

SOME METRIC INEQUALITIES IN THE SPACE OF MATRICES¹

KY FAN AND A. J. HOFFMAN

1. In this note we shall prove three inequalities suggested by the well-known analogy between matrices and complex numbers. These are the matricial analogues of the following simple numerical inequalities:

(a) If $z = |z| \cdot e^{i\theta}$, θ real, and if α is any real number, then

$$|z - e^{i\theta}| \leq |z - e^{i\alpha}| \leq |z + e^{i\theta}|.$$

(b) If z is complex, x real, then

$$|z - \operatorname{Re} z| \leq |z - x|.$$

(c) If x and y are real, then

$$\frac{1}{2} \left| \frac{x-i}{x+i} - \frac{y-i}{y+i} \right| \leq |x-y|.$$

In developing the matricial statements corresponding to (a), (b), (c), we must replace the modulus of a complex number by a suitably chosen norm for matrices. Let M_n denote the linear space of all square matrices of order n with complex coefficients. A norm on M_n is a real-valued function $\|\cdot\|$ defined on M_n such that: (i) $\|A\| \geq 0$; (ii) $\|A\| = 0$ if and only if $A = 0$; (iii) $\|cA\| = |c| \cdot \|A\|$ for any complex number c ; (iv) $\|A+B\| \leq \|A\| + \|B\|$. A norm $\|\cdot\|$ is *unitarily invariant* if it satisfies the additional condition: (v) $\|A\| = \|UA\| = \|AU\|$ for every unitary matrix U of order n . It is rather noteworthy that *the matricial analogues of (a), (b), (c) hold for any norm that is unitarily invariant*.

For any matrix $A \in M_n$, the non-negative square roots of the eigenvalues of A^*A will be called *singular values* of A . The following result of J. von Neumann [3] characterizes all unitarily invariant norms on M_n . For any symmetric gauge function² Φ of n real variables, the function $\|\cdot\|$ defined on M_n by

Received by the editors, January 25, 1954.

¹ The preparation of this paper was sponsored in part by the Office of Scientific Research, USAF; and in part by the Office of Naval Research, USN.

² Following von Neumann, a gauge function Φ (in the sense of Minkowski) is called symmetric if $\Phi(t_1, t_2, \dots, t_n) = \Phi(\epsilon_1 t_{j_1}, \epsilon_2 t_{j_2}, \dots, \epsilon_n t_{j_n})$ for any combination of signs $\epsilon_i = \pm 1$ and for any permutation (j_1, j_2, \dots, j_n) of $(1, 2, \dots, n)$. For general properties of symmetric gauge functions, see [4, pp. 84-92].

$$(1) \quad \|A\| = \Phi(\alpha_1, \alpha_2, \dots, \alpha_n) \quad (A \in M_n),$$

where $\alpha_1, \alpha_2, \dots, \alpha_n$ are the singular values of A , Φ is a unitarily invariant norm. Conversely, every unitarily invariant norm on M_n can be obtained in this way; let $\Phi(\alpha_1, \alpha_2, \dots, \alpha_n) = \|A\|$, where A is a diagonal matrix with diagonal entries $\alpha_1, \alpha_2, \dots, \alpha_n$.

Let $A, B \in M_n$. Let $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$ and $\beta_1 \geq \beta_2 \geq \dots \geq \beta_n$ be the singular values of A and B respectively. Then it is known [1, Theorem 4] that $\|A\| \leq \|B\|$ for every unitarily invariant norm $\|\cdot\|$ if and only if

$$(2) \quad \sum_{i=1}^k \alpha_i \leq \sum_{i=1}^k \beta_i \quad (1 \leq k \leq n).$$

According to these known results, the proof of the matricial analogues of (a), (b), (c) amounts to showing certain inequalities involving singular values.

In our proof of the matricial analogue of (a), we shall need the following theorem:³ If X, Y, Z are Hermitian matrices of order n , with eigenvalues

$$x_1 \geq x_2 \geq \dots \geq x_n, \quad y_1 \geq y_2 \geq \dots \geq y_n, \quad z_1 \geq z_2 \geq \dots \geq z_n$$

respectively, and if $X - Y = Z$, then

$$(3) \quad \text{Max}_{i_1 < i_2 < \dots < i_k} \sum_{i=1}^k (x_{i_i} - y_{i_i}) \leq \sum_{i=1}^k z_i \quad (1 \leq k \leq n).$$

2. THEOREM 1. Let $A \in M_n$ and $A = UH$, where U is unitary and H is Hermitian positive semi-definite. Then for any unitary matrix $W \in M_n$,

$$(4) \quad \|A - U\| \leq \|A - W\| \leq \|A + U\|$$

holds for every unitarily invariant norm.

