MAXIMUM PRINCIPLES IN MATRIX THEORY

by L. MIRSKY

(Received 29th August, 1958)

Notation

Unless the contrary is stated, all matrices are understood to be complex and of type $n \times n$. The transposed conjugate of A is denoted by A^*. The non-negative square roots of the characteristic roots of A^*A are called the *singular values* of A; they will be denoted by $s_i(A)$, $i=1, \ldots, n$, where $s_1(A) \geqslant \ldots \geqslant s_n(A)$. The symbol $[A]_k$ denotes the $k \times k$ submatrix standing in the upper left-hand corner of A. We shall write $E_j(z_1, \ldots, z_n)$ for the j-th elementary symmetric function of $z_1, \ldots, z_n$, and $E_j(A)$ for the j-th elementary symmetric function of the characteristic roots of A. It is understood that, throughout, $1 \leqslant j \leqslant k \leqslant n$.

Introduction

The object of the present note is to generalize a number of extremal properties involving characteristic roots and singular values of matrices, which were discovered by Ky Fan. Thus, for example, two maximum principles for completely continuous operators in Hilbert space [**4**, Theorem 1] can be stated, for the case of finite matrices, in the following form. Let $A_1, \ldots, A_m$ be given matrices, and write

$$\sigma_i = s_i(A_1) \ldots s_i(A_m) \quad (i=1, \ldots, n). \qquad \text{(1)}$$

Then

$$\sup | \operatorname{tr} ([U_1A_1 \ldots U_mA_mU_{m+1}]_k) | = \sigma_1 + \ldots + \sigma_k,$$

$$\sup | \det ([U_1A_1 \ldots U_mA_mU_{m+1}]_k) = \sigma_1 \ldots \sigma_k,$$

where both upper bounds are taken with respect to all sets of unitary matrices $U_1, \ldots, U_{m+1}$. We shall obtain the following generalization of these formulae.

THEOREM 1. *Let $A_1, \ldots, A_m$ be any given matrices, and let $\sigma_1, \ldots, \sigma_n$ be defined by* (1). *Then*

$$\sup | E_j([U_1A_1 \ldots U_mA_mU_{m+1}]_k) | = E_j(\sigma_1, \ldots, \sigma_k),$$

where the upper bound is taken with respect to all sets of unitary matrices $U_1, \ldots, U_{m+1}$.

For the more special case of normal matrices, we shall establish

THEOREM 2. *Let N be a normal matrix with characteristic roots $\omega_1, \ldots, \omega_n$, where $|\omega_1| \geqslant \ldots \geqslant |\omega_n|$; and let r be a positive integer. Then*

$$\sup E_j([V^*(N^*U^*)^r(UN)^rV]_k) = E_j(|\omega_1|^{2r}, \ldots, |\omega_k|^{2r}),$$

where the upper bound is taken with respect to all pairs of unitary matrices U, V.

When $j=1$, this theorem can be stated in the form

$$\sup \sum_{i=1}^{k} \| (UN)^r x_i \|^2 = \sum_{i=1}^{k} |\omega_i|^{2r};$$

here the upper bound is taken with respect to all unitary matrices U and all sets of orthonormal vectors $x_1, \ldots, x_k$. This result is due to Ky Fan [**2**, Theorem 2].

Finally, we shall obtain a relation for hermitian matrices.

https://doi.org/10.1017/S2040618500033827 Published online by Cambridge University Press

THEOREM 3. *If H is a non-negative hermitian matrix with characteristic roots $\omega_1 \geqslant \dots \geqslant \omega_n$, then*

$$\sup E_j([U^* H U]_k) = E_j(\omega_1, \dots, \omega_k),$$

where the upper bound is taken with respect to all unitary matrices U.

For $j=1$, this result reduces to [**2**, Theorem 1]†; for $j=k$ it reduces to [**3**, Lemma 3].

