Remarks on Hermitian matrices

Remarks on Hermitian Matrices

N. G. de Bruijn
Eindhoven University of Technology
Department of Mathematics
Eindhoven, The Netherlands

Submitted by Olga Taussky Todd

1. INTRODUCTION

The following research problem was proposed in 1959 by Olga Taussky Todd [5]: If the eigenvalues $\alpha_1,\dots,\alpha_n$ and $\beta_1,\dots,\beta_n$ of the hermitian matrices A and B are given, what can be said about the eigenvalues of the Jordan product $AB+BA$?

W. Gilbert Strang [4] gave what can be considered to be a complete solution, indicating best possible bounds for the eigenvalues of $AB+BA$. And quite recently D. W. Nicholson [3], unaware of Strang's work, published an incomplete solution for the same problem.

The present paper obtains Strang's results by an alternative method, and in a slightly more manageable form. The key is Theorem 2.1 on the values of the inner product (Ax,x); this theorem might have some independent interest.

Getting bounds for the eigenvalues of $AB+BA$ means getting bounds for $\mathrm{Re}(Ax,Bx)$ if x runs through the set of vectors of length 1. Strang also obtained a result for the possible values of the complex number (Ax,Bx) itself, if bounds for the eigenvalues of A and B are given, but he restricted himself to the special case $0\leqslant\alpha_1\leqslant\cdots\leqslant\alpha_n\leqslant 1$, $0\leqslant\beta_1\leqslant\cdots\leqslant\beta_n\leqslant 1$. We treat the general case here (Theorem 3.1), and we show that the result specializes to Strang's if the bounds are 0 and 1.

The author is indebted to Professor Olga Taussky Todd for stimulating him to publish this paper. It is essentially the material he communicated to her in 1962 as a comment on Strang's paper and as an alternative solution to her problem.

NOTATION. All matrices are complex n by n matrices, where n is a fixed positive integer, and all vectors are column vectors with n complex entries.

If p and q are real numbers, the notation $[p,q]$ indicates the point with coordinates p and q in some fixed coordinate plane.

0024-3795/80/040201+8$01.75

If p and q are real, $p \leqslant q$, then $C(p,q)$ is the circle with diameter from $[p,0]$ to $[q,0]$ [center $\left[\frac{1}{2}(p+q),0\right]$ and radius $\frac{1}{2}(q-p)$]. The interior is denoted by $C^i(p,q)$, and the closed disk (the union of C and C^i) by $C^c(p,q)$. In the case $p=q$, C^i is empty, and both C and C^c degenerate to a single point.

As usual, (x,y) is the inner product of x and y, and $\|x\|=(x,x)^{1/2}$.

2. KEY THEOREM

Let A be a hermitian matrix, and let x be a unit vector. We can split Ax in a vector along x and a vector orthogonal to x:

$$Ax=\alpha x+\gamma e, \tag{2.1}$$

where $\|e\|=1$, $(e,x)=0$, and γ is real. It follows that α is real,

$$\alpha=(Ax,x), \qquad \gamma=\pm\|Ax-\alpha x\|.$$

What can be said about the point $[\alpha,\gamma]$ if the eigenvalues $\alpha_1,\ldots,\alpha_n$ (with $\alpha_1 \leqslant \cdots \leqslant \alpha_n$) are given? The following theorem says that the range of this point is the set K, depicted in Fig. 1 and defined by

$$K=C^c(\alpha_1,\alpha_n)\backslash\{C^i(\alpha_1,\alpha_2)\cup C^i(\alpha_2,\alpha_3)\cup\cdots\cup C^i(\alpha_{n-1},\alpha_n)\}. \tag{2.2}$$

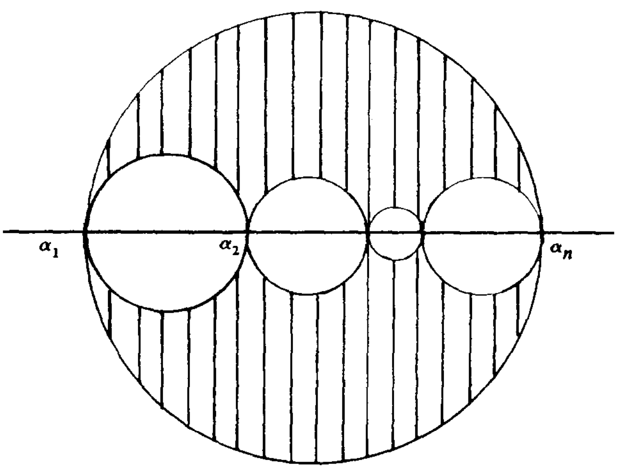

FIG. 1.

