TOPOLOGICAL ASPECTS OF SYLVESTER'S THEOREM ON THE INERTIA OF HERMITIAN MATRICES

HANS SCHNEIDER, University of Wisconsin, Madison

1. A set of $n\times n$ matrices with complex elements has a natural topology associated with it. One may therefore look for a topological interpretation of some results in the theory of matrices. We shall show that Sylvester's classical theorem on the inertia (signature) of Hermitian matrices concerns the connected components of the space of all Hermitian matrices of fixed rank r.

Most of the arguments used in the proof of our theorem are elementary and familiar. Yet our result does not appear in the literature. The reason may well be that matrix theorists tend to use "continuity properties" as they arise, without formalizing them, while topologists do not usually study equivalence relations on matrices. This note is offered as an illustration that even on a fairly elementary level, something is gained by looking for inter-connections between different mathematical fields.

2. Let n be a positive integer and let $\omega=(\pi, \nu, \delta)$ be an ordered triple of nonnegative integers with $\pi+\nu+\delta=n$. Let E_ω be the diagonal matrix $E_\omega=\operatorname{diag}(1, \cdots, 1, -1, \cdots, -1, 0, \cdots, 0)$ with π diagonal elements 1, ν diagonal elements -1 and δ elements 0. Sylvester's theorem ([7] p. 100, [3] p. 83) on the inertia of Hermitian matrices asserts that for each $n\times n$ Hermitian matrix H there exists just one matrix E_ω for which there exists a nonsingular matrix X such that $X^*HX=E_\omega$.

But can we pick out the triple ω that occurs in Sylvester's theorem, without use of that theorem and directly in terms of the matrix H? One possibility, which we mention merely because of its intrinsic interest, is to proceed geometrically. Let V be the space of all positive n-tuples and associate with H the quadratic form Δ: $(x, x)=x^*Hx$. If it should happen that for some $y\in V$, $(y, y)>0$ then also $(x, x)>0$ for $x=\alpha y$ if $\alpha\neq 0$. Thus we have found a subspace W of V of which $(x, x)>0$ for all $x\neq 0$. In other words, Δ is positive definite on W. Now suppose that π is the dimension of a subspace W of largest dimension on which Δ is positive and similarly suppose ν is the dimension of a subspace W' of largest dimension on which Δ is negative definite (i.e., $(x, x)<0$, if $0\neq x\in W'$). If $\delta=n-\pi-\nu$, then it may be proved that $\omega=(\pi, \nu, \delta)$ is the ω of Sylvester's theorem, (see [1] pp. 148–150).

3. We shall use an entirely different approach. Since H is Hermitian all its eigenvalues are real. We shall define π, ν, δ in terms of the eigenvalues of H.

DEFINITION 1. *Let $\pi(H)=\pi$ be the number of positive eigenvalues of H, $\nu(H)=\nu$ the number of negative eigenvalues of H, and $\delta(H)=\delta$, the number of zero eigenvalues of H. Then the ordered triple $\omega=(\pi, \nu, \delta)$ will be called the inertia of H. We shall write $\omega=\operatorname{In} H$.*

DEFINITION 2. *Two Hermitian matrices H, K are inertially equivalent if* In $H=$ In K. *We shall write* $H \overset{i}{\sim} K$.

The next definition is standard.

DEFINITION 3. (e.g. [6] p. 99, p. 84). *Two Hermitian matrices H and K are conjunctive (conjunctively equivalent) if there exists a nonsingular X such that $X^*HX=K$. We shall write $H \overset{c}{\sim} K$.*

It is evident that $\overset{i}{\sim}$ and $\overset{c}{\sim}$ are equivalence relations on any set of Hermitian matrices.

4. Our next two equivalence relations are of a different kind, since they may be defined on any topological space.

If E is a topological space, the space is called connected if the empty set and E are the only subsets of E which are both open and closed (see [5] p. 117). A subset U of E is connected if and only if it is a connected space in the topology induced on U by E. Thus U is connected if and only if, for any set $F \subseteq E$ which is both open and closed, either $U \subseteq F$ or $U \subseteq E \backslash F$, the complement of F in E. This motivates

DEFINITION 4. *Let E be a topological space. We call $x, y \in E$ connectable in E if for every open and closed set U in E both $x, y \in U$ or both $x, y \notin U$. We shall write $x \overset{u}{\sim} y$.*

An *arc* in a topological space E is a continuous image of the interval (0, 1) on the real line in the space E ([5] p. 139).

