Spectral inequalities and equalities involving products of matrices

Chi-Kwong Li[1]

Department of Mathematics, College of William & Mary, Williamsburg, Virginia 23187
(ckli@math.wm.edu)

Yiu-Tung Poon

Department of Mathematics, Iowa State University, Ames, Iowa 50011
(poon@math.iastate.edu)

Abstract

Let $A_1, \dots, A_k$ be $n \times n$ matrices. We studied inequalities and equalities involving eigenvalues, diagonal entries, and singular values of $A_0 = A_1 \cdots A_k$ and those of $A_1, \dots, A_k$. It is shown that the matrices attaining equalities often have special reducible structure. The results are then applied to study normality and reducibility of matrices, extending some results and answering some questions of Miranda, Wang and Zhang.

Keywords: matrices, eigenvalues, singular values.
AMS Classifications: 15A42, 15A18

1 Introduction

Let $A \in M_n$. Denote by $s_1(A) \geq \cdots \geq s_n(A)$ the singular values of A, $\lambda_1(A), \dots, \lambda_n(A)$ the eigenvalues of A with $|\lambda_1(A)| \geq \cdots \geq |\lambda_n(A)|$, and $d_1(A), \dots, d_n(A)$ the diagonal entries of A. Let $A_1, \dots, A_k \in M_n$ and $A = A_1 \cdots A_k$. It is known [1, 6] that for $r = 1, \dots, n$, one has

$$\left|\prod_{j=1}^{r} \lambda_j(A)\right| \leq \prod_{j=1}^{r} s_j(A) \leq \prod_{j=1}^{r}\prod_{i=1}^{k} s_j(A_i), \tag{1.1}$$

$$\left|\sum_{j=1}^{r} \lambda_j(A)\right| \leq \sum_{j=1}^{r} |\lambda_j(A)| \leq \sum_{j=1}^{r}\prod_{i=1}^{k} s_j(A_i), \tag{1.2}$$

$$\left|\sum_{j=1}^{r} d_j(A)\right| \leq \sum_{j=1}^{r} |d_j(A)| \leq \sum_{j=1}^{r} s_j(A) \leq \sum_{j=1}^{r}\prod_{i=1}^{k} s_j(A_i). \tag{1.3}$$

In this paper, we characterize those matrices $A_1, \dots, A_k$ for which any one of the equalities in (1.1) – (1.3) holds. If the equality under consideration does not involve $\prod_{j=1}^{r}\prod_{i=1}^{k} s_j(A_i)$ or $\sum_{j=1}^{r}\prod_{i=1}^{k} s_j(A_i)$, then one can only deduce conditions on the matrix A, and such conditions have been determined in [3]. Thus, we will focus on equalities that always involve the quantities $\prod_{j=1}^{r}\prod_{i=1}^{k} s_j(A_i)$ or $\sum_{j=1}^{r}\prod_{i=1}^{k} s_j(A_i)$. It turns out that the extreme matrices $A_1, \dots, A_k$ attaining the equalities often have special reducible structure. The results are then used to study normality and reducibility of matrices, extending the results in [7, 10] and answering

[1]Research supported by an NSF grant.

some questions in [7]. For example, in [7, Lemma 3 and Theorem 2], characterizations were given to those matrices A_1, A_2 and $A = A_1A_2$ in M_n such that

$$\sum_{j=1}^{r} s_j(A) = \sum_{j=1}^{r}\prod_{i=1}^{2} s_j(A_i) \qquad \text{and} \qquad \sum_{j=1}^{r} |d_j(A)| = \sum_{j=1}^{r}\prod_{i=1}^{2} s_j(A_i).$$

Furthermore, it was suggested (Comment 2 in [7]) that the results can be extended to more than two matrices. We show in Section 2 (the discussion after Theorem 2.2) that this comment is not accurate and extend the results utilizing the information of the ranks of the matrices $A_1, \ldots, A_k$ (Theorem 2.4). Furthermore, our results (Theorem 2.2 and Corollary 2.7) answer the questions raised in Comments 3 and 5 in [7]. In [10], the authors studied the equality cases in (1.2) when $r = n$ and $A_i = B$ or B^* for a fixed B. Some sufficient conditions for the matrix B to be normal were given. However, those conditions are not necessary in general. A complete understanding of these conditions will follow readily from our results in Section 4 (see Corollary 4.2).

We shall use the following notation of majorization in our discussion, see [6]. For two real vectors x and y in $\mathbf{R}^n$, if the sum of the m largest entries of x is not larger than that of y for $m = 1, \ldots, n$, we write

$$x \prec_w y;$$

if in addition that the sum of all the entries of x is the same as that of y, we write

$$x \prec y.$$

For two nonnegative vectors x and y in $\mathbf{R}^n$, if the product of the m largest entries of x is not larger than that of y for $m = 1, \ldots, n$, we write

$$\ln x \prec_w \ln y;$$

if in addition that the product of all the entries of x is the same as that of y, we write

$$\ln x \prec \ln y.$$

For a complex vector $x = (x_1, \ldots, x_n)$ we write $|x| = (|x_1|, \ldots, |x_n|)$.

