arXiv:2601.15789v1 [math.OC] 22 Jan 2026

互补特征值的定位

ANTONIO SASAKI[1], SOPHIE DEMASSEY[1], VALENTINA SESSA[1]

[1]数学中心（CMA），´巴黎矿业学院（Mines Paris），巴黎文理研究大学（PSL），索菲亚·安蒂波利斯，法国

谨以此文献给 Alfredo N. Iusem 教授，值此他 75 岁生日之际。

摘要。设 A、B 为对称的 n ×n 实矩阵，其中 B 正定且严格对角占优。我们为 (A, B) 的互补特征值推导出两个定位集，其中更紧的一个还额外假设 A 共正。这将 He-Liu-Shen 定位集推广到 B 非单位矩阵的情形。此外，我们将根据这些新定位集得到的可计算界与经典广义特征值的极值进行比较。

关键词。互补问题；Gershgorin 型定位集。
2020 年数学主题分类。90C33，15A18。

1. 引言

设 A、B 为对称的 n ×n 实矩阵，且 B 正定。对称特征值互补问题（EiCP）是求解满足以下条件的 x ∈ $\mathbb{R}^n \setminus \{\mathbf{0}\}$ 和 λ ∈ ℝ：

$$(\mathbf{A} - \lambda \mathbf{B})\mathbf{x} \geq \mathbf{0} \tag{1.1a}$$

$$\mathbf{x} \geq \mathbf{0} \tag{1.1b}$$

$$\mathbf{x}^\top (\mathbf{A} - \lambda \mathbf{B})\mathbf{x} = 0 \tag{1.1c}$$

$$\mathbf{e}^\top \mathbf{x} = 1, \tag{1.1d}$$

其中 e :=(1, . . . , 1)$^\top$ ∈ $\mathbb{R}^n$为全 1 向量，不等式均按分量理解。对于满足 (1.1) 的一对 (x λ)， λ ∈ ℝ 是互补特征值，x ∈ $\mathbb{R}^n$ 是互补特征向量。正交条件 (1.1c) 与非负性要求 (1.1a) 和 (1.1b) 一起推出，对于所有 i = 1，...，n，$x_i = 0$ 或 $w_i = 0$，其中 w =(A −λ B) x。

Seeger 在 [24] 针对**B** 等于单位矩阵的情形引入了这一问题。在此背景下，这一问题通常称为Pareto 特征值问题，因为此时互补谱对应于单个矩阵**A** 在非负锥（也称为 Pareto 锥）下的情形。Queiroz 等人[22] 随后将这一表述推广到一般的对称矩阵对 (**A**,**B**)，其动机是分析具有摩擦接触的机械系统的稳定性。EiCP 的实际重要性还体现在其在共正矩阵分析中的应用[12], 、Fucik 曲线的渐近研究 [13], 、图论 [9, 25],以及由互补条件支配的线性动力系统中[24]。

近年来，人们研究了对称 EiCP 的理论和数值性质，以及各种推广 [4, 5, 7, 15, 16]。EiCP 的数值求解

电子邮箱： {antonio.sasaki, sophie.demassey, valentina.sessa}@minesparis.psl.eu 1

通常基于将其重新表述为更易处理的问题，并据此设计专用算法。例如，在 [17], 中，EiCP 被表述为标准单纯形上的差凸规划，随后采用交替方向乘子法求解，该方法全局收敛到对称 EiCP 的一个解。在 [10], 中，对称 EiCP 被重新表述为求解一个驻点问题，即最小化一个凸函数与一个光滑但非凸函数之和。文中提出了一种算法，通过将非凸项线性化，依次求解由此得到的凸二次问题；在温和的假设下，该算法被证明收敛到 EiCP 的一个解。另一种重新表述方式是将其作为分式规划，即在标准单纯形上最小化两个二次函数之比，也就是广义瑞利商 [18, 22]。在 [19], 中，通过对当前迭代点处目标函数的分子进行线性化，再采用高效实现的 Dinkelbach 方法 [23] 计算一个驻点。在 [2], 中，EiCP 被表述为一个基于互补函数的等价非光滑方程组 [6] ，并采用半光滑牛顿法求解。在 [1], 中，EiCP 被表述为一个非线性方程组，并采用内点法的两个版本求解。对于中等规模的实例（$n \approx 100$），这两种方法的数值效率似乎高于平滑法，但对于较小规模的实例则并非如此。

EiCP (1.1) 扩展了广义特征值问题 [14], ，其定义为 $Ax = \lambda Bx$，即 (1.1a) 取等号成立，且去掉 (1.1b) 和 (1.1c)。注意，如果 $(x\ \lambda)$ 是 (1.1) 的解，且 $x_i > 0$，对所有 $i = 1, \ldots n$，正交约束 (1.1c) 都会迫使 $Ax - \lambda Bx = \mathbf{0}$。更一般地，若索引集 $S \subset \{1, \ldots, n\}$ 上的子向量 $x_S > \mathbf{0}$，则 (x_S, λ) 是子矩阵 (A_{SS}, B_{SS}) [8] 的广义特征对。这种子矩阵视角提供了一条构造性路径：通过扫描主子矩阵并计算经典广义特征对，可以枚举互补特征值 [24]。只有在规模非常小的实例中（$n \leq 15$），这一过程才具有计算可行性。与至多有 n 个解的广义特征值问题不同，系统 (1.1) 可能随着 n 增大而呈现指数级数量的互补特征对，至多为 $2^n - 1$，因此穷举枚举会变得难以承受 [8]。

最近，He 等人 [11] 解决了 Seeger [24], 提出的一个开放性问题，该问题涉及当 B 为单位矩阵时互补特征值的位置。作者推导出了可计算的定位集，即 $\mathbb{R}$中的区间；这些区间仅取决于 A 的行和，其方式与普通特征值的格尔圆盘定理 [14] 相关。这些结果为λ提供了明确的代数界。

本文针对对称 EiCP（1.1）且 **B** 严格对角占优的情形，探讨同一问题。值得一提的是，在此背景下，将 **B** 从单位矩阵推广为任意 **B**正定且严格对角占优的矩阵，是一项贡献。尽管这看似有所限制，但 EiCP 中通常假设 **B** 正定，以确保解的存在 [8, 10, 17, 19, 22], ；而对角占优假设也很常见，并经常用于通过 Gershgorin 型集合对各种特征值类进行定位的理论中 [20, 21, 26]。我们扩展 [11], 中的方法，在额外假设后一种情形下 **A** 共正的前提下，提出基于单行或行对的定位集。当 **B** 为单位矩阵时，我们的公式便化为 [11], 中的公式，从而恢复 Pareto 框架。我们还证明，双行细化所得的定位集比单行分析所得的定位集更紧，并且这些定位集给出的界与极端（最大和最小）广义特征值之间不存在支配关系。

本文结构安排如下：首先介绍必要的符号和预备概念。在第 2 节中，在 A 共正这一附加假设下，推导单行定位集 K_1 和双行定位集 K_2。随后，我们在第 3 节证明 K_2 包含于 K_1中。第 4 节从这些集合中提取可计算的下界和上界。最后，第 5 节将我们的定位集与标准广义特征值谱进行比较；第 6 节和第 7 节则给出总结性评论并讨论可能的扩展。

记号与预备知识

我们用 $[n]$ 表示 1 到 n 之间的自然数集合，即 $[n]:=\{1,\ldots,n\}$。令 $\mathbb{R}^n$、$\mathbb{R}^n_+$ 和 $\mathbb{S}^n$ 分别表示 n 维欧几里得空间、非负正交锥以及 $n\times n$ 阶实对称矩阵空间。

对于向量 $x\in\mathbb{R}^n$，其第 j 个分量记为 x_j，其中 $j\in[n]$。矩阵 $M\in\mathbb{R}^{n\times n}$ 的第 (i, j) 个元素记为 m_{ij}，$M_i^\top$表示 M 的第 i 行。

$\mathbb{R}^n$ 中的第 j 个标准基向量记为 e_j。我们用 $e\in\mathbb{R}^n$ 表示 n 维全 1 向量，用 $E=ee^\top\in\mathbb{S}^n$ 表示所有元素均为 1 的矩阵。

定义 1.1 (对角占优)。设 $M\in\mathbb{R}^{n\times n}$。若 M 是对角占优 (DD)，则称 $|m_{ii}|\geq\sum_{j\neq i}|m_{ij}|$ 对每个 $i\in[n]$均成立。若严格对角占优 (SDD)，则$|m_{ii}|>\sum_{j\neq i}|m_{ij}|$ 对每个 $i\in[n]$。

