

NEW PROOFS AND EXTENSIONS
OF SYLVESTER'S AND JOHNSON'S INERTIA THEOREMS
TO NON-HERMITIAN MATRICES

MAN KAM KWONG AND ANTON ZETTL

(Communicated by Ken Ono)

ABSTRACT. We present a new proof and extension of the classical Sylvester Inertia Theorem to a pair of non-Hermitian matrices which satisfies the property that any real linear combination of the pair has only real eigenvalues. In the proof, we embed the given problem in a one-parameter family of related problems and examine the eigencurves of the family. The proof requires only elementary matrix theory and the Intermediate Value Theorem. The same technique is then used to extend Johnson's extension of Sylvester's Theorem on possible values of the inertia of a product of two matrices.

1. INTRODUCTION

For $n > 1$, let $M_n(\mathbb{C})$ denote the $n \times n$ matrices over the complex numbers \mathbb{C} . For $A \in M_n(\mathbb{C})$ the inertia $i(A)$ of A is defined to be the integer triple $i(A) = (\pi(A), \nu(A), \delta(A))$ denoting the number of eigenvalues of A with positive, negative, or zero real parts, respectively, counting multiplicity. The real inertia $i_r(A) = (\pi_r(A), \nu_r(A), \delta_r(A))$ is defined in a similar way, but only real eigenvalues are counted. Obviously,

$$(1.1) \quad \pi_r(A) \leq \pi(A), \quad \nu_r(A) \leq \nu(A), \quad \delta_r(A) \leq \delta(A),$$

and

$$(1.2) \quad \pi(A) + \nu(A) + \delta(A) = n.$$

In this paper, we are mostly interested in matrices with only real eigenvalues. In this situation, the inertia and real inertia of A are identical.

The best known classical inertia theorem is that of Sylvester [6]:

Theorem 1 (Sylvester 1852). *If P is positive definite and A is Hermitian, then all eigenvalues of PA are real and $i(PA) = i(A)$.*

Received by the editors August 21, 2010.

2010 *Mathematics Subject Classification*. Primary 05C38, 15A15; Secondary 05A15, 15A18.

Key words and phrases. Matrix eigenvalues, positive and negative eigenvalues, eigenvalue curves.

Research of the first author is supported by the Hong Kong Research Grant Council grant B-Q21F.

The second author thanks the Department of Applied Mathematics of the Hong Kong Polytechnic University and especially the first author for the opportunity to visit the department in June 2010 when this project was completed.

©2011 American Mathematical Society
Reverts to public domain 28 years from publication

When A is invertible, Sylvester's Theorem is subsumed in a result of Johnson [5] which characterizes all possible values of the inertia of the product of two invertible Hermitian matrices. Johnson's Theorem is a special case of Theorem 3 in Section 4 for non-Hermitian matrices.

In this paper, we extend Sylvester's and Johnson's Theorems to certain pairs of non-Hermitian matrices, but our main emphasis is on the proofs. These are new and novel and consist of embedding the problem in a one-parameter family of problems and then using the eigencurves of these families to prove the theorems. A similar technique has been used before by Kong, Wu and Zettl, [4], to study left-definite Sturm-Liouville problems of second-order ordinary differential equations, a work that inspired the current one.

Another novel feature of these proofs is that they are elementary but use both algebra and analysis. One of the well-known proofs of Sylvester's result, due to Wimmer [7], uses the minimax characterization of the eigenvalues of Hermitian matrices. Since our result deals with non-Hermitian matrices, this variational approach is not applicable.

Let us first define a certain property for a pair of matrices:

(R): *A pair of matrices (A, B) in $M_n(\mathbb{C})$ is said to satisfy hypothesis (R), or just simply (R), if for each $t \in \mathbb{R}$, all eigenvalues $\lambda_j(t)$, $j = 1, 2, \dots, n$ of the matrix $(A - tB)$ are real.*

Note that if (A, B) satisfies (R), then (B, A) and $(A + \alpha B + \beta I, B)$, for any $\alpha, \beta \in \mathbb{R}$, also satisfy (R).

Any pair of Hermitian matrices satisfies (R). The converse is not true, as the simple example below shows:

$$(1.3) \quad A = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

We refer to $\lambda_j(t)$ as the eigencurves of $A - tB$. When ordered in ascending order,

$$\lambda_1(t) \leq \lambda_2(t) \leq \dots \leq \lambda_n(t), \quad t \in \mathbb{R},$$

they are continuous functions of t , defined for all $t \in \mathbb{R}$. Two curves may touch but not cross each other, in the sense that one curve that starts above another one always remains above the latter, except possibly at points where the two curves touch each other.