We remark that our results are valid for real matrices as long as the statements do not involve complex numbers.

2 Main Theorems

We begin with the following lemma which plays a crucial role in proving our main theorems.

Lemma 2.1 *Let $A = A_1 \cdots A_k$, where $k \geq 2$ and $A_i \in M_n$ satisfies* $\operatorname{rank}(A_i) \geq r$ *for all $i = 1, \ldots, k$. Suppose*

$$\prod_{j=1}^{r} s_j(A) = \prod_{j=1}^{r}\prod_{i=1}^{k} s_j(A_i)$$

If U_0 and U_n are unitary matrices such that $U_0^*AU_k = B \oplus C$ with $B \in M_r$ satisfying $\det(B) = \prod_{j=1}^r \prod_{i=1}^k s_j(A_i)$*, then there exist $U_1, \ldots, U_{k-1}$ such that $U_{i-1}^* A_i U_i = B_i \oplus C_i$ with $B_i \in M_r$ with* $\det(B_i) = \prod_{j=1}^r s_j(A_i)$ *for $i = 1, \ldots, k$.*

Proof. We prove the result by induction on $k \geq 2$. Suppose $k = 2$. Let U_0 and U_2 be unitary so that $U_0^* A_1 A_2 U_2 = B \oplus C$ with $B \in M_r$ such that $\det(B) = \prod_{j=1}^r s_j(A_1) s_j(A_2)$. Let U_1 be unitary so that the last $(n-r)$ columns of U_1 is orthogonal to the first r columns of A_2U_2. Then $U_1^* A_2 U_2 = \begin{pmatrix} B_2 & X_2 \\ 0 & C_2 \end{pmatrix}$ for some $B_2 \in M_r$. If $U_0^* A_1 U_1 = \begin{pmatrix} B_1 & X_1 \\ Y_1 & C_1 \end{pmatrix}$ with $B_1 \in M_r$, then the leading $r \times r$ submatrix of $U_0^* A_1 A_2 U_2$ is just $B_1 B_2$. It is known [9] that $s_j(B_i) \leq s_j(A_i)$ for $j = 1, \ldots, r$, and $i = 1, 2$. Hence

$$\prod_{j=1}^r s_j(A_1) s_j(A_2) = \det(B) = \det(B_1 B_2) \leq \prod_{j=1}^r s_j(B_1) s_j(B_2) \leq \prod_{j=1}^r s_j(A_1) s_j(A_2).$$

It follows that $s_j(A_i) = s_j(B_i)$ for $j = 1, \ldots, r$, and $i = 1, 2$. By Lemma 2.1 in [3], we conclude that $A_i = B_i \oplus C_i$ for $i = 1, 2$ as asserted.

Now, suppose the result is valid for the product of $k-1$ matrices with $k > 2$. Let $A = A_1 \cdots A_k$ and $\tilde{A}_2 = A_2 \cdots A_k$. Then $\prod_{j=1}^p s_j(\tilde{A}_2) \leq \prod_{j=1}^p \prod_{i=2}^k s_j(A_i)$ for $p = 1, \ldots, n$. Let U_0 and U_k be unitary such that $U_0^* A_1 \tilde{A}_2 U_k = B \oplus C$ with $B \in M_r$ satisfying $\det(A) = \prod_{j=1}^r \prod_{i=1}^k s_j(A_i)$. By the induction assumption on the product $A_1 \tilde{A}_2$, we see that there exists U_1 such that $U_0^* A_1 U_1 = B_1 \oplus C_1$ and $U_1^* \tilde{A}_2 U_k = \tilde{B}_2 \oplus \tilde{C}_2$ with $\det(B_1) = \prod_{j=1}^r s_j(A_1)$, and hence $\det(\tilde{B}_2) = \prod_{j=1}^r s_j(A) / \det(B_1) = \prod_{j=1}^r \prod_{i=2}^k s_j(A_i)$. By induction assumption on $U_1 \tilde{A}_2 U_k$, there exist unitary $U_2, \ldots, U_{k-1}$ such that $U_{i-1}^* A_i U_i = B_i \oplus C_i$ with $B_i \in M_r$ with $\det(B_i) = \prod_{j=1}^r s_j(A_i)$ for $i = 2, \ldots, k$. □

Theorem 2.2 *Let $1 \leq r \leq n$. Suppose $A_1, \ldots, A_k \in M_n$, where $k > 1$, and $A = \prod_{j=1}^k A_j$. Then*

$$\prod_{j=1}^r s_j(A) = \prod_{j=1}^r \prod_{i=1}^k s_j(A_i) \tag{2.1}$$

if and only if one of the following is satisfied.