For the definition of K the multiplicity of an eigenvalue does not matter, since $C^i(\alpha_k,\alpha_{k+1})$ is empty if $\alpha_k=\alpha_{k+1}$. Also note that if there are just *two* distinct eigenvalues α_1 and α_n, then K equals the circumference $C(\alpha_1,\alpha_n)$.

THEOREM 2.1. *We have* $\lceil \alpha,\gamma \rceil \in K$, *and every point of* K *can be attained by suitable selection of* x *if* A *is fixed.*

Proof. By unitary transformation we can get A and x in the form

$$A=\begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{bmatrix}, \qquad x=\begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}.$$

This transformation does not affect the inner products (x,x), (Ax,x), (Ax,Ax). We have

$$\|x\|^2=|\xi_1|^2+\cdots+|\xi_n|^2=1,$$

$$(Ax,x)=\alpha_1|\xi_1|^2+\cdots+\alpha_n|\xi_n|^2,$$

$$(Ax,Ax)=\alpha_1^2|\xi_1|^2+\cdots+\alpha_n^2|\xi_n|^2.$$

If we define Q by $Q=\lceil (Ax,x),(Ax,Ax) \rceil$, then the point Q belongs to the convex hull of the set

$$\{\lceil \alpha_1,\alpha_1^2 \rceil,\ldots,\lceil \alpha_n,\alpha_n^2 \rceil\}.$$

Moreover, every point of this convex hull is a possibility for Q: we just have to take appropriate nonnegative weights with sum 1, and to take $|\xi_1|^2,\ldots,|\xi_n|^2$ equal to those weights.

In the u,v-plane we consider the transformation f given by

$$f(\lceil u,v \rceil)=\lceil u,u^2+v^2 \rceil.$$

The inverse image of the convex hull turns out to be K. The theorem now follows from the identity

$$(Ax,Ax)=(Ax,x)^2+\|Ax-(Ax,x)x\|^2. \qquad \blacksquare$$

REMARK 1. By unitary transformation it follows that if x is fixed, then every point of K can be attained by suitable selection of A (with the prescribed eigenvalues).

REMARK 2. Theorem 2.1 generalizes a very well-known theorem in elasticity theory, where it is restricted to real vectors and matrices, and $n=3$. There the A is a (symmetrical) stress tensor; αx and γe are the normal and shear stress vectors associated with a surface element orthogonal to x. The figure with the circles is named after O. Mohr, who discussed it in 1882. It is treated in many textbooks on elasticity theory (e.g. [2]). For a derivation different from the usual ones see [1].

REMARK 3. The only thing we shall need about K in the next sections is

$$C(\alpha_1,\alpha_n) \subset K \subset C^c(\alpha_1,\alpha_n).$$

3. VALUES OF (Ax,Bx)

We next consider two hermitian matrices A and B, both with prescribed eigenvalues, $\alpha_1,\dots,\alpha_n$ for A (with $\alpha_1 \leqslant \dots \leqslant \alpha_n$), $\beta_1,\dots,\beta_n$ for B (with $\beta_1 \leqslant \dots \leqslant \beta_n$). We ask for the possible values of (Ax,Bx) with $\|x\|=1$. In particular we are interested in the values of $\operatorname{Re}(Ax,Bx)$ for the following reason. The so-called Jordan product $AB+BA$ is hermitian, we have $\frac{1}{2}((AB+BA)x,x)=\operatorname{Re}(Ax,Bx)$, and the set of all values of $((AB+BA)x,x)$ (with $\|x\|=1$) is the interval whose endpoints are the smallest and the largest eigenvalue of $AB+BA$.

According to Sec. 2 we have

$$Ax=\alpha x+\gamma e, \qquad Bx=\beta x+\delta e'$$

with $(x,e)=(x,e')=1$, $\|x\|=\|e\|=\|e'\|=1$, and

$$[\alpha,\gamma]\in C^c(\alpha_1,\alpha_n), \qquad [\beta,\delta]\in C^c(\beta_1,\beta_n). \tag{3.1}$$

Since $(Ax,Bx)=\alpha\beta+\gamma\delta(e,e')$, we are interested in the products $\alpha\beta$ and $\gamma\delta$. We can get those products just using the circumferences of the circles, as shown in the following lemma.

LEMMA. *If α, β, γ, δ satisfy* (3.1), *and if* $-1\leqslant\varepsilon\leqslant 1$, *then* α', β', γ', δ' *exist with*

$$\alpha'\beta'=\alpha\beta,\qquad \gamma'\delta'=\varepsilon\gamma\delta,$$

$$\lceil\alpha',\gamma'\rceil\in C(\alpha_1,\alpha_n),\qquad \lceil\beta',\delta'\rceil\in C(\beta_1,\beta_n).$$

Proof. For reasons of symmetry we may assume that γ, δ, ε are all $\geqslant 0$.