DEFINITION 5. *Let E be a topological space. We call $x, y \in E$ arc connectable in E if there exists an arc in E joining x, y, i.e., if there exists a continuous function f from the unit interval* (0, 1) *on the real line into E with $f(0)=x$ and $f(1)=y$. We shall write $x \overset{a}{\sim} y$.*

The following lemma is a restatement of a well-known result (see [5] p. 141).

LEMMA 1. *If E is a topological space, then $x \overset{a}{\sim} y$ implies that $x \overset{u}{\sim} y$.*

Proof. Let U be any open and closed set containing x and let f be a continuous function of (0, 1) into E with $f(0)=x$ and $f(1)=y$. Then $f^{-1}(U)$ is an open and closed subset of (0, 1) which contains 0, and the only such set is (0, 1) itself. Thus $y=f(1) \in U$ and so $x \overset{u}{\sim} y$.

5. Let S be any set of $n \times n$ matrices. We can norm S in many ways. For example we can put $\|A\| = \max_{i,j} |a_{ij}|$ for $A \in S$. To turn S into a topological space we choose as the open sets arbitrary unions of finite intersections of all cubes $N(A, \epsilon) = \{B \in S: \max_{i,j} |b_{ij} - a_{ij}| = \|B-A\| < \epsilon\}$, with $A \in S$ and $\epsilon > 0$. Thus a subset T of S is open if and only if for $A \in T$ we can find on $\epsilon = 0$ such that $N(A, \epsilon) \subseteq T$. Observe that S need not be a linear space, nor will our topological space S necessarily be complete.

In this section, we shall consider the space N of all nonsingular complex matrices normed as above.

LEMMA 2. *For all $A, B \in N$, A is arc connectable to B in N.*

Proof. It is enough to prove that for all $A \in N$, $A \overset{a}{\sim} I$, the identity matrix. Choose σ so that $e^{i\sigma}A$ has no negative eigenvalue, and set $f(t) = e^{i\sigma t}A$, $0 \leqq t \leqq 1$. If A belongs to N, so does $e^{i\sigma t}A$ and clearly f is continuous. Thus $A \overset{a}{\sim} C$. Next set $g(t) = (1-t)C + tI$, $0 \leqq t \leqq 1$. Evidently g is again continuous, and since the eigenvalues of $g(t)$ are of the form $\gamma(t) = (1-t)\gamma + t$ where γ is an eigenvalue of C and here γ is not negative, it follows that $\gamma(t) \neq 0$ and so $g(t)$ is nonsingular. Thus $C \overset{a}{\sim} I$. Hence $A \overset{a}{\sim} I$, and this completes the proof.

6. We now require a lemma of a different type. Usually it is expressed by asserting that the eigenvalues of a matrix are continuous functions of the elements of the matrix. We shall state the result precisely:

LEMMA 3. *Let A be a matrix with distinct eigenvalues $\alpha_1, \cdots, \alpha_s$ of multiplicities $m_1, \cdots, m_s$ respectively. Let $\epsilon > 0$ and let $\Gamma(\alpha_i, \epsilon)$ be the circle with center α_i and radius ϵ. Then there is a positive σ, such that every matrix B, for which $\|B - A\| < \sigma$, has exactly m_i eigenvalues in the circle $\Gamma(\alpha_i, \epsilon)$.*

Proof. Let $p(t) = \lambda^n + p_{n-1}\lambda^{n-1} + \cdots + p_0$, and $q(t) = \lambda^{n-1} + q_{n-1}\lambda^{n-1} + \cdots + q_0$. We use a theorem on the zeros of a polynomial (see [4], p. 3). If the zeros of $p(\lambda)$ are α_i with multiplicity m_i, $i = 1, \cdots, s$, $\alpha_i \neq \alpha_j$, if $i \neq j$, and if $(q_j - p_j) < \eta$, $j = 0, 1, \cdots, n-1$, where η is sufficiently small, then m_i zeros of $q(t)$ lie in the circle $\Gamma(\alpha_i, \epsilon)$. Now the eigenvalues of A and B are simply the zeros of the characteristic polynomials det $(\lambda I - A)$ whose coefficients are sums of products of elements of A, and similarly for B. Since addition and multiplication of complex numbers is continuous, we deduce that for sufficiently small $\sigma > 0$, $\|B - A\| < \sigma$, implies that $|q_i - p_i| < \eta$, $j = 0, \cdots, n-1$, and the result follows.