Since the norm is unitarily invariant, we have

$$\begin{aligned} \|A \mp U\| &= \|U(H \mp I)\| = \|H \mp I\|, \\ \|A - W\| &= \|U(H - U^*W)\| = \|H - U^*W\|. \end{aligned}$$

It follows that Theorem 1 is equivalent to the following apparently less general theorem:

THEOREM 1'. Let $H, V \in M_n$. If H is Hermitian positive semi-definite,

³ See [5, Theorem 2]. An equivalent geometric formulation of this result is stated in [2, Theorem 1].

nite, and V is unitary, then

$$(5) \quad \|H - I\| \leq \|H - V\|,$$

$$(6) \quad \|H - V\| \leq \|H + I\|$$

hold for every unitarily invariant norm.

PROOF OF (5). We first digress to define, for any matrix M of order n , the Hermitian matrix

$$\tilde{M} = \begin{pmatrix} 0 & M \\ M^* & 0 \end{pmatrix}$$

of order $2n$. Then it is easy to see that the eigenvalues of \tilde{M} are precisely the singular values of M and their negatives.⁴

Let $A = H - I$, $B = H - V$. Let $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$ be the singular values of A ; $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n$ be the singular values of B ; $\eta_1 \geq \eta_2 \geq \cdots \geq \eta_n$ be the eigenvalues (also singular values) of H . Then

$$\sum_{i=1}^k \alpha_i = \text{Max}_{i_1 < i_2 < \cdots < i_k} \sum_{i=1}^k |\eta_{i_i} - 1| \quad (1 \leq k \leq n).$$

To prove (2), which will imply (5), we must show

$$(7) \quad \text{Max}_{i_1 < i_2 < \cdots < i_k} \sum_{i=1}^k |\eta_{i_i} - 1| \leq \sum_{i=1}^k \beta_i \quad (1 \leq k \leq n).$$

This inequality (7) will be obtained, if we apply the theorem mentioned above to $\tilde{H} - \tilde{V} = \tilde{B}$. In fact, according to the remark made at the beginning of this proof, the eigenvalues of \tilde{H} , \tilde{V} , and \tilde{B} are

$$\begin{aligned} &\eta_1, \eta_2, \cdots, \eta_n, -\eta_n, -\eta_{n-1}, \cdots, -\eta_1, \\ &1, 1, \cdots, 1, -1, -1, \cdots, -1, \\ &\beta_1, \beta_2, \cdots, \beta_n, -\beta_n, -\beta_{n-1}, \cdots, -\beta_1 \end{aligned}$$

respectively. Thus (5) is proved.

PROOF OF (6). Let $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$ and $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n$ be the singular values of $H - V$ and $H + I$ respectively. Let $\eta_1 \geq \eta_2 \geq \cdots \geq \eta_n$ be the eigenvalues (also singular values) of H . We are to prove (2).

It is known [1, Theorem 2] that if X, Y, Z are any three matrices of order n , with singular values

$$x_1 \geq x_2 \geq \cdots \geq x_n, \quad y_1 \geq y_2 \geq \cdots \geq y_n, \quad z_1 \geq z_2 \geq \cdots \geq z_n$$

⁴ The authors are grateful to H. Wielandt for calling this useful fact to their attention.

respectively, and if $X + Y = Z$, then

$$z_{i+i+1} \leq x_{i+1} + y_{i+1},$$

and in particular:

$$z_i \leq x_i + y_1 \quad (1 \leq i \leq n).$$

If we apply this fact to $H - V = H + (-V)$, then

$$\alpha_i \leq \eta_i + 1 \quad (1 \leq i \leq n).$$

As $\eta_i + 1 = \beta_i$, we have not only (2), but actually

$$(8) \quad \alpha_i \leq \beta_i \quad (1 \leq i \leq n).$$

THEOREM 2. *Let $A, H \in M_n$. If H is Hermitian, then*

$$(9) \quad \left\| A - \frac{A + A^*}{2} \right\| \leq \|A - H\|$$

holds for every unitarily invariant norm.

PROOF. Observe first that the singular values of a matrix X are the same as those of X^* . Combining this fact with von Neumann's characterization of all unitarily invariant norms on M_n , it follows that $\|X\| = \|X^*\|$ for every unitarily invariant norm.

We write

$$A - \frac{A + A^*}{2} = \frac{A - H}{2} + \frac{H - A^*}{2},$$

which implies $\|A - (A + A^*)/2\| \leq \|A - H\|/2 + \|H - A^*\|/2$. This is precisely (9), since $\|H - A^*\| = \|A - H\|$.

REMARK 1. Corresponding to the inequality $|\operatorname{Re} z| \leq |z|$ for complex numbers z , we have the trivial inequality $\|(A + A^*)/2\| \leq \|A\|$ for matrices. In this connection, we mention the following less trivial proposition: *Let $A \in M_n$. If $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ are the eigenvalues of $(A + A^*)/2$, and if $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$ are the singular values of A , then*

$$(10) \quad \lambda_i \leq \alpha_i \quad (1 \leq i \leq n).$$

Observe that $\|(A + A^*)/2\| \leq \|A\|$ insures only that $\sum_{i=1}^k \lambda_i \leq \sum_{i=1}^k \alpha_i$ ($1 \leq k \leq n$). To prove (10), let $A = UH$, where U is unitary and H is Hermitian positive semi-definite. Let x_1, x_2, \dots, x_n be n orthonormal eigenvectors of A^*A such that $A^*Ax_i = \alpha_i^2 x_i$ ($1 \leq i \leq n$). Let μ_i denote the maximum of the inner product $((A + A^*)y/2, y)$, when the vector y varies under the conditions

$$(11) \quad \|y\| = 1; \quad (x_j, y) = 0 \quad \text{for } 1 \leq j \leq i-1.$$

