Linear Algebra and its Applications 438 (2013) 3085–3106

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Trace minimization principles for positive semi-definite pencils

Xin Liang [a,1], Ren-Cang Li [b,*,2], Zhaojun Bai [c,d,3]

[a] *School of Mathematical Sciences, Peking University, Beijing 100871, PR China*
[b] *Department of Mathematics, University of Texas at Arlington, P.O. Box 19408, Arlington, TX 76019, United States*
[c] *Department of Computer Science, University of California, Davis, CA 95616, United States*
[d] *Department of Mathematics, University of California, Davis, CA 95616, United States*

ARTICLE INFO

Article history:
Received 7 September 2012
Accepted 4 December 2012
Available online 11 January 2013

Submitted by P. Šemrl

AMS classification:
15A18
15A22
65F15

Keywords:
Hermitian matrix pencil
Positive semi-definite
Trace minimization
Eigenvalue
Eigenvector

ABSTRACT

This paper is concerned with inf trace($X^{\mathrm{H}}AX$) subject to $X^{\mathrm{H}}BX = J$ for a Hermitian matrix pencil $A - \lambda B$, where J is diagonal and $J^2 = I$ (the identity matrix of apt size). The same problem was investigated earlier by Kovač-Striko and Veselić (*Linear Algebra Appl.* 216 (1995) 139–158) for the case in which B is assumed nonsingular. But in this paper, B is no longer assumed nonsingular, and in fact $A - \lambda B$ is even allowed to be a singular pencil. It is proved, among others, that the infimum is finite if and only if $A - \lambda B$ is a positive semi-definite pencil (in the sense that there is a real number λ_0 such that $A - \lambda_0 B$ is positive semi-definite). The infimum, when finite, can be expressed in terms of the finite eigenvalues of $A - \lambda B$. Sufficient and necessary conditions for the attainability of the infimum are also obtained.

1. Introduction

Consider Hermitian matrix $A \in \mathbb{C}^{n\times n}$. Denote its eigenvalues by λ_i $(i = 1, 2, \ldots, n)$ in the ascending order:

$$\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n. \tag{1.1}$$

* Corresponding author.
E-mail addresses: liangxinslm@pku.edu.cn (X. Liang), rcli@uta.edu (R.-C. Li), bai@cs.ucdavis.edu (Z. Bai).
1 Supported in part by China Scholarship Council. This author is currently a visiting student at Department of Mathematics, University of Texas at Arlington, Arlington, TX 76019, United States.
2 Supported in part by NSF grants DMS-0810506 and DMS-1115834.
3 Supported in part by NSF grants OCI-0749217 and DMS-1115817, and DOE grant DE-FC02-06ER25794.

http://dx.doi.org/10.1016/j.laa.2012.12.003

One, among numerous others, well-known result for a Hermitian matrix is the following trace minimization principle [1, p. 191]

$$\min_{X^{\mathrm{H}}X=I_k} \operatorname{trace}(X^{\mathrm{H}}AX) = \sum_{i=1}^{k} \lambda_i, \tag{1.2}$$

where I_k is the $k \times k$ identity matrix, and $X \in \mathbb{C}^{n\times k}$ is implied by size compatibility in matrix multiplications. Moreover for any minimizer $X_{\min}$ of (1.2), i.e., $\operatorname{trace}(X_{\min}^{\mathrm{H}}AX_{\min}) = \sum_{i=1}^{k} \lambda_i$, its columns span A's invariant subspace[4] associated with the first k eigenvalues λ_i, $i = 1, 2, \ldots, k$. Eq. (1.2) can be proved by using Cauchy's interlacing property, for example, and is also a simple consequence of the more general Wielandt's theorem [2, p. 199].

This minimization principle (1.2) can be extended to the *generalized eigenvalue problem* for a matrix pencil $A - \lambda B$, where $A,\ B \in \mathbb{C}^{n\times n}$ are Hermitian and B is positive definite. Abusing the notation, we still denote the eigenvalues of $A - \lambda B$ by λ_i ($i = 1, 2, \ldots, n$) in the ascending order as in (1.1). The extended result reads [3]

$$\min_{X^{\mathrm{H}}BX=I_k} \operatorname{trace}(X^{\mathrm{H}}AX) = \sum_{i=1}^{k} \lambda_i. \tag{1.3}$$

Moreover for any minimizer $X_{\min}$ of (1.3), there is a Hermitian $A_0 \in \mathbb{C}^{k\times k}$ whose eigenvalues are λ_i, $i = 1, 2, \ldots, k$ such that $AX_{\min} = BX_{\min}A_0$. The result (1.3), seemingly more general than (1.2), is in fact implied by (1.2) by noticing that the eigenvalue problem for $A - \lambda B$ is equivalent to the standard eigenvalue problem for $B^{-1/2}AB^{-1/2}$, where $B^{-1/2} = (B^{1/2})^{-1}$ and $B^{1/2}$ is the unique positive definite square root of B.

The next question is how far we can go in extending (1.2). In 1995, Kovač-Striko and Veselić [4] obtained a few surprising results in this regard. To explain their results, we first give the following definition.

Definition 1.1. $A - \lambda B$ is a *Hermitian pencil* of order n if both $A,\ B \in \mathbb{C}^{n\times n}$ are Hermitian. $A - \lambda B$ is a *positive (semi-)definite matrix pencil* of order n if it is a Hermitian pencil of order n and if there exists $\lambda_0 \in \mathbb{R}$ such that $A - \lambda_0 B$ is positive (semi-)definite.

Note that this definition does not demand anything on the regularity of $A - \lambda B$, i.e., a Hermitian pencil or a positive semi-definite matrix pencil can be either regular (meaning $\det(A - \lambda B) \not\equiv 0$) or singular (meaning $\det(A-\lambda B) \equiv 0$ for all $\lambda \in \mathbb{C}$). Kovač-Striko and Veselić [4] focused on a Hermitian[5] pencil $A - \lambda B$ with B always nonsingular but possibly indefinite. That B is invertible ensures

$$\det(A - \lambda B) \not\equiv 0$$

and thus the regularity of $A-\lambda B$. Denote by n_+ and n_- the numbers of positive and negative eigenvalues of B, respectively, and let k_+ and k_- be two nonnegative integers such that $k_+ \leqslant n_+$, $k_- \leqslant n_-$, and $k_+ + k_- \geqslant 1$, and set

$$J_k = \begin{bmatrix} I_{k_+} & \\ & -I_{k_-} \end{bmatrix} \in \mathbb{C}^{k\times k}, \quad k = k_+ + k_-. \tag{1.4}$$

[4] This invariant subspace is unique if $\lambda_k < \lambda_{k+1}$. This is also true for the deflating subspace spanned by the columns of the minimizer for (1.3).

[5] Although Kovač-Striko and Veselić [4] were concerned about real symmetric matrices, but their arguments can be easily modified to work for Hermitian matrices.

J_k will have this assignment for the rest of this paper. Since B is nonsingular, $n = n_+ + n_-$. The following remarkable results are obtained in [4].

Theorem 1.1 (Kovač-Striko and Veselić [4]). *Let $A - \lambda B$ be a Hermitian pencil of order n and suppose that B is nonsingular.*

1. *Suppose that $A - \lambda B$ is positive semi-definite, and denote by $\lambda_i^{\pm}$ the eigenvalues*[6] *of $A - \lambda B$ arranged in the order:*

$$\lambda_{n_-}^- \leqslant \cdots \leqslant \lambda_1^- \leqslant \lambda_1^+ \leqslant \cdots \leqslant \lambda_{n_+}^+. \tag{1.5}$$

Let $X \in \mathbb{C}^{k\times k}$ satisfying $X^{\mathrm{H}}BX = J_k$, and denote by $\mu_i^{\pm}$ the eigenvalues of $X^{\mathrm{H}}AX - \lambda X^{\mathrm{H}}BX$ arranged in the order:

$$\mu_{k_-}^- \leqslant \cdots \leqslant \mu_1^- \leqslant \mu_1^+ \leqslant \cdots \leqslant \mu_{k_+}^+. \tag{1.6}$$

Then

$$\lambda_i^+ \leqslant \mu_i^+ \leqslant \lambda_{i+n-k}^+, \quad \text{for } 1 \leqslant i \leqslant k_+, \tag{1.7}$$

$$\lambda_{j+n-k}^- \leqslant \mu_i^- \leqslant \lambda_i^-, \quad \text{for } 1 \leqslant j \leqslant k_-, \tag{1.8}$$

where we set $\lambda_i^+ = \infty$ for $i > n_+$ and $\lambda_j^- = -\infty$ for $j > n_-$.

2. *If $A - \lambda B$ is positive semi-definite, then*

$$\inf_{X^{\mathrm{H}}BX=J_k} \operatorname{trace}(X^{\mathrm{H}}AX) = \sum_{i=1}^{k_+} \lambda_i^+ - \sum_{i=1}^{k_-} \lambda_i^-. \tag{1.9}$$

   (a) *The infimum is attainable, if there exists a matrix $X_{\min}$ that satisfies $X_{\min}^{\mathrm{H}}BX_{\min} = J_k$ and whose first k_+ columns consist of the eigenvectors associated with the eigenvalues λ_j^+ for $1 \leqslant j \leqslant k_+$ and whose last k_- columns consist of the eigenvectors associated with the eigenvalues λ_i^- for $1 \leqslant i \leqslant k_-$.*
   (b) *If $A - \lambda B$ is positive definite or positive semi-definite but diagonalizable,*[7] *then the infimum is attainable.*
   (c) *When the infimum is attained by $X_{\min}$, there is a Hermitian $A_0 \in \mathbb{C}^{k\times k}$ whose eigenvalues are $\lambda_i^{\pm}$, $i = 1, 2, \ldots, k_{\pm}$ such that*

$$X_{\min}^{\mathrm{H}}BX_{\min} = J_k, \quad AX_{\min} = BX_{\min}A_0.$$

3. *$A - \lambda B$ is a positive semi-definite pencil if and only if*

$$\inf_{X^{\mathrm{H}}BX=J_k} \operatorname{trace}(X^{\mathrm{H}}AX) > -\infty. \tag{1.10}$$

4. *If $\operatorname{trace}(X^{\mathrm{H}}AX)$ as a function of X subject to $X^{\mathrm{H}}BX = J_k$ has a local minimum, then $A - \lambda B$ is a positive semi-definite pencil and the minimum is global.*

[6] Positive semi-definite pencil $A - \lambda B$ with nonsingular B always has only real eigenvalues implied by [4, Proposition 4.1, 5, Theorem 5.10.1]. See also Lemma 3.8 later.

[7] Hermitian pencil $A - \lambda B$ of order n is *diagonalizable* if there exists a nonsingular $n \times n$ matrix W such that both $W^{\mathrm{H}}AW$ and $W^{\mathrm{H}}BW$ are diagonal.

Item 1 of this theorem is [4, Theorem 2.1], item 2 is [4, Theorem 3.1 and Corollary 3.4], item 3 is [4, Corollary 3.8], and item 4 is [4, Theorem 3.5]. They are proved with the prerequisite that B is nonsingular. In [4, Footnote 1 on p. 140], Kovač-Striko and Veselić wrote

> "it seems plausible that many results of this paper are extendable to pencils with B singular, but $\det(A - \lambda B)$ not identically zero. As yet we know of no simple way of doing it."

One of the aims of this paper is to confirm this suspicion that the nonsingularity assumption is indeed not necessary. Moreover in an attempt of being even more general, we cover singular pencils, as well.

We point out that the Courant–Fischer min–max principle [2, p. 201] (for a single eigenvalue, instead of sums of several eigenvalues like traces) has been generalized to arbitrary Hermitian pencils, include semi-definite ones [6–10]. Eq. (1.9) for $k = 1$ can be considered as a special case of those.

Lancaster and Ye [8, Theorem 1.2] defined a positive definite pencil by requiring that $\beta_0 A - \alpha_0 B$ be positive definite for some $\alpha_0, \beta_0 \in \mathbb{R}$. This definition is less restrictive than ours.