(a) *$r = n$.*

(b) *One of the matrix A_j has rank less than r.*

(c) *There exist unitary matrices $U_0, U_1, \ldots, U_k$ such that $U_{i-1}^* A_i U_i = B_i \oplus C_i$ so that $B_i \in M_r$ has singular values $s_1(A_i), \ldots, s_r(A_i)$.*

Consequently,

$$\left|\prod_{j=1}^{r} \lambda_j(A)\right| = \prod_{j=1}^{r}\prod_{i=1}^{k} s_j(A_i). \tag{2.2}$$

if and only if (a), (b), *or* (c) *with* $U_0 = U_k$ *holds* .

Proof. If (a) or (b) holds then clearly we have (2.1). If (c) holds, then for $X = [I_r|0_{r,n-r}]$ we have

$$\prod_{j=1}^{r}\prod_{i=1}^{k} s_j(A_i) = |\det(XU_0AU_kX^t)| \le \prod_{j=1}^{r} s_j(A).$$

By (1.1), we see that (2.1) holds.

Now, suppose $A = A_1 \cdots A_k$ satisfies (2.1). Assume that neither (a) nor (b) holds. Let U_0 and U_k be unitary such that $U_0AU_k = \mathrm{diag}\,(s_1(A),\ldots,s_n(A))$. By Lemma 2.1, we get condition (c).

The proof of the last assertion is similar. The only difference is in the last part. Suppose (2.2) is true. Then $|\prod_{j=1}^r \lambda_j(A)| = \prod_{j=1}^r s_j(A)$. If neither (a) nor (b) holds, then by Theorem 2.2 in [3] there exists a unitary matrix U_k such that $U_k^*AU_k = B \oplus C$, where $B \in M_r$ has singular values $s_1(A),\ldots,s_r(A)$. By Lemma 2.1, we get condition (c). □

Consider matrices $A_1,\ldots,A_k$ and $A = A_1\cdots A_k$ in M_n. When $k = 2$, it was shown in [7, Lemma 3, Theorem 2] that

$$\sum_{j=1}^{r} s_j(A) = \sum_{j=1}^{r}\prod_{i=1}^{k} s_j(A_i) \tag{2.3}$$

if and only if there exist unitary U_0, U_1, U_2 such that

$$U_{i-1}^*A_iU_i = \mathrm{diag}\,(s_1(A_i),\ldots,s_r(A_i)) \oplus B_i, \tag{2.4}$$

for $i = 1, 2$; and

$$\sum_{j=1}^{r} |d_j(A)| = \sum_{j=1}^{r}\prod_{i=1}^{k} s_j(A_i) \tag{2.5}$$

if and only if $A = B \oplus C$ with $B \in M_r$ and there exists a diagonal unitary matrix $D \in M_r$ such that DB is positive semi-definite with eigenvalues

$$\prod_{i=1}^{k} s_1(A_i) \ge \cdots \ge \prod_{i=1}^{k} s_r(A_i),$$

and (2.4) holds. At the end of the paper, the author said (in Comment 2) that the same result holds for $k > 2$. However, this comment is not accurate. In fact, if $\mathrm{rank}(A_1) = p < r$, then neither (2.3) nor (2.4) convey much information about $s_j(A)$ for $j > p$. To see an extreme example, let $A_1 = 0$. Then (2.3) and (2.5) hold, but nothing can be said about $A_2,\ldots,A_k$. Such a problem does not arise when $k = 2$ because of the following reason: if A_1 has rank $p < r$, one can first find unitary X_0, X_1, X_2 so that $X_{i-1}^* A_i X_i = \mathrm{diag}\,(s_1(A_i), \ldots, s_p(A_i)) \oplus B_i$, where $B_1 = 0$. Then one can find unitary Y_1 and Y_2 so that $Y_1^* B_2 Y_2 = \mathrm{diag}\,(s_{p+1}(A_2), \ldots, s_n(A_2))$. As a result, $U_0 = X_0$, $U_1 = X_1(I_p \oplus Y_1)$ and $U_2 = X_2(I_p \oplus Y_2)$ satisfy (2.4) for $i = 1, 2$.

To get around the problem mentioned above, we make use of the quantity

$$m = \begin{cases} r & \text{if } k = 2, \\ \min\left(\{r\} \cup \{\mathrm{rank}(A_i) : 1 \le i \le k\}\right) & \text{otherwise}, \end{cases} \tag{2.6}$$

in our results. In particular, it follows from Theorem 2.4 below that if $\mathrm{rank}(A_i) \ge r$ for all i, then (2.3) holds if and only if there exist unitary $U_0, \ldots, U_k$ such that (2.4) holds for $i = 1, \ldots, k$. Actually, (2.3) is just one of the many conditions leading to (2.4) as shown in the theorem following the next definition.

Definition 2.3 *Let $\mathbf{P} = [0, \infty)$. For $r \ge 1$ let $\mathcal{F}_r$ be the set of functions $f : \mathbf{P}^r \to \mathbf{R}$ such that for any $x, y \in \mathbf{P}^r$ with $\ln x \prec_w \ln y$ we have $f(x) = f(y)$ if and only if $x = Qy$ for a permutation matrix Q.*

The set $\mathcal{F}_r$ contains many different functions, see [6, Chapter 3]. For examples, the mth elementary symmetric function $E_m(x_1, \ldots, x_r)$ with $1 \le m < r$, and the ℓ_p-norm on $\mathbf{R}^r$ with $1 \le p$. In particular, the function $f(x_1, \cdots, x_r) = x_1 + \cdots + x_r$ is $\mathcal{F}_r$. Thus it follows from (1.1) that conditions (a)–(c) in the following theorem are equivalent.