Theorem 1 (or, more precisely, a result equivalent to it) was established recently by Marcus and Moyls [**7**, Theorem 3]. Their statement and proof of the theorem involve concepts (such as those of exterior products and compound matrices) which are proper to multilinear algebra. It may, however, be of some interest to observe that an entirely elementary treatment is possible, and that the results stated above follow almost immediately from known inequalities. Our discussion is based largely on the work of de Bruijn [**1**].

Preliminary results

We shall need to quote a few results from the literature.

LEMMA 1. *Let $\lambda_1, \dots, \lambda_n$ be complex numbers and $\mu_1, \dots, \mu_n$ real numbers, and suppose that*

$$|\lambda_1| \geqslant \dots \geqslant |\lambda_n|, \quad \mu_1 \geqslant \dots \geqslant \mu_n$$

and

$$|\lambda_1 \dots \lambda_i| \leqslant \mu_1 \dots \mu_i \quad (i=1, \dots, n). \qquad \text{.......(2)}$$

Then

$$|E_j(\lambda_1, \dots, \lambda_n)| \leqslant E_j(\mu_1, \dots, \mu_n) \quad (j=1, \dots, n).$$

This result was noted by de Bruijn [**1**, p. 27] as an immediate consequence of the theorems of Horn [**6**, Theorem 3] and Weyl [**9**, p. 410].‡ It should be pointed out that, in de Bruijn's statement of the hypothesis, there is equality in (2) for $i=n$. To obtain our form of the lemma, we simply put $\lambda_{n+1} = \mu_{n+1} = 0$.

LEMMA 2. *For any matrices A and B, we have*

$$\prod_{i=1}^{r} s_i(AB) \leqslant \prod_{i=1}^{r} s_i(A)\, s_i(B) \quad (r=1, \dots, n).$$

This is a special case of an inequality due to Horn [**5**, Theorem 3]. Other proofs have been given by Visser and Zaanen [**8**, Theorem 2] and by de Bruijn [**1**, Theorem 6.2].

LEMMA 3. *Let A be any matrix and denote by $\alpha_1, \dots, \alpha_k$ the characteristic roots of $[A]_k$. If $|\alpha_1| \geqslant \dots \geqslant |\alpha_k|$, then*

$$|\alpha_1 \dots \alpha_i| \leqslant s_1(A) \dots s_i(A) \quad (i=1, \dots, k).$$

This result is due to de Bruijn [**1**, Theorem 8.1].

Proofs of the theorems

For any matrix A, we have by Lemmas 3 and 1,

$$|E_j([A]_k)| \leqslant E_j\{s_1(A), \dots, s_k(A)\}. \qquad \text{.......(3)}$$

(Alternatively, we can establish this inequality by making use of the Fischer-Courant minimax principle and a theorem of Weyl [**9**, equation (4)]).

Now let $U_1, \dots, U_{m+1}$ be any unitary matrices, and write $B = U_1 A_1 \dots U_m A_m U_{m+1}$. Then, by Lemma 2,

† With the inessential difference that Ky Fan's statement is not restricted to *non-negative* hermitian matrices.

‡ A direct proof of the lemma (i.e. one independent of matrix theory) can also be given. One such proof has, in fact, been communicated to the author by Professor R. Rado.

$$\prod_{i=1}^{r} s_i(B) \leqslant \prod_{i=1}^{r} \sigma_i \quad (r=1, \dots, k);$$

and therefore, by Lemma 1,

$$E_j\{s_1(B), \dots, s_k(B)\} \leqslant E_j(\sigma_1, \dots, \sigma_k).$$

Using (3) with B in place of A, we therefore infer that, for any unitary matrices $U_1, \dots, U_{m+1}$,

$$| E_j([U_1 A_1 \dots U_m A_m U_{m+1}]_k) | \leqslant E_j(\sigma_1, \dots, \sigma_k). \quad \dots\dots\dots(4)$$

Moreover, there exist unitary matrices V_r, W_r such that

$$V_r A_r W_r = \operatorname{diag}\{s_1(A_r), \dots, s_n(A_r)\} \quad (r=1, \dots, m).$$

Hence

$$V_1 A_1 W_1 \dots V_m A_m W_m = \operatorname{diag}(\sigma_1, \dots, \sigma_n);$$

and this implies that, for a special choice of U's, the relation (4) reduces to an equality. The proof of Theorem 1 is therefore complete.