定义 1.2（正性）。设 $A\in\mathbb{S}^n$。若对所有 $x\in\mathbb{R}^n$，都有 $x^\top Ax\geq 0$，则称 A 为半正定（PSD）。若对所有满足 $x\neq\mathbf{0}$ 的 $x\in\mathbb{R}^n$，都有 $x^\top Ax>0$，则称 A 为正定（PD）。

定义 1.3（余正性）。设 $A\in\mathbb{S}^n$。若对所有满足 $x\geq\mathbf{0}$ 的 $x\in\mathbb{R}^n$，都有 $x^\top Ax\geq 0$，则称 A 为共正。若对所有满足 $x>\mathbf{0}$ 的 $x\in\mathbb{R}^n$，都有 $x^\top Ax>0$，则称 A 为严格共正。

显然，任何非负矩阵或正定矩阵都是共正的。众所周知，对称的严格对角占优矩阵若对角元素为正，则为正定矩阵（参见 [14]中的定理 6.1.10）。

性质 1.4。EiCP(A, B) 有解 $(x\lambda)$ 当且仅当 EiCP$(A+\mu B,B)$ 有解 $(x\ \lambda+\mu)$，其中 $\mu\geq 0$。

注意，由于 **B** 是正定的，因此对于某个 $\mu>0$ 矩阵 $\mathbf{A}+\mu\mathbf{B}$ 也为正定 [14]。因此，不失一般性，我们可以假设 EiCP（**A** 是正定（或至少是共正）矩阵（1.1）。

性质 1.5。设 A、$B\in\mathbb{S}^n$ 和 B 均为正定矩阵。若 A 是共正矩阵，则 EiCP(A, B) 的任意解 $(x\lambda)$ 中，互补特征值 λ 均非负。

一个解 满足正交条件 $(\mathbf{x}\ \lambda)$ 的 EiCP 解满足正交条件（1.1c），则 $\mathbf{x}\geq\mathbf{0}$ 一个 d

$$\lambda=\frac{\mathbf{x}^\top\mathbf{A}\mathbf{x}}{\mathbf{x}^\top\mathbf{B}\mathbf{x}} \tag{1.2}$$

是广义瑞利商 $(\mathbf{A},\mathbf{B})$ [22]。若 **A** 是共正的，且 **B** 是正定的，则根据定义 1.2和 1.3，可得性质 1.5中的结论。

2. 互补特征值问题解的定位集

若 **A** 和 **B** 是对角矩阵，则很容易确定互补特征值的位置。事实上，它们是实空间 $_{ii}/\ _{ii}$中的点 $i \in [n],$。与复空间中经典特征值问题的格尔圆盘定理 [24]类似，我们考虑当矩阵 [14] 非对角时，相对于点 a_{ii}/b_{ii} 来定位互补特征值。第一个定位集通过逐行考察非对角元，评估矩阵距离对角矩阵有多远。**A** 和 **B**。

定义 2.1. 对任意矩阵 $\mathrm{M} \in \mathbb{R}^{n\times n}$ 和行 $\mathrm{i} \in [\mathrm{n}],$，正、负非对角元的行和及相应的对角偏移量记为：

$$r_i^+(\mathbf{M}) := \sum_{j\neq i} \max\{m_{ij}, 0\}, \quad r_i^-(\mathbf{M}) := -\sum_{j\neq i} \min\{m_{ij}, 0\}.$$

$$m_i^+ := m_{ii} + r_i^+(\mathbf{M}), \quad m_i^- := m_{ii} - r_i^-(\mathbf{M}).$$

注意，对于正定矩阵，$\mathrm{m}_i^+ > 0$。然而，对于这些类别的矩阵，m_i^- 的符号无法先验确定。

定理 2.2（单行定位）。设 $\mathrm{A,B} \in \mathbb{S}_n$，并假设 B 正定且严格对角占优。设 $\lambda \in \mathbb{R}$ 是 (A,B) 的互补特征值，则

$$\lambda \in \bigcup_{i\in[n]} \left[\min\left(\frac{a_i^-}{b_i^-}, \frac{a_i^-}{b_i^+}\right), \max\left(\frac{a_i^+}{b_i^-}, \frac{a_i^+}{b_i^+}\right)\right] =: K_1.$$

为证明该定理，我们将用到以下结果。

引理 2.3。设 $\mathrm{M} \in \mathbb{R}^{n\times n}$，$\mathrm{x} \in \mathbb{R}^n$，其中 $\mathrm{x} \geq \mathbf{0}$，且任意 $\mathrm{p} \in \arg\max^{i\in[n]} \mathrm{x}_i$，则：

(i) $\mathrm{m}_p^- \mathrm{x}_p \leq \mathrm{M}_p^\top \mathrm{x} \leq \mathrm{m}_p^+ \mathrm{x}_p$(ii) 若 M 是严格对角占优且 m $_{ii} \geq 0$，则 $\mathrm{m}_i^+ \geq \mathrm{m}_i^- > 0.$

证明。由于 $\mathrm{x}_p \geq \mathrm{x}_j \geq 0$，对于 $\mathrm{j} \neq \mathrm{p}$，

$$-r_p^-(\mathbf{M})x_p \leq \sum_{\substack{j\neq p \\ m_{pj}<0}} m_{pj}x_j \leq \sum_{j\neq p} m_{pj}x_j \leq \sum_{\substack{j\neq p \\ m_{pj}>0}} m_{pj}x_j \leq r_p^+(\mathbf{M})x_p.$$

相加$m_{pp}x_p$可得(i)。若$\mathbf{M}$是严格对角占优且$m_{ii} \geq 0$，则$m_{ii} = |m_{ii}| > \sum |m_{ij}| = r_i^+(\mathbf{M})$ +j≠i

$r_i^-(\mathbf{M}) \geq r_i^-(\mathbf{M}) \geq 0$, and $m_i^+ \geq m_i^- = m_{ii} - r_i^-(\mathbf{M}) > 0.$

□

定理 2.2 的证明。设 (x λ) 是 EiCP(A, B) 的一个解。则 $\mathrm{x} \geq \mathbf{0}$，且 $\mathrm{e}^\top \mathrm{x} = 1$，并且存在 p，使得 $\mathrm{x}_p = \max^{i\in[n]} \mathrm{x}_i > 0$。令 $\mathrm{w} := (\mathrm{A} - \lambda \mathrm{B})\mathrm{x} \geq 0$ 且 $\mathrm{x}^\top \mathrm{w} = 0$，所有项 $\mathrm{x}_i \mathrm{w}_i$，对于 $\mathrm{i} \in [\mathrm{n}],$均为非负数且和为零。因此，$\mathrm{x}_p \mathrm{w}_p = 0$，从而 $\mathrm{w}_p = 0$ 且 $\lambda = \mathrm{A}_p^\top \mathrm{x}/\mathrm{B}_p^\top \mathrm{x}$。

引理 2.3 (i) 给出了操作数的界：$\mathrm{a}_p^- \mathrm{x}_p \leq \mathrm{A}_p^\top \mathrm{x} \leq \mathrm{a}_p^+ \mathrm{x}_p$ 以及 $\mathrm{b}_p^- \mathrm{x}_p \leq \mathrm{B}_p^\top \mathrm{x} \leq \mathrm{b}_p^+ \mathrm{x}_p$，由于 B 是严格对角占优且正定，因此分母的界均为正，即 $\mathrm{b}_{pp} = \mathrm{e}_p^\top \mathrm{B}\mathrm{e}_p > 0$。因此，$\mathrm{a}_p^- \mathrm{x}_p/\mathrm{B}_p^\top \mathrm{x} \leq \lambda \leq \mathrm{a}_p^+ \mathrm{x}_p/\mathrm{B}_p^\top \mathrm{x}$。

映射 $u \to v/u$ 在 $\mathbb{R}_+$ 上是单调的，因此 $a_p^- x_p / B_p^\top x \geq \min(a_p^- x_p / b_p^- x_p,\ a_p^- x_p / b_p^+ x_p) = \min(a_p^- / b_p^-,\ a_p^- / b_p^+)$，同理，$a_p^+ x_p / B_p^\top x \leq \max(a_p^+ / b_p^-,\ a_p^+ / b_p^+)$ as $x_p \leq 1$。

定位集 K_1 表示为所有可能的 $p \in [n]$, 取值对应的此类区间的并集；由于对所有 $i \in [n]$均有 $b_i^+ \geq b_i^- > 0$，因此所有区间都有明确定义。

□

推论 2.4. 设 $A, B \in \mathbb{S}_n$。假设 B 正定且严格对角占优，A 共正。设 $\lambda \in \mathbb{R}$ 是 (A,B) 的一个互补特征值，则

$$\lambda \in \bigcup_{i \in [n]} \left[\max\left(0, \frac{a_i^-}{b_i^+}\right), \frac{a_i^+}{b_i^-} \right] =: K_1'.$$

证明. 矩阵 A 的余正性使我们能够通过考察以下性质来细化定位集 K_1 。

(i) 性质 1.5 为 λ给出了显然的零下界。 K_1 中各区间的下界取决于 a_i^-的符号。当 a_i^-为负时，相应的下界也为负，因此被零所支配。相反，当 $a_i^- > 0$时，可以得到更紧的下界。由于仅凭矩阵的余正性无法确定 a_i^-的符号，我们重新定义 K_1'中各区间的下界。

(ii) 由于 $a_i^+ \geq 0$ 且 $b_i^+ \geq b_i^- > 0$ K_1 中每个区间的上界就是 a_i^+ / b_i^-.