Theorem 2.4 *Let $A_1, \ldots, A_k \in M_n$ with $1 < k$. If $1 \le r \le n$, then the following conditions are equivalent.*

(a) $f(s_1(A), \ldots, s_r(A)) = f(\prod_{i=1}^k s_1(A_i), \ldots, \prod_{i=1}^k s_r(A_i))$ *for all (or some)* $f \in \mathcal{F}_r$.

(b) $s_1(A) + \cdots + s_r(A) = \prod_{i=1}^k s_1(A_i) + \cdots + \prod_{i=1}^k s_r(A_i)$.

(c) $s_j(A) = \prod_{i=1}^k s_j(A_i)$ *for* $j = 1, \ldots, r$.

(d) *There exist unitary $U_0, \ldots, U_k$ such that $U_{i-1}^* A_i U_i = \mathrm{diag}\,(s_1(A_i), \ldots, s_m(A_i)) \oplus B_i$ for $i = 1, \ldots, k$, where m is defined as in (2.6).*

Proof. By the remark before the theorem, conditions (a) – (c) are equivalent. The implication (d) $\Rightarrow$ (a) is clear.

Now, suppose (c) holds. If $k = 2$, then condition (d) holds by the result in [7]. Suppose $k > 2$. Then $s_j(A) = \prod_{i=1}^k s_j(A_i)$ for $j = 1, \ldots, m$. We prove condition (d) by induction on m as follows. If $m = 1$, the result follows from Theorem 2.2. Suppose that the result is valid for the $m - 1$ singular values. Since $s_1(A) = \prod_{i=1} s_1(A_i)$, by Theorem 2.2, there exist unitary $X_0, \ldots, X_k$ such that $X_{i-1}^* A_i X_i = [s_1(A_i)] \oplus \tilde{A}_i$. By induction assumption on $\tilde{A}_1, \ldots, \tilde{A}_k$, there exist unitary $Y_0, \ldots, Y_k$ such that $Y_{i-1}^* \tilde{A}_i Y_i = \mathrm{diag}\,(s_2(A_i), \ldots, s_m(A_i)) \oplus B_i$ for $i = 1, \ldots, k$. Setting $U_i = ([1] \oplus Y_i) X_i$ for $i = 0, \ldots, k$, we get condition (d). □

Corollary 2.5 *Suppose $A_1, \dots, A_k \in M_n$ with $1 < k$, and $1 \le r \le n$. Let m be defined as in (2.6). The following conditions* (a.i) – (a.iii) *are equivalent, and conditions* (b.i) – (b.iii) *are equivalent.*

(a.i) $\sum_{j=1}^{r} |d_j(A)| = \sum_{j=1}^{r} \prod_{i=1}^{k} s_j(A_i)$.

(a.ii) *$A = B \oplus C$ with $B \in M_r$ and there exists a diagonal unitary matrix $D \in M_r$ satisfying DB is positive semi-definite with eigenvalues*

$$\prod_{i=1}^{k} s_1(A_i) \ge \cdots \ge \prod_{i=1}^{k} s_r(A_i).$$

(a.iii) *There exist unitary matrices $U_0, \cdots, U_k \in M_n$ such that $U_0\tilde{D} = U_k = V \oplus I_{n-r}$ for some diagonal unitary matrix $\tilde{D}$ and $U_{i-1}^* A_i U_i = \mathrm{diag}\,(s_1(A_i), \dots, s_m(A_i)) \oplus C_i$ for $i = 1, \dots, k$.*

(b.i) $\left|\sum_{j=1}^{r} d_j(A)\right| = \sum_{j=1}^{r} \prod_{i=1}^{k} s_j(A_i)$

(b.ii) *$A = B \oplus C$ where $B \in M_r$ is a unit multiple of a positive semi-definite matrix with eigenvalues*

$$\prod_{i=1}^{k} s_1(A_i) \ge \cdots \ge \prod_{i=1}^{k} s_r(A_i).$$

(b.iii) *There exist unitary matrices $U_0, \cdots, U_k \in M_n$ such that $U_0 = e^{\sqrt{-1}t} U_k = V \oplus I_{n-r}$ with $t \in \mathbf{R}$, and $U_{i-1}^* A_i U_i = \mathrm{diag}\,(s_1(A_i), \dots, s_m(A_i)) \oplus C_i$ for $i = 1, \dots, k$.*

Proof. The implication (a.i) $\Rightarrow$ (a.ii) follows from the fact that (2.5) holds if and only if the last two inequalities in (1.3) become equalities, and Theorem 3.1 in [3]. The implication (a.ii) $\Rightarrow$ (a.iii) follows from Theorem 2.2. The implication (a.iii) $\Rightarrow$ (a.i) is clear.