To prove Theorem 2, we note that if $\omega_1, \dots, \omega_n$ are the characteristic roots of N, then $|\omega_1|, \dots, |\omega_n|$ are the singular values of N and also of N^*. Hence, if U and V are unitary, we have, by (4),

$$E_j([V^*(N^*U^*)^r(UN)^r V]_k) \leqslant E_j(|\omega_1|^{2r}, \dots, |\omega_k|^{2r}). \quad \dots\dots\dots(5)$$

Now N can be written in the form

$$N = W^*.\operatorname{diag}(\omega_1, \dots, \omega_n).W,$$

where W is unitary. Hence (5) reduces to an equality for $V=W^*$, $U=I$; and Theorem 2 is therefore proved. We may note that Theorem 1 leads, in fact, to the following more general result. Let $N_1, \dots, N_m$ be normal matrices; denote by $\omega_1^{(s)}, \dots, \omega_n^{(s)}$ the characteristic roots of N_s, where $|\omega_1^{(s)}| \geqslant \dots \geqslant |\omega_n^{(s)}|$; and put

$$\rho_i = |\omega_i^{(1)} \dots \omega_i^{(m)}| \quad (i=1, \dots, n).$$

Then
$$\sup | E_j([U_1 N_1 \dots U_m N_m U_{m+1}]_k) | = E_j(\rho_1, \dots, \rho_k),$$

where the upper bound is taken with respect to all sets of unitary matrices $U_1, \dots, U_{m+1}$.

Finally, Theorem 3 follows at once from the case $r=1$ of Theorem 2. Alternatively, we can derive it by observing that, if $\xi_1 \geqslant \dots \geqslant \xi_k$ are the characteristic roots of $[H]_k$, then $\xi_i \leqslant \omega_i$ $(i=1, \dots, k)$. Hence

$$E_j([H]_k) \leqslant E_j(\omega_1, \dots, \omega_k)$$

and so, for any unitary matrix U,

$$E_j([U^*HU]_k) \leqslant E_j(\omega_1, \dots, \omega_k).$$

This implies our assertion.

REFERENCES

1. N. G. de Bruijn, Inequalities concerning minors and eigenvalues, *Nieuw Archief v. Wiskunde* (3), **4** (1956), 18–35.

2. Ky Fan, On a theorem of Weyl concerning eigenvalues of linear transformations (I), *Proc. Nat. Acad. Sci.*, **35** (1949), 652–655.

3. Ky Fan, On a theorem of Weyl concerning eigenvalues of linear transformations (II), *Proc. Nat. Acad. Sci.*, **36** (1950), 31–35.

4. Ky Fan, Maximum properties and inequalities for the eigenvalues of completely continuous operators, *Proc. Nat. Acad. Sci.*, **37** (1951), 760–766.

5. A. Horn, On the singular values of a product of completely continuous operators, *Proc. Nat. Acad. Sci.*, **36** (1950), 374–375.

6. A. Horn, On the eigenvalues of a matrix with prescribed singular values, *Proc. Amer. Math. Soc.*, **5** (1954), 4–7.

7. M. Marcus and B. N. Moyls, On the maximum principle of Ky Fan, *Canad. J. Math.*, **9** (1957), 313–320.

8. C. Visser and A. C. Zaanen, On the eigenvalues of compact linear transformations, *Proc. Kon. Ned. Akad. Wetensch.*, Ser. A, **14** (1952), 71–78.

9. H. Weyl, Inequalities between the two kinds of eigenvalues of a linear transformation, *Proc. Nat. Acad. Sci.*, **35** (1949), 408–411.

THE UNIVERSITY
SHEFFIELD