i(t) \end{bmatrix}).$$

考察最后几行，可得

$$(D_2 - \lambda_i(t)I)r_i(t) = tB_2(D_1 - \lambda_i(t)I)x_i^{(top)}(t) \tag{5.2}$$

这里我们利用 $E_2$ 的定义改写了 $D_2B_2$。这意味着

$$\|r_i(t)\| \le \frac{t^2\|B_2\|\|E\|}{gap - \|E\|}$$

这里我们使用 Bauer-Fike 对 $\|D_1 - \lambda_i(t)I\|$ 和 $\|D_2 - \lambda_i(t)I\|$进行界定。我们证明了以下定理。

THEOREM 5.1. 考虑一个形式如下的对角矩阵 $D \in \mathbb{C}^{n\times n}$

$$D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}$$

其中 $D_1 = \lambda_1 I$，并令

$$gap = \min_{\lambda\in\Lambda(D_2)} |\lambda_1 - \lambda| > 0$$

为谱间隙。设 $E$ 和 $B$ 具有如下分块

$$E = \begin{bmatrix} 0 & E_1 \\ E_2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & B_1 \\ B_2 & 0 \end{bmatrix}$$

其中各块的大小与 $A$相匹配，且 $B_1$ 满足

$$E_1 = D_1B_1 - B_1D_2, \quad E_2 = D_2B_2 - B_2D_1.$$

设 $\mathcal{X}$ 为

$$\mathcal{X} = span\left(\begin{bmatrix} I \\ -B_2 \end{bmatrix}\right)$$

那么，对于 $D+tE$ 的每一个相对于 $D$的左上块的扰动特征值的归一化右特征向量 $x_i(t)$，我们有

$$\|\Pi_{\mathcal{X}^\perp} x_i(t)\| \leq \frac{\|B_2\|\|E\|}{gap - \|E\|}$$

其中 $\Pi_{\mathcal{X}^\perp}$ 是 $\mathcal{X}^\perp$ 上的投影。

**5.1. $D_1$ 不是单位矩阵的倍数**。这里，假设 $D_1$ 是对角矩阵，且其特征值位于闭球内，我们给出一个类似的界。

$$\bar{B}(\lambda_1; \Delta/2)$$

也就是说，我们允许 $D_1$ 不是单位矩阵的倍数，但假设 $\Delta$ 与 $\|E\|$ 相当或更小。由 方程 (5.2)

$$(D_2 - \lambda_i(t))r_i(t) = tB_2(D_1 - \lambda_i(t))x_i^{(top)}(t)$$

其中 $r_i(t)$ 和 $x_i^{(top)}$ ($t$的定义与 公式 (5.1)中相同，我们得到

$$\|r_i(t)\| \leq \frac{t^2\|B_2\|(\|E\| + \Delta)}{gap - \|E\|}$$

这里我们用到了 Bauer-Fike。注意，在这种情况下，即使分子中的 $\|E_1\|$ 通过二次项得到改进（即我们使用了比 Bauer-Fike 更强的结果），该项

$$t^2\|B_2\|\Delta$$

无法进一步改进。如果矩阵是厄米矩阵，则我们在第 6.2 小节中给出更强的结果。

**6. 埃尔米特矩阵**。

**6.1. 厄米矩阵的一个简单初步界**。本节将说明，当矩阵 $A$ 为厄米矩阵时，情况会如何简化。以下简单结果无需假设 $A$ 为对角矩阵，也无需假设其谱具有分离性。

THEOREM 6.1. 设 $A \in \mathbb{C}^{n\times n}$ 为厄米矩阵， $B \in \mathbb{C}^{n\times n}$ 为斜埃尔米特矩阵。定义

$$E := AB - BA.$$

由于 $B$ 为斜埃尔米特矩阵， $E$ 为厄米矩阵。因此

$$|\lambda_i(A+E) - \lambda_i(A)| \leq \|B\|_2\|AB - BA\|_2 = \|B\|_2\|E\|_2.$$

证明。 假设

$$(A + t(AB - BA))x_i(t) = \lambda_i(t)x_i(t) \tag{6.1}$$

其中，对于所有 $t$， $x_i(t)$ 是与 $\lambda_i(t)$ 相关联的归一化特征向量（当 $\lambda_i(t)$ 为重特征值时，特征向量不唯一似乎会带来问题，但实际上并非如此 [25]）。然后，根据标准的特征值扰动理论，我们有