The proof of the equivalence of (b.i) – (b.iii) is similar. $\square$

We now turn to the inequalities in (1.2). Clearly, the first inequality becomes an equality if and only if all $\lambda_i(A)$ has the same argument for $i = 1, \dots, r$. The equality case of the second inequality can be treated in the same way as the equality case of $\left|\prod_{j=1}^{r} d_j(A)\right| = \prod_{j=1}^{r} \prod_{i=1}^{k} s_j(A_i)$ once we put A in triangular form by a suitable unitary similarity.

Corollary 2.6 *Suppose $A_1, \dots, A_k \in M_n$ with $1 < k$, and $1 \le r \le n$. Let m be defined as in (2.6). The following conditions are equivalent.*

(a) *$f(|\lambda_1(A)|, \dots, |\lambda_r(A)|) = f(\prod_{i=1}^{k} s_1(A_i), \dots, \prod_{i=1}^{k} s_r(A_i))$ for all (or some) $f \in \mathcal{F}_r$.*

(b) $|\lambda_1(A)| + \cdots + |\lambda_r(A)| = \prod_{i=1}^{k} s_1(A_i) + \cdots + \prod_{i=1}^{k} s_r(A_i)$.

(c) $|\lambda_j(A)| = \prod_{i=1}^k s_j(A_i)$ *for* $j = 1, \ldots, r$.

(d) *There exist unitary* $U_0, \ldots, U_k$ *such that* $U_k U_0^*$ *is a diagonal matrix and* $U_{i-1}^* A_i U_i = \mathrm{diag}\,(s_1(A_i), \ldots, s_m(A_i)) \oplus B_i$ $i = 1, \ldots, k$.

Corollary 2.7 *Suppose* $A_1, \ldots, A_k \in M_n$ *with* $1 < k$, *and* $1 \le r \le n$. *Let* m *be defined as in (2.6). Then*

$$\left|\sum_{j=1}^r \lambda_j(A)\right| = \sum_{j=1}^r \prod_{i=1}^k s_j(A_i) \tag{2.7}$$

if and only if there exist unitary $U_0, \ldots, U_k$ *such that* $U_k U_0^*$ *is a scalar matrix and* $U_{i-1}^* A_i U_i = \mathrm{diag}\,(s_1(A_i), \ldots, s_m(A_i)) \oplus B_i$ *for* $i = 1, \ldots, k$.

Define the *generalized spectral radius* by

$$\rho(A_1, \ldots, A_k) = \max\left\{\left|\mathrm{tr}\left(\prod_{j=1}^k (U_j^* A_j U_j\right)\right| : U_j \text{ is unitary and}\right.$$
$$\left. U_j^* A_j U_j \text{ is triangular for } j = 1, \ldots, k\right\}.$$

The quantity $\rho(C, A)$ is known as the C-*spectral radius* of A in the literature (see [4] and [5]) and has been studied as a generalization of the spectral radius of A (when $C = \mathrm{diag}\,(1, 0, \ldots, 0)$). It is known that $\rho(C, A) \le \sum_{j=1}^n s_j(C) s_j(A)$ and the equality case has been determined. Here we study the equality case for the inequality

$$\rho(A_1, \ldots, A_k) \le \sum_{j=1}^n \prod_{i=1}^k s_j(A_i).$$

Proposition 2.8 *Let* $A_1, \ldots, A_k \in M_n$ *such that* $\min\{\mathrm{rank}\,(A_i) : 1 \le i \le k\} = r$. *Then*

$$\rho(A_1, \ldots, A_k) = \sum_{j=1}^n \prod_{i=1}^k s_j(A_i) \tag{2.8}$$

if and only if A_i *is unitarily similar to* $\mathrm{diag}\,(\lambda_1(A_i), \ldots, \lambda_r(A_i)) \oplus B_i$ *with* $B_i \in M_{n-r}$ *for all* i *so that* $\prod_{i=1}^k \lambda_j(A_i) = e^{\sqrt{-1}t} \prod_{i=1}^k s_j(A_i)$, $t \in \mathbf{R}$, *for all* $j = 1, \ldots, r$.

Proof. The ($\Leftarrow$) part is clear. Suppose (2.8) holds. Let U_i be unitary such that $U_i^* A_i U_i$ is in upper triangular form satisfying

$$\left|\mathrm{tr}\left(\prod_{i=1}^k (U_i^* A_i U_i\right)\right| = \rho(A_1, \ldots, A_k) = \sum_{j=1}^n \prod_{i=1}^k s_j(A_i) = \sum_{j=1}^r \prod_{i=1}^k s_j(A_i).$$