To help visualize these curves, Figure 1 shows the plot of a special example:

$$(1.4) \quad A = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -5 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

The two dashed lines are the t and λ axes. The four solid curves are the eigencurves. At first sight, one tends to think that there is one curve AH that diagonally cuts across the other three curves, BIE , CJF , and DKG . In fact, the correct interpretation is that λ_1 is made up of the two parts DK and KH , λ_2 is $CJKG$, λ_3 is $BIJF$, and λ_4 is AIE .

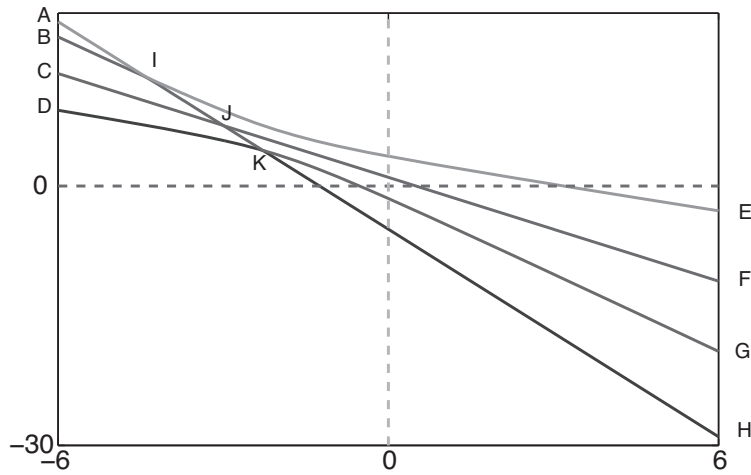


FIGURE 1. Eigenvalue curves $\lambda_1(t) \leq \lambda_2(t) \leq \lambda_3(t) \leq \lambda_4(t)$.

2. EXTENDING SYLVESTER'S THEOREM TO NON-HERMITIAN MATRICES

Our first result extends Sylvester's Theorem to pairs of matrices that satisfy (R). Indeed, if we take A and B in the next theorem to be Hermitian and $B^{-1} = P$, then it reduces to Sylvester's result.

Theorem 2. *Let $A, B \in M_n(\mathbb{C})$. If (A, B) satisfies (R) and all eigenvalues of B are positive, then all eigenvalues of $B^{-1}A$ are real and*

$$(2.1) \quad i(B^{-1}A) = i(AB^{-1}) = i(A).$$

Proof. $B^{-1}A$ and $AB^{-1} = B(B^{-1}A)B^{-1}$ have the same eigenvalues because they are similar matrices.

For each $t \in \mathbb{R}$ consider the family of matrix eigenvalue problems

$$(2.2) \quad (A - tB)u = \lambda u,$$

where u is a column vector of length n . Let $\lambda_j(t)$, $j = 1, 2, \dots, n$, denote the eigencurves of (2.2), arranged in nondecreasing order.

When one of these curves, say, $\lambda_1(t)$, crosses the t -axis, say, at $t = t_1$, it means that $\lambda_1(t_1) = 0$. Hence, there exists an eigenvector u such that $(A - t_1B)u = 0$, and hence $(B^{-1}A - t_1I)u = 0$. In other words, t_1 is an eigenvalue of $B^{-1}A$. The same arguments show that, conversely, every real eigenvalue of $B^{-1}A$ is given by a zero of one of the curves $\lambda_i(t)$. In other words, the set of all real eigenvalues of $B^{-1}A$ is equal to the set of all zeros of the functions $\lambda_i(t)$, $i = 1, \dots, n$.

When t is negatively very large, $A - tB$ is dominated by $-tB$ and so the eigenvalues of the former are dominated by those of the latter, which are all positive. This implies that the $\lambda_i(t)$ are positive when $t \rightarrow -\infty$. Likewise, when $t \rightarrow \infty$, the $\lambda_i(t)$ are dominated by the eigenvalues of $-tB$ and so are negative. By the Intermediate Value Theorem, each curve $\lambda_i(t)$ must intersect the t -axis at least once. It follows that $B^{-1}A$ must have at least n real eigenvalues. Since $B^{-1}A$ cannot have more than n eigenvalues, we conclude that it must have exactly n real eigenvalues and each $\lambda_i(t)$ must intersect the t -axis exactly once.