$$\frac{d\lambda_i(t)}{dt} = x_i(t)^*(AB - BA)x_i(t).$$

现在，利用 方程（6.1），我们有 $Ax_i(t) = (\lambda_i(t)I - t(AB-BA))x_i(t)$，因此

$$\begin{aligned}\frac{d\lambda_i(t)}{dt} &= x_i(t)^*(AB-BA)x_i(t)\\ &= ((\lambda_i(t)I - t(AB-BA))x_i(t))^*Bx_i(t) - x_i(t)^*B(\lambda_i(t)I - t(AB-BA))x_i(t)\\ &= x_i(t)^*B((\lambda_i(t)I + t(AB-BA))x_i(t)) - x_i(t)^*B(\lambda_i(t)I - t(AB-BA))x_i(t)\\ &= tx_i^*(t)B(AB-BA)x_i(t) + tx_i^*(t)(AB-BA)Bx_i(t).\end{aligned}$$

因此有 $|\frac{d\lambda_i(t)}{dt}| \le 2|t|\|B\|_2\|AB-BA\|_2$. 因此，

$$|\lambda_i(1)-\lambda_i(0)| \le \int_0^1 2|t|\|B\|_2\|AB-BA\|_2 dt = \|B\|_2\|AB-BA\|_2 = \|B\|_2\|E\|_2. \qquad \square$$

这个界既简洁优雅，又无需对矩阵进行对角化，也不依赖任何关于特征向量的结果。此外，该界是二次的，因为

$$\|B\|_2 \le \frac{\|E\|_F}{\eta}$$

其中 $\eta$ 定义为全局谱间隙

$$\eta = \min_{\lambda_i,\ \lambda_j\in\Lambda(A),\lambda_i\neq\lambda_j} |\lambda_i-\lambda_j|$$

且在所有可能的 $B$ 取值中， $B$ 的范数最小（回想一下， $B$ 有 $n$个自由度）。另一方面，这个界有一个重大局限：假设 $A$ 是对角的，则元素 $B(i,\ j)$ 随 $\frac{1}{\lambda_i-\lambda_j}$ 变化。这意味着，如果存在一个特征值簇，即使 $\lambda_i$ 与其余部分充分分离，这个界也可能偏松。不仅如此，如果 $A$ 具有重特征值，我们实际上还假设 $E$不仅在对角线上，而且在其他主子式中也有零元素（即 $E$必须满足额外的约束）。

在继续之前，值得注意的是，如果矩阵 $A$ 是厄米矩阵，那么等谱流形的切空间总是包含斜埃尔米特矩阵。这是因为如果 $E$ 是斜埃尔米特矩阵，那么

$$x^*Ex = 0 \qquad \forall x \in \mathbb{C}^n$$

特别是，当 $x$ 是 $A$的特征向量时，该结论成立。由 定理 2.1可得该结论。

**6.2.** $2\times 2$ **分块对角厄米矩阵**。本节继续 第 5 节中的论证，并在埃尔米特性的假设下给出更强的结果。我们考虑厄米对角矩阵

$$D = \begin{bmatrix} D_1 & 0\\ 0 & D_2\end{bmatrix}$$

其中 $D_1 = \lambda_1 I$。此外，矩阵 $E$

$$E = \begin{bmatrix} 0 & E_1\\ E_2 & 0\end{bmatrix}$$

是厄米的，因此

$$E_2 = E_1^* \quad \text{and} \quad B_2 = -B_1^*.$$

考虑导数 引理 3.7

$$\frac{d\lambda_i(t)}{dt} = x_i^*(t)Ex_i(t).$$

利用公式 (5.1)中的分解，该表达式等于

$$x_i^*(t)Ex_i(t) = (x_i^{(top)}(t)^* \begin{bmatrix} tB_1E_2 & E_1 \end{bmatrix} + \begin{bmatrix} l_i(t)^*E_2 & 0 \end{bmatrix})(\begin{bmatrix} I \\ -tB_2 \end{bmatrix} x_i^{(top)}(t) + \begin{bmatrix} 0 \\ r_i(t) \end{bmatrix}) =$$

$$= tx_i^{(top)}(t)^*(B_1E_2 - E_1B_2)x_i^{(top)}(t) + x_i^{(top)}(t)^*E_1r_i(t) + l_i(t)^*E_2x_i^{(top)}(t).$$