Our proof of Lemma 3 is not really much of a proof, since it refers the result for the spectra of a matrices back to the corresponding theorem for the zeros of polynomials ("continuity of zeros of polynomials"). This latter result is deeper than any other theorem we have used in this note, and we shall not attempt to prove it here. We may note that in the application of Lemma 3, the matrices A and B are both Hermitian. For normal, and therefore for Hermitian matrices, a more precise result is given in [2]:

LEMMA 4. *If A and B are normal matrices with eigenvalues $\alpha_1, \cdots, \alpha_n$ and $\beta_1, \cdots, \beta_n$ respectively, then there exists a suitable numbering of the eigenvalues such that*

$$\sum_i |\alpha_i - \beta_i|^2 \leqq \sum_{i,j} |a_{ij} - b_{ij}|^2.$$

This result rests on the famous theorem of Birkhoff that the permutation matrices are the vertices of the convex polyhedron of doubly stochastic matrices. For a proof of Birkhoff's theorem see [3] p. 97, or [2], and for a proof of

Lemma 4 see [2]. Of course, Lemma 4 implies Lemma 3 since $\sum_{i,j} |a_{ij}-b_{ij}|^2 \leqq n^2\|A-B\|^2$.

7. From now on our space will be the space H_r^n of all $n\times n$ Hermitian matrices of *fixed rank r*. We shall first examine a trivial situation. The space of all Hermitian 1×1 matrices is just the real line R and hence H_1^1 is the real line with the origin removed. The connectivity properties of R and H_1^1 are quite different. R is connected and H_1^1 is not. Similarly the connectivity properties of H_r^n will be quite different from these of the space of all $n\times n$ Hermitian matrices. The reason for focusing on H_r^n is that this space yields an interesting theorem.

Notation. Let $H\in H_r^n$. Then the set of all $K\in H_r^n$ such that $H\overset{i}{\sim}K$ will be denoted by $\mathbf{I}(H)$, and the eigenvalue class $\mathbf{I}(H)$ will be called an inertial component of H_r^n. Similarly we define $\mathbf{C}(H)$, $\mathbf{U}(H)$, $\mathbf{A}(H)$ to be the equivalence classes of H for $\overset{c}{\sim}$, $\overset{u}{\sim}$, $\overset{a}{\sim}$, respectively, and we call $\mathbf{C}(H)$ a conjunctive component, $\mathbf{U}(H)$ a connected component and $\mathbf{A}(H)$ an arc component of H_r^n.

THEOREM. *Let H_r^n be the topological space of all $n\times n$ Hermitian matrices of rank r. Then the four equivalence relations $\overset{i}{\sim}$, $\overset{c}{\sim}$, $\overset{a}{\sim}$, $\overset{u}{\sim}$, coincide on H_r^n. Equivalently, for any $H\in H_r^n$, $\mathbf{I}(H)=\mathbf{C}(H)=\mathbf{A}(H)=\mathbf{U}(H)$.*

Proof. We shall prove that $H\overset{i}{\sim}K$ implies $H\overset{c}{\sim}K$, $H\overset{c}{\sim}K$ implies $H\overset{a}{\sim}K$, $H\overset{a}{\sim}K$ implies $H\overset{u}{\sim}K$, and $H\overset{u}{\sim}K$ implies $H\overset{i}{\sim}K$.

(a) $H\overset{i}{\sim}K$ implies $H\overset{c}{\sim}K$: Suppose $\operatorname{In} H=\operatorname{In} K=\omega=(\pi,\ \nu,\ \delta)$ say. It is enough to prove that $H\overset{c}{\sim}E_\omega$, where E_ω is defined in Section 1. (Note that $\pi+\nu=r$ and that $E_\omega\in H_r^n$.) Since H is Hermitian there exists a unitary Y such that $Y^*HY=\operatorname{diag}(\alpha_1,\cdots,\alpha_n)$. By definition of H, and since the α_i are the eigenvalues of H, it follows that π of the α_i are positive, ν are negative and the remaining δ are zero. Replacing Y, if necessary, by the unitary matrix YP, where P is a permutation matrix, we may suppose that $Y^*HY=\operatorname{diag}(\alpha_1,\cdots,\alpha_n)$ where $\alpha_i>0$, $i=1,\cdots,\pi$, $\alpha_{i+1}<0$, $i=\pi+1,\cdots,\pi+\nu$, and $\alpha_i=0$, $i=\pi+\nu+1,\cdots,n$.

Thus if $D=\operatorname{diag}(\sqrt{\alpha_1},\cdots,\sqrt{\alpha_\pi},\sqrt{-\alpha_{\pi+1}},\cdots,\sqrt{-\alpha_{\pi+\nu}},1,\cdots,1)$ and $X=YD^{-1}$, then $X^*HX=E_\omega$.