In order to show that $i(B^{-1}A) = i(A)$, we need to exhibit a one-to-one correspondence between the negative, zero, and positive eigenvalues, respectively, of the two matrices. The eigencurves intersect the λ -axis at the points $\lambda_1(0), \dots, \lambda_n(0)$. These are the eigenvalues of A . Suppose $\lambda_i(0)$ is negative. Since $\lambda_i(t) > 0$ for t near $-\infty$, the Intermediate Value Theorem shows that the curve $\lambda_i(t)$ must intersect the t -axis at some negative t , which is a negative eigenvalue of $B^{-1}A$. Likewise, if $\lambda_i(0)$ is positive, since $\lambda_i(t) < 0$ for t near ∞ , the curve $\lambda_i(t)$ must intersect the t -axis at some positive t , giving rise to a positive eigenvalue of $B^{-1}A$. Finally, if $\lambda_i(0) = 0$, then obviously 0 itself is an eigenvalue of $B^{-1}A$. \square

Corollary 1. *If (A, B) satisfies (R), and β_0 is the smallest eigenvalue of B , then*

$$(2.3) \quad i((B + \beta I)^{-1}A) = i(A) \quad \text{for all } \beta > \beta_0.$$

3. REMARKS ON THE PROPERTY (R)

The following simple example shows that if property (R) is not satisfied, the conclusion of Theorem 2 may not hold:

$$(3.1) \quad A = \begin{bmatrix} 10 & 5 \\ 8 & 7 \end{bmatrix}, \quad i(A) = (2, 0, 0)$$

because the two eigenvalues of A are 2 and 15;

$$(3.2) \quad B = \begin{bmatrix} 1 & 10 \\ 0 & 1 \end{bmatrix}, \quad B^{-1}A = \begin{bmatrix} -70 & -65 \\ 8 & 7 \end{bmatrix}, \quad i(B^{-1}A) = (0, 2, 0)$$

because the two eigenvalues of $B^{-1}A$ are -0.4798 and -62.5202 .

It is a natural and interesting question to ask for easily verifiable criteria for pairs of matrices to satisfy (R). Since eigenvalues are invariant under similarity transformation, namely, that $A - tB$ and $T^{-1}(A - tB)T$ have the same eigenvalues for any invertible matrix T , we may assume without loss of generality that B has been reduced to its Jordan form. In the case of real 2×2 matrices, the above question can be answered completely.

Lemma 1. *Suppose that A and B are 2×2 matrices and B is either in diagonal or upper triangular Jordan form. The pair (A, B) satisfies (R) if and only if one of the following three cases holds:*

- (1) B is a multiple of the identity matrix and both eigenvalues of A are real.
- (2) B is a diagonal matrix with distinct diagonal elements and $a_{12}a_{21} \geq 0$.
- (3) $B = \begin{bmatrix} b & 1 \\ 0 & b \end{bmatrix}$ and A is upper triangular.

Proof. This follows from the elementary fact that a 2×2 matrix $M = (m_{ij})$ has real eigenvalues if and only if $(m_{11} - m_{22})^2 + 4m_{12}m_{21} \geq 0$. We leave the details of the proof to the readers. \square

The problem of characterizing property (R) for higher dimensions is open. Some simple sufficient conditions can be derived from the following known result.

Lemma 2. For $n > 2$, consider a real tridiagonal matrix:

$$(3.3) \quad J = \begin{bmatrix} a_1 & \bar{a}_1 & 0 & & & \\ \underline{a}_1 & a_2 & \bar{a}_2 & & & \\ 0 & \underline{a}_2 & a_3 & \bar{a}_3 & & \\ & & & \ddots & & \\ & & & & a_{n-1} & \bar{a}_{n-1} \\ & & & & \underline{a}_{n-1} & a_n \end{bmatrix}.$$

Assume either

$$(3.4) \quad \bar{a}_k \underline{a}_k > 0, \quad \forall k = 1, 2, \dots, n - 1,$$

or

$$(3.5) \quad \bar{a}_k \underline{a}_k < 0, \quad \forall k = 1, 2, \dots, n - 1.$$

Then all the eigenvalues of J are real and distinct.

Proof. See Gantmacher and Krein [2], pp. 67-68. The proof makes use of the characteristic equations of successive principal submatrices of J . \square

Matrices of the form (3.3) are studied in [2] in connection with the oscillation of mechanical systems and are closely related to the Sturm-Liouville theory of second-order ordinary differential equations.