利用定理 5.1 的结果（并注意到，在此情况下， $\|E\| = \|E_1\|$）we have

$$|x_i(t)^*Ex_i(t)| \leq t\|B_1E_1^* + E_1B_1^*\| + \frac{2t^2\|E_1\|^2\|B_1\|}{gap - \|E_1\|}$$

因此

$$|\lambda_i(1) - \lambda_i| \leq \frac{1}{2}\|B_1E_1^* + E_1B_1^*\| + \frac{2}{3}\frac{\|E_1\|^2\|B_1\|}{gap - \|E_1\|}.$$

我们证明了以下结果：

THEOREM 6.2. 给定一个形式如下的厄米矩阵 $D$

$$D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}$$

其中 $D_1 = \lambda_1 I$。令 $E$ 为如下形式的厄米扰动

$$E = \begin{bmatrix} 0 & E_1 \\ E_1^* & 0 \end{bmatrix}$$

且 $B$ 等于

$$B = \begin{bmatrix} 0 & B_1 \\ -B_1^* & 0 \end{bmatrix}$$

其中 $B_1$ 满足

$$E_1 = D_1B_1 - B_1D_2.$$

令 $\lambda_i$（1）为 $D+E$的第 $i$ 个特征值，并令 $\lambda_i$为 $D$的第 $i$ 个特征值，且假设 $\lambda_i$ 是 $D_1$ 的一个特征值，则

$$|\lambda_i(1) - \lambda_i| \leq \frac{1}{2}\|B_1E_1^* + E_1B_1^*\| + \frac{2}{3}\frac{\|E_1\|^2\|B_1\|}{gap - \|E_1\|}.$$

注意，项

$$\frac{1}{2}\|B_1E_1^* + E_1B_1^*\|$$

在 $B_1$ 中隐藏着关于整个 $D_2$ 的信息，而不仅仅是间隙。当 $D_2$ 的某些特征值远离 $\lambda_1$，但 $gap$ 很小时，这就能保证得到精确得多的界。

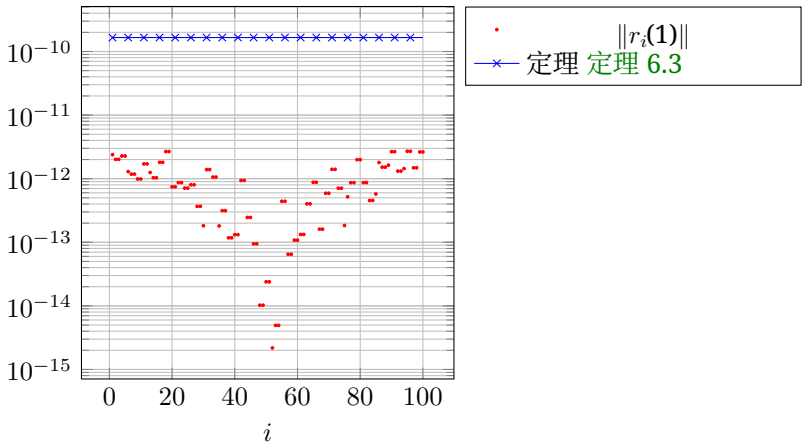


FIG. 4 *Bound in Theorem 6 3 where* $D \in \mathbb{R}^{200\ 200}$ $\lambda_1 = 1$ *and* $D_2$ *is diagonal with eigenvalues* 。 。 。 , ( ) *equispaced from* 2 *to* 101. *The matrix* $B_1$ *is random and scaled such that* $\|E\| = 10^{-2}$. *Note that this bound is much tighter than* $\|E\|/gap\ ^3 = 10^{-6}$.

**6.3. 进一步改进。** 再考虑 方程 (5.2)

$$(D_2 - \lambda_i(t))r_i(t) = tB_2(D_1 - \lambda_i(t))x_i^{(top)}(t).$$

现在，我们可以使用 定理 6.2 来界定特征值扰动。这得出

$$\|r_i(t)\| \leq \frac{1}{gap - \|E_1\|}\left(\frac{t^2}{2}\|B_1\|\|B_1E_1^* + E_1B_1^*\| + \frac{2}{3}t^3\frac{\|B_1\|^2\|E_1\|^2}{gap - \|E_1\|}\right).$$

THEOREM 6.3. 在与 定理 6.2 相同的假设下，并使用与方程 (5.1)中相同的记号，我们有

$$\|r_i(t)\| \leq \frac{1}{gap - \|E_1\|}\left(\frac{t^2}{2}\|B_1\|\|B_1E_1^* + E_1B_1^*\| + \frac{2}{3}t^3\frac{\|B_1\|^2\|E_1\|^2}{gap - \|E_1\|}\right).$$

