A USEFUL THEOREM IN MATRIX THEORY

S. N. ROY

1. Introduction and notation. In this paper a theorem on the characteristic roots of (square) matrices, believed to be new, is obtained, which has proved extremely useful in a sector of mathematical statistics and which, the author hopes, might be useful in other sectors of applied mathematics. It will be assumed here that the elements of the matrices considered are real or complex numbers. Any capital letter, say M, will stand for a matrix, M' for its transpose, M^* for its conjugate transpose, m_{ij} for its (ij)th element, $\bar{m}_{ij}$ for the conjugate of m_{ij}, and $M(p\times q)$ will denote that the matrix consists of p rows and q columns. In this paper, so far as the new results are concerned, only square matrices will be discussed for which we have $p=q\geqq 1$. For any matrix M, the rank will be denoted by $r(M)$, any characteristic root by $c(M)$. If M is $p\times p$, there will be, of course, p such roots, say $c_1, c_2, \cdots, c_p$. $I(p)$ will stand for a $p\times p$ unit or identity matrix and $D_\lambda(p)$ will stand for a $p\times p$ diagonal matrix whose diagonal elements are, say $\lambda_1, \lambda_2, \cdots, \lambda_p$. In course of this paper, some well known results in matrix algebra will be used [1; 2], including the following:

$$(1.1)\qquad c_i(AB) = c_i(BA), \qquad i = 1, 2, \cdots, p,$$

where A and B are two $p\times p$ matrices. It may be noted that this result can be easily generalized to

$$(1.2)\qquad c[A(p\times q)B(q\times p)] = c[B(q\times p)A(p\times q)], \qquad p \leqq q,$$

with the meaning that any nonzero root of AB is a nonzero root of BA and vice versa. Notice that, out of the q $c(BA)$'s, $q-p$ must be necessarily zero.

(1.3) If M is a $p\times p$ hermitian and positive-definite or positive-semidefinite matrix (to be called respectively p.d. or p.s.d.), i.e., if all $c(M)$'s are positive or non-negative, then there is a unitary matrix P such that $M=PD_{c(M)}P^*$.

$$(1.4)\qquad r[A(p\times q)] = r[A^*(q\times p)] = r(AA^*).$$

(1.5) $A(p\times q)A^*(q\times p)$ is hermitian and at least p.s.d., so that all $c(AA^*)$ will be at least non-negative. The case of $q=p$ is the one that will be actually used in this paper.

Received by the editors January 15, 1954.

2.1. **The bound theorem on characteristic roots.** If A and B are two $p\times p$ matrices of which at least one is nonsingular, then for all the characteristic roots $c(AB)$ we have

$$c_{\min}(AA^*)c_{\min}(BB^*) \leqq c(AB)\bar{c}(AB) \leqq c_{\max}(AA^*)c_{\max}(BB^*), \tag{2.1.1}$$

where $c_{\min}$ and $c_{\max}$ stand respectively for the smallest and the largest characteristic root (each, of course, non-negative).

PROOF. By (1.3) there are unitary matrices, say P_A and P_B, such that

$$AA^* = P_A D_{c(AA^*)} P_A^* \quad \text{and} \quad BB^* = P_B D_{c(BB^*)} P_B^*. \tag{2.1.2}$$

Notice that all $c(AA^*)$ and $c(BB^*)$ are real and non-negative and, if we further assume that A is nonsingular, then all $c(AA^*)$ are positive.

From (2.1.2) it follows that it is possible to find unitary matrices, say Q_A and Q_B, such that

$$A = P_A D_{c^{1/2}(AA^*)} Q_A \quad \text{and} \quad B = P_B D_{c^{1/2}(BB^*)} Q_B. \tag{2.1.3}$$

We have now, for any characteristic root of AB,

$$\begin{aligned} c(AB) &= c(P_A D_{c^{1/2}(AA^*)} Q_A P_B D_{c^{1/2}(BB^*)} Q_B) \\ &= c(D_{c^{1/2}(AA^*)} Q_A P_B D_{c^{1/2}(BB^*)} Q_B P_A) \qquad \text{(using (1.2))} \\ &= c(D_\lambda R D_\mu S), \end{aligned}$$

where

$$R = Q_A P_B, \quad S = Q_B P_A, \quad \lambda_i = c_i^{1/2}(AA^*), \quad \mu_i = c_i^{1/2}(BB^*) \quad (i = 1, 2, \cdots, p). \tag{2.1.4}$$

Notice that, since P_A, P_B, Q_A, Q_B are all unitary, therefore, $Q_A P_B$ and $Q_B P_A$ are also each unitary.