The assertion that all eigenvalues of J are real can also be deduced from the classical Sylvester Theorem and the easily proven fact that J can be written as the product DH of an appropriate diagonal matrix D and a Hermitian tridiagonal matrix H .

Corollary 2. If the strict inequalities in (3.4) and (3.5) are replaced respectively by \geq and \leq , i.e.,

$$(3.6) \quad \bar{a}_k \underline{a}_k \geq 0, \quad \forall k = 1, 2, \dots, n - 1,$$

or

$$(3.7) \quad \bar{a}_k \underline{a}_k \leq 0, \quad \forall k = 1, 2, \dots, n - 1.$$

then all the eigenvalues of J are real (but may not be distinct).

Proof. This follows easily from Lemma 1 by a continuity argument. \square

Corollary 3. Suppose that J has the form (3.3) satisfying either (3.6) or (3.7), and B has an analogous form satisfying

$$(3.8) \quad \bar{a}_k \underline{b}_k = \underline{a}_k \bar{b}_k, \quad \forall k = 1, 2, \dots, n - 1.$$

Then (J, B) satisfies (R).

Note that (3.8) is satisfied if $\bar{b}_k = \underline{b}_k = 0$. In particular, it is satisfied for all k if B is diagonal. In other words, if J satisfies (3.6) or (3.7) and B is diagonal, then (J, B) satisfies (R).

Corollary 1 has the following extension.

Lemma 3. Let J satisfy the hypothesis of Corollary 2, i.e., it has the form (3.3) and either (3.6) or (3.7) holds. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be any scalar function and K any matrix such that $K - f(J)$ is diagonal. Then all eigenvalues of K are real.

Proof. First let us assume that the more stringent condition (3.4) or (3.5) holds. As mentioned above, $J = DH$ can be represented as the product of a nonsingular diagonal matrix D and a Hermitian matrix H . It is easy to see that any positive power of J (and hence, any polynomial and any function of J) is D times a Hermitian matrix. Let $f(J) = DH_1$ for some Hermitian matrix H_1 . By assumption, $K = f(J) + D_1 = D(H_1 + D_2) = DH_2$, where D_1 is diagonal, $D_2 = D^{-1}D_1$ is also diagonal, and $H_2 = H_1 + D_2$ is Hermitian. By the classical Sylvester Theorem, all eigenvalues of DH_2 are real. The general case in which only the weaker condition (3.6) or (3.7) is assumed to hold can be deduced by a continuity argument. \square

In the above proof we notice the simple fact that if a pair of matrices (A, B) has the representation (PH_1, PH_2) , where P is a positive definite matrix and H_1 and H_2 are Hermitian matrices, then the pair satisfies (R). In such a situation, we do not really need to rely on Theorem 1 to arrive at the conclusion (2.1) because now $B^{-1}A = (PH_2)^{-1}PH_1 = H_2^{-1}H_1$ and we can invoke the classical Sylvester's Theorem instead.

This observation raises the question: Are all pairs of matrices (A, B) that satisfy (R) representable in this way? If the answer to this question is affirmative, the extension of Sylvester's result would have actually been vacuous.

Referring to Lemma 1, in the 2×2 case, apart from the more or less trivial cases (1) and (3), the generic case (2) does lead to an affirmative answer to the above question. Hence, for dimension 2, Theorem 2 has actually not added much value to the classical Theorem 1.

For higher dimensions, the next example demonstrates that Theorem 2 is a true extension of Theorem 1.

Lemma 4. *The pair of matrices*

$$(3.9) \quad A = \begin{bmatrix} 0 & 1 & \alpha \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

satisfies (R) if and only if $|\alpha| \leq \sqrt{32/27}$. Except when $\alpha = 0$, the pair is not representable in the form (PH_1, PH_2) described above.

Proof. We have used MAPLE to help us with the necessary symbolic computations.

The characteristic polynomial of $A - tB$ is

$$(3.10) \quad \lambda^3 + 6t\lambda^2 - (2 - 11t^2)\lambda + (6t^3 - 4t - \alpha).$$

One method to determine if a cubic equation $\lambda^3 + p\lambda^2 + q\lambda + r = 0$ has complex roots or not is to check whether the discriminant $\Delta = p^2q^2 - 4q^3 - 4p^3r - 27r^2 + 18pqr$ is negative or not. For our equation,

$$(3.11) \quad \Delta = 4t^6 + 24t^4 + 48t^2 + 32 - 27\alpha^2.$$

Since Δ is positive, when $|\alpha| \leq \sqrt{32/27}$, the eigenvalues of $A - tB$, being the roots of the characteristic polynomial, are real. When $|\alpha| > \sqrt{32/27}$, $\Delta < 0$ for $t = 0$ and so the eigenvalues of A are not all real.