(b) $H\overset{c}{\sim}K$ implies $H\overset{a}{\sim}K$: Let $X^*HX=K$, where X is nonsingular. Since by Lemma 2 any two nonsingular matrices are connected in the space N of nonsingular complex matrices, there is a continuous function $t\rightarrow X(t)$ of $(0,1)$ into N such that $X(0)=I$ and $X(1)=X$. Further, transposition and matrix multiplication are continuous operations. Hence the function $f(t)=X^*(t)HX(t)$ is continuous. But rank $X^*(t)HX(t)=r$, since $X(t)$ is nonsingular, whence $X(t)HX(t)\in H_r^n$. Further $f(0)=H$ and $f(1)=K$, and so $H\overset{a}{\sim}K$.

(c) $H\overset{c}{\sim}K$ implies $H\overset{u}{\sim}K$. This is just Lemma 1.

(d) $H\overset{u}{\sim}K$ implies $H\overset{i}{\sim}K$. We shall prove the equivalent result that $\mathbf{I}(H)$ contains $\mathbf{U}(H)$, for $H\in H_r^n$: Let $K\in\mathbf{I}(H)$, and let $\operatorname{In} K=\omega=(\pi,\ \nu,\ \delta)$, and suppose the eigenvalues α_i, $i=1,\cdots,\pi$, of K are positive, the eigenvalues α_i,

$i=\pi+1, \cdots, \pi+\nu$ are negative and the eigenvalues α_i, $i=\pi+\nu+1, \cdots, n$ are zero.

Let $0<\epsilon<\min\{|\alpha_i|:\alpha_i\neq 0\}$. By Lemma 3 there exists a neighborhood $N(K, \sigma)$ of K in H_r^n (thus each $L\in N(K, \sigma)$ is, by assumption, Hermitian of rank r) so that the spectrum of each $L\in N(K, \sigma)$ is contained in the union of the n circles $\Gamma(\alpha_i, \epsilon)$. It follows that each L in this neighborhood of K has at least as many positive (negative) eigenvalues as K has positive (negative) eigenvalues, i.e. $\pi(L)\geqq\pi$, and $\nu(L)\geqq\nu$. But as $L\in H_r^n$, $\pi(L)+\nu(L)=r=\pi+\nu$, whence $\pi(L)=\pi$ and $\nu(L)=\nu$. Thus $\text{In } L=\text{In } K$. It follows for each $K\in \mathbf{I}(H)$, there exists an $N(K, \sigma)\subseteq\mathbf{I}(H)$, and so $\mathbf{I}(H)$ is open. Now $H_r^n\backslash\mathbf{I}(H)=\bigcup\{\mathbf{I}(M): M\in H_r^n, M\notin\mathbf{I}(H)\}$ and a union of open sets is open. Hence $\mathbf{I}(H)$ is also closed. By the definition of $\mathbf{U}(H)$, we see that $\mathbf{U}(H)$ is contained in every open and closed set containing H, whence $\mathbf{U}(H)\subseteq\mathbf{I}(H)$. This completes the proof of the theorem.

COROLLARY. *The topological space H_r^n has precisely $r+1$ distinct inertial components (or conjunctive components, or connected components, or arc components).*

Proof. Obviously, each $\omega=(\pi, \nu, \delta)$ with $\pi+\nu=r$ corresponds to one inertial component of H_r^n, and there are just $r+1$ such ω. By the theorem, each inertial component is a conjunctive component (and a connected component and an arc component).

8. It should be noted that in the proof of our theorem (a) is just a standard proof that there exists an ω such that $H\overset{c}{\sim}E_\omega$. The direct proof that this ω is unique is simple (see [7] pp. 92, 100), but we do not need this proof. For distinct ω, the corresponding E_ω obviously lie in distinct inertial components and the uniqueness now follows from the equality of $\overset{c}{\sim}$ and $\overset{i}{\sim}$ on H_r^n.

The concept of inertia may be extended to matrices that are not Hermitian. For some results in this direction see [6].

I wish to thank Marvin Marcus, Judith Molinar, and Robert Thompson for comments which have helped to improve this paper.

References

1. E. Artin, Geometric Algebra, Interscience, New York, 1957.

2. A. J. Hoffman and H. Wielandt, The variation of the spectrum of a normal matrix, Duke M. J., 20 (1953) 37–40.

3. M. Marcus and H. Minc, A survey of matrix theory and matrix inequalities, Allyn and Bacon, Boston, 1964.

4. M. Marden, The geometry of the zeros of a polynomial in a complex variable, Math. Surveys III, American Math. Soc., 1949.

5. B. Mendelson, Introduction to Topology, Allyn and Bacon, Boston, 1962.

6. A. Ostrowski and H. Schneider, Some theorems on the inertia of matrices, J. Math. Analysis Appl., 4(1962) 72–84.

7. S. Perlis, Theory of matrices, Addison-Wesley, Reading, Mass., 1952.