Let $A = \prod_{j=1}^r U_j^* A_j U_j$. Since

$$\left|\sum_{j=1}^r \lambda_j(A)\right| \leq \sum_{j=1}^r s_j(A) \leq \sum_{j=1}^r \prod_{i=1}^k (U_i^* A_i U_i),$$

we see that all the inequalities become equalities. It follows that the first r diagonal entries of A must equal $e^{\sqrt{-1}t}\prod_{i=1}^k s_1(A_i), \ldots, e^{\sqrt{-1}t}\prod_{i=1}^k s_r(A_i)$ for some $t \in \mathbf{R}$. Hence, the first r diagonal entries of each $U_j^* A_j U_j$ must be $\lambda_j(A_i)$ with $|\lambda_j(A_i)| = s_j(A_i)$ for $j = 1, \ldots, r$. As a result, $U_j^* A_j U_j = D_j \oplus B_j$ for some diagonal matrix $D_j \in M_r$. Applying a suitable permutation similarity to all D_j yields the conclusion. □

3 Powers of a Single Matrix

The results in Section 2 can be used to study normality of matrices. We begin with the following theorem.

Theorem 3.1 *Let $A \in M_n$. The following conditions are equivalent.*

(a) *A is normal.*

(b) *There exists $m > 1$ such that $s_j(A^m) = s_j(A)^m$ for $j = 1, \ldots, n$.*

(c) *There exists $m > 1$ such that $f(s_1(A^m), \ldots, s_n(A^m)) = f(s_1(A)^m, \ldots, s_n(A)^m)$ for all (or some) $f \in \mathcal{F}_n$.*

(d) *There exists $m > 1$ such that $s_1(A^m) + \cdots + s_n(A^m) = s_1(A)^m + \ldots + s_n(A)^m$.*

(e) *There exists $m \geq 1$ such that $|\lambda_j(A^m)| = s_j(A)^m$ for $j = 1, \ldots, n$.*

(f) *There exists $m \geq 1$ such that $f(|\lambda_1(A^m)|, \ldots, |\lambda_n(A^m)|) = f(s_1(A)^m, \ldots, s_n(A)^m)$ for all (or some) $f \in \mathcal{F}_n$.*

(g) *There exists $m \geq 1$ such that $|\lambda_1(A^m)| + \cdots + |\lambda_n(A^m)| = s_1(A)^m + \ldots + s_n(A)^m$.*

(h) *There exists $m > 1$ and a unitary U such that $\sum_{j=1}^n |d_j(U^* A^m U)| = \sum_{j=1}^n s_j(A)^m$.*

(i) *There exists $m > 1$ and a unitary U such that $\sum_{j=1}^n |\mathrm{tr}\,(A^m U)| = \sum_{j=1}^n s_j(A)^m$.*

Proof. If (a) holds, then all other conditions hold. By Theorem 2.4, Corollary 2.5, and Corollary 2.6, if any of (b) – (i) holds, then $s_j(A^2) = s_j(A)^2$ for all $j = 1, \ldots, n$. As a result, $\mathrm{tr}\, A^2(A^*)^2 = \mathrm{tr}\,(AA^*)^2$ and hence $\mathrm{tr}\,(AA^* - A^*A)(AA^* - A^*A) = 0$. Thus A is normal. □

Corollary 3.2 *Let $A \in M_n$, and $m \geq 1$. Then $|\mathrm{tr}\, A^m| = \sum_{j=1}^n s_j(A)^m$ if and only if A is normal and the unitary part of A in its polar decomposition $A = PU$ satisfies $U^m = e^{\sqrt{-1}t}I$ for some $t \in \mathbf{R}$.*

Proof. The sufficiency part is clear. For the necessity part, note that

$$|\operatorname{tr} A^m| = \left|\sum_{j=1}^n \lambda_j(A^m)\right| \le \sum_{j=1}^n |\lambda_j(A^m)| \le \sum_{j=1}^n s_j(A)^m.$$

Since inequalities become equalities, we see that A is normal by Proposition 3.1 and λ_j^m have the same argument for all j. Thus U^m is a scalar matrix. □

Sometimes, we can use equality cases involving part of the singular values of A^m to derive reducibility for A.

Corollary 3.3 *Suppose $A \in M_n$. The following conditions are equivalent.*

(a) *A is unitarily similar to $B \oplus C$, where $B \in M_r$ is normal with singular values $s_j(A)$ for $j = 1, \dots, r$.*

(b) *There exists $m \ge 1$ such that $|\lambda_j(A^m)| = s_j(A)^m$ for $j = 1, \dots, r$.*

(c) *There exists $m \ge 1$ such that $|\lambda_1(A^m)| + \cdots + |\lambda_r(A^m)| = s_1(A)^m + \cdots + s_r(A)^m$.*

(d) *There exists $m \ge 1$ such that $f(|\lambda_1(A^m)|, \dots, |\lambda_r(A^m)|) = f(s_1(A)^m, \dots, s_r(A)^m)$ for all (or some) $f \in \mathcal{F}_n$.*

Consequently, $|\sum_{j=1}^r \lambda(A^m)| = \sum_{j=1}^r s_j(A)^m$ if and only if (a) holds and the unitary part of the polar decomposition of $B = PU$ satisfies $U^m = e^{\sqrt{-1}t} I_r$.