1. If $\beta_0 \neq 0$, we let $\lambda_0 = \alpha_0/\beta_0$ to see that Lancaster's and Ye's definition of a definite pencil includes $A - \lambda_0 B$ being either positive or negative definite. Definition 1.1, on the other hand, requires $A - \lambda_0 B$ be positive definite.
2. If $\beta_0 = 0$, then $\alpha_0 \neq 0$ and thus Lancaster and Ye [8] require that B be either positive or negative definite. In this case, $A - \lambda B$ is also positive definite by Definition 1.1 because we can always pick some $\lambda_0 \in \mathbb{R}$ so that $A - \lambda_0 B$ is positive definite.

Even more general but closely related is the concept of a *definite pencil* which is defined by the existence of a *complex* linear combination of A and B being positive definite [11–14]. But to serve our purpose in this paper, we will stick to Definition 1.1.

The rest of this paper is organized as follows. Section 2 presents our first set of main results which are essentially those summarized in Theorem 1.1 but without the nonsingularity assumption on B, while another main result of ours will be given in Section 4 and it is about a sufficient and necessary condition on the attainability for the infimum of the trace function in terms of the eigen-structure of $A - \lambda B$. All proofs related to the main results in Section 2 are grouped in Section 3 for readability. Conclusions are given in Section 5.

Notation. Throughout this paper, $\mathbb{C}^{n\times m}$ is the set of all $n \times m$ complex matrices, $\mathbb{C}^n = \mathbb{C}^{n\times 1}$, and $\mathbb{C} = \mathbb{C}^1$. $\mathbb{R}$ is set of all real numbers. I_n (or simply I if its dimension is clear from the context) is the $n \times n$ identity matrix, and e_j is its jth column. For a matrix X, $\mathcal{N}(X) = \{x : Xx = 0\}$ denotes X's null space and $\mathcal{R}(X)$ denotes X's column space, the subspace spanned by its columns. X^{H} is the conjugate transpose of a vector or matrix. $A \succ 0$ ($A \succeq 0$) means that A is Hermitian positive (semi-)definite, and $A \prec 0$ ($A \preceq 0$) if $-A \succ 0$ ($-A \succeq 0$). $\mathrm{R{\scriptsize E}}(\alpha)$ is the real part of $\alpha \in \mathbb{C}$. For matrices or scalars X_i, both $\mathrm{diag}(X_1, \ldots, X_k)$ and $X_1 \oplus \cdots \oplus X_k$ denote the same matrix

$$\begin{bmatrix} X_1 & & \\ & \ddots & \\ & & X_k \end{bmatrix}.$$

2. Main results

Throughout the rest of this paper, $A - \lambda B$ is always a Hermitian pencil of order n. It may even be singular, i.e., possibly $\det(A - \lambda B) \equiv 0$ for all $\lambda \in \mathbb{C}$. In particular, B is possibly indefinite and singular. The integer triplet (n_+, n_0, n_-) is the *inertia* of B, meaning B has n_+ positive, n_0 0, and n_- negative eigenvalues, respectively. Necessarily

$$r := \mathrm{rank}(B) = n_+ + n_-. \tag{2.1}$$

We say $\mu \neq \infty$ is a *finite eigenvalue* of $A - \lambda B$ if

$$\mathrm{rank}(A - \mu B) < \max_{\lambda \in \mathbb{C}} \mathrm{rank}(A - \lambda B), \tag{2.2}$$

and $x \in \mathbb{C}^n$ is a corresponding *eigenvector* if $0 \neq x \notin \mathcal{N}(A) \cap \mathcal{N}(B)$ satisfies

$$Ax = \mu Bx, \tag{2.3}$$

or equivalently, $0 \neq x \in \mathcal{N}(A - \mu B) \backslash (\mathcal{N}(A) \cap \mathcal{N}(B))$.

To state our main results, for the moment we will take it for granted that *a positive semi-definite pencil $A - \lambda B$ has only $r = \mathrm{rank}(B)$ finite eigenvalues all of which are real*, but we will prove this claim later in Lemma 3.8. Denote these finite eigenvalues by the same notations $\lambda_i^{\pm}$ as in Section 1 for the case of a nonsingular B and arrange them in the order as (1.5):

$$\lambda_{n_-}^- \leqslant \cdots \leqslant \lambda_1^- \leqslant \lambda_1^+ \leqslant \cdots \leqslant \lambda_{n_+}^+. \tag{1.5}$$

throughout the rest of this paper. What we have to keep in mind that now $n_+ + n_-$ may possibly be less than n. Also in Lemma 3.8, we will see that if $\lambda_0 \in \mathbb{R}$ such that $A - \lambda_0 B \succeq 0$ as in Definition 1.1, then for all $i,\ j$

$$\lambda_i^- \leqslant \lambda_0 \leqslant \lambda_j^+. \tag{2.4}$$

Theorem 2.1. *In Theorem 1.1, the condition that B is nonsingular can be removed.*

We emphasize again that Theorem 2.1 covers not only the case when $A - \lambda B$ is a regular pencil and B is singular but also $A - \lambda B$ is a singular pencil.

Remark 2.1. In both Theorems 1.1 and 2.1, the infimum is taken subject to $X^{\mathrm{H}}BX = J_k$. It is not difficult to see this restriction can be relaxed to that $X^{\mathrm{H}}BX$ is unitarily similar to J_k, or equivalently $X^{\mathrm{H}}BX$ is unitary and has the eigenvalue 1 with multiplicity k_+ and -1 with multiplicity k_-. Furthermore for item 1, this restriction can be relaxed to that the inertia of $X^{\mathrm{H}}BX$ is $(k_+, 0, k_-)$.

A necessary condition for a Hermitian pencil $A - \lambda B$ to be definite is that it must be regular. The next theorem extends two other results: Corollary 3.7 and Theorem 3.10 of [4] to a regular pencil.

Theorem 2.2. *Let $A - \lambda B$ be a Hermitian matrix pencil of order n, and suppose it is regular, i.e.,* $\det(A - \lambda B) \not\equiv 0$. *Suppose also that $n_+ \geqslant 1$ and $n_- \geqslant 1$.*

1. *A necessary and sufficient condition for $A - \lambda B$ to be positive definite is that both infimums*

$$t_0^+ = \inf_{x^{\mathrm{H}}Bx=1} x^{\mathrm{H}}Ax, \quad t_0^- = \inf_{x^{\mathrm{H}}Bx=-1} x^{\mathrm{H}}Ax \tag{2.5}$$

are attainable and $t_0^+ + t_0^- > 0$. In this case $\left(-t_0^-, t_0^+\right)$ is the positive definiteness interval of $A - \lambda B$, i.e., $A - \mu B \succ 0$ for any $\mu \in \left(-t_0^-, t_0^+\right)$.

2. *Suppose $1 \leqslant k_+ \leqslant n_+$ and $1 \leqslant k_- \leqslant n_-$ and that the positive definiteness intervals of pencils $X^{\mathrm{H}}AX - \lambda J_k$, taken for all X satisfying $X^{\mathrm{H}}BX = J_k$, have a nonvoid intersection $\mathscr{I}$. Then $A - \lambda B$ is positive definite, and $\mathscr{I}$ is the definiteness interval of $A - \lambda B$.*

Another main result of ours to be given in Section 4 is a sufficient and necessary condition for the attainability of the infimum in the terms of the eigen-structure of the pencil $A - \lambda B$.

Remark 2.2. Lancaster and Ye [8, Theorem 1.2] gave a different characterization of a positive definite pencil with nonsingular B. To state their result, we will characterize each finite real eigenvalue μ of regular Hermitian pencil $A - \lambda B$ as of the *positive type* or the *negative type* according to whether $x^H Bx > 0$ or $x^H Bx < 0$, where x is a corresponding eigenvector. For a multiple eigenvalue μ with the same algebraic and geometric multiplicity, we can choose a basis of the associated eigenspace and pair each copy of μ with one basis vector and define the type of each copy accordingly. Theorem 1.2 of [8] says that $A - \lambda B$ *is positive definite if and only if it is diagonalizable, has all eigenvalues real, and the smallest finite eigenvalue of the positive type is bigger than the largest finite eigenvalue of the negative type*. This result, too, can be extended to include the case when B is singular, using our proving techniques here.

3. Proofs

All notations in Section 2 will be adopted in whole. We will also use integer triplet

$$(i_+(H), i_0(H), i_-(H))$$

for the inertia of a Hermitian matrix H, where $i_+(H)$, $i_0(H)$, and $i_-(H)$ are the number of positive, zero, and negative eigenvalues of H, respectively. In particular,

$$i_+(B) = n_+, \quad i_0(B) = n - r, \quad i_-(B) = n_-.$$

The eventual proofs of Theorems 2.1 and 2.2 relay on a series of lemmas below.

Lemma 3.1. *There is a unitary $U \in \mathbb{C}^{n\times n}$ such that*

$$U^{\mathrm{H}} BU = \begin{array}{c} \\ {\scriptstyle r} \\ {\scriptstyle n-r} \end{array}\!\!\begin{array}{c} \begin{array}{cc} {\scriptstyle r} & {\scriptstyle n-r} \end{array} \\ \begin{bmatrix} B_1 & \\ & 0 \end{bmatrix} \end{array}, \quad U^{\mathrm{H}} AU = \begin{array}{c} \\ {\scriptstyle r} \\ {\scriptstyle n-r} \end{array}\!\!\begin{array}{c} \begin{array}{cc} {\scriptstyle r} & {\scriptstyle n-r} \end{array} \\ \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \end{array}. \tag{3.1}$$

where $A_{ij}^{\mathrm{H}} = A_{ji}$, and $B_1^{\mathrm{H}} = B_1 \in \mathbb{C}^{r\times r}$ is nonsingular.

Lemma 3.1 can be proved by noticing that there is a unitary $U \in \mathbb{C}^{n\times n}$ to transform B as in the first equation in (3.1). The second equation there is simply due to partition $U^{\mathrm{H}}AU$ accordingly for the convenience of our later use.

Now if $A_{21} = A_{12}^{\mathrm{H}}$ in (3.1) can be somehow annihilated, the situation is then very much reduced to the case studied by Kovač-Striko and Veselić [4], namely a nonsingular B. Finding a way to annihilate $A_{21} = A_{12}^{\mathrm{H}}$ is the key to our whole proofs in this section.

Lemma 3.2. *Let $A - \lambda B$ be a Hermitian matrix pencil of order n, and let $\boldsymbol{P}_B$ be the orthogonal projection onto $\mathcal{R}(B)$. If*

$$\mathcal{R}([I - \boldsymbol{P}_B]A\boldsymbol{P}_B) \subseteq \mathcal{R}([I - \boldsymbol{P}_B]A[I - \boldsymbol{P}_B]), \tag{3.2}$$

then there exists a nonsingular $Y \in \mathbb{C}^{n\times n}$ such that

$$
Y^{\rm H}AY = \begin{array}{c c} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r \\ n-r \end{array} & \begin{bmatrix} A_1 & \\ & A_2 \end{bmatrix} \end{array}, \quad Y^{\rm H}BY = \begin{array}{c c} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r \\ n-r \end{array} & \begin{bmatrix} B_1 & \\ & 0 \end{bmatrix} \end{array}, \tag{3.3}
$$

where $B_1^{\rm H} = B_1$ is invertible, and $A_i^{\rm H} = A_i$. Moreover $A - \lambda B$ has r finite eigenvalues which are the same as the eigenvalues of $A_1 - \lambda B_1$.

Proof. We have (3.1) by Lemma 3.1. The condition (3.2) is equivalent to

$$
\mathcal{R}(A_{21}) \subseteq \mathcal{R}(A_{22}).
$$

Thus $A_{22}Z = A_{21} = A_{12}^{\rm H}$ has solutions one of which is $Z = A_{22}^{\dagger}A_{21}$, where $A_{22}^{\dagger}$ is the Moore–Penrose inverse of A_{22}. Define

$$
C = \begin{bmatrix} I_r & 0 \\ -Z & I_{n-r} \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ -A_{22}^{\dagger}A_{21} & I_{n-r} \end{bmatrix}. \tag{3.4}
$$

It can be verified that

$$
C^{\rm H}U^{\rm H}AUC = \begin{bmatrix} A_{11} - A_{12}A_{22}^{\dagger}A_{21} & 0 \\ 0 & A_{22} \end{bmatrix}, \quad C^{\rm H}U^{\rm H}BUC = \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix}. \tag{3.5}
$$

Take $A_1 = A_{11} - A_{12}A_{22}^{\dagger}A_{21}$, $A_2 = A_{22}$, and $Y = UC$ to get (3.3). $\square$

Although the condition (3.2) seems a bit of mysterious, it is always true for positive semi-definite matrix pencils as confirmed by the next lemma.