Suppose $B = PH_2$. Then $H_2 = QB$, where $Q = P^{-1}$. It is a well-known fact that the product of two Hermitian matrices, here Q and B , is Hermitian if and only if the two Hermitian matrices commute with each other. Since B is diagonal with distinct diagonal elements, only diagonal matrices can commute with it. Hence, Q must be diagonal. Suppose also that $A = PH_1$. Then $H_1 = QA$. But it is easy

to see that when $\alpha \neq 0$, no diagonal matrix Q can make QA Hermitian. We thus arrive at a contradiction. \square

Given (A, B) satisfying (R), we can generate other pairs satisfying (R) using the following simple observation.

Lemma 5. *If B is nonsingular, then (A, B) satisfies (R) if and only if $(B^{-1}A, B^{-1})$ satisfies (R).*

Proof. For the “if” part, let $t \in \mathbb{R}$. The assumption that (A, B) satisfies (R) implies that $(A + tI, B)$ satisfies (R), where I is the $n \times n$ identity matrix. By Theorem 2, all eigenvalues of $B^{-1}(A + tI) = B^{-1}A + tB^{-1}$ are real. Since this is true for all t , $(B^{-1}A, B^{-1})$ satisfies (R). The converse is obtained by applying the “if” part to $(B^{-1}A, B^{-1})$. \square

Start with a given pair of (A, B) that satisfies (R) and choose any $\beta > \beta_0$, the smallest eigenvalue of B . Then $(A_1, B_1) = ((B + \beta I)^{-1}A, (B + \beta)^{-1})$ also satisfies (R). Repeating the algorithm on (A_1, B_1) we can get other pairs. We can also apply the algorithm to (B, A) or (B_1, A_1) instead.

4. EXTENDING JOHNSON'S THEOREM

When the positive definiteness assumption on P in Sylvester's Theorem is dropped, the inertia of PA is no longer exactly determinable; only lower bounds can be assigned. However, only certain configurations satisfying the bounds are realizable. Johnson [5] has determined all possible configurations of the inertia of the product of two nonsingular Hermitian matrices. A related result can be found in Dancis [1]. In this section we show that our technique can also be used to extend Johnson's result to matrices satisfying (R). Our result also covers the case when one of the factor matrices is not invertible.

Before describing the new result, it is instructive to retrace the proof of Theorem 2 when the roles of A and B are switched; namely, we now assume that the eigenvalues of A are all positive, while the eigenvalues of B can assume both positive and negative values.

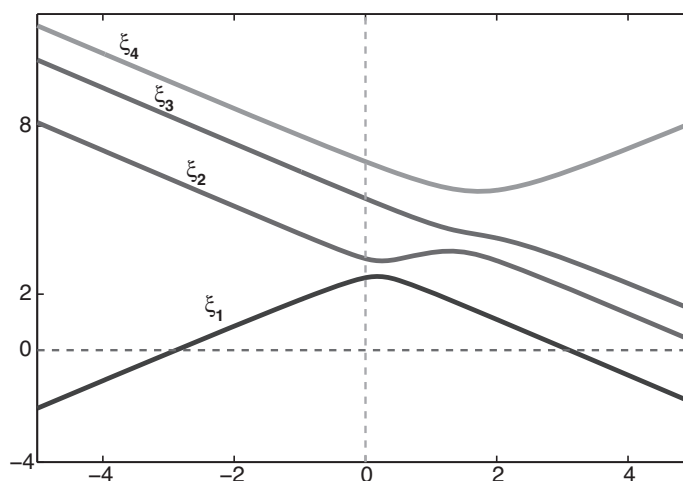
We illustrate this by using the particular example:

$$(4.1) \quad A = \begin{bmatrix} 5 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 \\ 0 & 1 & 6 & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

By Theorem 2, $i(A^{-1}B) = i(B) = (3, 0, 1)$. Since the eigenvalues of a matrix T are the reciprocals of the corresponding eigenvalues of its inverse T^{-1} , we have $i(T) = i(T^{-1})$. Applying this to $T = A^{-1}B$, we see that $i(B^{-1}A) = i(A^{-1}B) = (3, 0, 1)$. Let us see if we can arrive at the same conclusion by examining the eigencurves of $A - tB$, which are shown in Figure 2, instead of resorting to Theorem 2.