换句话说，

$$\|r_i(t)\| \leq C\left(\frac{\|E\|}{gap}\right)^3$$

如果 $\frac{\|E\|}{\eta_i}$ 足够小于 1，且 $C$ 是一个常数。这表明，特征子空间扰动不是二次的，而是三次的。在特征值扰动的背景下， [21] 也指出了这一现象。

注意，根据 小节 5.1，如果$D_1$ 不是单位矩阵的倍数，而是具有宽度与 $\|E\|$相当的特征值簇，则这种改进并不能推广。

**7. 奇异值分解。** 在本节中，我们希望将此前用于特征值的同一思想应用于奇异值的情形。设 $A \in \mathbb{C}^{m\times n}$，并设 $A = U\Sigma V^*$ 为其奇异值分解。具有相同奇异值的矩阵流形可以参数化为

$$\mathbb{U}(m) \times \mathbb{U}(n) \longrightarrow \mathbb{C}^{m\times n}$$
$$[P, Q] \longrightarrow P\Sigma Q^*.$$

为求出切空间，我们按先前的方式进行：设 $P_1$ 和 $P_2$ 为两个小扰动。首先，我们必须确保分别沿着 $\mathbb{U}(m)$和 $\mathbb{U}(n)$的切空间移动，这意味着 $U+P_1$ 和 $V+P_2$在一阶意义下是正交归一的，从而可得

$$(U+P_1)^*(U+P_1) = I + O(\|P_1\|^2) \Rightarrow (U^*P_1)^* = -(U^*P_1)$$

$$(V+P_2)^*(V+P_2) = I + O(\|P_2\|^2) \Rightarrow (V^*P_2)^* = -(V^*P_2).$$

现在，扰动后的奇异值分解具有如下形式

$$(U+P_1)\Sigma(V+P_2)^* = A + (P_1U^*)A + A(VP_2^*).$$

由于 $P_1$ 和 $P_2$ 是一般矩阵，切空间可描述为

$$SkewHerm(m) \times SkewHerm(n) \longrightarrow \mathbb{C}^{m\times n}$$

$$[S_1, S_2] \longrightarrow S_1A + AS_2.$$

同样，该矩阵满足

$$\mathscr{R}(u_i^*(S_1A + AS_2)v_i) = 0,$$

这是第 $i$个奇异值的一阶扰动。约束的数量与奇异向量（与非零奇异值相关联）的自由度相匹配，因为右奇异向量和左奇异向量可以用同一个酉对角矩阵进行缩放。

设 $A \in \mathbb{C}^{m\times n}$ 是一个具有 $m \geq n$ 的复矩阵，并设 $S_1 \in \mathbb{C}^{m\times m}$、 $S_2 \in \mathbb{C}^{n\times n}$ 为斜厄米矩阵。我们感兴趣的是研究该矩阵的奇异值

$$A + S_1A + AS_2.$$

注7.1. 假设 $A$ 是对角的（即仅有 $A$（$i$，$i$）$\neq 0$，对于 $i = 1$，.....，$n$），并且对角线上包含奇异值。在这种情况下， $E$ 的对角线必须为零，并且 $S_1$ 和 $S_2$ 可以用 $E$ 按如下方式表示:

$$S_1(i,j) = \frac{\sigma_j\, E(i,j) + \sigma_i\, E(j,i)}{\sigma_j^2 - \sigma_i^2}, \quad S_2(i,j) = -\frac{\sigma_j\, E(j,i) + \sigma_i\, E(i,j)}{\sigma_j^2 - \sigma_i^2}, \quad i \neq j, i \leq n, j \leq n.$$