Lemma 3.3. *If $A - \lambda B$ is a positive semi-definite matrix pencil of order n, then the condition (3.2) is satisfied and thus the equations in (3.3) hold for some nonsingular $Y \in \mathbb{C}^{n\times n}$, and moreover, $A_2 \succeq 0$ and $A_1 - \lambda B_1$ is a positive semi-definite matrix pencil of order $n - r$.*

Proof. There exists $\lambda_0 \in \mathbb{R}$ such that $\widehat{A} := A - \lambda_0 B \succeq 0$. We have (3.1) by Lemma 3.1, and then

$$
U^{\rm H}\widehat{A}U = U^{\rm H}(A - \lambda_0 B)U = \begin{array}{c c} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r \\ n-r \end{array} & \begin{bmatrix} A_{11} - \lambda_0 B_1 & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \end{array} \succeq 0.
$$

Thus $\mathcal{R}(A_{21}) \subseteq \mathcal{R}(A_{22})$ which is (3.2), as expected. Finally, $A_2 \succeq 0$ and that $A_1 - \lambda B_1$ is positive semi-definite are due to $Y^{\rm H}(A - \lambda_0 B)Y \succeq 0$. $\square$

The decompositions in (3.3), if exist, are certainly not unique. The next lemma says the reduced pencils $A_1 - \lambda B_1$ and $A_2 - \lambda \cdot 0$ are unique, up to nonsingular congruence transformation.

Lemma 3.4. *Let $A - \lambda B$ be a Hermitian matrix pencil of order n, and suppose it admits decompositions in* (3.3)*, where* $r = \operatorname{rank}(B)$. *Suppose it also admits*

$$\widetilde{Y}^{\mathrm{H}} A \widetilde{Y} = \begin{array}{c} \\ r \\ n-r \end{array} \!\!\begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \begin{bmatrix} \widetilde{A}_1 & \\ & \widetilde{A}_2 \end{bmatrix} \end{array}, \quad \widetilde{Y}^{\mathrm{H}} B \widetilde{Y} = \begin{array}{c} \\ r \\ n-r \end{array} \!\!\begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \begin{bmatrix} \widetilde{B}_1 & \\ & 0 \end{bmatrix} \end{array}, \tag{3.6}$$

where $\widetilde{Y} \in \mathbb{C}^{n\times n}$ is nonsingular. Then there exist nonsingular $M_1 \in \mathbb{C}^{r\times r}$ and $M_2 \in \mathbb{C}^{(n-r)\times(n-r)}$ such that

$$\widetilde{A}_1 - \lambda \widetilde{B}_1 = M_1^{\mathrm{H}}(A_1 - \lambda B_1)M_1, \quad \widetilde{A}_2 = M_2^{\mathrm{H}} A_2 M_2.$$

Proof. Partition $Y = [Y_1, Y_2]$ and $\widetilde{Y} = [\widetilde{Y}_1, \widetilde{Y}_2]$ with $Y_1,\ \widetilde{Y}_1 \in \mathbb{C}^{n\times r}$. Since $BY_2 = B\widetilde{Y}_2 = 0$, we have $\mathcal{R}(\widetilde{Y}_2) = \mathcal{N}(B) = \mathcal{R}(Y_2)$ and thus $\widetilde{Y}_2 = Y_2 M_2$ for some nonsingular $M_2 \in \mathbb{C}^{(n-r)\times(n-r)}$. Set $M = Y^{-1}\widetilde{Y}_1$ and partition M to get

$$\widetilde{Y}_1 = YM, \quad M = \begin{array}{c} \\ r \\ n-r \end{array} \!\!\begin{array}{c} r \\ \begin{bmatrix} M_1 \\ Z \end{bmatrix} \end{array}.$$

Hence $\widetilde{Y} = [\widetilde{Y}_1, \widetilde{Y}_2] = [Y_1, Y_2]\begin{bmatrix} M_1 & 0 \\ Z & M_2 \end{bmatrix}$ which implies M_1 must be nonsingular. We have by (3.3) and (3.6)

$$\begin{aligned}
0 &= \widetilde{Y}_1^{\mathrm{H}} A \widetilde{Y}_2 = M^{\mathrm{H}} Y^{\mathrm{H}} A Y_2 M_2 = M^{\mathrm{H}} \begin{bmatrix} 0 \\ A_2 \end{bmatrix} M_2 \quad \Rightarrow \quad M^{\mathrm{H}} \begin{bmatrix} 0 \\ A_2 \end{bmatrix} = 0, \\
\widetilde{A}_1 &= \widetilde{Y}_1^{\mathrm{H}} A \widetilde{Y}_1 = M^{\mathrm{H}} Y^{\mathrm{H}} A Y M = M^{\mathrm{H}} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} M = M^{\mathrm{H}} \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} M = M_1^{\mathrm{H}} A_1 M_1, \\
\widetilde{B}_1 &= \widetilde{Y}_1^{\mathrm{H}} B \widetilde{Y}_1 = M^{\mathrm{H}} Y^{\mathrm{H}} B Y M = M^{\mathrm{H}} \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix} M = M_1^{\mathrm{H}} B_1 M_1, \\
\widetilde{A}_2 &= \widetilde{Y}_2^{\mathrm{H}} A \widetilde{Y}_2 = M_2^{\mathrm{H}} Y_2^{\mathrm{H}} A Y_2 M_2 = M_2^{\mathrm{H}} A_2 M_2,
\end{aligned}$$

as expected. □

Lemma 3.5. *Let $M \in \mathbb{C}^{\ell\times\ell}$ be Hermitian and nonsingular, and let $0 \neq y \in \mathbb{C}^{\ell}$. Then there exists $x \in \mathbb{C}^{\ell}$ such that both $x^{\mathrm{H}}Mx \neq 0$ and $x^{\mathrm{H}}y \neq 0$. In the case when M is indefinite, the chosen x can be made either $x^{\mathrm{H}}Mx > 0$ or $x^{\mathrm{H}}Mx < 0$ as needed.*

Proof. If M is positive or negative definite, taking $x = y$ will do. Suppose M is indefinite. There is a nonsingular matrix $Z \in \mathbb{C}^{\ell\times\ell}$ such that $Z^{\mathrm{H}}MZ = \operatorname{diag}(I_{\ell_+}, -I_{\ell_-})$, where $\ell_\pm \geqslant 1$. Partition

$Z^{\mathrm{H}}y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, where $y_1 \in \mathbb{C}^{\ell_+}$. We may take x by

$$\text{either} \quad Z^{-1}x = \begin{bmatrix} y_1 \\ 0 \end{bmatrix} \quad \text{or} \quad Z^{-1}x = \begin{bmatrix} 0 \\ y_2 \end{bmatrix}, \tag{3.7}$$

depending on if $y_i = 0$ or not. Because at least one of y_i is nonzero, one of the choices in (3.7) will make both $x^{\mathrm{H}}Mx \neq 0$ and $x^{\mathrm{H}}y \neq 0$.

It can also be done to ensure $x^{\mathrm{H}}Mx > 0$ regardless. In fact, if $y_1 \neq 0$, the first choice in (3.7) will do. But if $y_1 = 0$, then $y_2 \neq 0$. Take

$$Z^{-1}x = \begin{bmatrix} (y_2^{\mathrm{H}}y_2 + 1)^{1/2}e_1 \\ y_2 \end{bmatrix}.$$

Then $x^{\mathrm{H}}Mx = 1$ and $x^{\mathrm{H}}y = y_2^{\mathrm{H}}y_2$. Similarly we can ensure $x^{\mathrm{H}}Mx < 0$ if needed. $\square$

Lemma 3.6. *Let $A - \lambda B$ be a Hermitian matrix pencil of order n. If*

$$\inf_{X^{\mathrm{H}}BX=J_k} \operatorname{trace}(X^{\mathrm{H}}AX) > -\infty,$$

then the condition (3.2) *holds.*

Proof. We have (3.1) by Lemma 3.1. Now for any $X \in \mathbb{C}^{n\times k}$, write

$$\widetilde{X} = U^{\mathrm{H}}X = \begin{array}{c} {\scriptstyle r} \\ {\scriptstyle n-r} \end{array} \begin{bmatrix} \widetilde{X}_1 \\ \widetilde{X}_2 \end{bmatrix}. \tag{3.8}$$

We have

$$X^{\mathrm{H}}BX = \widetilde{X}^{\mathrm{H}}U^{\mathrm{H}}BU\widetilde{X} = \widetilde{X}_1^{\mathrm{H}}B_1\widetilde{X}_1, \tag{3.9}$$

$$\operatorname{trace}(X^{\mathrm{H}}AX) = \operatorname{trace}(\widetilde{X}_1^{\mathrm{H}}A_{11}\widetilde{X}_1) + 2\mathrm{R}\textsc{e}\big(\operatorname{trace}(\widetilde{X}_1^{\mathrm{H}}A_{12}\widetilde{X}_2)\big) + \operatorname{trace}(\widetilde{X}_2^{\mathrm{H}}A_{22}\widetilde{X}_2). \tag{3.10}$$

The condition (3.2) is equivalent to $\mathcal{R}(A_{21}) \subseteq \mathcal{R}(A_{22})$ which we will prove.

Assume to the contrary that $\mathcal{R}(A_{21}) \nsubseteq \mathcal{R}(A_{22})$, or equivalently

$$\mathcal{N}(A_{12}) = \mathcal{N}(A_{21}^{\mathrm{H}}) = \mathcal{R}(A_{21})^{\perp} \nsupseteq \mathcal{R}(A_{22})^{\perp} = \mathcal{N}(A_{22}^{\mathrm{H}}) = \mathcal{N}(A_{22}),$$

i.e., there exists $0 \neq x_2 \in \mathbb{C}^{n-r}$ such that $A_{22}x_2 = 0$ but $y := A_{12}x_2 \neq 0$. By Lemma 3.5, there is $x_1 \in \mathbb{C}^r$ such that $x_1^{\mathrm{H}}B_1x_1 \neq 0$ and $x_1^{\mathrm{H}}y \neq 0$. For our purpose, we will make $x_1^{\mathrm{H}}B_1x_1 > 0$ if $k_+ > 0$ and $x_1^{\mathrm{H}}B_1x_1 < 0$ otherwise. Scale x_1 so that $|x_1^{\mathrm{H}}B_1x_1| = 1$. B_1 induces an indefinite-inner product in $\in \mathbb{C}^r$ and since $|x_1^{\mathrm{H}}B_1x_1| = 1$, we can extend x_1 to an orthonormal basis with respect to this B_1-indefinite-inner product [5, p. 10]: $x_1, x_2, \ldots, x_r$, i.e., $x_i^{\mathrm{H}}B_1x_j = 0$ for $i \neq j$ and $x_i^{\mathrm{H}}B_1x_i = \pm 1$. Suppose for the moment $x_1^{\mathrm{H}}B_1x_1 = 1$. Pick k x_i out of all: $x_{j_1}, x_{j_2}, \ldots, x_{j_k}$ with $j_1 = 1$ (i.e., x_1 is included in), such that among $x_{i_j}^{\mathrm{H}}B_1x_{i_j}$ for $1 \leqslant j \leqslant k$ there are k_+ of them $+1$s and k_- of them -1s. Now consider those $\widetilde{X}$ in (3.8) with

$$\widetilde{X}_1 = [x_{j_1}, x_{j_2}, \ldots, x_{j_k}]\Pi, \quad \widetilde{X}_2 = \xi\,[y, 0, \ldots, 0],$$

where $\xi \in \mathbb{C}$, and Π is the $r \times r$ permutation matrix such that $\widetilde{X}_1^{\mathrm{H}} B_1 \widetilde{X}_1 = J_k$ and x_1 is in the first column of $\widetilde{X}_1$. Then by (3.10),

$$\mathrm{trace}(X^{\mathrm{H}}AX) = \mathrm{trace}\left(\widetilde{X}_1^{\mathrm{H}} A_{11} \widetilde{X}_1\right) + 2\mathrm{Re}\left(\xi x_1^{\mathrm{H}} y\right)$$

which can be made arbitrarily small towards $-\infty$, contradicting that $\mathrm{trace}(X^{\mathrm{H}}AX)$ as a function of X restricted to $X^{\mathrm{H}}BX = J_k$ is bounded from below. Therefore $\mathcal{R}(A_{21}) \subseteq \mathcal{R}(A_{22})$. The case for $x_1^{\mathrm{H}} B_1 x_1 = -1$ is similar. The proof is completed. $\square$

The standard involutory permutation matrix (SIP) of size n is the $n \times n$ identity matrix with its columns rearranged from the last to the first:

$$\begin{bmatrix} & & & & 1 \\ & & & 1 & \\ & & \iddots & & \\ & 1 & & & \\ 1 & & & & \end{bmatrix}. \tag{3.11}$$

The next lemma presents the well-known canonical form of a Hermitian pencil $A - \lambda B$ with a nonsingular B under nonsingular congruence transformations.