As we have seen in the proof of Theorem 1, the set of real eigenvalues of $B^{-1}A$ is equal to the set of points of intersection of the eigencurves with the t -axis.

When $t \rightarrow -\infty$, $A - tB$ is dominated by $-tB$, which has three positive eigenvalues and one negative eigenvalue. Hence, when t is negatively large, three of the eigencurves are above the t -axis while one is below. On the other hand, when $t \rightarrow \infty$, three of the eigencurves will be below the t -axis and one is above. Since

FIGURE 2. Eigenvalue curves of $A - tB$.

all the eigenvalues of A are positive, when the eigencurves intersect the λ -axis, the intersection points are all above the t -axis.

Let us look at each of the curves. The lowest curve λ_1 starts (when t is near $-\infty$) below the t -axis, becomes positive at $t = 0$, and eventually (for t near ∞) dips below the t -axis again. Invoking the Intermediate Value Theorem, we see that λ_1 must cut the t -axis at (at least) two points, one positive and the other negative. The second curve λ_2 starts above the t -axis and eventually dips below it. Hence, it must intersect the t -axis at some positive point (albeit beyond the right edge of the figure). Likewise λ_3 must intersect the t -axis at some positive point. Up to now, we have already obtained four real eigenvalues of $B^{-1}A$ and that exhausts all its eigenvalues. We can thus infer that λ_1 must have exactly one negative and one positive zero and each of λ_2 and λ_3 must have exactly one positive zero, while λ_4 cannot have a zero. This completes the proof that $i(B^{-1}A) = i(A^{-1}B) = (3, 1, 0)$.

It can be shown that the above arguments can be formalized to give an alternative proof of Theorem 2. There is, however, a slight technical difficulty in dealing with the situation when some eigenvalues of B are zero. Chronologically, we first discovered this alternative proof of Theorem 2 before we realized that the proof we presented earlier was simpler.

What happens when A also has negative eigenvalues. Let $A_\tau = A - \tau I$, where I is the identity matrix. It is easy to see that the the graphs of the eigencurves of $A_\tau - tB$ look exactly the same as those of $A - tB$, except that the t -axis for the former can be obtained by moving the t -axis for the latter (i.e. the horizontal dotted line in Figure 2) up by a distance of τ . To fix our ideas, let us choose $\tau = 2.9$, which is between the first (2.5858) and second (3.2679) eigenvalues of A . The graphs of the eigencurves of $A_\tau - tB$ are shown in Figure 3.

Now A_τ has one negative eigenvalue. The lowest eigencurve λ_1 is below the t -axis, and the picture shows that the set of intersection points of the eigencurves with the t -axis consists of two positive values. We have lost one positive and one negative value when compared to that of Figure 2. This is because the Intermediate

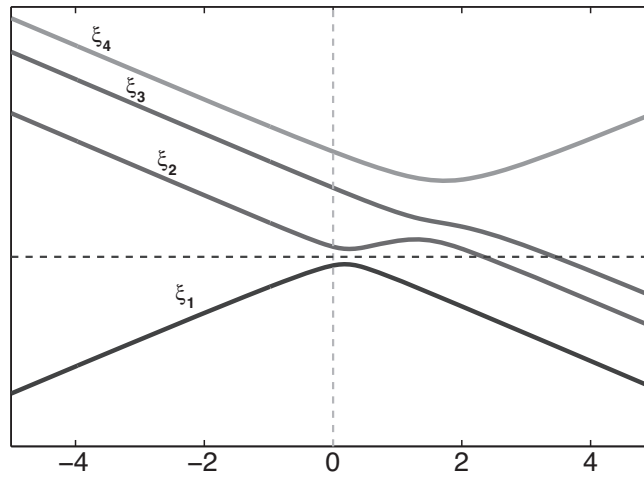


FIGURE 3. Eigenvalue curves of $B_\tau - tA$, $\tau = 2.9$.

Value Theorem can no longer be used to guarantee that λ_1 must intersect the t -axis twice. Neither can we use the argument of having exhausted all four eigenvalues to draw a conclusion about the exact number of intersection points of the individual curves with the t -axis.