Now, if c is to be a characteristic root of AB, there exists a set of (complex) numbers $z_1, \cdots, z_p$, not all of which are zero, such that the following set of equations is satisfied:

$$\sum_{j,k=1}^{p} \lambda_i r_{ij} \mu_j s_{jk} z_k = c z_i \qquad (i = 1, 2, \cdots, p). \tag{2.1.5}$$

Remembering that λ_i's and μ_j's are real, and (here) μ_j's are non-negative and λ_i's positive, and dividing by λ_i, and taking the conjugate of (2.1.5), multiplying the two and then summing over $i=1, 2, \cdots, p$, we have

$$\sum_i \sum_{j,j',k,k'} r_{ij}\bar{r}_{ij'}\mu_j\mu_{j'}s_{jk}\bar{s}_{j'k'}z_k\bar{z}_{k'} = c\bar{c}\sum_i z_i\bar{z}_i/\lambda_i^2. \tag{2.1.6}$$

Now, since R is unitary, we have

$$\sum_i r_{ij}\bar{r}_{ij'} = \delta_{jj'} \qquad \text{(where } \delta \text{ is the Kronecker symbol).} \tag{2.1.7}$$

Thus (2.1.6) reduces to

$$\begin{aligned} c\bar{c}\sum_i z_i\bar{z}_i/\lambda_i^2 &= \sum_{j,k,k'} \mu_j^2 s_{jk}\bar{s}_{jk'}z_k\bar{z}_{k'} \\ &= \sum_j \mu_j^2 \sum_k (s_{jk}z_k)\cdot \sum_{k'} (\bar{s}_{jk'}\bar{z}_{k'}). \end{aligned} \tag{2.1.8}$$

It is easy to see that the coefficients of $1/\lambda_i^2$'s on the left-hand side and those of μ_j^2's on the right-hand side are each at least non-negative. Hence, if we replace all μ_j's by $\mu_{\max}$ and all λ_i's by $\lambda_{\max}$, the right-hand side is increased (or at least not diminished) and the left-hand side is diminished (or at least not increased). We have thus

$$\begin{aligned} (c\bar{c}/\lambda_{\max}^2)\sum_i z_i\bar{z}_i &\leqq \mu_{\max}^2 \sum_j \sum_{k,k'} s_{jk}\bar{s}_{jk'}z_k\bar{z}_{k'}, \\ \text{i.e.,} \qquad &\leqq \mu_{\max}^2 \sum_j \delta_{kk'}z_k\bar{z}_{k'} \text{ (since } S \text{ is unitary)}, \\ \text{i.e.,} \qquad &\leqq \mu_{\max}^2 \sum_i z_i\bar{z}_i. \end{aligned} \tag{2.1.9}$$

Since $\sum_i z_i\bar{z}_i$ is positive, it follows that

$$c\bar{c} \leqq \lambda_{\max}^2\mu_{\max}^2,$$

i.e.,

$$c(AB)\bar{c}(AB) \leqq c_{\max}(AA^*)c_{\max}(BB^*). \tag{2.1.10}$$

Likewise in (2.1.8), replacing all λ_i's by $\lambda_{\min}$ and all μ_j's by $\mu_{\min}$ and arguing in a similar manner, we have

$$c_{\min}(AA^*)c_{\min}(BB^*) \leqq c(AB)\bar{c}(AB). \tag{2.1.11}$$

Combining (2.1.10) and (2.1.11), we have the theorem (2.1.1). It is easy to see that (2.1.1) can be generalized to the case of the product of any finite number of matrices $A_1, A_2, \cdots, A_n$, provided that not more than one of them is possibly singular, the rest being all non-singular. Thus we have the theorem

$$\prod_{i=1}^n c_{\min}(A_iA_i^*) \leqq c\left(\prod_{i=1}^n A_i\right)\bar{c}\left(\prod_{i=1}^n A_i\right) \leqq \prod_{i=1}^n c_{\max}(A_iA_i^*). \tag{2.1.12}$$

2.2. Some useful special cases of (2.1.1). Putting $A=I(p)$ in

(2.1.1), we have

$$c_{\min}(AA^*) \leqq c(A)\bar{c}(A) \leqq c_{\max}(AA^*), \tag{2.2.1}$$

a result due to Professor E. T. Browne [1].

If A and B are both hermitian, then it is well known that all $c(A)$'s and $c(B)$'s are real and also that

$$c(AA^*) = c(A^2) = c^2(A) \quad \text{and} \quad c(BB^*) = c(B^2) = c^2(B). \tag{2.2.2}$$

Hence, in this case, (2.1.1) will reduce to

$$c^2_{\min}(A)c^2_{\min}(B) \leqq c(AB)\bar{c}(AB) \leqq c^2_{\max}(A)c^2_{\max}(B). \tag{2.2.3}$$

Suppose that A and B are both hermitian and one of them, say A, is p.d. and the other, i.e. B, at least p.s.d. Then all $c(A)$'s are real and positive, all $c(B)$'s real and at least non-negative. Also AB is at least p.s.d., so that all $c(AB)$ are real and at least non-negative. In this case, (2.1.1) reduces to

$$c_{\min}(A)c_{\min}(B) \leqq c(AB) \leqq c_{\max}(A)c_{\max}(B). \tag{2.2.4}$$

References

1. E. T. Browne, Bull. Amer. Math. Soc. vol. 34 (1928) pp. 363–368.
2. C. C. MacDuffee, *The theory of matrices*, New York, Chelsea, 1946, pp. 17–29.

University of North Carolina