Lemma 3.7 [5, Theorem 5.10.1]. *Let $A - \lambda B$ be a Hermitian matrix pencil of order n, and suppose that B is nonsingular. Then there exists a nonsingular $W \in \mathbb{C}^{n\times n}$ such that*

$$W^{\mathrm{H}}AW = s_1K_1 \oplus \cdots \oplus s_pK_p \oplus \begin{bmatrix} 0 & K_{p+1} \\ K_{p+1}^{\mathrm{H}} & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & K_q \\ K_q^{\mathrm{H}} & 0 \end{bmatrix}, \tag{3.12a}$$

$$W^{\mathrm{H}}BW = s_1S_1 \oplus \cdots \oplus s_pS_p \oplus \begin{bmatrix} 0 & S_{p+1} \\ S_{p+1}^{\mathrm{H}} & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & S_q \\ S_q^{\mathrm{H}} & 0 \end{bmatrix}, \tag{3.12b}$$

where

$$K_i = \begin{bmatrix} & & & & \alpha_i \\ & & & \alpha_i & 1 \\ & & \iddots & 1 & \\ & \iddots & \iddots & & \\ \alpha_i & 1 & & & \\ \alpha_i & 1 & & & \end{bmatrix}, \tag{3.13}$$

$\alpha_i \in \mathbb{R}$ for $1 \leqslant i \leqslant p$; $\alpha_i \in \mathbb{C}$ is nonreal for $p+1 \leqslant i \leqslant q$, and $s_i = \pm 1$ for $1 \leqslant i \leqslant p$; S_i is a SIP whose size is the same as that of K_i for all i. The representations in (3.12) are uniquely determined by the pencil $A - \lambda B$, up to a simultaneous permutation of the corresponding diagonal block pairs.

Lemma 3.8. *Let $A - \lambda B$ be a positive semi-definite matrix pencil of order n, and suppose that $\lambda_0 \in \mathbb{R}$ such that $A - \lambda_0 B \succeq 0$.*

1. *There exists a nonsingular* $W \in \mathbb{C}^{n\times n}$ *such that*

$$W^{\mathrm{H}}AW = \begin{array}{c} \\ {\scriptstyle n_1} \\ {\scriptstyle r-n_1} \\ {\scriptstyle n-r} \end{array}\!\!\begin{array}{c} \begin{array}{ccc} {\scriptstyle n_1} & {\scriptstyle r-n_1} & {\scriptstyle n-r} \end{array} \\ \begin{bmatrix} \Lambda_1 & & \\ & \Lambda_0 & \\ & & \Lambda_\infty \end{bmatrix} \end{array}, \quad W^{\mathrm{H}}BW = \begin{array}{c} \\ {\scriptstyle n_1} \\ {\scriptstyle r-n_1} \\ {\scriptstyle n-r} \end{array}\!\!\begin{array}{c} \begin{array}{ccc} {\scriptstyle n_1} & {\scriptstyle r-n_1} & {\scriptstyle n-r} \end{array} \\ \begin{bmatrix} \Omega_1 & & \\ & \Omega_0 & \\ & & 0 \end{bmatrix} \end{array}, \tag{3.14}$$

where $r = \operatorname{rank}(B) = n_+ + n_-$, *and*
(a) $\Lambda_1 = \operatorname{diag}(s_1\alpha_1, \ldots, s_\ell\alpha_\ell)$, $\Omega_1 = \operatorname{diag}(s_1, \ldots, s_\ell)$, $s_i = \pm 1$, *and* $\Lambda_1 - \lambda_0\Omega_1 \succ 0$,
(b) $\Lambda_0 = \operatorname{diag}(\Lambda_{0,1}, \ldots, \Lambda_{0,m+m_0})$ *and* $\Omega_0 = \operatorname{diag}(\Omega_{0,1}, \ldots, \Omega_{0,m+m_0})$ *with*

$$\Lambda_{0,i} = t_i\lambda_0, \qquad \Omega_{0,i} = t_i = \pm 1, \quad \text{for } 1 \leqslant i \leqslant m,$$
$$\Lambda_{0,i} = \begin{bmatrix} 0 & \lambda_0 \\ \lambda_0 & 1 \end{bmatrix}, \ \Omega_{0,i} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \text{for } m+1 \leqslant i \leqslant m+m_0,$$

(c) $\Lambda_\infty = \operatorname{diag}(\alpha_{r+1}, \ldots, \alpha_n) \succeq 0$ *with* $\alpha_i \in \{1, 0\}$ *for* $r+1 \leqslant i \leqslant n$.
The representations in (3.14) *are uniquely determined by* $A - \lambda B$, *up to a simultaneous permutation of the corresponding* 1×1 *and* 2×2 *diagonal block pairs* $(s_i\alpha_i, s_i)$ *for* $1 \leqslant i \leqslant \ell$, $(\Lambda_{0,i}, \Omega_{0,i})$ *for* $1 \leqslant i \leqslant m+m_0$, *and* $(\alpha_i, 0)$ *for* $r+1 \leqslant i \leqslant n$.

2. $A - \lambda B$ *has* $n_+ + n_-$ *finite eigenvalues all of which are real. Denote these finite eigenvalues by* $\lambda_i^\pm$ *and arrange them in the order as in* (1.5). *Write* $m = m_+ + m_-$, *where* m_+ *is the number of those* 1×1 *diagonal blocks in* Λ_0 *with* $s_i = 1$ *and* m_- *is that of those with* $s_i = -1$. *The respective sources of these finite eigenvalues are*
   source 1. *the* 1×1 *block pairs* $(\Lambda_{0,j}, \Omega_{0,j})$ *with* $t_j = -1$ *produce* $\lambda_i^- = \lambda_0$ *for* $1 \leqslant i \leqslant m_-$;
   source 2. *the* 1×1 *block pairs* $(\Lambda_{0,j}, \Omega_{0,j})$ *with* $t_j = +1$ *produce* $\lambda_i^+ = \lambda_0$ *for* $1 \leqslant i \leqslant m_+$;
   source 3. *the* 2×2 *block pairs* $(\Lambda_{0,m+i}, \Omega_{0,m+i})$ *for* $1 \leqslant i \leqslant m_0$ *produce* $\lambda_{m_-+i}^- = \lambda_0$ *and* $\lambda_{m_++i}^+ = \lambda_0$;
   source 4. *the diagonal matrix pair* (Λ_1, Ω_1) *produces* $\lambda_i^\pm$ *(according to* $s_j = \pm 1$*) for* $m_0 + m_\pm \leqslant i \leqslant n_\pm$.

   Each eigenvalue from sources other than **source 3** *has an eigenvector* x *that satisfies* $x^{\mathrm{H}}Bx = +1$ *for* λ_i^+ *and* $x^{\mathrm{H}}Bx = -1$ *for* λ_j^-, *while for* **source 3**, *each pair* $\left(\lambda_{m_-+i}^-, \lambda_{m_++i}^+\right)$ *of eigenvalues shares one eigenvector* x *that satisfies* $x^{\mathrm{H}}Bx = 0$. *To be more specific than* (1.5), *we can order these finite eigenvalues as*

$$\begin{aligned} \lambda_{n_-}^- \leqslant \cdots \leqslant \lambda_{m_0+m_-+1}^- < \underbrace{\lambda_0 = \cdots = \lambda_0}_{m_0} &= \underbrace{\lambda_0 = \cdots = \lambda_0}_{m_-} \\ &= \underbrace{\lambda_0 = \cdots = \lambda_0}_{m_+} = \underbrace{\lambda_0 = \cdots = \lambda_0}_{m_0} < \lambda_{m_0+m_++1}^+ \\ &\leqslant \cdots \leqslant \lambda_{n_+}^+. \end{aligned} \tag{3.15}$$

   In particular $\lambda_i^- = \lambda_0$ *for* $1 \leqslant i \leqslant m_0 + m_-$ *and* $\lambda_i^+ = \lambda_0$ *for* $1 \leqslant i \leqslant m_0 + m_+$.

3. $\{\gamma \in \mathbb{R} \mid A - \gamma B \succeq 0\} = [\lambda_1^-, \lambda_1^+]$. *Moreover, if* $A - \lambda B$ *is regular, then* $A - \lambda B$ *is a positive definite pencil if and only if* $\lambda_1^- < \lambda_1^+$, *in which case* $\{\gamma \in \mathbb{R} \mid A - \gamma B \succ 0\} = (\lambda_1^-, \lambda_1^+)$.

4. *Let* $\mu = (\lambda_1^- + \lambda_1^+)/2$. *For* $\gamma > \mu$, *let* $\mathfrak{n}(\gamma)$ *be the number of the eigenvalues of the matrix pencil* $A - \lambda B$ *in* $[\mu, \gamma)$, *where* μ, *if an eigenvalue, is counted* $i_+(\Omega_0)$ *times. For* $\gamma < \mu$, *let* $\mathfrak{n}(\gamma)$ *be the number of the eigenvalues of the matrix pencil* $A - \lambda B$ *in* $(\lambda, \mu]$, *where* μ, *if an eigenvalue, is counted* $i_-(\Omega_0)$ *times. Then*

$$\mathfrak{n}(\gamma) = i_-(A - \gamma B).$$

Proof. In Lemma 3.3, $A_1 - \lambda B_1$ is a positive semi-definite matrix pencil with B_1 nonsingular. Such a pencil can be transformed by congruence so that $Y_1^{\mathrm{H}} A_1 Y_1$ and $Y_1^{\mathrm{H}} B_1 Y_1$ are in their canonical forms as given in the right-hand sides of (3.12a) and (3.12b), respectively, where $Y_1 \in \mathbb{C}^{r\times r}$ is nonsingular. We now use the positive semi-definiteness to describe all possible diagonal blocks in the right-hand sides. There are a few cases to deal with:

Case 1. *No K_i ($1 \leqslant i \leqslant p$) is 3×3 or larger*. For a 3×3 K_i with $\alpha_i \in \mathbb{R}$, the right-bottom corner 2×2 submatrix of $K_i - \mu S_i$

$$\begin{bmatrix} \alpha_i - \mu & 1 \\ 1 & 0 \end{bmatrix} \not\succeq 0 \quad \text{nor} \quad \begin{bmatrix} \alpha_i - \mu & 1 \\ 1 & 0 \end{bmatrix} \not\preceq 0$$

for any $\mu \in \mathbb{R}$. For a $k \times k$ K_i with $\alpha_i \in \mathbb{R}$ and $k \geqslant 4$, the submatrix of $K_i - \mu S_i$, consisting of the intersections of its row 2 and k and its column 2 and k is always the 2×2 SIP which is indefinite.

Case 2. *No 2×2 K_i ($1 \leqslant i \leqslant p$) is with $s_i = -1$*. This is because for $s_i = -1$

$$s_i \begin{bmatrix} 0 & \alpha_i \\ \alpha_i & 1 \end{bmatrix} - \mu s_i \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -\alpha_i + \mu \\ -\alpha_i + \mu & -1 \end{bmatrix} \not\succeq 0 \quad \text{for any } \mu \in \mathbb{R}.$$

Case 3. *The α_i for any 2×2 K_i ($1 \leqslant i \leqslant p$), if any, is λ_0*. This is because

$$\begin{bmatrix} 0 & \alpha_i \\ \alpha_i & 1 \end{bmatrix} - \mu \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \alpha_i - \mu \\ \alpha_i - \mu & 1 \end{bmatrix} \succeq 0 \quad \text{if and only if } \mu = \alpha_i.$$

Case 4. *K_i ($1 \leqslant i \leqslant p$) with $\alpha_i \neq \lambda_0$ is 1×1*. This is a result of **Case 1** and **Case 3** above.