而

$$S_1(i,j) = \frac{E(i,j)}{\sigma_j}, \qquad i > n,\ j \leq n.$$

$S_1$ 的右下块是自由的。

Let

$$A(t) = A + t(S_1A + AS_2)$$

并令 $\sigma_i(t)$ $u_i(t)$ $v_i(t)$ 为 $A(t)$ 的第 $i$ 个奇异值/左奇异向量/右奇异向量，即

$$A(t)v_i(t) = \sigma_i(t)u_i(t), \quad u_i(t)^*A(t) = \sigma_i(t)v_i(t)^*.$$

利用已知公式 [11], 可得

$$\begin{aligned}\frac{d\sigma_i(t)}{dt} &= \mathscr{R}(u_i(t)^*(S_1A + AS_2)v_i(t)) = \\ &= \mathscr{R}(u_i(t)^*S_1(\sigma_i(t)u_i(t) - t(S_1A + AS_2)v_i(t)) + (\sigma_i(t)v_i(t)^* - tu_i(t)^*(S_1A + AS_2))S_2v_i(t)) = \\ &= \mathscr{R}(-tu_i(t)^*S_1(S_1A + AS_2)v_i(t) - tu_i(t)^*(S_1A + AS_2)S_2v_i(t) = \\ &= \mathscr{R}(-tu_i(t)^*(S_1(S_1A + AS_2) + (S_1A + AS_2)S_2)v_i(t))).\end{aligned}$$

由此可得

$$|\sigma(1) - \sigma(0)| \le \frac{1}{2}\|S_1^2A + S_1AS_2 + S_2AS_1 + AS_2^2\|_2 \le \frac{1}{2}\|E\|(\|S_1\| + \|S_2\|).$$

对于特征值，使用整个 $S_1$ 和 $S_2$ 意味着包含所有差值$\sigma_j^2 - \sigma_i^2$。因此，如果某一对奇异值之间的间隙很小，该界可能会比较宽松。

利用先前关于特征值的界，可以推导出不同的界。假设矩阵 $A \in \mathbb{C}^{m\times n}$, $m \ge n$ 具有如下形式

$$A = \begin{bmatrix}\Sigma \\ 0\end{bmatrix}$$

其中 $\Sigma \in \mathbb{R}^{n\times n}$ 为对角矩阵，包含奇异值。Jordan-Wielandt 矩阵

$$M = \begin{bmatrix}0 & A \\ A^* & 0\end{bmatrix}$$

其特征值等于 $A$的奇异值，外加一些零。该矩阵可由以下矩阵对角化

$$Q = \frac{1}{\sqrt{2}}\begin{bmatrix}I_n & I_n & 0 \\ 0 & 0 & \sqrt{2}I_{m-n} \\ I_n & -I_n & 0\end{bmatrix}.$$

此外，如果将 $A$ 扰动为

$$A + E = A + S_1A + AS_2$$

矩阵 $M$ 被扰动为

$$M + \begin{bmatrix}0 & E \\ E^* & 0\end{bmatrix} = \begin{bmatrix}0 & A+E \\ (A+E)^* & 0\end{bmatrix}.$$

最后，如果从左侧将此矩阵乘以 $Q^*$，并从右侧乘以 $Q$，我们得到 t

$$Q^*\begin{bmatrix}0 & A+E \\ (A+E)^* & 0\end{bmatrix}Q = \begin{bmatrix}\Sigma + \frac{E_1+E_1^*}{2} & \frac{-E_1+E_1^*}{2} & \frac{1}{\sqrt{2}}E_2^* \\ \frac{E_1-E_1^*}{2} & -\Sigma - \frac{E_1+E_1^*}{2} & -\frac{1}{\sqrt{2}}E_2^* \\ \frac{1}{\sqrt{2}}E_2 & -\frac{1}{\sqrt{2}}E_2 & 0\end{bmatrix}$$

其中

$$E_1 = S_1(1:n, 1:n)\Sigma + \Sigma S_2, \quad E_2 = S_1(n+1:m, 1:n)\Sigma$$

分别是 $E$ 的第一行和最后一行。我们已将原问题化简为对角矩阵加非对角微扰的情形。

注意到这里出现了一个由零组成的分块，这表明，如果我们只扰动 $A$ 的最后几行（即由零组成的行），那么奇异值的扰动就不依赖于其他奇异值之间的间隙，而只取决于它与 0 之间的间隙，也就是该奇异值本身。可以通过以下方式验证这一点。考虑

$$A + E = \begin{bmatrix} \Sigma \\ E_2 \end{bmatrix}$$

then

$$(A+E)^*(A+E) = \Sigma^2 + E_2^* E_2.$$

若将 $\tilde{\sigma}_i$ 称为 $A+E$ 的第 $i$ 个奇异值，根据 Bauer-Fike 定理，我们有

$$|\tilde{\sigma}_i^2 - \sigma_i^2| \leq \|E_2^* E_2\| = \|E_2\|^2.$$

这意味着

$$|\tilde{\sigma}_i - \sigma_i| \lesssim \frac{\|E_2\|^2}{\sigma_i} = \frac{\|E\|^2}{\sigma_i}.$$