□

现在我们引入另一种定位集，它考虑矩阵i、$j \in [n]$ 的**A** 和**B**的行对。与之前一样，我们根据非对角元计算相关量。此外，为了推导这个新的定位集，我们额外假设**A** 是共正的；根据性质1.5 ，有 $\lambda \geq 0$。注意，对**A** 的余正性假设并非限制性条件，因为可以通过平移使其成立，如性质1.4所示。

我们从以下定义开始。

定义 2.5。对于任意 $M \in \mathbb{R}^{n \times n}$，以及两行 $i, j \in [n]$, , 记为:

$$m_{ij}^+ := m_{ii} m_{jj} - r_i^+(\mathbf{M}) r_j^+(\mathbf{M})$$
$$m_{ij}^- := m_{ii} m_{jj} - r_i^-(\mathbf{M}) r_j^-(\mathbf{M}).$$

定义 2.6。对于任意一对 $A,B \in \mathbb{R}^{n \times n}$ 以及两行 $i, j \in [n]$, , 记为:

$$s_{ij}^+(\mathbf{A},\mathbf{B}) := a_{ii} b_{jj} + a_{jj} b_{ii} + r_i^+(\mathbf{A}) r_j^-(\mathbf{B}) + r_j^+(\mathbf{A}) r_i^-(\mathbf{B})$$
$$s_{ij}^-(\mathbf{A},\mathbf{B}) := a_{ii} b_{jj} + a_{jj} b_{ii} + r_i^-(\mathbf{A}) r_j^+(\mathbf{B}) + r_j^-(\mathbf{A}) r_i^+(\mathbf{B})$$
$$P_{ij}^{\mathrm{up}}(y) := b_{ij}^- y^2 - s_{ij}^+(\mathbf{A},\mathbf{B}) y + a_{ij}^+, \quad P_{ij}^{\mathrm{low}}(y) := b_{ij}^+ y^2 - s_{ij}^-(\mathbf{A},\mathbf{B}) y + a_{ij}^- \quad \forall y \in \mathbb{R}.$$

定理 2.7（双行局部化）。设 $A, B \in \mathbb{S}_n$。假设 B 正定且严格对角占优，A 共正。设 $\lambda \in \mathbb{R}$ 是 (A, B) 的互补特征值，则

$$\lambda \in \bigcup_{\substack{i,j \in [n] \\ i \neq j}} \left[\max\{0, \min\{y \in \mathbb{R} \mid P_{ij}^{\mathrm{low}}(y) = 0\}\},\ \max\{y \in \mathbb{R} \mid P_{ij}^{\mathrm{up}}(y) = 0\} \right] =: K_2.$$

定理 2.7 的证明使用了以下结果。

引理 2。 8. 设 $\mathrm{M} \in \mathbb{S}_n$ 和 $\mathrm{x} \in \mathbb{R}^n$，其中 $\mathrm{x} \geq \mathbf{0}$ 且 $\mathrm{e}^\top \mathrm{x} = 1$，以及两个最大的元素 $_{ts}$

$$p \in \operatorname{argmax}_{i\in[n]} x_i \ \textit{and}\ q \in \operatorname{argmax}_{i\in[n]\setminus\{p\}} x_i$$

那么，对于 $\mathrm{i} = \mathrm{p}$ 和 $\mathrm{i} = \mathrm{q}$，以下界成立：

$$-r_i^-(\mathbf{M})x_p x_q \ \leq \sum_{j\neq i} m_{ij}x_i x_j \ \leq \ r_i^+(\mathbf{M})x_p x_q,$$

证明。由于对于 $\mathrm{j} \notin \{\mathrm{p}, \mathrm{q}\}$，$\mathrm{x} \geq \mathbf{0} \text{and } \mathrm{x}_p \geq \mathrm{x}_q \geq \mathrm{x}_j$，因此对于 p 行有

$$\sum_{j\neq p} m_{pj}x_p x_j = \sum_{\substack{j\neq p\\ m_{pj}>0}} m_{pj}x_p x_j + \sum_{\substack{j\neq p\\ m_{pj}<0}} m_{pj}x_p x_j \ \leq \ \sum_{\substack{j\neq p\\ m_{pj}>0}} m_{pj}x_p x_q = r_p^+(\mathbf{M})x_p x_q,$$

并且，对称地，

$$\sum_{j\neq p} m_{pj}x_p x_j \ \geq \ \sum_{\substack{j\neq p\\ m_{pj}<0}} m_{pj}x_p x_j \ \geq \ -\sum_{\substack{j\neq p\\ m_{pj}<0}} |m_{pj}|x_p x_q = -r_p^-(\mathbf{M})x_p x_q.$$

对于 $\mathrm{i} = \mathrm{q}$ 的情况，我们利用对于所有 $\mathrm{j} \neq \mathrm{q}$，$\mathrm{x}_j \leq \mathrm{x}_p$，于是

$$\sum_{j\neq q} m_{qj}x_q x_j \leq \sum_{\substack{j\neq q\\ m_{qj}>0}} m_{qj}x_q x_p = r_q^+(\mathbf{M})x_p x_q, \quad \sum_{j\neq q} m_{qj}x_q x_j \geq \sum_{\substack{j\neq q\\ m_{qj}<0}} m_{qj}x_q x_p = -r_q^-(\mathbf{M})x_p x_q.$$

□

引理 2.9。 二次函数 $\mathrm{P}_i^{\mathrm{low}}{}_j$ 和 $\mathrm{P}^{\mathrm{up}}{}_{ij}$ 在定义 2.6 中给出，对所有 $\mathrm{y} \in \mathbb{R}$,

$$P_{ij}^{\mathrm{up}}(y) = (yb_{ii} - a_{ii})(yb_{jj} - a_{jj}) - (r_i^+(\mathbf{A}) + y r_i^-(\mathbf{B}))\,(r_j^+(\mathbf{A}) + y r_j^-(\mathbf{B})), \tag{2.1}$$

$$P_{ij}^{\mathrm{low}}(y) = (yb_{ii} - a_{ii})(yb_{jj} - a_{jj}) - (r_i^-(\mathbf{A}) + y r_i^+(\mathbf{B}))\,(r_j^-(\mathbf{A}) + y r_j^+(\mathbf{B})). \tag{2.2}$$

*Pro*证明。 $\mathrm{P}^{\mathrm{up}}{}_{ij}(y)$ 的定义可通过展开表达式（2.1 ）：

$$\begin{aligned}
&\left(b_{ii}b_{jj} - r_i^-(\mathbf{B})r_j^-(\mathbf{B})\right)y^2 + \left(a_{ii}a_{jj} - r_i^+(\mathbf{A})r_j^+(\mathbf{A})\right)\\
&\quad - \left(a_{ii}b_{jj} + a_{jj}b_{ii} + r_i^+(\mathbf{A})r_j^-(\mathbf{B}) + r_j^+(\mathbf{A})r_i^-(\mathbf{B})\right)y\\
&= b_{ij}^- y^2 - s_{ij}^+(\mathbf{A},\mathbf{B})\,y + a_{ij}^+ = P_{ij}^{\mathrm{up}}(y)
\end{aligned}$$

注意，由于定义 2.6中的对称性，$\mathrm{P}_i^{\mathrm{low}}{}_j(y)$ 在（2.2）中由 $P^{\mathrm{up}}{}_{ij}(y)$ 在（2.1）中得到，只需将项 r^+ 替换为r^- ，反之亦然。

□

引理 2.10。在定理 2.7 的假设下，二次函数 $\mathrm{P}_i^{\mathrm{low}}\mathrm{j}$ 和 $\mathrm{P}_{\mathrm{up}}\mathrm{i\,j}$ 严格凸且具有实根。因此，对于任意 $\mathrm{y} \in \mathbb{R}$，若 $\mathrm{P}^{\mathrm{up}}\,\mathrm{ij}\,(\mathrm{y}) \leq 0$，则 $\mathrm{y} \leq \max\{\mathrm{y} \in \mathbb{R} \mid \mathrm{P}^{\mathrm{up}}\mathrm{i\,j}\,(\mathrm{y}) = 0\}$；或者若 $\mathrm{P}_i^{\mathrm{low}}$ $\mathrm{j}\,(\mathrm{y}) \leq 0$，则 $\mathrm{y} \geq \min\{\mathrm{y} \in \mathbb{R} \mid \mathrm{P}_i^{\mathrm{low}}\mathrm{j}\,(\mathrm{y}) = 0\}$。