Case 5. *The blocks associated with nonreal α_i cannot exist*. This is because the submatrix consisting of the intersections of the first and last row and the first and last column of

$$\begin{bmatrix} 0 & K_i \\ K_i^{\mathrm{H}} & 0 \end{bmatrix} - \mu \begin{bmatrix} 0 & S_i \\ S_i^{\mathrm{H}} & 0 \end{bmatrix}$$

is $\begin{bmatrix} 0 & \alpha_i - \mu \\ \bar{\alpha}_i - \mu & 0 \end{bmatrix}$ which is never semi-definite for any $\mu \in \mathbb{R}$.

Together, they imply

$$Y_1^{\mathrm{H}} A_1 Y_1 = \operatorname{diag}(\Lambda_1, \Lambda_0), \quad Y_1^{\mathrm{H}} B_1 Y_1 = \operatorname{diag}(\Omega_1, \Omega_0), \tag{3.16}$$

where $\Lambda_1, \Lambda_0, \Omega_1, \Omega_0$ as described in the lemma. Since $A_2 \succeq 0$, there exists a nonsingular $Y_2 \in \mathbb{C}^{(n-r)\times(n-r)}$ such that

$$Y_2^{\mathrm{H}} A_2 Y_2 = \operatorname{diag}(\alpha_{r+1}, \ldots, \alpha_n)$$

with $\alpha_i \in \{1, 0\}$ for $r + 1 \leqslant i \leqslant n$. Now set $W = Y \operatorname{diag}(Y_1, Y_2)$ to get (3.14).

The uniqueness of the representations in (3.14), up to simultaneous permutation, is a consequence of the uniqueness claims in Lemma 3.7 and that in Lemma 3.4 up to congruence transformation.

For item 2, we note that the finite eigenvalues of $A-\lambda B$ are the union of the eigenvalues of $\Lambda_1-\lambda\Omega_1$ and these of $\Lambda_0-\lambda\Omega_0$. The rest are a simple consequence of item 1.

For item 3, we note $\Lambda_1-\lambda_0\Omega_1=\operatorname{diag}(s_i(\alpha_i-\lambda_0))\succ 0$. Obviously α_i, $i=1,\dots,\ell$ are some eigenvalues of $A-\lambda B$. If $s_i=1$, $\alpha_i>\lambda_0$, and thus $\alpha_i=\lambda_j^+$ for some $j>m_++m_0$. Similarly, if $s_i=-1$, $\alpha_i=\lambda_k^-$ for some $k>m_-+m_0$. Hence

$$\begin{aligned}&\Lambda_1-\gamma\Omega_1=\operatorname{diag}(s_i(\alpha_i-\gamma))\succeq 0\\ \Leftrightarrow\ &\lambda_k^-\leqslant\gamma\leqslant\lambda_j^+\quad\text{for all}\quad k>m_-+m_0,\, j>m_++m_0.\end{aligned}\tag{3.17}$$

Also,

$$\begin{aligned}&\Lambda_{0,i}-\gamma\Omega_{0,i}=t_i(\lambda_0-\gamma)\succeq 0\quad\text{for}\quad i=1,\dots,m\\ \Leftrightarrow\ &\lambda_k^-\leqslant\gamma\leqslant\lambda_j^+\quad\text{for all}\quad 1\leqslant k\leqslant m_-,\, 1\leqslant j\leqslant m_+,\end{aligned}\tag{3.18}$$

and

$$\Lambda_{0,i}-\gamma\Omega_{0,i}\succeq 0\quad\text{for}\quad i=m+1,\dots,m+m_0\quad\Leftrightarrow\quad\gamma=\lambda_0.\tag{3.19}$$

Putting all together, we have $A-\gamma B\succeq 0\quad\Leftrightarrow\quad\lambda_1^-\leqslant\gamma\leqslant\lambda_1^+$.

For $A-\lambda B$ to be regular and positive semi-definite, $\Lambda_\infty\succ 0$. Now if $A-\lambda B$ is a positive definite pencil, then there exists γ such that the inequalities in (3.17), (3.18) and (3.19) are strict. This can only happen when $m_0=0$ and $\lambda_1^-<\lambda_1^+$, in which case $A-\gamma B\succ 0\Leftrightarrow\lambda_1^-<\gamma<\lambda_1^+$. On the other hand, if $\lambda_1^-<\lambda_1^+$, then $m_0=0$ and only one of m_+ and m_- can be bigger than 0, or equivalently only one of λ_1^- and λ_1^+ can possibly be λ_0 but not both. So for $\lambda_1^-<\gamma<\lambda_1^+$, the inequalities in (3.17) and (3.18) are strictly, and therefore $A-\gamma B\succ 0$.

Item 4 can be proved by separately considering four cases: (1) $\lambda_1^-<\lambda_0<\lambda_1^+$; (2) $\lambda_1^-<\lambda_0=\lambda_1^+$; (3) $\lambda_1^-=\lambda_0<\lambda_1^+$; and (4) $\lambda_1^-=\lambda_0=\lambda_1^+$. Detail is omitted. $\square$

Lemma 3.9. *Suppose B is nonsingular. $A-\lambda B$ is a positive semi-definite matrix pencil if*

$$\inf_{X^{\mathrm{H}}BX=J_k}\operatorname{trace}(X^{\mathrm{H}}AX)>-\infty.$$

Proof. This is part of [4, Corollary 3.8], where the proof is rather sketchy with claims that, though true, were not obvious and substantiated. What follows is a more detailed proof.

If either $B\prec 0$ or $B\succ 0$, then there is $\lambda_0\in\mathbb{R}$ such that $A-\lambda_0B\succ 0$, and thus no proof is necessary. Suppose in what follows that B is indefinite.

If the infimum is attainable, then $\operatorname{trace}(X^{\mathrm{H}}AX)$ as a function of X restricted to $X^{\mathrm{H}}BX=J_k$ has a (local) minimum. By item 2 of Theorem 1.1, $A-\lambda B$ is a positive semi-definite matrix pencil.

Consider the case when the infimum is not attainable. Perturb A to $A_\epsilon:=A+\epsilon I$, where $\epsilon>0$, and define

$$f_\epsilon(X):=\operatorname{trace}(X^{\mathrm{H}}A_\epsilon X)=\operatorname{trace}(X^{\mathrm{H}}AX)+\epsilon\|X\|_{\mathrm{F}}^2\geqslant\operatorname{trace}(X^{\mathrm{H}}AX),$$

where $\|X\|_{\mathrm{F}}$ is X's Frobenius norm. We have for any given $\epsilon>0$

$$\inf_{X^{\mathrm{H}}BX=J_k}f_\epsilon(X)\geqslant\inf_{X^{\mathrm{H}}BX=J_k}\operatorname{trace}(X^{\mathrm{H}}AX)>-\infty.\tag{3.20}$$

We claim $\inf f_\epsilon(X)$ subject to $X^{\mathrm{H}}BX = J_k$ can be attained. In fact, let $X^{(i)}$ be a sequence such that

$$(X^{(i)})^{\mathrm{H}}BX^{(i)} = J_k, \quad \lim_{i\to\infty} f_\epsilon(X^{(i)}) = \inf_{X^{\mathrm{H}}BX=J_k} f_\epsilon(X). \tag{3.21}$$

$\{X^{(i)}\}$ is a bounded sequence; otherwise

$$\lim_{i\to\infty} f_\epsilon(X^{(i)}) \geqslant \inf_{X^{\mathrm{H}}BX=J_k} \operatorname{trace}(X^{\mathrm{H}}AX) + \limsup_{i\to\infty} \epsilon\|X^{(i)}\|_{\mathrm{F}}^2 = +\infty,$$

contradicting (3.20) and (3.21). So for any given $\epsilon > 0$, $A_\epsilon - \lambda B$ is a positive semi-definite pencil, which means for every $\epsilon > 0$, there is $\lambda_\epsilon \in \mathbb{R}$ such that $A_\epsilon - \lambda_\epsilon B \succeq 0$. Pick a sequence $\{\epsilon_i > 0\}$ that converges to 0 as $i \to \infty$. We claim that $\{\lambda_{\epsilon_i}\}$ is a bounded sequence which then must have a convergent subsequence converging to, say λ_0. Through renaming, we may assume the sequence itself is the subsequence. Then let $i \to \infty$ on $A_{\epsilon_i} - \lambda_{\epsilon_i}B \succeq 0$ to conclude that $A - \lambda_0 B \succeq 0$, i.e., $A - \lambda B$ is a positive semi-definite matrix pencil. We have to show that $\{\lambda_{\epsilon_i}\}$ is bounded. To this end, it suffices to show $\{\lambda_\epsilon : 0 < \epsilon \leqslant 1\}$ is bounded. Since $A_\epsilon - \lambda B$ is a positive semi-definite matrix pencil of order n, its eigenvalues are real and can be ordered as, by Lemma 3.8,

$$\lambda_{n_-}^-(\epsilon) \leqslant \cdots \leqslant \lambda_1^-(\epsilon) \leqslant \lambda_1^+(\epsilon) \leqslant \cdots \leqslant \lambda_{n_+}^+(\epsilon),$$

and $\lambda_1^-(\epsilon) \leqslant \lambda_\epsilon \leqslant \lambda_1^+(\epsilon)$. Therefore for $0 < \epsilon \leqslant 1$

$$|\lambda_\epsilon| \leqslant \|B^{-1}A_\epsilon\|_{\mathrm{F}} \leqslant \|B^{-1}A\|_{\mathrm{F}} + \|B^{-1}\|_{\mathrm{F}},$$

as was to be shown. □

Proof of Theorem 2.1. To prove item 1 (which is the item 1 of Theorem 1.1 without assuming $A - \lambda B$ is regular, let alone B is nonsingular), we complement[8] X by X_{c} to a nonsingular $X_1 = [X, X_{\mathrm{c}}] \in \mathbb{C}^{n\times n}$ such that

$$X_1^{\mathrm{H}}BX_1 = \begin{array}{c|cc} & k & n-k \\ \hline k & J_k & 0 \\ n-k & 0 & B_{\mathrm{c}} \end{array}, \quad X_1^{\mathrm{H}}AX_1 = \begin{array}{c|cc} & k & n-k \\ \hline k & A_{11} & A_{12} \\ n-k & A_{21} & A_{22} \end{array}. \tag{3.22}$$

For any $\gamma \in \mathbb{R}$ that makes $A_{11} - \gamma J_k$ nonsingular, let

$$Z = \begin{bmatrix} I_k & -(A_{11} - \gamma J_k)^{-1}A_{12} \\ 0 & I_{n-k} \end{bmatrix},$$

[8] X_{c} can be found as follows. Since X has full column rank, we can expand it to a nonsingular $Y = [X, \hat{X}_{\mathrm{c}}] \in \mathbb{C}^{n\times n}$. Partition $Y^{\mathrm{H}}BY = \begin{bmatrix} J_k & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ and let $Y_1^{\mathrm{H}} = \begin{bmatrix} I_k & 0 \\ -B_{21}J_k & I_{n-k} \end{bmatrix}$ to get

$$Y_1^{\mathrm{H}}Y^{\mathrm{H}}BYY_1 = \begin{bmatrix} J_k & 0 \\ 0 & B_{22} - B_{21}J_kB_{12} \end{bmatrix}.$$

Notice $YY_1 = [X, \hat{X}_{\mathrm{c}} - XJ_kB_{12}]$. Set $B_{\mathrm{c}} = B_{22} - B_{21}J_kB_{12}$ and $X_{\mathrm{c}} = \hat{X}_{\mathrm{c}} - XJ_kB_{12}$ and thus $X_1 = YY_1$ to get the first equation in (3.22). The second equation is simply obtained by partitioning $X_1^{\mathrm{H}}AX_1$ accordingly.

then

$$Z^{\mathrm{H}}X_1^{\mathrm{H}}(A-\gamma B)X_1Z=\mathrm{diag}(A_{11}-\gamma J_k,\widehat{A}_{22}),$$

where $\widehat{A}_{22}=-A_{21}(A_{11}-\gamma J_k)^{-1}A_{12}+A_{22}-\gamma B_{\mathrm{c}}$. Thus,

$$\begin{aligned} i_-(A_{11}-\gamma J_k)\leqslant i_-(A-\gamma B)&=i_-(A_{11}-\gamma J_k)+i_-(\widehat{A}_{22}) &(3.23)\\ &\leqslant i_-(A_{11}-\gamma J_k)+n-k. &(3.24)\end{aligned}$$

Assume $\mu_i^+<\lambda_i^+$ for some i. Then there exists $\gamma\in(\mu_i^+,\lambda_i^+)$ such that $A_{11}-\gamma J_k$ is nonsingular. The number $\mathfrak{n}(\gamma)$ for $A_{11}-\lambda J_k$ as defined in item 3 of Lemma 3.8 is at least i, and therefore $i_-(A_{11}-\gamma J_k)\geqslant i$, and $\mathfrak{n}(\gamma)$ for $A-\lambda B$ is at most $i-1$, and therefore $i_-(A-\gamma B)\leqslant i-1$. This contradicts the inequality in (3.23).