我们可以在这个特定情形下计算 $B$，即我们在第3节中用到的矩阵。由于 $M$ 是对称矩阵， $B$ 将是反对称矩阵，并且我们可以将其分块为

$$B = \begin{bmatrix} B_1 & B_2 & B_3 \\ B_2^* & B_1 & B_3 \\ B_3^* & B_3^* & 0 \end{bmatrix}$$

其中

$$B_1(i,j) = \frac{1}{2}\frac{E(i,j) + \bar{E}(j,i)}{\sigma_i - \sigma_j} \quad B_2(i,j) = \frac{1}{2}\frac{\bar{E}(j,i) - E(i,j)}{\sigma_i + \sigma_j} \quad B_3(i,j) = \frac{1}{\sqrt{2}}\frac{E(n+1+i,j)}{\sigma_i}.$$

这些表达式与渐近展开一致。如果我们假设 $E$ 为实数，且在切空间中有 $E_2 = 0$，那么渐近展开为

$$\sigma_i(D+E) = \sigma_i + \frac{1}{2\sigma_i}\left[\sum_{j\neq i} E(j,i)^2 + \sum_{j\neq i} \frac{(\sigma_i E(j,i) + \sigma_j E(i,j))^2}{\sigma_i^2 - \sigma_j^2}\right] + \mathcal{O}(\|E\|^3). \tag{7.1}$$

如果 $E_1$ 是对称的，公式（7.1）等于

$$\frac{1}{2\sigma_i}\left[\sum_{j\neq i} \frac{(\sigma_j + \sigma_i + \sigma_i - \sigma_j)\, E(i,j)^2}{\sigma_i - \sigma_j}\right] = \sum_{j\neq i} \frac{E(i,j)^2}{\sigma_i - \sigma_j}.$$

如果 $E_1$ 是反对称的，公式（7.1）等于

$$\frac{1}{2\sigma_i}\left[\sum_{j\neq i} \frac{(2\sigma_i^2 - 2\sigma_i\sigma_j) E(i,j)^2}{\sigma_i^2 - \sigma_j^2}\right] = \sum_{j\neq i} \frac{E(i,j)^2}{\sigma_i + \sigma_j}$$

**8. 若尔当块**。考虑幂零乔丹块

$$J = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & \ddots & \ddots \end{bmatrix} \in \mathbb{C}^{n\times n}$$

以及形如 $E$ 的扰动矩阵

$$E = JB - BJ$$

其中 $B \in \mathbb{C}^{n\times n}$。注意， $E$ 满足每条次对角线（包括主对角线）上的元素之和为 0。特别地，这排除了位置 $(n,\ 1)$ 处的元素非零的情形；这种情形通常会得到揭示一般 $\mathcal{O}(\varepsilon^{1/n})$ 扰动的结果，而这是 $n\times n$矩阵的特征值所能表现出的最敏感行为。给定满足次对角线元素和为零这一条件的矩阵 $E$，矩阵 $B$ 在一个上三角托普利茨矩阵（同样有 $n^2 - n$ 个自由度）的意义下是唯一的。

对于 Jordan 块情形，我们只给出特征值的渐近展开。

THEOREM 8.1. 设 $J \in \mathbb{C}^{n\times n}$ 为维数 $n\times n$ 、特征值均为0的 *Jordan* 块。令 $E \in \mathbb{C}^{n\times n}$ 等于

$$E = JB - BJ$$

对于某个 $B \in \mathbb{C}^{n\times n}$。等价地， $E$ 的所有次对角线之和均为零（包括 $g$ 主对角线）。如果

$$\sum_{r=1}^{n-1}(-1)^r E(n, r+1)E(r, 1) \neq 0$$

那么 $J + \varepsilon E$ 的特征值扰动的阶为 $\mathcal{O}$（$\varepsilon^{2/n}$）。

证明。 定义 $p_\varepsilon$（$t$）为 的特征多项式。根据 [5, 定理4] ，应用于我们的情形，在将 $p_\varepsilon$（$t$按 $\varepsilon$展开时， $\varepsilon$中的线性项消失，即

$$p_\varepsilon(t) = t^n + \varepsilon^2 p^{(2)}(t) + \mathcal{O}(\varepsilon^3).$$