证明。根据引理 2.3(ii)，二次多项式的首项系数 $\mathrm{b}_{i\,j}^-$ 和 b_{ij}^+ 均为正，因此 $\mathrm{P}_i^{\mathrm{low}}{}_j$ 和 $\mathrm{P}^{\mathrm{up}}{}_{ij}$ 均为严格凸。对于一个有两个实根（不一定互异）的严格凸二次多项式 P，两根之间的区间恰好是$\{\mathrm{y} \in \mathbb{R} : \mathrm{P(y)} \leq 0\}$。因此，接下来只需证明 $\mathrm{P}_{ij}^{\mathrm{low}}$ 和 $\mathrm{P}^{\mathrm{up}}{}_{ij}$ 的判别式 $\Delta_{ij}^{\mathrm{low}}$ 和 $\Delta^{\mathrm{up}}{}_{ij}$ 均非负。

由于 $\mathbf{A}$ 是共正的，且 $\mathbf{B}$ 是正定的，因此它们的对角项均非负，定义 r^+ 和 r^- 中的单行和也均非负2.1，因此：

$$a_{ii}b_{jj}+a_{jj}b_{ii} \geq 2\sqrt{a_{ii}a_{jj}b_{ii}b_{jj}}, \quad r_i^-(\mathbf{A})r_j^+(\mathbf{B})+r_j^-(\mathbf{A})r_i^+(\mathbf{B}) \geq 2\sqrt{r_i^-(\mathbf{A})r_j^-(\mathbf{A})\,r_i^+(\mathbf{B})r_j^+(\mathbf{B})}.$$

将两边分别求和并平方，得到

$$s_{ij}^-(\mathbf{A},\mathbf{B})^2 \geq \left(2\sqrt{a_{ii}a_{jj}b_{ii}b_{jj}}+2\sqrt{r_i^-(\mathbf{A})r_j^-(\mathbf{A})\,r_i^+(\mathbf{B})r_j^+(\mathbf{B})}\right)^2.$$

减去 $4\,b_{ij}^+ a_{ij}^- = 4\left(b_{ii}b_{jj}-r_i^+(\mathbf{B})r_j^+(\mathbf{B})\right)\left(a_{ii}a_{jj}-r_i^-(\mathbf{A})r_j^-(\mathbf{A})\right)$，可得

$$\Delta_{ij}^{\text{low}} := s_{ij}^-(\mathbf{A},\mathbf{B})^2 - 4b_{ij}^+a_{ij}^- \geq 4\left(\sqrt{a_{ii}a_{jj}\,r_i^+(\mathbf{B})r_j^+(\mathbf{B})}+\sqrt{b_{ii}b_{jj}\,r_i^-(\mathbf{A})r_j^-(\mathbf{A})}\right)^2 \geq 0.$$

对于 $\Delta^{\text{up}}{}_{ij}$，通过求逆 r^+ 和 r^-，证明过程是对称的，正如引理 2.9 中所述，即

$$\Delta_{ij}^{\text{up}} := s_{ij}^+(\mathbf{A},\mathbf{B})^2 - 4b_{ij}^-a_{ij}^+ \geq 4\left(\sqrt{a_{ii}a_{jj}\,r_i^-(\mathbf{B})r_j^-(\mathbf{B})}+\sqrt{b_{ii}b_{jj}\,r_i^+(\mathbf{A})r_j^+(\mathbf{A})}\right)^2 \geq 0.$$

□

定理 2.7 的证明。 设 $(\mathbf{x}\ \lambda)$ 是 $(\mathbf{A},\mathbf{B})$ 的一个互补特征对。根据向量 $\mathbf{x}$ 是否只有一个非零元素（即 $\mathbf{x}=\mathbf{e}_p$，其中某个指标 $p\in[n]$）或至少有两个非零元素，证明分为两种情形。

情形 1：$\mathrm{x}=\mathrm{e}_p, \mathrm{p}\in[\mathrm{n}]$。根据互补约束，(1.1a) 的第 p 行为 $\lambda b_{pp}-a_{pp}=0$，即 $\lambda = a_{pp}/b_{pp}$；由于 A 和 B 满足正性假设，该式非负。对于任意指标 $\mathrm{j}\in[\mathrm{n}]$，且 $\mathrm{j}\neq\mathrm{p}$，根据引理 2.9，$\mathrm{P}^{\text{up}}{}_{pj}$ 和 $\mathrm{P}_p^{\text{low}}{}_j$ 在 $\lambda = a_{pp}/b_{pp}$ 处的取值为

$$P_{pj}^{\text{up}}(\lambda) = -\left(r_p^+(\mathbf{A})+\lambda\, r_p^-(\mathbf{B})\right)\left(r_j^+(\mathbf{A})+\lambda\, r_j^-(\mathbf{B})\right)$$
$$P_{pj}^{\text{low}}(\lambda) = -\left(r_p^-(\mathbf{A})+\lambda\, r_p^+(\mathbf{B})\right)\left(r_j^-(\mathbf{A})+\lambda\, r_j^+(\mathbf{B})\right),$$

它们均为非正数，因为 λ, r^+ 和 r^- 均为非负数。根据引理 2.10，λ至少属于 $n-1$ 个区间（其中 $i=p$），这些区间定义了 K_2。

情形 2：$\mathrm{x}_p \geq \mathbf{x}_q > 0$，$p\neq q$。 从现在起，假设 $\mathbf{x}$ 至少有两个正元素，且 p 和 q 是 $[n]$ 中对应于 $\mathbf{x}$ 最大和第二大元素的不同下标，如引理 2.8所示。互补条件（1.1c）可写为

$$(\lambda b_{ii}-a_{ii})x_i^2 = \sum_{j\neq i}(a_{ij}-\lambda b_{ij})x_ix_j, \text{ for } i\in[n]. \tag{2.3}$$

将引理 2.8 给出的界应用于关于 $i=p$ 或 $i=q$的这个恒等式，并利用 $\lambda\geq 0$ 性质 1.5，得到

$$(\lambda b_{ii}-a_{ii})x_i^2 \leq \left(r_i^+(\mathbf{A})+\lambda r_i^-(\mathbf{B})\right)x_px_q, \text{ for } i\in\{p,q\}. \tag{2.4}$$

如果 $\lambda > a_{ii}/b_{ii}$ 对两个 $i\in\{p,\ q\}$均成立，那么不等式（2.4）左端为正。将两个不等式相乘，再除以 $x_p^2x_q^2>0$，得到

$$(\lambda b_{pp}-a_{pp})(\lambda b_{qq}-a_{qq}) \leq (r_p^+(\mathbf{A})+\lambda\, r_p^-(\mathbf{B}))\,(r_q^+(\mathbf{A})+\lambda\, r_q^-(\mathbf{B})).$$

通过重新 将所有项移到左侧，并利用（2.1），得到 $P_{pq}^{\text{up}}(\lambda)\leq 0$。

现在考虑 $\lambda \le a_{ii}/b_{ii}$，其中对某个 i 有 $\in \{p, q\}$。根据情形 1，我们有 $P^{up}_{pq}(a_{ii}/b_{ii}) \le 0$，因此可以得出 $\lambda \le a_{ii}/b_{ii} \le \max\{y \in \mathbb{R} \mid P^{up}_{pq}(y) = 0\}$。

对称论证可用于证明 $\lambda \ge \min\{y \in \mathbb{R} \mid P^{low}_{pq}(y) = 0\}$。我们先将恒等式（2.3）两边乘以 -1，然后应用引理 2.8，其中 $\lambda \ge 0$，得到

$$(a_{ii} - \lambda b_{ii}) x_i^2 \le \left(\lambda r_i^+(\mathbf{B}) + r_i^-(\mathbf{A})\right) x_p x_q, \text{ for } i \in \{p, q\}. \tag{2.5}$$

若对于两个 $i \in \{p, q\}$，$\lambda < a_{ii}/b_{ii}$ 均成立于（2.5），则这些不等式的左边均为正；将它们相乘可得 $P^{low}_{pq}(\lambda) \le 0$，进而 $\lambda \ge \min\{y \in \mathbb{R} \mid P^{low}_{pq}(y) = 0\}$。若对于某个 $i \in \{p, q\}$，$\lambda \ge a_{ii}/b_{ii}$ 成立，则 $P^{low}_{pq}(a_{ii}/b_{ii}) \le 0$，由传递性可得 $\lambda \ge a_{ii}/b_{ii} \ge \min\{y \in \mathbb{R} \mid P^{low}_{pq}(y) = 0\}$。