If we know exactly the value of τ , we may be able to determine all the eigenvalues of $B^{-1}A$ either numerically or theoretically. Such is the case when $\tau = 2.9$ and, as Figure 2 shows, there are exactly two positive eigenvalues. If, however, we only know that τ is between the smallest and second smallest eigenvalues of A , then all we can conclude is that there are at least two positive eigenvalues, the intersection points of λ_2 and λ_3 with the t -axis. It is possible to show (at least visually) for our particular example of A and B that with suitable choices of τ (restricted between the first and second eigenvalues of A), we may be able to get four eigenvalues. The two extra ones can come either from λ_1 or from λ_2 intersecting the t -axis. For our example, the two extra eigenvalues must be both positive. With other choices of A and B , the two extra eigenvalues can be both negative. These extra eigenvalues cannot consist of one positive and one negative number, because they must come from the same positive or negative branch of an eigencurve.

To summarize, from the assumptions $i(A_\tau) = (3, 1, 0)$ and $i(B) = (3, 1, 0)$, we conclude that $i_r(B^{-1}A_\tau)$ has only three possible configurations: $(2, 0, 0)$, $(4, 0, 0)$, and $(2, 2, 0)$, and all three are realizable.

The above arguments can be extended to obtain the following results. We omit the straightforward proofs. The Hermitian version of these results has been established by Johnson [5].

Theorem 3. *Suppose that (A, B) satisfies (R) and*

$$(4.2) \quad i(A) = (\pi(A), \nu(A), 0), \quad i(B) = (\pi(B), \nu(B), 0).$$

Then

$$(4.3) \quad \pi_r(B^{-1}A) = \pi_r(A^{-1}B) \geq |\pi(A) + \pi(B) - n|,$$

and the discrepancy, i.e., the difference between the left-hand side and the right-hand side, must be a nonnegative even number. For the negative eigenvalues,

$$(4.4) \quad \nu_r(B^{-1}A) = \nu_r(A^{-1}B) \geq |\nu(A) - \nu(B)|,$$

and the discrepancy must be a nonnegative even number.

Note that condition (4.2) requires that none of the eigenvalues of A and B are zero; hence, both A and B are nonsingular. The result can be modified to cover the case when A is not invertible. The main difficulty in treating this degenerate case lies in counting the extra eigenvalues. In the regular case, these come in pairs of numbers having the same sign. In the degenerate case, it is possible that one or both of these in the pair can be 0.

Theorem 4. *Suppose that (A, B) satisfies (R) and*

$$(4.5) \quad i(A) = (\pi(A), \nu(A), \delta(A)), \quad i(B) = (\pi(B), \nu(B), 0),$$

with $\delta(A) > 0$. Then

$$(4.6) \quad \delta_r(B^{-1}A) \geq \delta(A).$$

We also have

$$(4.7) \quad \pi_r(B^{-1}A) \geq |\pi(A) + \pi(B) - n|.$$

If equality holds in (4.6), then the discrepancy must be a nonnegative even number. For the negative eigenvalues,

$$(4.8) \quad \nu_r(B^{-1}A) \geq |\nu(A) + \delta(A) - \nu(B)|.$$

If equality holds in (4.6), then the discrepancy must be a nonnegative even number.

5. FURTHER REMARKS

It is surprising to us that the simple technique of extending a single eigenvalue problem into a one-parameter family of problems and then tracing the eigencurves can lead to some nontrivial results. Yet we believe that the power of this technique has not been fully exploited. We state one simple result concerning 2×2 matrices to illustrate this point. It certainly begs for extension to higher dimensions. We leave that effort to others or to a future work.

For simplicity, let us restrict ourselves to matrices with real entries.

Theorem 5. *Let $A, B \in M_2(\mathbb{R})$ be such that A has no real eigenvalues and B has one positive and one negative eigenvalue. Then AB (hence, also BA) has exactly one positive and one negative eigenvalue.*

If, instead, we assume that B has two positive (or two negative) eigenvalues, then AB (and BA) has either no real eigenvalues or two eigenvalues of the same sign.