为了确定特征值扰动的大小，我们必须判断多项式 $p_\varepsilon(t)$ 中的项 $\varepsilon^2 t^0$ 是否为零。我们知道，特征多项式中 $t^0$ 的系数等于

$$det(J + \varepsilon E).$$

若将行列式改写为

$$det(J + \varepsilon E) = \sum_{\sigma\in S_m} sgn(\sigma)\prod_{i=1}^{n}(J(i, \sigma(i)) + \varepsilon E(i, \sigma(i)))$$

我们可以看出，为了使 $\varepsilon^2$ 出现，我们需要

$$J(i, \sigma(i)) = 1$$

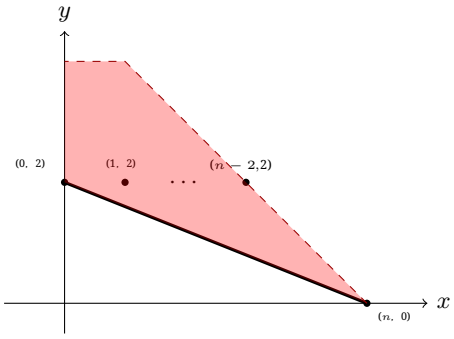


FIG. 5 *Netwon diagram in the generic case where the coefficient of* $\varepsilon^2 t^0$ *is not zero. The slope . of the solid line indicates the power of* $\varepsilon$ *in the perturbation of the eigenvalues* $(2/n)$.

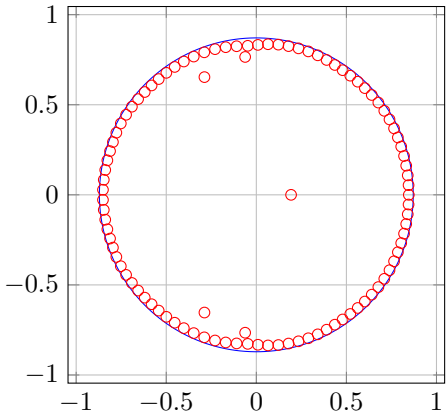


FIG. 6。 *Perturbed eigenvalues for a nilpotent Jordan block of size* $100 \times 100$ *and a perturbatio n E in the tangent with norm* $\|E\| = 10^{-3}$. *The circle has radius* $\|E\|^{2/100}$.

至少出现 $n-2$ 次。 $\sigma$ 唯一可能的情况是排列，其中存在一个指标 $r$ ，使得

$$\sigma(r) = 1$$

$$\sigma(n) = r + 1$$

而对于所有其他指标， $\sigma(k) = k+1$ 。最后， $\varepsilon^2 t^0$ 的系数为

$$\sum_{r=1}^{n-1} (-1)^r E(n, r+1) E(r, 1). \tag{8.1}$$

最后，我们使用牛顿图技术 [22], [4, §8.3]。在我们的情形中，该图见图 5，从中可以看出，导致抵消的第一个指数是 $2/n$ 若方程（8.1）不为零。

□

请注意， $2/n$ 是牛顿图中直线可能具有的最低系数，因为点（$n$，0）始终位于图中。因此， $\mathcal{O}(\varepsilon^{2/n})$ 是特征值可能产生的最大扰动。在图 6 中，我们绘出了幂零若尔当块的扰动特征值。

**9. 结论**。我们分析了矩阵 $A$ 在等谱矩阵切空间中的加性扰动 $E$ 下的特征值和特征向量扰动，即 $E = AB - BA$。我们说明了如何用 $E$ 和 $B$简洁地表示特征向量和特征值的渐近展开，并阐明了 $B$ 与特征向量的一阶扰动密切相关。除了渐近展开之外，我们还给出了特征向量近似的高阶界。此外，我们将 [20] 中提出的特征值二次界推广到了这一情形；与埃尔米特分块非对角扰动相比，这一情形更为一般。我们将分析拓展到了分块对角情形，给出了扰动特征子空间的界，并在厄米假设下进一步改进了这些界。最后，我们将分析拓展到了若尔当块的奇异向量、奇异值和特征值。