Corollary 3.4 *Let $A \in M_n$ and $m > 1$.*

(i) *We have $\sum_{j=1}^r |d_j(A^m)| = \sum_{j=1}^r s_j(A)^m$ if and only if $A = B \oplus C$, where $B \in M_r$ is normal with singular values $s_1(A), \dots, s_r(A)$.*

(ii) *We have $|\sum_{j=1}^r d_j(A^m)| = \sum_{j=1}^r s_j(A)^m$ if and only if $A = B \oplus C$, where $B \in M_r$ is normal with singular values $s_1(A), \dots, s_r(A)$ and the unitary part of the polar decomposition of $B = PU$ satisfies $U^m = e^{\sqrt{-1}t} I_r$.*

Corollary 3.5 *Let $A \in M_n$ have rank at least r. The following conditions are equivalent.*

(a) *A is unitarily similar $B \oplus C$ with $B \in M_r$ satisfying $\det(B) = \prod_{j=1}^r s_j(A)$.*

(b) *There exists $m \ge 1$ such that $|\prod_{j=1}^r \lambda_j(A^m)| = \prod_{j=1}^r s_j(A)^m$.*

(c) *A is unitarily similar to a matrix with diagonal entries $d_1, \dots, d_n$ such that $\sum_{j=1}^r |d_j| = \sum_{j=1}^r s_j(A)$.*

Apart from the above corollaries, for a given $r < n$, even if there exist $m > 1$ such that $s_j(A^m) = s_j(A)^m$ for all $j = 1, \ldots, r$, we may not be able to get reducibility for A. For example, let $A \in M_n$ be the upper triangular Jordan block of zero. If $r \leq n-2$, then $s_j(A^2) = s_j(A)^2$ for all $j = 1, \ldots, r$. Clearly, A has no reducing subspace. Nonetheless, one can use the argument in [8] (see also [2, pp.44-45]) to get the following.

Theorem 3.6 *Let $A \in M_n$, and let p be the degree of the minimal polynomial of A. The following conditions are equivalent.*

(a) *A is normal.*

(b) *There exists an integer $m \geq p$ such that $s_j(A^m) = s_j(A)^m$ for all $j = 1, \ldots, m$.*

(c) *There exists an integer $m \geq p$ such that $s_1(A^m) + \cdots + s_n(A^m) = s_1(A)^m + \cdots + s_n(A)^m$.*

(d) *There exists an integer $m \geq p$ such that for all (or some) $f \in \mathcal{F}_n$,*

$$f(s_1(A^m), \ldots, s_n(A^m)) = f(s_1(A)^m, \ldots, s_n(A)^m).$$

4 Words involving A and A^*

In [10], the authors used some trace equalities of matrices of the form $A_1 \cdots A_k$ with $A_i \in \{A, A^*\}$ to give sufficient condition for the normality of A. Such a product is denoted by $W(A, A^*)$ and referred to as a word with letters A or A^*. If $W(A, A^*) = A^m$ or $(A^*)^m$, then we are back to the study in Section 3. We also exclude the words $W(A, A^*) = (AA^*)^m$ or $(A^*A)^m$, which reduce to the problem of studying $tr(X^m)$ with $X = AA^*$ or A^*A. In the following, we show that one can get necessary and sufficient conditions for the trace equalities considered in [10], and obtain other equivalent conditions for normality in terms of other trace equalities.

Theorem 4.1 *Let $A \in M_n$, and let $W(A, A^*)$ be a word of length $m > 1$ not equal to A^m, $(A^*)^m$, $(AA^*)^{m/2}$ or $(A^*A)^{m/2}$. Then the following conditions* (a.i) – (a.iii) *are equivalent, and conditions* (b.i) – (b.iii) *are equivalent.*

(a.i) *$f\left(s_1(W(A, A^*)), \ldots, s_n(W(A, A^*))\right) = f\left(s_1(A)^m, \ldots, s_n(A)^m\right)$ for all (or some) $f \in \mathcal{F}_n$.*

(a.ii) *$s_j(W(A, A^*)) = s_j(A)^m$ for all $j = 1, \ldots, n$.*

(a.iii) *A is normal or $W(A, A^*)$ is of the form $A(A^*A)^{(m-1)/2}$ or $A^*(AA^*)^{(m-1)/2}$.*

(b.i) *$f\left(|\lambda_1(W(A, A^*))|, \ldots, |\lambda_n(W(A, A^*))|\right) = f\left(s_1(A)^m, \ldots, s_n(A)^m\right)$ for all (or some) $f \in \mathcal{F}_n$.*

(b.ii) *$|\lambda_j(W(A, A^*))| = s_j(A)^m$ for all $j = 1, \ldots, n$.*

(b.iii) *A is normal.*

Proof. (a.i) $\Rightarrow$ (a.ii) follows from the definition of f and the fact that

$$\ln\left(s_1(W(A,A^*)),\ldots,s_n(W(A,A^*))\right) \prec \ln\left(s_1(A)^m,\ldots,s_n(A)^m\right).$$

(a.ii) $\Rightarrow$ (a.iii) Suppose (a.ii) holds. If there exists 2 consecutive A in $W(A,A^*)$, then there exist unitary matrices X,Y,Z such that $X^*AY = Y^*AZ = \mathrm{diag}\,(s_1(A),\ldots,s_n(A))$ by Theorem 2.3 (c). Thus, we see that A^2 have singular values $s_1(A)^2,\ldots,s_n(A)^2$, and hence A is normal by Proposition 3.1. Similarly, if there exist 2 consecutive A^* in $W(A,A^*)$, then we are done. If none of the above two cases holds, then $W(A,A^*)$ must be of the form $(AA^*)^rA$ or $(A^*A)^rA^*$ for some r.