因此，我们证明了 λ 至少属于 K_2 的一个区间，且该区间对应的 $(i, j) = (p, q)$。

□

注 2.11. 从上述证明中的情形 1 可知，定义 K_2 的并集中的每个区间，对任意一对指标 $i, j \in [n]$，都是良定义的，因为它包含值 a_{ii}/b_{ii} 和 a_{jj}/b_{jj}。事实上，在 $y = a_{ii}/b_{ii}$ 处应用引理 2.9，由假设可知该点非负，可得 $P^{low}_{ij}(a_{ii}/b_{ii}) \le 0$ 和 $P^{up}{}_{ij}(a_{ii}/b_{ii}) \le 0$，进而由引理 2.10，

$$\min\{y \in \mathbb{R} \mid P^{low}_{ij}(y) = 0\} \le a_{ii}/b_{ii} \le \max\{y \in \mathbb{R} \mid P^{up}_{ij}(y) = 0\}.$$

3. 定位集的比较

在本节中，我们将证明定理 2.7 中给出的双行集 K_2 小于定理 2.2 中给出的单行集 K_1。为此，必须对这两个定理采用相同的假设（尤其是 K_1 中 A 的余正性）。注意，$K_1' \subset K_1$（由构造可知），然后直接证明 $K_2 \subset K_1'$ 可推出 $K_2 \subset K_1$。因此，我们在下文引入记号 $C^{up}{}_{ij}$ 和 C^{low}_{ij}，表示推论 2.4 中定义 K_1' 的各区间端点的双行版本。

引理 3.1。设 $A, B \in \mathbb{S}_n$。假设 B 正定且严格对角占优，而 A 共正。对于任意 $i, j \in [n]$，，满足 $i \ne j$，定义

$$C^{up}_{ij} := \max\left\{\frac{a_i^+}{b_i^-}, \frac{a_j^+}{b_j^-}\right\}, \quad C^{low}_{ij} := \min\left\{\frac{a_i^-}{b_i^+}, \frac{a_j^-}{b_j^+}\right\}, \tag{3.1}$$

$$y^{\star,up}_{ij} := \frac{s^+_{ij}(\mathbf{A}, \mathbf{B})}{2b^-_{ij}}, \quad y^{\star,low}_{ij} := \frac{s^-_{ij}(\mathbf{A}, \mathbf{B})}{2b^+_{ij}}. \tag{3.2}$$

则有以下性质

(a) $y^{\star,\ up}{}_{ij}$ 和 $y^{\star,\ low}{}_{ij}$ 分别是 $P^{up}{}_{ij}$ 和 P^{low}_{ij} 的顶点。(b) 对每个满足 $i, j \in [n]$，且 $i \ne j$ 的 i, j，$y^{\star,\ up}{}_{ij} \le C^{up}{}_{ij}$，且 $y^{\star,\ low}{}_{ij} \ge C^{low}_{ij}$。

证明。为证明 (a) 中的结论，回顾引理 2.10 指出，$P^{up}{}_{ij}$ 是一个严格凸二次函数。其最小值出现在该抛物线的顶点处，可通过令一阶导数等于零求得。由此得到 $y^{\star,\ up}{}_{ij}$（3.2）。类似地，$y^{\star,\ low}$（3.2）是 P^{low}_{ij}的顶点。

(b) 中的两个不等式可通过以下推理得到。

(i) 根据 $C^{up}{}_{ij}$ 以及 $b_i^- > 0$ 和 $b_j^- > 0$ 的定义，并由于 B 是正定且严格对角占优的：

$$a_i^+ \le C_{ij}^{up} b_i^-, \quad a_j^+ \le C_{ij}^{up} b_j^-.$$

将第一个不等式乘以 $b_{jj} + r_j^-(B)$，第二个不等式乘以 $b_{ii} + r_i^-(B)$，并将所得的两个不等式相加，得到

$$a_i^+(b_{jj} + r_j^-(\mathbf{B})) + a_j^+(b_{ii} + r_i^-(\mathbf{B})) \le C_{ij}^{up}\left(b_i^-(b_{jj} + r_j^-(\mathbf{B})) + b_j^-(b_{ii} + r_i^-(\mathbf{B}))\right).$$

右侧等于 $2C^{up}{}_{ij}b_{ij}^-$，展开左侧得

$$s_{ij}^+(\mathbf{A},\mathbf{B}) + a_{ii}r_j^-(\mathbf{B}) + a_{jj}r_i^-(\mathbf{B}) + r_i^+(\mathbf{A})b_{jj} + r_j^+(\mathbf{A})b_{ii},$$

其中所有项均非负。因此 $s_{ij}^+(A,B) \le 2C^{up}{}_{ij}b_{ij}^-$，由此得到

$$y_{ij}^{\star,up} := \frac{s_{ij}^+(\mathbf{A},\mathbf{B})}{2b_{ij}^-} \le C_{ij}^{up}.$$

(ii) 同理，由于 $b^- > 0$，我们有

$$a_i^-(b_{jj} - r_j^+(\mathbf{B})) + a_j^-(b_{ii} - r_i^+(\mathbf{B})) \ge C_{ij}^{low}\left(b_i^+(b_{jj} - r_j^+(\mathbf{B})) + b_j^+(b_{ii} - r_i^+(\mathbf{B}))\right)$$

由此得到

$$s_{ij}^-(\mathbf{A},\mathbf{B}) \ge s_{ij}^-(\mathbf{A},\mathbf{B}) - a_{ii}r_j^+(\mathbf{B}) - a_{jj}r_i^+(\mathbf{B}) - r_i^-(\mathbf{A})b_{jj} - r_j^-(\mathbf{A})b_{ii} \ge 2C_{ij}^{low}b_{ij}^+,$$

hence $y_{ij}^{\star,low} \ge C_{ij}^{low}$.

□

在以下定理的证明中，不失一般性，我们可以假设 $C_{ij}^{low} > 0$因为，否则，根据推论 K_1' 2.4 中 与 (i,j) 相关联的区间的下界为零（它总是小于或等于定义 K_2 的任一区间端点）。

定理 3.2（K_2 比 K_1更紧）。设 A、B $\in \mathbb{S}_n$。假设 B 正定且严格对角占优，A 共正，则 $K_2 \subset K_1$ （准确地说，$K_2 \subset K_1' \subset K_1$）。

证明。对于任意 $i \in [n]$ 和 $y \in \mathbb{R}$，记

$$X_i(y) = b_{ii}y - a_{ii}, \quad U_i(y) = r_i^+(\mathbf{A}) + yr_i^-(\mathbf{B}), \quad L_i(y) = r_i^-(\mathbf{A}) + yr_i^+(\mathbf{B}).$$

根据引理 2.9，对于 i、$j \in [n], i \ne j$，有

$$P_{ij}^{up}(y) = X_i(y)X_j(y) - U_i(y)U_j(y), \quad P_{ij}^{low}(y) = X_i(y)X_j(y) - L_i(y)L_j(y). \tag{3.3}$$

此外，对于 $y \ge 0$，$U_i(y) \ge 0$ 且 $L_i(y) \ge 0$，并且有以下等价关系：

$$X_i(y) > U_i(y) \iff (b_{ii} - r_i^-(\mathbf{B}))y > a_{ii} + r_i^+(\mathbf{A}) \iff y > a_i^+/b_i^-, \tag{3.4}$$

$$-X_i(y) > L_i(y) \iff (b_{ii} + r_i^+(\mathbf{B}))y < a_{ii} - r_i^-(\mathbf{A}) \iff y < a_i^-/b_i^+. \tag{3.5}$$

- 若 $y > \max\left\{a^+{}_i b^-{}_i,\ a^+{}_j b^-{}_j\right\} = C\,up\,{}_{ij}$，则 $X_i(y) > U_i(y) \ge 0$ 且 $X_j(y) > U_j(y) \ge 0$。根据式（3.3）中 $Pup\,{}_{ij}$的定义，这些条件可得 $P\,up\,{}_{ij}(y) > 0$。由于 $Pup\,{}_{ij}$是严格凸二次函数，$Pup\,{}_{ij}$为正的 y 值位于其最小根之外

和最大根。然而，根据引理 3.1，可得 $y > C^{up}{}_{ij} \geq y^{\star}$，$^{up}{}_{ij}$，也就是说，y 位于 $P^{up}{}_{ij}$ 最大根的右侧。

- 若 $0 \leq y < \min\left\{a^{-}{}_{i}b^{+}{}_{i},\ a^{-}{}_{j}b^{+}{}_{j}\right\} = Clow_{ij}$，则 $-X_i(y) > L_i(y) \geq 0$ 且 $-X_j(y) > L_j(y) \geq 0$，因此 $Plow_{ij}(y) > 0$。同样，由引理 3.1 可知，y 位于 Plow$_{ij}$的顶点 $y^{\star}$,low $_{ij}$左侧，且位于其最小根处。