Assume $\mu_i^+>\lambda_{i+n-k}^+$ for some i. Then there exists $\gamma\in\left(\lambda_{i+n-k}^+,\mu_i^+\right)$ such that $A_{11}-\gamma J_k$ is nonsingular. The number $\mathfrak{n}(\gamma)$ for $A_{11}-\lambda J_k$ as defined in item 3 of Lemma 3.8 is at most $i-1$, and therefore $i_-(A_{11}-\gamma J_k)\leqslant i-1$, and $\mathfrak{n}(\gamma)$ for $A-\lambda B$ is at least $i+n-k$, and therefore $i_-(A-\gamma B)\leqslant i+n-k$. This contradicts the inequality in (3.24).

This proves (1.7), and (1.8) can be proved in a similar way.

For item 2, the condition of Lemma 3.3 is satisfied by $A-\lambda B$ here. So we have (3.3) in which $A_2\succeq 0$ and $A_1-\lambda B_1$ is a positive semi-definite pencil with B_1 nonsingular. Now for any $X\in\mathbb{C}^{n\times k}$, write

$$\widehat{X}=Y^{-1}X=\begin{array}{c}{\scriptstyle r}\\{\scriptstyle n-r}\end{array}\!\!\begin{bmatrix}\widehat{X}_1\\ \widehat{X}_2\end{bmatrix}, \tag{3.25}$$

which gives $X^{\mathrm{H}}BX=\widehat{X}^{\mathrm{H}}Y^{\mathrm{H}}BY\widehat{X}=\widehat{X}_1^{\mathrm{H}}B_1\widehat{X}_1$, having nothing to do with $\widehat{X}_2$. Since the mapping $X\to\widehat{X}$ is one-one, we have

$$\begin{aligned}\inf_{X^{\mathrm{H}}BX=J_k}\mathrm{trace}(X^{\mathrm{H}}AX)&=\inf_{\widehat{X}_1^{\mathrm{H}}B_1\widehat{X}_1=J_k}\mathrm{trace}\left(\begin{bmatrix}\widehat{X}_1\\ \widehat{X}_2\end{bmatrix}^{\mathrm{H}}\begin{bmatrix}A_1&0\\0&A_2\end{bmatrix}\begin{bmatrix}\widehat{X}_1\\ \widehat{X}_2\end{bmatrix}\right)\\ &=\inf_{\widehat{X}_1^{\mathrm{H}}B_1\widehat{X}_1=J_k}\left[\mathrm{trace}(\widehat{X}_1^{\mathrm{H}}A_1\widehat{X}_1)+\mathrm{trace}(\widehat{X}_2^{\mathrm{H}}A_2\widehat{X}_2)\right]\\ &=\inf_{\widehat{X}_1^{\mathrm{H}}B_1\widehat{X}_1=J_k}\mathrm{trace}(\widehat{X}_1^{\mathrm{H}}A_1\widehat{X}_1)+\inf_{\widehat{X}_2}\mathrm{trace}(\widehat{X}_2^{\mathrm{H}}A_2\widehat{X}_2)\\ &=\inf_{\widehat{X}_1^{\mathrm{H}}B_1\widehat{X}_1=J_k}\mathrm{trace}(\widehat{X}_1^{\mathrm{H}}A_1\widehat{X}_1). \end{aligned}\tag{3.26}$$

The last equality is due to $A_2\succeq 0$ and is attained by any $\widehat{X}_2$ satisfying $\mathcal{R}(\widehat{X}_2)\subseteq\mathcal{N}(A_2)$. Theorem 1.1 is applicable to $A_1-\lambda B_1$ and the application gives, by (3.26),

$$\begin{aligned}\inf_{X^{\mathrm{H}}BX=J_k}\mathrm{trace}(X^{\mathrm{H}}AX)&=\inf_{\widehat{X}_1^{\mathrm{H}}B_1\widehat{X}_1=J_k}\mathrm{trace}\left(\widehat{X}_1^{\mathrm{H}}A_1\widehat{X}_1\right)\\ &=\sum_{i=1}^{k_+}\lambda_i^+-\sum_{i=1}^{k_-}\lambda_i^-,\end{aligned}$$

as expected. Track each equal sign in the above equations to conclude the claims in items 2(a,b,c). This proved item 2.

For item 3, item 2 implies that the condition (1.10) is necessary. We have to prove that it is sufficient, too. Suppose (1.10) is true. By Lemma 3.6, the condition (3.2) of Lemma 3.2 is satisfied. So we have (3.3), (3.25), and

$$\inf_{X^{\mathrm{H}}BX=J_k} \operatorname{trace}(X^{\mathrm{H}}AX) = \inf_{\widehat{X}_1^{\mathrm{H}}B_1\widehat{X}_1=J_k} \operatorname{trace}\left(\widehat{X}_1^{\mathrm{H}}A_1\widehat{X}_1\right) + \inf_{\widehat{X}_2} \operatorname{trace}\left(\widehat{X}_2^{\mathrm{H}}A_2\widehat{X}_2\right)$$

which is bounded from below. Therefore

$$A_2 \succeq 0, \quad \inf_{\widehat{X}_1^{\mathrm{H}}B_1\widehat{X}_1=J_k} \operatorname{trace}(\widehat{X}_1^{\mathrm{H}}A_1\widehat{X}_1) > -\infty. \tag{3.27}$$

Since B_1 is nonsingular, Lemma 3.9 says that $A_1 - \lambda B_1$ is a positive semi-definite matrix pencil by the second inequality in (3.27). Therefore $Y^{\mathrm{H}}AY - \lambda Y^{\mathrm{H}}BY$ is, too; so is $A - \lambda B$.

Now we turn to item 4. In what follows, we first use Lagrange's multiplier method, similar to [4] in proving its Theorem 3.5 there, to show that $A - \lambda B$ is a positive semi-definite pencil. Since $X^{\mathrm{H}}BX = J_k$ provides k^2 independent constraints on X (in $\mathbb{R}$), we can use a $k \times k$ Hermitian matrix Λ which has k^2 degrees of freedom to express Lagrange's function as [9]

$$\mathscr{L}(X) = \operatorname{trace}(X^{\mathrm{H}}AX) - \langle \Lambda, X^{\mathrm{H}}BX - J_k \rangle.$$

The gradient of $\mathscr{L}$ at X is

$$\nabla \mathscr{L}(X) = 2(AX - BX\Lambda).$$

Therefore for any local minimal point X_0, there exists a group of Lagrange's multipliers, i.e., some Hermitian $\Lambda_0 \in \mathbb{C}^{k\times k}$ such that

$$AX_0 = BX_0\Lambda_0, \quad X_0^{\mathrm{H}}BX_0 = J_k. \tag{3.28}$$

Without loss of generality, we may assume that Λ_0 is diagonal. Here is why. Pre-multiply the first equation in (3.28) by X_0^{H} to get $X_0^{\mathrm{H}}AX_0 = J_k\Lambda_0$. Therefore $J_k\Lambda_0 = (J_k\Lambda_0)^{\mathrm{H}} = \Lambda_0 J_k$ which implies Λ_0 is block diagonal, i.e., $\Lambda_0 = \Lambda_{0+} \oplus \Lambda_{0-}$, where $\Lambda_{0\pm} \in \mathbb{C}^{k_\pm \times k_\pm}$ are Hermitian. Hence there exists a block diagonal unitary $V = V_{0+} \oplus V_{0-}$ such that $V^{\mathrm{H}}\Lambda_0 V$ is diagonal, where $V_{0\pm} \in \mathbb{C}^{k_\pm\times k_\pm}$ are unitary. So $V^{\mathrm{H}}J_kV = J_k$, and thus we have by (3.28)

$$A(X_0V) = B(X_0V)(V^{\mathrm{H}}\Lambda_0V), \quad (VX_0)^{\mathrm{H}}B(X_0V) = J_k.$$

It can also be seen that X_0V is a local minimal point, too. Assume Λ_0 is diagonal, and write

$$\Lambda_0 = \operatorname{diag}(\omega_1^+, \ldots, \omega_{k_+}^+, \omega_{k_-}^-, \ldots, \omega_1^-), \tag{3.29a}$$

$$\omega_{k_-}^- \leqslant \cdots \leqslant \omega_1^-, \quad \omega_1^+ \leqslant \cdots \leqslant \omega_{k_+}^+. \tag{3.29b}$$

Since X_0 is a local minimal point as assumed, the second derivative $D^2\mathscr{L}(X)$ at X_0, taken as a quadratic form and restricted to the tangent space of

$$\mathbb{S} = \{X \in \mathbb{C}^{n\times k} \mid X^{\mathrm{H}}BX = J_k\},$$

[9] The standard inner product $\langle X, Y\rangle$ for matrices of compatible sizes is defined as $\langle X, Y\rangle = \mathrm{Re}(\operatorname{trace}(X^{\mathrm{H}}Y))$, the real part of $\operatorname{trace}(X^{\mathrm{H}}Y)$.

must be nonnegative, i.e.,

$$\operatorname{trace}(W^{\mathrm{H}}AW) - \langle \Lambda_0, W^{\mathrm{H}}BW\rangle \geqslant 0 \tag{3.30}$$

for any $W \in \mathbb{C}^{n\times k}$ satisfying

$$X_0^{\mathrm{H}}BW + W^{\mathrm{H}}BX_0 = 0. \tag{3.31}$$

Complement X_0 by X_{c} to a nonsingular $X_1 = [X_0, X_{\mathrm{c}}] \in \mathbb{C}^{n\times n}$ such that

$$X_1^{\mathrm{H}}BX_1 = \begin{bmatrix} J_k & 0 \\ 0 & B_{\mathrm{c}} \end{bmatrix}. \tag{3.32}$$

Thus $X_{\mathrm{c}}^{\mathrm{H}}BX_0 = 0$ and $X_{\mathrm{c}}^{\mathrm{H}}AX_0 = X_{\mathrm{c}}^{\mathrm{H}}BX_0\Lambda_0 = 0$ by (3.28). Therefore

$$X_1^{\mathrm{H}}AX_1 = \begin{bmatrix} X_0^{\mathrm{H}}AX_0 & 0 \\ 0 & X_{\mathrm{c}}^{\mathrm{H}}AX_{\mathrm{c}} \end{bmatrix} = \begin{bmatrix} J_k\Lambda_0 & 0 \\ 0 & X_{\mathrm{c}}^{\mathrm{H}}AX_{\mathrm{c}} \end{bmatrix}. \tag{3.33}$$