Let ϕ denote the nonnegative function such that $\lceil u,\phi(u)\rceil\in C(\alpha_1,\alpha_n)$ for $\alpha_1\leqslant u\leqslant\alpha_n$, and ψ the similar function for the other circle. We note that $\gamma\leqslant\phi(\alpha)$, $\delta\leqslant\psi(\beta)$.

In the u,v-plane we consider the rectangle $\alpha_1\leqslant u\leqslant\alpha_n$, $\beta_1\leqslant v\leqslant\beta_n$, and the hyperbola $uv=\alpha\beta$. Since $\lceil\alpha,\beta\rceil$ belongs to the rectangle, the hyperbola has a point on the boundary too. If $\lceil u,v\rceil$ moves along the hyperbola from $\lceil\alpha,\beta\rceil$ to such a boundary point, the product $\phi(u)\psi(v)$ moves to zero. It follows that on this hyperbola $\phi(u)\psi(v)$ attains every value between $\phi(\alpha)\psi(\beta)$ and 0. In particular α' and β' exist such that $\phi(\alpha')\psi(\beta')=\varepsilon\gamma\delta$. Now take $\gamma'=\phi(\alpha')$, $\delta'=\psi(\beta')$. ∎

THEOREM 3.1. *If x is fixed,* $\|x\|=1$, *then the set of all possible values for* (Ax,Bx) (A *and* B *with the prescribed eigenvalues*) *is equal to the set of all* $\alpha\beta+\gamma\delta\zeta$ *with*

$$\lceil\alpha,\gamma\rceil\in C^c(\alpha_1,\alpha_n),\quad \lceil\beta,\delta\rceil\in C^c(\beta_1,\beta_n),\qquad \zeta\in\mathbb{C},\quad |\zeta|\leqslant 1 \tag{3.2}$$

and also identical to the set of all $\alpha\beta+\gamma\delta\zeta$ *with*

$$\lceil\alpha,\gamma\rceil\in C(\alpha_1,\alpha_n),\quad \lceil\beta,\delta\rceil\in C(\beta_1,\beta_n),\qquad \zeta\in\mathbb{C},\quad |\zeta|=1. \tag{3.3}$$

Proof. It follows from the lemma (take $\varepsilon=|\zeta|$) that the characterizations by means of (3.2) and (3.3) are equivalent.

Since $(Ax,Bx)=\alpha\beta+\gamma\delta(e,e')$ and $|(e,e')|\leqslant 1$, every (Ax,Bx) lies in the set given by (3.2).

Next let α, β, γ, δ, ζ satisfy (3.3). We can find hermitian matrices A and B_1 [cf. Remark 1 at the end of Sec. 2 and note that $C(\alpha_1,\alpha_n)\subset K$] such that $Ax=\alpha x+\gamma e$, $B_1x=\beta x+\delta e'$, where e and e' have length 1, $(x,e)=(x,e')=0$. Next we can find a unitary matrix U such that $Ux=x$ and $Ue'=\bar{\zeta}e$. Fixing B by $B=UB_1U^*$, we obtain $Bx=\beta x+\delta\bar{\zeta}e$, $(Ax,Bx)=\alpha\beta+\gamma\delta\zeta$. ∎

THEOREM 3.2. *The set of all possible values for* Re(Ax,Bx) *is equal to the set of all* $\alpha\beta+\gamma\delta$ *with*

$$\lceil \alpha,\gamma \rceil \in C(\alpha_1,\alpha_n), \qquad \lceil \beta,\delta \rceil \in C(\beta_1,\beta_n). \tag{3.4}$$

Proof. We can repeat the proof of Theorem 3.1, now with restriction to real values of ζ. In (3.3) we can write $\zeta=\pm 1$ instead of $|\zeta|=1$. The value $\zeta=-1$ can be replaced by $\zeta=+1$ if we just change the sign of δ. ■

4. PRELIMINARIES

As a preparation to Sec. 5, we take positive numbers a and b and we ask for the set

$$\left\{ \operatorname{Re}(1+\eta)(1+\bar{\zeta}) \,\middle|\, \eta\in\mathbb{C},\ \zeta\in\mathbb{C},\ |\eta|=a,\ |\zeta|=b \right\}. \tag{4.1}$$

We shall use the identity

$$2\operatorname{Re}(1+\eta)(1+\bar{\zeta}) = 1-|\eta|^2-|\zeta|^2+|1+\eta+\zeta|^2. \tag{4.2}$$