由于定义 K_1' 的端点为 $a_i^{-}b_i^{+}$和 $a_i^{+}b_i^{-}$，其中 $i \in [n]$，；而定义 K_2的端点是 $P_i{}^{low}{}_j$的最小根和 $P^{up}{}_{ij}$的最大根，其中 $i \neq j$，因此上述两点意味着，每个对 K_2 有贡献的区间都包含于某个对 K_1'有贡献的区间中。对所有满足 $i \neq j$ 的有序对 (i, j) 取并集，可得 $K_2 \subset K_1'$，从而推出 $K_2 \subset K_1$。

□

4. 下界与上界

定位集 K_1 和 K_2 约束每个互补特征值，其约束比仅给出全局下界或上界更强。对于实际筛选和基准测试，通常只需控制互补谱的最大值和最小值。因此，我们分别针对单行情形和双行情形，推导可计算的最大互补特征值与最小互补特征值的界，如推论 4.1 和 4.2所述。

推论 4.1（单行界）。设 A、B $\in \mathbb{S}_n$，并假设 B 正定且严格对角占优。设 λ 为 (A, B) 的一个互补特征值。则，

$$\min\left\{\min_{i\in[n]} \frac{a_i^-}{b_i^-},\ \min_{i\in[n]} \frac{a_i^-}{b_i^+}\right\} \leq \lambda \leq \max\left\{\max_{i\in[n]} \frac{a_i^+}{b_i^-},\ \max_{i\in[n]} \frac{a_i^+}{b_i^+}\right\}.$$

推论 4.2（双行界）。设 A、B $\in \mathbb{S}_n$。假设 B 正定且严格对角占优，且 A 共正。设 λ 为 (A, B) 的一个互补特征值。则，

$$\min_{\substack{i,j\in[n]\\ i\neq j}} \frac{s_{ij}^-(\mathbf{A},\mathbf{B}) - \sqrt{s_{ij}^-(\mathbf{A},\mathbf{B})^2 - 4b_{ij}^+a_{ij}^-}}{2b_{ij}^+} \leq \lambda \leq \max_{\substack{i,j\in[n]\\ i\neq j}} \frac{s_{ij}^+(\mathbf{A},\mathbf{B}) + \sqrt{s_{ij}^+(\mathbf{A},\mathbf{B})^2 - 4b_{ij}^-a_{ij}^+}}{2b_{ij}^-}.$$

可以将上述两个界都解释为凸包，即K_1 和K_2。换言之，推论4.1 和4.2 可分别改写为 $\lambda \in \mathrm{conv}(K_1)$ 和 $\lambda \in \mathrm{conv}(K_2)$。

例 4.3。考虑由 $\mathbb{R}^3$ 给出的矩阵：

$$\mathbf{A} = \begin{pmatrix} 14 & 1 & 1 \\ 1 & 11 & -2 \\ 1 & -2 & 13 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 10 & 2 \\ 0 & 2 & 10 \end{pmatrix}.$$

注意，这两个矩阵都是对角元为正的严格对角占优矩阵，因此它们都是正定矩阵。特别地，**A** 是共正的。因此，定理2.2 和定理2.7 的假设均成立。

如引言所述，我们可以按支集 $S \subset \{1,2,3\}$ 枚举互补特征对：对于每个非空 S，候选值 λ 是 (A_{SS},B_{SS}) 的广义特征值

对应 $\mathbf{x}_S \geq \mathbf{0}$，然后通过支集外不等式 $(\mathbf{A}_{iS} - \lambda \mathbf{B}_{iS})\mathbf{x} \geq \mathbf{0}$（其中 $i \notin S$）进行验证。可行支集为 $\{1\}, \{1,2\}, \{1,3\}, \{2,3\}$ 和 $\{1,2,3\}$，互补谱为

$$\Pi \approx \{0.822,\ 2.333,\ 2.347,\ 2.349,\ 2.352\}.$$

（i）推论 4.1 给出的一行包围为

$$\Pi \subset \operatorname{conv}(K_1) = \left[\frac{3}{4}, \frac{8}{3}\right] \approx [0.750,\ 2.667].$$

（ii）推论 4.2 给出的两行包围化简为

$$\Pi \subset \operatorname{conv}(K_2) = \left[\frac{31 - \sqrt{127}}{24},\ \frac{109 + \sqrt{1081}}{60}\right] \approx [0.822,\ 2.365].$$

如定理 3.2所示，我们有 conv（K_2）$\subset$ conv（K_1）。

让我们将其与经典广义特征值的谱 [$\mu_{\min}$（$\mathbf{A}$，$\mathbf{B}$）$\mu_{\max}$（$\mathbf{A}$，$\mathbf{B}$）] 进行比较。两个对称矩阵 （$\mathbf{A}$，$\mathbf{B}$）且 $\mathbf{B}$ 为正定时，其最小和最大广义特征值可定义为广义瑞利商（1.2）关于 $\mathbf{x} \neq \mathbf{0}$ 的最小化和最大化问题的解（见 [3]中的 A.5.3 节），即

$$\mu_{\min}(\mathbf{A}, \mathbf{B}) := \min_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^\top \mathbf{A}\mathbf{x}}{\mathbf{x}^\top \mathbf{B}\mathbf{x}}, \quad \mu_{\max}(\mathbf{A}, \mathbf{B}) := \max_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^\top \mathbf{A}\mathbf{x}}{\mathbf{x}^\top \mathbf{B}\mathbf{x}}. \tag{4.1}$$

于是，我们有 $\Pi \subset [\mu_{\min}(A, B)\ \mu_{\max}(A,B)] \approx [0.804, 2.352]$。因此，在此情况下，广义谱也是互补特征值的一个定位集，它包含于 conv(K_l) 中，但仅与 conv(K_2) 重叠。

5. 与广义谱的比较

在本节中，我们将说明，这个最后的观察并非偶然：最小和最大广义特征值总是分别为（1.1）的互补特征值提供下界和上界，且它们与第 4 节中的界不可比较。

命题 5.1（广义谱定位集）。设 A、B $\in \mathbb{S}_n$，并假设 B 正定。若 λ 是 (A, B) 的互补特征值，且 $\mu_{\min}$(A, B) 和 $\mu_{\max}$(A, B) 分别是 (A, B) 的最小和最大广义特征值，则

$$\lambda \in [\mu_{\min}(\mathbf{A}, \mathbf{B}), \mu_{\max}(\mathbf{A}, \mathbf{B})] := \Gamma$$

证明。 设 $(\mathbf{x}\lambda)$ 是 EiCP（1.1）的一个解。正交约束（1.1c）给出 $\lambda = \mathbf{x}^\top \mathbf{A}\mathbf{x}\mathbf{x}^\top \mathbf{B}\mathbf{x}$，该表达式定义良好，因为 $\mathbf{B}$ 是正定且 $\mathbf{x} \neq \mathbf{0}$。利用（4.1）中的定义，即得结论。

□

为了说明以下对象之间不存在支配关系： Γ 以及集合 K_1 和 K_2, 我们依赖以下引理，它刻画了$\mathbf{A}$ 和 $\mathbf{B}$ 可交换时的广义特征值。随后，我们推导出两类问题：其广义谱分别比我们的定位集严格宽松（命题 5.3）或严格收紧（命题5.4）。

引理 5.2。设 A, B $\in \mathbb{S}_n$，并假设 B 是正定的，且 AB $=$BA。设 $\mu_1^{\mathbf{A}}, \ldots, \mu_n^{\mathbf{A}}$ 是 A 的标准特征值，$\mu_1^{\mathbf{B}}, \ldots, \mu_n^{\mathbf{B}} > 0$ 是 B 的标准特征值，所有特征值均按重数计。则矩阵对 (A, B) 的广义特征值 μ_1, $\ldots, \mu_n$ 恰好是 n 个比值

$$\mu_1 = \mu_{\sigma(1)}^{\mathbf{A}}/\mu_{\tau(1)}^{\mathbf{B}}, \ \ldots, \ \mu_n = \mu_{\sigma(n)}^{\mathbf{A}}/\mu_{\tau(n)}^{\mathbf{B}}$$

对于某些置换 σ, τ of [n]。通过对公共标准正交特征基进行编号，可以令σ 和 τ为单位矩阵，从而得到 $\mu_i = \mu_i^{\mathbf{A}}/\mu_i^{\mathbf{B}}$，其中 i $\in$ [n]。