Rewrite (3.31) as $X_0^{\mathrm{H}}X_1^{-\mathrm{H}}X_1^{\mathrm{H}}BX_1X_1^{-1}W + W^{\mathrm{H}}X_1^{-\mathrm{H}}X_1^{\mathrm{H}}BX_1X_1^{-1}X_0 = 0$ and partition

$$X_1^{-1}W = \begin{array}{c} {\scriptstyle k} \\ {\scriptstyle n-k} \end{array}\begin{bmatrix} \widehat{W}_1 \\ \widehat{W}_2 \end{bmatrix}, \quad X_1^{-1}X_0 = \begin{bmatrix} I_k \\ 0 \end{bmatrix}.$$

to get

$$J_k\widehat{W}_1 + \widehat{W}_1^{\mathrm{H}}J_k = 0 \tag{3.34}$$

which says $S := J_k\widehat{W}_1$ is skew-Hermitian. We have $\widehat{W}_1 = J_kS$ which gives all possible $\widehat{W}_1$ that satisfies (3.34) as S runs through all possible $k\times k$ skew-Hermitian matrices. From (3.30), we have for any $\widehat{W}_2$ and $S = -S^{\mathrm{H}}$

$$\begin{aligned} 0 &\leqslant \operatorname{trace}(W^{\mathrm{H}}AW) - \langle \Lambda_0, W^{\mathrm{H}}BW\rangle \\ &= \operatorname{trace}\left(W^{\mathrm{H}}X_1^{-\mathrm{H}}X_1^{\mathrm{H}}AX_1X_1^{-1}W\right) - \langle \Lambda_0, W^{\mathrm{H}}X_1^{-\mathrm{H}}X_1^{\mathrm{H}}BX_1X_1^{-1}W\rangle \\ &= \operatorname{trace}\left(\widehat{W}_1^{\mathrm{H}}(J_k\Lambda_0)\widehat{W}_1\right) + \operatorname{trace}\left(\widehat{W}_2^{\mathrm{H}}\left(X_{\mathrm{c}}^{\mathrm{H}}AX_{\mathrm{c}}\right)\widehat{W}_2\right) - \langle \Lambda_0, \widehat{W}_1^{\mathrm{H}}J_k\widehat{W}_1 + \widehat{W}_2^{\mathrm{H}}B_{\mathrm{c}}\widehat{W}_2\rangle \\ &= -\operatorname{trace}(S\Lambda_0J_kS) + \operatorname{trace}\left(\widehat{W}_2^{\mathrm{H}}X_{\mathrm{c}}^{\mathrm{H}}AX_{\mathrm{c}}\widehat{W}_2\right) - \langle \Lambda_0, -SJ_kS + \widehat{W}_2^{\mathrm{H}}B_{\mathrm{c}}\widehat{W}_2\rangle. \end{aligned} \tag{3.35}$$

This is true for any $\widehat{W}_2$ and $S = -S^{\mathrm{H}}$. Recall (3.29). For any given $i \leqslant k_+$ and $j \leqslant k_-$, set $\widehat{W}_2 = 0$ and $S = e_ie_{k+1-j}^{\mathrm{H}} - e_{k+1-j}e_i^{\mathrm{H}}$ in (3.35) to get

$$0 \leqslant -\operatorname{trace}(S\Lambda_0J_kS) + \operatorname{trace}(\Lambda_0SJ_kS) = 2\left(\omega_i^+ - \omega_j^-\right).$$

Therefore for any ω_0 such that $\omega_1^- \leqslant \omega_0 \leqslant \omega_1^+$,

$$X_0^{\mathrm{H}}AX_0 - \omega_0J_k = J_k\Lambda_0 - \omega_0J_k = J_k(\Lambda_0 - \omega_0I) \succeq 0.$$

On the other hand, for any given $w \in \mathbb{C}^{n-k}$ and $i \leqslant k$, set $S = 0$ and $\widehat{W}_2 = we_i^{\mathrm{H}}$ in (3.35) to get

$$0 \leqslant \operatorname{trace}\left(e_i w^{\mathrm{H}} X_{\mathrm{c}}^{\mathrm{H}} A X_{\mathrm{c}} w e_i^{\mathrm{H}}\right) - \operatorname{Re}\left(\operatorname{trace}\left(\Lambda_0 e_i w^{\mathrm{H}} B_{\mathrm{c}} w e_i^{\mathrm{H}}\right)\right) = w^{\mathrm{H}}\left(X_{\mathrm{c}}^{\mathrm{H}} A X_{\mathrm{c}} - \omega B_{\mathrm{c}}\right) w,$$

where $\omega = e_i^{\mathrm{H}} \Lambda_0 e_i$ which is one of $\omega_j^{\pm}$. Since i and w are arbitrary, $X_{\mathrm{c}}^{\mathrm{H}} A X_{\mathrm{c}} - \omega B_{\mathrm{c}} \succeq 0$ for any $\omega \in \{\omega_j^{\pm},\ 1 \leqslant j \leqslant k_{\pm}\}$. This[10] implies $X_{\mathrm{c}}^{\mathrm{H}} A X_{\mathrm{c}} - \omega B_{\mathrm{c}} \succeq 0$ for any $\omega_{k_-}^{-} \leqslant \omega \leqslant \omega_{k_+}^{+}$. In particular, $X_{\mathrm{c}}^{\mathrm{H}} A X_{\mathrm{c}} - \omega_0 B_{\mathrm{c}} \succeq 0$. By (3.32) and (3.33), we conclude that $A - \omega_0 B \succeq 0$ for $\omega_1^{-} \leqslant \omega_0 \leqslant \omega_1^{+}$. That means $A - \lambda B$ is a positive semi-definite pencil.

It remains to show that X_0 is also a global minimizer. Since $A - \lambda B$ is a positive semi-definite pencil, by Lemma 3.3, we have (3.3). Define the one-one mapping between X and $\widehat{X}$ by (3.25). We have

$$\operatorname{trace}(X^{\mathrm{H}} A X) = \operatorname{trace}\left(\widehat{X}_1^{\mathrm{H}} A_1 \widehat{X}_1\right) + \operatorname{trace}\left(\widehat{X}_2^{\mathrm{H}} A_2 \widehat{X}_2\right).$$

Notice

$$\{X \in \mathbb{C}^{n \times k} \ : \ X^{\mathrm{H}} B X = J_k\} = Y \cdot \left\{\begin{bmatrix} \widehat{X}_1 \\ \widehat{X}_2 \end{bmatrix} \in \mathbb{C}^{n \times k} \ : \ \widehat{X}_1^{\mathrm{H}} B_1 \widehat{X}_1 = J_k\right\}$$

which places no constraint on $\widehat{X}_2$. If $\operatorname{trace}(X^{\mathrm{H}} A X)$ as a function of X restricted to $X^{\mathrm{H}} B X = J_k$ has a local minimum, then either $r = n$ or $r < n$ and $A_2 \succeq 0$. In the case $r = n$, B is invertible and the theorem is already proved in [4] (see Theorem 1.1). Suppose $r < n$ and thus $A_2 \succeq 0$. At any local minimizer $X_{\min}$, the corresponding $\widehat{X}_{\min}$ is

$$\widehat{X}_{\min} = Y^{-1} X_{\min} = \begin{array}{c} r \\ n-r \end{array}\begin{bmatrix} \widehat{X}_{\min,1} \\ \widehat{X}_{\min,2} \end{bmatrix}.$$

We have $\widehat{X}_{\min,2}^{\mathrm{H}} A_2 \widehat{X}_{\min,2} = 0$. Consequently $\widehat{X}_{\min,1}$ is a local minimizer of $\operatorname{trace}(\widehat{X}_1^{\mathrm{H}} A_1 \widehat{X}_1)$ as a function of $\widehat{X}_1$ restricted to $\widehat{X}_1^{\mathrm{H}} B_1 \widehat{X}_1 = J_k$. Since B_1 is nonsingular, item 4 of Theorem 1.1 is applicable and leads to that $\widehat{X}_{\min,1}$ is a global minimizer for $\operatorname{trace}(\widehat{X}_1^{\mathrm{H}} A_1 \widehat{X}_1)$. This in turn implies that $X_{\min}$ is a global minimizer for $\operatorname{trace}(X^{\mathrm{H}} A X)$ as a function of X restricted to $X^{\mathrm{H}} B X = J_k$. $\square$

Proof of Theorem 2.2. The basic idea is to essentially reduce the current case to the case in which B is nonsingular.

For item 1, we note that if $A - \lambda B$ is positive definite, then we have (3.3) with $A_2 \succ 0$. For any $x \in \mathbb{C}^n$, write

$$\hat{x} = Y^{-1} x = \begin{array}{c} r \\ n-r \end{array}\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}, \tag{3.36}$$

[10] For two Hermitian matrices M and N of the same size and $\alpha < \beta$, if $M - \gamma N \succeq 0$ for $\gamma = \alpha$ and $\gamma = \beta$, then $M - \gamma N \succeq 0$ for any $\alpha \leqslant \gamma \leqslant \beta$. In fact, any $\alpha \leqslant \gamma \leqslant \beta$ can be written as $\gamma = t\alpha + (1-t)\beta$ for some $0 \leqslant t \leqslant 1$ and therefore

$$M - \gamma N = t(M - \alpha N) + (1-t)(M - \beta N) \succeq 0.$$

which gives $x^{\mathrm{H}}Bx = \hat{x}^{\mathrm{H}}Y^{\mathrm{H}}BY\hat{x} = \hat{x}_1^{\mathrm{H}}B_1\hat{x}_1$. Since the mapping $x \to \hat{x}$ is one-one and since $A_2 \succ 0$, we have

$$\inf_{x^{\mathrm{H}}Bx=1} x^{\mathrm{H}}Ax = \inf_{\hat{x}_1^{\mathrm{H}}B_1\hat{x}_1=1} \hat{x}_1^{\mathrm{H}}A_1\hat{x}_1, \quad \inf_{x^{\mathrm{H}}Bx=-1} x^{\mathrm{H}}Ax = \inf_{\hat{x}_1^{\mathrm{H}}B_1\hat{x}_1=-1} \hat{x}_1^{\mathrm{H}}A_1\hat{x}_1. \tag{3.37}$$

On the other hand, if the infimums in (2.5) are attainable, then $A - \lambda B$ is positive semi-definite by Theorem 2.1 and thus we also have (3.3) with $A_2 \succeq 0$ and thus (3.37). But $A - \lambda B$ is assumed regular; A_2 must not be singular and so $A_2 \succ 0$. Either way, the problem is reduced to the one about $A_1 - \lambda B_1$. Apply [4, Corollary 3.7] to conclude the proof.

For item 2, pick a $\lambda_0 \in \mathscr{I}$, then $X^{\mathrm{H}}AX - \lambda_0 J_k \succ 0$ for all X satisfying $X^{\mathrm{H}}BX = J_k$. Therefore

$$\begin{aligned} &\inf_{X^{\mathrm{H}}BX=J_k} \operatorname{trace}(X^{\mathrm{H}}AX - \lambda J_k) \geqslant 0 \\ \Rightarrow \quad &\inf_{X^{\mathrm{H}}BX=J_k} \operatorname{trace}(X^{\mathrm{H}}AX) \geqslant \lambda_0(k_+ - k_-) > -\infty, \end{aligned}$$

implying that $A - \lambda B$ is positive semi-definite by Theorem 2.1. Hence we have (3.3) with $A_2 \succeq 0$. But $A - \lambda B$ is assumed regular; A_2 must not be singular and so $A_2 \succ 0$. Again the problem is reduced to the one about $A_1 - \lambda B_1$. Apply [4, Theorem 3.10] to conclude the proof. □

4. A sufficient and necessary condition for infimum attainability

Both Theorems 1.1 and 2.1 imply that for a positive semi-definite pencil $A - \lambda B$ the infimum is attainable if and only there is an eigenvector matrix $X_{\min} \in \mathbb{C}^{n\times k}$ such that

$$X_{\min}^{\mathrm{H}}BX_{\min} = J_k, \quad AX_{\min} = BX_{\min} \operatorname{diag}\left(\lambda_{k_+}^+, \ldots, \lambda_1^+, \lambda_1^-, \ldots, \lambda_{k_-}^-\right).$$

In this section, we shall use the indices in the canonical form of $A - \lambda B$ as given in Lemma 3.8 to derive another sufficient and necessary condition.

Throughout this section, $A - \lambda B$ is a Hermitian positive semi-definite pencil of order n. Recall, in Lemma 3.8, the finite eigenvalues of $A - \lambda B$ are

$$\begin{aligned} \lambda_{n_-}^- \leqslant \cdots \leqslant \lambda_{m_0+m_-+1}^- < \underbrace{\lambda_0 = \cdots = \lambda_0}_{m_0} = \underbrace{\lambda_0 = \cdots = \lambda_0}_{m_-} = \\ = \underbrace{\lambda_0 = \cdots = \lambda_0}_{m_+} = \underbrace{\lambda_0 = \cdots = \lambda_0}_{m_0} < \lambda_{m_0+m_++1}^+ \leqslant \cdots \leqslant \lambda_{n_+}^+. \end{aligned} \tag{3.15}$$

In particular $\lambda_i^- = \lambda_0$ for $1 \leqslant i \leqslant m_0 + m_-$ and $\lambda_i^+ = \lambda_0$ for $1 \leqslant i \leqslant m_0 + m_+$. By Lemma 3.8, m_0 and $m_\pm$ are uniquely determined by $A - \lambda B$.