Then by the minimum-maximum principle:

$$(12) \quad \lambda_i \leq \mu_i \quad (1 \leq i \leq n).$$

On the other hand, since $A = UH$, we have

$$\begin{aligned} \left(\frac{A + A^*}{2} y, y \right) &= \operatorname{Re} (Ay, y) = \operatorname{Re} (Hy, U^*y) \\ &\leq \|Hy\| \cdot \|U^*y\| = \|Hy\| \cdot \|y\|. \end{aligned}$$

If $\|y\| = 1$, then

$$\left(\frac{A + A^*}{2} y, y \right) \leq \|Hy\| = (A^*Ay, y)^{1/2}.$$

Hence μ_i is not greater than the maximum of $(A^*Ay, y)^{1/2}$, when y varies under conditions (11). But this maximum is precisely α_i , so we have $\mu_i \leq \alpha_i$, which together with (12) proves (10).

REMARK 2. If $\rho_1 \geq \rho_2 \geq \dots \geq \rho_n$ and $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$ denote the singular values of $(A + A^*)/2$ and A respectively, then inequalities $\rho_i \leq \alpha_i$ ($1 \leq i \leq n$) are generally false. This can be seen by taking

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

THEOREM 3. Let $H, K \in M_n$ be both Hermitian, and let U, V be their Cayley transforms:

$$U = (H - iI)(H + iI)^{-1}, \quad V = (K - iI)(K + iI)^{-1}.$$

If $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$ and $\beta_1 \geq \beta_2 \geq \dots \geq \beta_n$ are the singular values of $(U - V)/2$ and $H - K$ respectively, then

$$(13) \quad \alpha_i \leq \beta_i \quad (1 \leq i \leq n).$$

Consequently, we have

$$(14) \quad \|U - V\| \leq 2\|H - K\|$$

for every unitarily invariant norm.

PROOF. We write

$$U = I - 2i(H + iI)^{-1}, \quad V = I - 2i(K + iI)^{-1}$$

so that

$$\begin{aligned} (U - V)/2i &= (K + iI)^{-1} - (H + iI)^{-1} \\ &= (K + iI)^{-1}[(H + iI) - (K + iI)](H + iI)^{-1}, \end{aligned}$$

or

$$(15) \quad (U - V)/2i = (K + iI)^{-1}(H - K)(H + iI)^{-1}.$$

It is known [1, Theorem 2] that if X , Y , Z are any three matrices of order n , with singular values

$$x_1 \geq x_2 \geq \cdots \geq x_n, \quad y_1 \geq y_2 \geq \cdots \geq y_n, \quad z_1 \geq z_2 \geq \cdots \geq z_n$$

respectively, and if $XY = Z$, then

$$z_{i+j+1} \leq x_{i+1} \cdot y_{j+1}.$$

The singular values of $(U - V)/2i$ are obviously also those of $(U - V)/2$. Let $\eta_1 \geq \eta_2 \geq \cdots \geq \eta_n$ and $\kappa_1 \geq \kappa_2 \geq \cdots \geq \kappa_n$ be the singular values of $(H + iI)^{-1}$ and $(K + iI)^{-1}$ respectively. Applying the inequality just mentioned to (15), we get

$$\alpha_{i+j+k+1} \leq \kappa_{i+1} \beta_{j+1} \eta_{k+1}.$$

In particular:

$$(16) \quad \alpha_i \leq \kappa_1 \beta_1 \eta_1 \quad (1 \leq i \leq n).$$

On the other hand, from

$$(H + iI)^{-1*}(H + iI)^{-1} = (H^2 + I)^{-1},$$

we infer that $\eta_1 \leq 1$. Similarly, $\kappa_1 \leq 1$. Hence (16) implies (13).

REFERENCES

1. K. Fan, *Maximum properties and inequalities for the eigenvalues of completely continuous operators*, Proc. Nat. Acad. Sci. U.S.A. vol. 37 (1951) pp. 760-766.
2. V. B. Lidskiĭ, *The proper values of the sum and product of symmetric matrices* (in Russian), Doklady Akad. Nauk SSSR vol. 75 (1950) pp. 769-772.
3. J. von Neumann, *Some matrix-inequalities and metrization of matrix-space*, Tomsk Univ. Rev. vol. 1 (1937) pp. 286-300.
4. R. Schatten, *A theory of cross-spaces*, Princeton University Press, 1950.
5. H. Wielandt, *An extremum property of sums of eigenvalues*, Proc. Amer. Math. Soc. vol. 6 (1955) pp. 106-110.

UNIVERSITY OF NOTRE DAME,
AMERICAN UNIVERSITY, AND
NATIONAL BUREAU OF STANDARDS