证明。 根据交换的实对称矩阵可由一个正交矩阵同时对角化这一事实（参见 [14]中的定理 2.5.5），存在一个正交 $\mathbf{V} \in \mathbb{R}^{n\times n}$ 使得

$$\mathbf{V}^\top \mathbf{A}\, \mathbf{V} = \mathrm{diag}(\mu_1^{\mathbf{A}}, \ldots, \mu_n^{\mathbf{A}}), \quad \mathbf{V}^\top \mathbf{B}\, \mathbf{V} = \mathrm{diag}(\mu_1^{\mathbf{B}}, \ldots, \mu_n^{\mathbf{B}}).$$

考虑方程 Ax $=$ μBx，其中 μ为任意广义特征值。令 y $=$V$^\top$x，得到

$$(\mathbf{V}^\top \mathbf{A}\, \mathbf{V})\mathbf{y} = \mu\, (\mathbf{V}^\top \mathbf{B}\, \mathbf{V})\mathbf{y} \iff (\mu_i^{\mathbf{A}} - \mu\, \mu_i^{\mathbf{B}})\, y_i = 0 \quad \text{for all } i \in [n].$$

若 y $\neq \mathbf{0}$，则存在某个 $y_i \neq 0$ en], 使得 $\mu = \mu_i^{\mathbf{A}}/\mu_i^{\mathbf{B}}$。反过来，对于每个 i $\in$ [n], , 取 y $=$ e_i（因此，x $=$ Ve_i）即可得到一个非零解 $\mu_i = \mu_i^{\mathbf{A}}/\mu_i^{\mathbf{B}}$。因此，广义特征值恰好是这 n 个比值 $\mu_i^{\mathbf{A}}/\mu_i^{\mathbf{B}}$（按重数计）。

□

命题 5.3 （一族满足 $K_1 = K_2 \subsetneq \Gamma$ 的矩阵）。固定 n $\geq$2 和 $\varepsilon > 1$。令 E $=$e e$^\top$，并定义 e

$$\mathbf{A} := \mathbf{E} + \varepsilon\, \mathbf{I}, \quad \mathbf{B} := ((n-1)+\varepsilon)\, \mathbf{I} - \mathbf{E}.$$

那么，**B** 是 严格对角占优 且 正定，而且 **A** 是共正的（因此定理 *2.2* 和 *2.7* 均适用）。此外，

$$\mathrm{conv}(K_1) = \mathrm{conv}(K_2) = \left[\frac{1+\varepsilon}{n-2+\varepsilon}\,,\ \frac{n+\varepsilon}{\varepsilon-1}\right], \quad [\mu_{\min}(\mathbf{A},\mathbf{B}), \mu_{\max}(\mathbf{A},\mathbf{B})] = \left[\frac{\varepsilon}{n-1+\varepsilon}\,,\ \frac{n+\varepsilon}{\varepsilon-1}\right].$$

因此，$K_1 = K_2 \subsetneq$ [$\mu_{\min}$(A, B) $\mu_{\max}$(A,B)]。

证明。B 是严格对角占优的，因为 $b_{ii} =$ n $-2+\varepsilon > \sum |b_{ij}| =$ n -1 对 $\varepsilon > 1$成立；又由于 B 是对称矩阵，且 j$\neq$i 的对角元素为正，因此 B 是正定的。此外，A 非负，因此它是共正的。矩阵 A 和 B 可交换（因为对于某个常数 δ，有 AB $=$BA $=$ δI $-$E），根据引理 5.2，我们可以通过计算 A 和 B 的最大、最小特征值，求出 $\mu_{\min}$ 和 $\mu_{\max}$ 的值。为此，我们先回顾 E 的谱性质。由于 E $=$ ee$^\mathsf{T}$，可得 Ee $=$ne，且对于所有满足 v$^\mathsf{T}$e $=0$ 的 v，都有 Ev $=0$。因此，E 的特征值 n 对应特征向量 e，而特征值 0 的重数为 n -1，对应于任意与 e 正交的向量。利用 E 的这一分解，我们可以在下文的 (i) 和 (ii) 中确定 A 和 B 的特征值。

(i) 对于 A $=$E+εI，有

$$\mathbf{A}\mathbf{e} = (\mathbf{E}+\varepsilon\mathbf{I})\mathbf{e} = (n+\varepsilon)\mathbf{e}, \quad \mathbf{A}\mathbf{v} = (\mathbf{E}+\varepsilon\mathbf{I})\mathbf{v} = \varepsilon\mathbf{v} \quad \text{for every } \mathbf{v} \text{ such that } \mathbf{v}^\mathsf{T}\mathbf{e} = 0.$$

因此，根据上述特征值方程，**A** 的经典谱等于$\{n+\varepsilon, \ \varepsilon, \ \underbrace{\ldots\ \varepsilon}_{\text{n}-1\ \text{次}}\}$，然后是 $\mu_{\min}^{\mathbf{A}} = \varepsilon$，以及$\mu_{\max}^{\mathbf{A}} = n+\varepsilon$。

（ii）对于 $B=((n-1)+\varepsilon)I-E$,

$$\mathbf{Be}=\big((n-1)+\varepsilon\big)\mathbf{e}-\mathbf{Ee}=\big((n-1)+\varepsilon-n\big)\mathbf{e}=(\varepsilon-1)\mathbf{e},$$

并且对于任意满足 $v^{T}e=0$ 的 v,

$$\mathbf{Bv}=\big((n-1)+\varepsilon\big)\mathbf{v}-\mathbf{Ev}=\big((n-1)+\varepsilon\big)\mathbf{v}.$$

类似地，**B**的谱为 $\{\varepsilon-1,\underbrace{n-1+\varepsilon\ \ldots\ n-1+\varepsilon}_{n-1\text{ times}}\}$，因此我们有

$$\mu_{\min}^{\mathbf{B}}=\varepsilon-1 \text{ and } \mu_{\max}^{\mathbf{B}}=n-1+\varepsilon.$$

因此，$\mu_{\min}(A,B)=\dfrac{n-1+\varepsilon}{\varepsilon-1}$ 且 $\mu_{\max}(A,B)=n+\varepsilon$。

对于 K_2，将 $r_i^{-}(\mathbf{A})=0$、$r_i^{+}(\mathbf{A})=n-1$、$r_i^{-}(\mathbf{B})=n-1$、$r_i^{+}(\mathbf{B})=0$ 代入引理 2.9的多项式表达式中，可得

$$P_{ij}^{\text{low}}(y)=\big((n-2+\varepsilon)y-(1+\varepsilon)\big)^2,\quad P_{ij}^{\text{up}}(y)=\big((n-2+\varepsilon)y-(1+\varepsilon)\big)^2-(n-1)^2(1+y)^2,$$

因此，K_2 中的区间均等于 $\left[\dfrac{1+\varepsilon}{n-2+\varepsilon},\ n\dfrac{\varepsilon}{\varepsilon-1}\right]$，与 $(i,\ j)$无关。由推论 2.4可知，K_1 中的所有区间也与此一致。上界与$\mu_{\max}$一致，但下界严格优于 $\mu_{\min}$，因为：

$$\frac{1+\varepsilon}{n-2+\varepsilon}-\frac{\varepsilon}{n-1+\varepsilon}=\frac{(n-1)+2\varepsilon}{(n-2+\varepsilon)(n-1+\varepsilon)}>0,$$

□

命题 5.4（一个满足 $\Gamma\subsetneq K_1=K_2$ 的族）。固定 $n\geq 2$ 和参数 $\beta>R>0$ 以及 $c>0$。令 $E=e\,e^{\top}$，设 $\rho:=R/(n-1)$，并定义

$$\mathbf{B}:=\beta\,\mathbf{I}+\rho\,(\mathbf{E}-\mathbf{I}),\quad \mathbf{A}:=c\,\mathbf{B}.$$

那么**B** 是严格对角占优且正定，而**A** 是共正的（因此定理2.2 和2.7 适用）。此外

$$\operatorname{conv}(K_1)=\operatorname{conv}(K_2)=\left[\frac{c\beta}{\beta+R},\ \frac{c(\beta+R)}{\beta}\right],\quad [\mu_{\min}(\mathbf{A},\mathbf{B}),\mu_{\max}(\mathbf{A},\mathbf{B})]=\{c\}.$$

因此，$[\mu_{\min}(A,B)\ \mu_{\max}(A,B)]\subsetneq K_1=K_2$。

证明。B 的每一行的对角元素为 $\beta>0$，非对角元素之和为 $(n-1)\rho$，根据定义该值为 R。此外，由于 $\beta>R$，因此 B 是严格对角占优的，进而是正定的。由于 $c>0$ 且 B 是正定的，矩阵 A 是共正的。因此，定理 2.2 和 2.7 成立。对于广义谱，$B^{-1}A=cI$，因此 $\mu_{\min}(A,B)=\mu_{\max}(A,B)=c$，所以区间 Γ 是点 $\{c\}$。