(a.iii) $\Rightarrow$ (a.i) is clear.

The proof of (b.iii) $\Rightarrow$ (b.i) $\Rightarrow$ (b.ii) is similar. Suppose (b.ii) holds. Then (a.ii) holds, and then (a.iii) follows. Since XY and YX have the same eigenvalues, if $W(A,A^*)$ has the form $A(A^*A)^{(m-1)/2}$, we see that $|\lambda_j((A^*A)^{(m-1)/2}A)| = |\lambda_j(W(A,A^*))| = s_j(A)^m$ for all j. Now, the last two letters of the word $(A^*A)^{(m-1)/2}A$ are equal to A. We conclude that A is normal. Similarly, we can show that A is normal if $W(A,A^*) = A^*(AA^*)^{(m-1)/2}$. $\square$

The following corollary follows easily.

Corollary 4.2 *Let $A \in M_n$, and let $W(A,A^*)$ be a word of length $m > 1$ not equal to A^m, $(A^*)^m$, $(AA^*)^{m/2}$ or $(A^*A)^{m/2}$. Suppose p of the letter in $W(A,A^*)$ equal A and $q = m-p$. Then*

$$|\mathrm{tr}\, W(A,A^*)| = \sum_{j=1}^n s_j(A)^m$$

if and only if A is unitarily similar to the diagonal matrix DB such that D is a diagonal unitary matrix and $B = \mathrm{diag}\,(s_1(A),\ldots,s_n(A))$ satisfying $D^{p-q} = \mu I$; in particular,

(i) *A is simply normal without additional condition if $p = q$,*

(ii) *A is a multiple of a positive semi-definite matrix if $|p-q| = 1$,*

(iii) *A is a multiple of a Hermitian matrix if $|p-q| = 2$.*

In [10], the authors proved that A is normal if

$$\mathrm{tr}\, W(A,A^*) = \sum_{j=1}^n s_j(A)^m. \tag{4.1}$$

By Corollary 4.2, we see that one can get more precise information about A if (4.1) holds. Furthermore, we have the following corollary.

Corollary 4.3 *Let $A \in M_n$, and let $W(A, A^*)$ be a word with $2m$ letters so that m of the letters equal to A, but $W(A, A^*) \neq (AA^*)^m$ or $(A^*A)^m$. Then A is normal if and only if one or both of the following equalities holds: $W(A, A^*) = A^m(A^*)^m$, $W(A, A^*) = (A^*)^m A^m$.*

Remark 4.4 *In general, one cannot get reducibility condition on A if the equality involves only part of the singular values of $W(A, A^*)$. For example, let $A \in M_n$ be the upper triangular elementary Jordan block of zero, and let $W(A, A^*) = AAA^*A^*$. Then $\sum_{j=1}^k \lambda_j(W(A, A^*)) = \sum_{j=1}^k s_j(W(A, A^*)) = \sum_{j=1}^k s_j(A)^4$ for $k = 1, \ldots, n-2$, but A has no reducing subspace.*

References

[1] A. Horn, On the singular values of a product of completely continuous operators, *Proc. Nat. Acad. Sci. USA* 36 (1950), 374-375.

[2] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis,* Cambridge University Press, 1991.

[3] C.K. Li, Matrices with some extremal properties, *Linear Algebra Appl.* 101 (1988), 255-267.

[4] C.K. Li, C-numerical ranges and C-numerical radii, *Linear and Multilinear Algebra* 371 (1994), 51-82.

[5] C.K. Li, T.Y. Tam and N.K. Tsing, The generalized spectral radius, numerical radius and spectral norm, *Linear and Multilinear Linear Algebra* 16 (1984), 215-237.

[6] A.W. Marshall and I. Olkin, *Inequalities: The Theory of Majorizations and Its Applications,* Academic Press, 1979.

[7] H.F. Miranda, Singular values, diagonal elements, and extreme matrices, *Linear Algebra Appl.* 305 (2000), 151-159.

[8] V. Pták, Isometric parts of an operator and the critical exponent, *Časopis pro Pěstovani Matematiky* 101 (1976), 383-388.

[9] R.C. Thompson, Principal submatrices IX: Interlacing inequalities for singular values of submatrices, *Linear Algebra Appl.* 5 (1972), 1-12.

[10] B.-Y. Wang and F. Zhang, Words and normality of matrices, *Linear and Multilinear Algebra* 40 (1995), 111-118.