In other words, assuming $i_r(A) = (0, 0, 0)$, then

$$(5.1) \quad i_r(B) = (1, 1, 0) \implies i_r(AB) = i_r(BA) = (1, 1, 0)$$

$$(5.2) \quad i_r(B) = (2, 0, 0) \text{ or } (0, 2, 0) \implies \\ i_r(AB) = i_r(BA) = (0, 0, 0) \text{ or } (2, 0, 0) \text{ or } (0, 2, 0).$$

Proof. We consider the family of eigenvalue problems

$$(5.3) \quad (A - tB^{-1})u = \lambda u$$

and try to trace the eigencurves. To help to visualize them, Figure 4 is the plot for the special choice of

$$(5.4) \quad A = \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}.$$

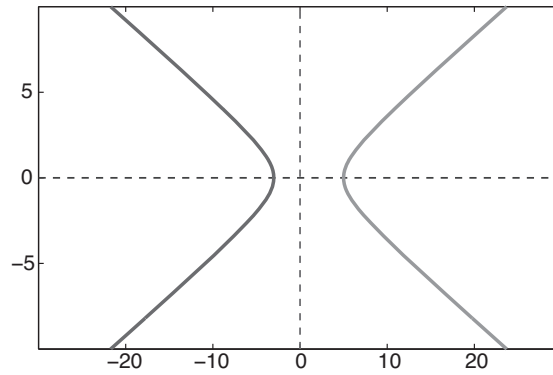


FIGURE 4. Eigenvalue curves $\lambda_1(t) \leq \lambda_2(t) \leq \lambda_3(t) \leq \lambda_4(t)$.

We see two eigencurves (because $n = 2$), but in contrast to the situation depicted in previous figures, each eigencurve does not flow from $t = -\infty$ to $t = \infty$ and they cannot be ordered in ascending order of λ (i.e., from bottom to top). When t is near $-\infty$, the two eigenvalues of $A - tB^{-1}$ are close to those of $-tB^{-1}$, and so one of them must be negatively large and the other positively large. As t increases, these two eigenvalues move closer to each other. Because A has no real eigenvalues, the eigencurves cannot intersect the λ -axis. So the eigencurves that start near $t = -\infty$ must be contained strictly to the left of the λ -axis. The only way that this can happen is that the curves traced out by the two eigenvalues of $A - tB^{-1}$, in the region $t < 0$, must join together somewhere before reaching $t = 0$ to form one continuous curve. We are waving our hand a little here by appealing to your visual intuition, but it should not be too difficult to justify the arguments rigorously. This left-hand eigencurve must cut the t -axis at some negative t and this point corresponds to a negative eigenvalue of BA . Likewise, there is a right-hand eigencurve contained strictly to the right of the λ -axis and this curve must cut the t -axis at some positive t , yielding a positive eigenvalue of BA .

The other conclusion for the case when the eigenvalues of B have the same sign can be obtained in a similar way by observing that now the two ends (i.e. when t is negatively large) of the left-hand eigencurve are either both above or both below the t -axis. As t varies towards $t = 0$, the curve may or may not intersect the t -axis. In the former situation, AB has two negative eigenvalues. In the latter situation, AB has no negative eigenvalues. The same arguments apply to the right-hand eigencurve. \square

ACKNOWLEDGMENT

The authors would like to thank the referee for many useful suggestions.

REFERENCES

- [1] J. Dancis, A quantitative formulation of Sylvester's law of inertia, III, *Linear Alg. Appl.* 80 (1986), 141-158. MR851937 (88c:15012)
- [2] F. P. Gantmacher and M. G. Krein, "Oscillation Matrices and Kernels and Small Vibrations of Mechanical Systems", AMS Chelsea, 2002. MR1908601 (2003f:34161)
- [3] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985. MR832183 (87e:15001)
- [4] Q. Kong, H. Wu, and A. Zettl, Left-definite Sturm-Liouville problems, *J. Diff. Eq.* 177 (2001), 1-26. MR1867611 (2002h:34020)
- [5] C. R. Johnson, The inertia of a product of two Hermitian matrices, *J. Math. Anal. Appl.* 57 (1977), 85-90. MR0432679 (55:5665)
- [6] J. J. Sylvester, A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by real orthogonal substitutions to the form of a sum of positive and negative squares. *Phil. Mag.* (4), 4 (1852), 142.
- [7] H. K. Wimmer, On Ostrowski's generalization of Sylvester's law of inertia, *Linear Alg. Appl.* 52-53 (1983), 739-741. MR709383 (84j:15011)

DEPARTMENT OF APPLIED MATHEMATICS, THE HONG KONG POLYTECHNIC UNIVERSITY,
HUNG HOM, KOWLOON, HONG KONG

E-mail address: `mankwong@polyu.edu.hk`

DEPARTMENT OF MATHEMATICS, NORTHERN ILLINOIS UNIVERSITY, DEKALB, ILLINOIS 60115

E-mail address: `zettl@math.niu.edu`