计算 K_2，时，注意行和量对于所有 $i\in[n]$ 都是常数：

$$a_{ii}=c\beta,\quad r_i^{+}(\mathbf{A})=cR,\quad r_i^{-}(\mathbf{A})=0,\quad b_{ii}=\beta,\quad r_i^{+}(\mathbf{B})=R,\quad r_i^{-}(\mathbf{B})=0.$$

代入引理 2.9 中的多项式可得，对于任意 $i\neq j$,

$$P_{ij}^{\text{low}}(y)=(y\beta-c\beta)^2-(yR)^2 \text{ and } P_{ij}^{\text{up}}(y)=(y\beta-c\beta)^2-(cR)^2.$$

因此,

$$\min\{y:\ P_{ij}^{\text{low}}(y)=0\}=\frac{c\beta}{\beta+R},\quad \max\{y:\ P_{ij}^{\text{up}}(y)=0\}=\frac{c(\beta+R)}{\beta},$$

独立于 $(i,\ j)$，因此 $K_2 = [c\,\beta/(\beta+R),\ c\,(\beta+R)/\beta]$。应用推论 2.4中的公式得到完全相同的端点，因此 $K_1 = K_2$。

最后，由于 $\beta/(\beta+R) < 1 < (\beta+R)/\beta$ 对于 $\beta > R > 0$，我们有 $c\,\beta + R < c < c\,(\beta+R)/\beta$，因此 $\{c\} \subsetneq K_1 = K_2$ 严格成立。

□

6. 讨论

我们可以在共同假设（K_1 下，统一推论2.4 中的K_2 与引理2.9中的公式，如下所示：（**A** 为共正矩阵，且**B** 为严格对角占优且正定矩阵）对于$m = 1$ 和$m = 2$,

$$K_m = \bigcup_{\substack{S\subset[n]\\ |S|=m}} \left[\max\{0, \min\{y\in\mathbb{R} \mid P_S^{\mathrm{low}}(y)=0\}\}, \max\{y\in\mathbb{R} \mid P_S^{\mathrm{up}}(y)=0\}\right] \tag{6.1}$$

$$\text{given}\quad P_S^{\mathrm{up}}(y) := \prod_{i\in S}\left(b_{ii}y - a_{ii}\right) - \prod_{i\in S}\left(r_i^+(\mathbf{A}) + y\,r_i^-(\mathbf{B})\right), \tag{6.2}$$

$$P_S^{\mathrm{low}}(y) := \prod_{i\in S}\left(a_{ii} - b_{ii}y\right) - \prod_{i\in S}\left(r_i^-(\mathbf{A}) + y\,r_i^+(\mathbf{B})\right) \quad \forall y\in\mathbb{R}. \tag{6.3}$$

在这两种情形下，m 表示用于推导每个区间所考虑的行数。我们可以问，这个公式是否能推广到 $m \geq 3$。遗憾的是，例 4.3 给出了一个反例，说明该公式不适用于 $m = 3$。事实上，对三元组 (6.2) 和 (6.3) 取 $S = \{1,2,3\}$ 可得

$$P_S^{\mathrm{up}}(y) = (6y-14)(10y-11)(10y-13) - 2, \quad P_S^{\mathrm{low}}(y) = (14-6y)(11-10y)(13-10y).$$

P_S^{low}的最小实根为 1.1，然而互补特征值 $\lambda \approx 0.822 \in \Pi$ 位于其左侧。$P^{\mathrm{up}}\ s$ 的最大实根为 ≈ 2.336，然而 $\lambda \approx 2.347 \in \Pi$ 位于其右侧（事实上，$P^{\mathrm{up}}\ s\,(2.347) \approx 8.5 > 0$）。因此，如何根据所考虑子矩阵的大小构造一系列定位集，仍是一个悬而未决的问题。

7. 结论

我们分析了 B 为正定且严格对角占优矩阵时的对称 EiCP(A,B)。我们将 He-Liu-Shen 包围界推广至更一般的正定矩阵 B，并提出 Gershgorin 型定位集 K_1 和 K_2, ; 在 A 为共正矩阵的假设下，后者优于前者。这些定位集给出的界计算开销较低，可用于近似最小和最大互补特征值；最小和最大广义特征值也能提供此类近似值，且这些界之间不存在支配关系。

我们的结果还为未来研究指明了若干方向。在算法应用中，局部化集或其界可为互补特征对的候选值提供快速证书。讨论也引出了将该方法推广到多行耦合的问题。最后，该方法还可适用于 EiCP 的变体，例如张量特征值互补问题。

致谢
作者感谢 Jorma K. Merikoski、Pentti Haukkanen、TimoTossavainen 和 Mika Mattila 与他们进行了有益的交流，尤其是关于引理 5.2。

参考文献

[1] S. Adly、M. Haddou 和 M. H. Le。求解帕累托特征值互补问题的内点法。Optim. Methods Softw., 38:543–569, 2023.[2] S. Adly 和 A. Seeger。锥约束特征值问题的非光滑算法。Comput. Optim. Appl., 49:299–318, 2011.[3] S. Boyd 和 L. Vandenberghe。凸优化。剑桥大学出版社，2004.[4] C. P. Br´as、M. Fukushima、A. N. Iusem 和 J. J´udice。一般凸锥上的二次特征值互补问题。Appl. Math. Comput., 271:594–608, 2015.[5] C. P. Br´as、A. N. Iusem 和 J. J´udice。关于二次特征值互补问题。J. Glob. Optim., 66:153–171, 2016.[6] F. Facchinei、J.-S. Pang 和 R. Stone。有限维变分不等式与互补问题。施普林格，纽约，2003.[7] J. Fan、J. Nie 和 A. Zhou。张量特征值互补问题。Math. Program., 170:507–539, 2018.[8] L. M. Fernandes、J. J´udice、H. D. Sherali 和 M. Fukushima。关于计算特征值互补问题的所有特征值。J. Glob. Optim., 59(2):307–326, 2014.[9] R. Fernandes、J. J´udice 和 V. Trevisan。图的互补特征值。Linear Algebra Appl., 527:216–231, 2017.[10] M. Fukushima、J. J´udice、W. de Oliveira 和 V. Sessa。求解对称特征值互补问题的序贯部分线性化算法。Comput. Optim. Appl., 77:711–728, 2020.[11] J. He、Y. Liu 和 X. Shen。帕累托特征值的定位集及其应用。Appl. Math. Lett., 144, 2023.[12] J.-B. Hiriart-Urruty 和 A. Seeger。共正矩阵的变分方法。SIAM Rev., 52:593–629, 2010.[13] G. Holubov´a 和 P. Ne˘casl。关于 fu˘c´ık 谱与帕累托特征值之间关系的注记。J. Math. Anal. Appl., 427, 2015.[14] R. A. Horn 和 C. R. Johnson。矩阵分析。剑桥大学出版社，剑桥，第 2 版，2013.[15] A. N. Iusem、J. J´udice、V. Sessa 和 H. D. Sherali。关于二次特征值互补问题的数值解。Numer. Algorithms, 72:721–747, 2016.[16] A. N. Iusem、J. J´udice、H. D. Sherali 和 V. Sessa。二阶锥二次特征值互补问题。Pac. J. Optim., 13:475–500, 2017.[17] J. J´udice、M. Fukushima、A. N. Iusem、J. M. Mart´ınez 和 V. Sessa。用于特征值互补问题的交替方向乘子法。Optim. Methods Softw., 36(2-3):337–370, 2021.[18] J. J´udice、M. Raydan、S. S. Rosa 和 S. A. Santos。利用谱投影梯度算法求解对称特征值互补问题。Numer. Algorithms, 47:391–407, 2008.[19] J. J´udice、V. Sessa 和 M. Fukushima。单纯形上的分式二次规划及其在特征值互补问题中的应用。J. Optim. Theory Appl., 193:545–573, 2022.[20] V. Kosti´c。通过对角占优定位特征值的一般原理。Adv. Comput. Math., 41(1):55–75, 2015.[21] Y. Nakatsukasa。欧几里得度量下广义特征值问题的格尔施戈林定理。Math. Comp., 80(276):2127–2142, 2011.[22] M. G. Queiroz、J. J´udice 和 C. Humes Jr.。对称特征值互补问题。Math. Comp., 73:1849–1863, 2004.[23] S. Schaible。分式规划。II：关于 Dinkelbach算法。Manag. Sci., 22:868–873, 1976.[24] A. Seeger。由线性互补条件定义的均衡过程的特征值分析。Linear Algebra Appl., 292, 1999.[25] A. Seeger。连通图的互补特征值分析。Linear Algebra Appl., 543:205–225, 2018.[26] S. Singh 和 G. Panda。区间矩阵的广义特征值问题。Arch. Math. (Basel), 121:267–278, 2023.