## 参考文献

[1] A. E. Ashari and B. Labibi, *Application of matrix perturbation theory in robust control of large-scale systems* 48 2012 1868 1873, Automatica J. IFAC, ( ), 第 – 页。[2] F. L. Bauer and C. T. Fike, *Norms and exclusion theorems*, Numer. Math., 2 (1960), 第 137–141 页。[3] R. Bhatia, *Matrix Analysis*, 数学研究生教材, Springer New York, 2012年。[4] E. Brieskorn and H. Knörrer, *Plane Algebraic Curves*, Modern Birkh¨auser Classics, Springer Basel, 2012年。[5] J. V. Burke and M. L. Overton, *Stable perturbations of nonsymmetric matrices*, Linear Algebra Appl., 171 (1992), 第 249–273 页。[6] F. Chatelin, *Spectral approximation of linear operators*, 应用数学经典著作第 65 卷, 工业与应用数学学会（SIAM）, 费城, PA, 2011年。[7] M. T. Chu, *Linear algebra algorithms as dynamical systems*, Acta Numer., 17 (2008), 第 1–86 页。[8] M. T. Chu and K. R. Driessel, *The projected gradient method for least squares matrix approximations with spectral constraints* 27 1990 1050 1060, SIAM 数值分析杂志, ( ), 第 – 页。[9] K. Dasgupta, M. M. Sheikh-Jabbari, and M. Van Raamsdonk, *Matrix perturbation theory for M-theory on a pp-wave* 2002 56, 52, J. High Energy Phys., ( ), 第 号 .。[10] C. Davis and W. M. Kahan, *The rotation of eigenvectors by a perturbation. III*, SIAM 数值分析杂志, 7 (1970), 第 1–46 页。[11] M. B. Giles, *An extended collection of matrix derivative results for forward and reverse mode algorithmic differentiation* 2008, 技术报告, 牛津大学计算机实验室, 。[12] A. Greenbaum, R.-C. Li, and M. L. Overton, *First-order perturbation theory for eigenvalues and eigenvectors* 62 2020 463 482, SIAM 评论, ( ), 第 – 页。[13] I.-B. Haas and Y. Nakatsukasa, *Sharp error bounds for approximate eigenvalues and singular values from subspace methods*, SIAM 矩阵分析与应用杂志, 47 (2026), 第 308–329 页。[14] B. C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, 数学研究生教材第 222 卷, Springer, 第 2 版, 2015年。[15] M. Karow and D. Kressner, *On a perturbation bound for invariant subspaces of matrices*, SIAM 矩阵分析与应用杂志, 35 (2014), 第 599–618 页。[16] T. Kato, *Perturbation Theory for Linear Operators*, 第 132 卷, Springer, 1995年。[17] L. Landau and E. Lifshits, *Quantum Mechanics, Non-relativistic Theory*, A-W 高等物理学系列, Pergamon Press, 1958年。[18] P. D. Lax, *Integrals of nonlinear equations of evolution and solitary waves*, Comm. Pure Appl. Math., 21 (1968), 第 467–490 页。[19] L. Lazzarino, H. Al Daas, and Y. Nakatsukasa, *Matrix perturbation analysis of methods for extracting singular values from approximate singular subspaces*, SIAM 矩阵分析与应用杂志, 46 (2025), 第 2614–2634 页。[20] C.-K. Li and R.-C. Li, *A note on eigenvalues of perturbed Hermitian matrices*, Linear Algebra Appl., 395 (2005), 第 183–190 页。[21] R.-C. Li, Y. Nakatsukasa, N. Truhar, and W.-g. Wang, *Perturbation of multiple eigenvalues of Hermitian matrices* 437 2012 202 213, Linear Algebra Appl., ( ), 第 – 页。[22] Y. Ma and A. Edelman, *Nongeneric eigenvalue perturbations of Jordan blocks*, Linear Algebra Appl., 273 (1998), 第 45–63 页。

[23] R. Mathias, *Quadratic residual bounds for the Hermitian eigenvalue problem*, SIAM 矩阵分析与应用杂志, 19 (1998), 第 541–550 页。[24] C. D. Meyer and G. W. Stewart, *Derivatives and perturbations of eigenvectors*, SIAM 数值分析杂志, 25 (1988), 第 679–691 页。[25] Y. Nakatsukasa, *Eigenvalue perturbation bounds for Hermitian block tridiagonal matrices*, 应用数值数学, 62 (2012), 第 67–78 页。[26] J. J. Sakurai and J. Napolitano, *Modern quantum mechanics*, 剑桥大学出版社，第三版，2021年。[27] G. W. Stewart, *Matrix Algorithms Volume II: Eigensystems*, SIAM，费城，2001年。[28] G. W. Stewart and J. G. Sun, *Matrix perturbation theory*, 计算机科学与科学计算，学术出版社，波士顿，MA，1990年。[29] D. S. Watkins, *Isospectral flows*, SIAM 评论, 26 (1984), 第 379–391 页。[30] H. Weyl, *Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung)*, Math. Ann., 71 (1912), 第 441–479 页。