Lemma 4.1. *Suppose $A - \lambda B$ is regular. Let $Y \in \mathbb{C}^{n\times \ell}$ that satisfies $Y^{\mathrm{H}}BY = I_\ell$ be an eigenvector matrix of $A - \lambda B$ associated with λ_0 (i.e., each column of Y is an eigenvector). Then $\ell \leqslant m_+$.*

Proof. By Lemma 3.8, $A - \lambda B$ has $m_+ + m_- + m_0$ linearly independent eigenvectors associated with λ_0. One set of them can be chosen according to the three sources: $x_1^-, \ldots, x_{m_-}^-$ from **source 1**, $x_1^+, \ldots, x_{m_+}^+$ from **source 2**, and $x_1, \ldots, x_{m_0}$ from **source 3** such that

$$X^{\mathrm{H}}BX = I_{m_+} \oplus (-I_{m_-}) \oplus 0_{m_0\times m_0},$$

where $X = \left[x_1^+, \ldots, x_{m_+}^+, x_1^-, \ldots, x_{m_-}^- x_1, \ldots, x_{m_0}\right]$. Any eigenvector matrix $Y \in \mathbb{C}^{n\times\ell}$ associated with λ_0 can be expressed as $Y = XZ$ for some $Z \in \mathbb{C}^{(m_+ + m_- + m_0)\times\ell}$. Then $Y^{\mathrm{H}}BY = I_\ell$ is equivalent to

$$Z^{\mathrm{H}} \begin{bmatrix} I_{m_+} & 0 & 0 \\ 0 & -I_{m_-} & 0 \\ 0 & 0 & 0 \end{bmatrix} Z = I_\ell$$

which implies $\ell \leqslant m_+$. $\square$

Theorem 4.1. *Let $A - \lambda B$ be a Hermitian positive semi-definite pencil of order n. Then*

$$\inf_{X^{\mathrm{H}}BX=J_k} \operatorname{trace}(X^{\mathrm{H}}AX) = \sum_{i=1}^{k_+} \lambda_i^+ - \sum_{i=1}^{k_-} \lambda_i^-$$

is attainable if and only if $m_0 = 0$ or $k_\pm \leqslant m_\pm$ in the case of $m_0 > 0$.

Proof. We have (3.25) and (3.26). It can be seen that the infimums in

$$\inf_{X^{\mathrm{H}}BX=J_k} \operatorname{trace}(X^{\mathrm{H}}AX), \quad \inf_{\widehat{X}_1^{\mathrm{H}}B_1\widehat{X}_1=J_k} \operatorname{trace}\left(\widehat{X}_1^{\mathrm{H}}A_1\widehat{X}_1\right)$$

are either both attainable or neither is. Also m_0 and $m_\pm$ are the same for $A - \lambda B$ and the reduced $A_1 - \lambda B_1$. So without loss of generality, we assume B is nonsingular.

Suppose $m_0 = 0$ or $k_\pm \leqslant m_\pm$ in the case of $m_0 > 0$. The above analysis indicates that there are $k_+ + k_-$ eigenvectors associated with the eigenvalues λ_i^-, λ_j^+ for $1 \leqslant i \leqslant k_-$, $1 \leqslant j \leqslant k_+$. Put these eigenvectors side-by-side with those for λ_j^+ first and then those for λ_i^- to give a matrix X that satisfies $X^{\mathrm{H}}BX = J_k$ and at the same time

$$\operatorname{trace}(X^{\mathrm{H}}AX) = \sum_{i=1}^{k_+} \lambda_i^+ - \sum_{i=1}^{k_-} \lambda_i^-.$$

Suppose now the infimum is attainable. For any $X \in \mathbb{C}^{n\times k}$, partition $X = [X_+, X_-]$, where $X_\pm \in \mathbb{C}^{n\times k_\pm}$. $X^{\mathrm{H}}BX = J_k$ is equivalent to $X_+^{\mathrm{H}}BX_+ = I_{k_+}$, $X_-^{\mathrm{H}}BX_- = -I_{k_-}$, and $X_+^{\mathrm{H}}BX_- = 0$. We have

$$\begin{aligned}
\sum_{i=1}^{k_+} \lambda_i^+ - \sum_{i=1}^{k_-} \lambda_i^- &= \inf_{X^{\mathrm{H}}BX=J_k} \operatorname{trace}(X^{\mathrm{H}}AX) && (4.1)\\
&= \inf_{\substack{X_+^{\mathrm{H}}BX_+=I_{k_+},\, X_-^{\mathrm{H}}BX_-=-I_{k_-}\\ X_+^{\mathrm{H}}BX_-=0}} \left[\operatorname{trace}\left(X_+^{\mathrm{H}}AX_+\right) + \operatorname{trace}\left(X_-^{\mathrm{H}}AX_-\right)\right] \\
&\geqslant \inf_{X_+^{\mathrm{H}}BX_+=I_{k_+},\, X_-^{\mathrm{H}}BX_-=-I_{k_-}} \left[\operatorname{trace}\left(X_+^{\mathrm{H}}AX_+\right) + \operatorname{trace}\left(X_-^{\mathrm{H}}AX_-\right)\right] \\
&= \inf_{X_+^{\mathrm{H}}BX_+=I_{k_+}} \operatorname{trace}\left(X_+^{\mathrm{H}}AX_+\right) + \inf_{X_-^{\mathrm{H}}BX_-=-I_{k_-}} \operatorname{trace}\left(X_-^{\mathrm{H}}AX_-\right) && (4.2)\\
&= \sum_{i=1}^{k_+} \lambda_i^+ - \sum_{i=1}^{k_-} \lambda_i^-.
\end{aligned}$$

Therefore for the infimum in (4.1) to be attainable, both infimums in (4.2) must be attainable. We claim that when $m_0 > 0$, if $k_+ > m_+$, $\inf_{X_+^{\mathrm{H}}BX_+=I_{k_+}} \operatorname{trace}\left(X_+^{\mathrm{H}}AX_+\right)$ is not attainable; similarly when $m_0 > 0$, if $k_- > m_-$, $\inf_{X_-^{\mathrm{H}}BX_-=-I_{k_-}} \operatorname{trace}\left(X_-^{\mathrm{H}}AX_-\right)$ is not attainable. The claim implies the necessity of the condition $m_0 = 0$ or $k_\pm \leqslant m_\pm$ in the case of $m_0 > 0$.

We shall consider the "+" case only; the other one is similar. Suppose that $m_0 > 0$ and $k_+ > m_+$ and assume to the contrary that there existed an $X_+ \in \mathbb{C}^{k_+\times k_+}$ such that $X_+^{\mathrm{H}}BX_+ = I_{k_+}$ and $\operatorname{trace}\left(X_+^{\mathrm{H}}AX_+\right) = \sum_{i=1}^{k_+}\lambda_i^+$. Since X_+ is a global minimizer, by Theorem 2.1 there existed a Hermitian $\Lambda_+ \in \mathbb{C}^{k_+\times k_+}$ such that

$$AX_+ = BX_+\Lambda_+, \quad X_+^{\mathrm{H}}BX_+ = I_{k_+}.$$

As a result, $X_+^{\mathrm{H}}AX_+ = \Lambda_+$. Let $\Lambda_+ = U^{\mathrm{H}}\Omega U$ be its eigendecomposition, where U is unitary, $\Omega = \operatorname{diag}(\omega_1, \dots, \omega_{k_+})$, and $\omega_1 \leqslant \cdots \leqslant \omega_{k_+}$. Write $Y = X_+U = (y_1, \dots, y_{k_+})$. We have

$$AY = BY\Omega, \quad Y^{\mathrm{H}}BY = I_{k_+},$$

which implies ω_i is an eigenvalue of $A - \lambda B$ and y_i is a corresponding eigenvector. Since

$$\sum_{i=1}^{k_+}\lambda_i^+ = \operatorname{trace}\left(X_+^{\mathrm{H}}AX_+\right) = \operatorname{trace}(Y^{\mathrm{H}}AY) = \operatorname{trace}(\Omega) = \sum_{i=1}^{k_+}\omega_i$$

and $\lambda_i^+ \leqslant \omega_i$ for $1 \leqslant i \leqslant k_+$ by [4, Theorem 2.1], we have $\lambda_i^+ = \omega_i$ for $1 \leqslant i \leqslant k_+$. Let $\ell = \min\{k_+, m_+ + m_0\}$ and $Y_1 = Y_{(:,1:\ell)}$, the submatrix consisting the first ℓ columns of Y. Since $m_0 > 0$ and $k_+ > m_+$, $\ell > m_+$. Y_1 is an eigenvector matrix associated with λ_0 with more than m_+ columns, and $Y_1^{\mathrm{H}}B_1Y = I_\ell$, contradicting Lemma 4.1. Thus $\inf_{X_+^{\mathrm{H}}BX_+=I_{k_+}} \operatorname{trace}\left(X_+^{\mathrm{H}}AX_+\right)$ is not attainable if $m_0 > 0$ and $k_+ > m_+$. $\square$

5. Conclusions

Given a Hermitian matrix pencil $A - \lambda B$ of order n, we are interested in when

$$\inf_{X^{\mathrm{H}}BX=J_k} \operatorname{trace}(X^{\mathrm{H}}AX) \tag{5.1}$$

is finite, attainable, and what it is when it is finite. The same questions were investigated in detail with remarkable results by Kovač-Striko and Veselić [4] for the case when B is nonsingular. They suspected that their results would be true without the nonsingularity assumption on B but with $A - \lambda B$ being regular. Our first contribution here is to confirm that indeed the nonsingularity assumption on B is not needed, but we also have gone further to allow the singular pencil into the picture. Our second contribution is a sufficient necessary condition for the attainability of the infimum in (5.1) in terms of certain indices in the canonical representation of the pencil.

References

[1] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[2] G.W. Stewart, J.-G. Sun, Matrix Perturbation Theory, Academic Press, Boston, 1990.
[3] A.H. Sameh, J.A. Wisniewski, A trace, minimization algorithm for the generalized eigenvalue problem, SIAM J. Numer. Anal. 19 (6) (1982) 1243–1259.
[4] J. Kovač-Striko, K. Veselić, Trace minimization and definiteness of symmetric pencils, Linear Algebra Appl. 216 (1995) 139–158.

[5] I. Gohberg, P. Lancaster, L. Rodman, Indefinite Linear Algebra and Applications, Birkhäuser, Basel, Switzerland, 2005.
[6] P. Binding, B. Najman, Q. Ye, A variational principle for eigenvalues of pencils of Hermitian matrices, Integral Equations Operator Theory 35 (1999) 398–422.
[7] P. Binding, Q. Ye, Variational principles for indefinite eigenvalue problems, Linear Algebra Appl. 218 (1995) 251–262.
[8] P. Lancaster, Q. Ye, Variational properties and Rayleigh quotient algorithms for symmetric matrix pencils, Oper. Theory Adv. Appl. 40 (1989) 247–278.
[9] B. Najman, Q. Ye, A minimax characterization of eigenvalues of Hermitian pencils, Linear Algebra Appl. 144 (1991) 217–230.
[10] B. Najman, Q. Ye, A minimax characterization of eigenvalues of Hermitian pencils II, Linear Algebra Appl. 191 (1993) 183–197.
[11] R.-C. Li, On perturbations of matrix pencils with real spectra, a revisit, Math. Comp. 72 (2003) 715–728.
[12] G.W. Stewart, Perturbation bounds for the definite generalized eigenvalue problem, Linear Algebra Appl. 23 (1979) 69–86.
[13] J.-G. Sun, A note on Stewart's theorem for definite matrix pairs, Linear Algebra Appl. 48 (1982) 331–339.
[14] J.-G. Sun, Perturbation bounds for eigenspaces of a definite matrix pair, Numer. Math. 41 (1983) 321–343.