The range of $\eta+\zeta$ is the ring between the circles with center 0 and radii $|a-b|$ and $a+b$ (boundary included). Therefore, the range of $|1+\eta+\zeta|$ is the closed interval $[V,a+b+1]$, where $V=a-b-1$ if $a\geqslant b+1$, $V=0$ if $|a-b|\leqslant 1\leqslant a+b$, $V=a+b-1$ if $a+b\leqslant 1$, $V=b-a-1$ if $b\geqslant a+1$. By (4.2) we infer that the set (4.1) is the closed interval $[W,(1+a)(1+b)]$, where

$$W=\begin{cases} (1-a)(1+b) & (a\geqslant b+1), \\ \frac{1}{2}(1-a^2-b^2) & (|a-b|\leqslant 1\leqslant a+b), \\ (1-a)(1-b) & (a+b\leqslant 1), \\ (1+a)(1-b) & (b\geqslant a+1). \end{cases}$$

5. APPLICATION

The result of Sec. 4 can be applied to find the set of all possible values of Re(Ax,Bx), i.e., the set of all $\alpha\beta+\gamma\delta$ with (3.4). We have $\alpha\beta+\gamma\delta=$

$\mathrm{Re}[(\alpha+i\gamma)(\beta-i\delta)]$, and $\alpha+i\gamma$ runs through the circle $C(\alpha_1,\alpha_n)$, $\beta+i\delta$ through $C(\beta_1,\beta_n)$ (we now consider the plane of those circles as the complex plane).

If $\frac{1}{2}(\alpha_1+\alpha_n)$ happens to be zero, the range of $(\alpha+i\gamma)(\beta-i\delta)$ is obtained from $C(\beta_1,\beta_n)$ by multiplication, and the range of its real part is

$$\left[\tfrac{1}{2}(\alpha_n-\alpha_1)\beta_1,\ \tfrac{1}{2}(\alpha_n-\alpha_1)\beta_n\right].$$

If $\frac{1}{2}(\beta_1+\beta_n)=0$ we have a similar result. If both $\alpha_1+\alpha_n$ and $\beta_1+\beta_n$ are different from zero, we put

$$\frac{\alpha_n-\alpha_1}{|\alpha_1+\alpha_n|}=a, \qquad \frac{\beta_n-\beta_1}{|\beta_1+\beta_n|}=b,$$

and the answer is supplied by Sec. 4. The range is

$$\begin{aligned} &[\rho W,\rho(1+a)(1+b)] \quad (\rho>0),\\ &[\rho(1+a)(1+b),\rho W] \quad (\rho<0), \end{aligned}$$

where $\rho=(\alpha_1+\alpha_n)(\beta_1+\beta_n)/4$.

6. SPECIALIZATION TO STRANG'S RESULT

In [4] Strang gives the set of values of (Ax,Bx) in an explicit form in the case that $\alpha_1=\beta_1=0$, $\alpha_n=\beta_n=1$. We can do this here too, on the basis of (3.2). Now α and β run through the interval $[0,1]$, and $\gamma^2\leqslant\alpha(1-\alpha)$, $\delta^2\leqslant\beta(1-\beta)$. If $\alpha\beta$ is kept constant, $\alpha\beta=p$ say, the product $\gamma\delta$ runs between $-q$ and $+q$, where $q=p^{1/2}-p$. In the complex plane we now have to find the union of the closed circular disks with center p and radius $p^{1/2}-p$, if p varies from 0 to 1. Taking coordinates x,y in the complex plane we find, by classical analytic geometry, the region given by

$$-\tfrac{1}{8}\leqslant x\leqslant 1, \qquad 27y^2\leqslant(x-1)^2(8x+1). \tag{6.1}$$

Note that the circle $(x-p)^2+y^2=(p^{1/2}-p)^2$ lies entirely inside (6.1), and that as long as $p\geqslant\frac{1}{16}$, it touches the boundary at the points with $x=\frac{1}{2}(3p^{1/2}-1)$. For values of p less than $\frac{1}{16}$ the circles lie in the interior of the region.

It is easy to check that (6.1) is equivalent to Strang's remarkable representation in terms of polar coordinates r, ϕ:

$$0 \leqslant r \leqslant \left\{ \frac{\cos\frac{1}{3}\pi}{\cos\frac{1}{3}(\pi - |\phi|)} \right\}^3 \qquad (-\pi \leqslant \phi \leqslant \pi).$$

REFERENCES

1 W. L. Esmeijer, On a well-known theorem in the theory of the Mohr-circles of stress, Report, Dept. of Mech. Engineering, Eindhoven Univ. of Technology, 1977.

2 H. Leipholz, *Einführung in die Elastizitätstheorie*, Braun Karlsruhe, 1968.

3 D. W. Nicholson, Eigenvalue bounds for $AB+BA$, with A, B positive definite matrices, *Linear Algebra and Appl.* 24:173–183 (1979).

4 W. Gilbert Strang, Eigenvalues of Jordan products, *Amer. Math. Monthly* 69:37–40 (1962).

5 Olga Taussky, Research problem 2, *Bull. Amer. Math. Soc.* 66:275 (1960).

Received 13 September 1979; revised 21 September 1979