Inequalities for Mixed Schur Functions

R. B. Bapat
Indian Statistical Institute
7, S.J.S. Sansanwal Marg
New Delhi, 110016, India[*]

Submitted by Richard A. Brualdi

ABSTRACT

If $A^k = (a_{ij}^k)$, $k = 1, 2, \dots, n$, are $n \times n$ positive semidefinite matrices and if $\alpha : S_n \to \mathbf{C}$, where S_n is the symmetric group of degree n, an inequality is obtained for the "mixed Schur function,"

$$\sum_{\sigma, \tau \in S_n} \alpha(\sigma) \overline{\alpha(\tau)} \prod_{i=1}^{n} a^i_{\sigma(i)\tau(i)}.$$

When the matrices A^k, $k = 1, 2, \dots, n$, are all equal, we get some known results due to Schur as consequences of the inequality. It is also deduced that the mixed discriminant of a set of positive semidefinite matrices exceeds or equals the geometric mean of their determinants.

1. INTRODUCTION

If A is a hermitian positive definite (positive semidefinite) matrix, we write $A > 0$ ($A \geqslant 0$). Also, $A \geqslant B$ means that $A \geqslant 0$, $B \geqslant 0$, and $A - B \geqslant 0$. The determinant and the permanent of the matrix A will be denoted by $|A|$ and $\operatorname{per} A$, respectively. As usual, S_n denotes the symmetric group of degree n, and $\epsilon(\sigma) = 1$ or -1 according as $\sigma \in S_n$ is even or odd.

In this paper we prove some inequalities for "mixed" Schur functions (this concept will be made precise later). As an example, we state the following, which is a special case of a more general inequality which we prove.

[*]This work was done while the author was visiting California State University at Hayward.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/86/$3.50

THEOREM 1. *Let $A^k=(a^k_{ij})$, $k=1,2,\dots,n$, be $n\times n$ positive semidefinite matrices. Then*

$$\frac{1}{n!}\sum_{\sigma,\tau\in S_n}\prod_{i=1}^{n}a^i_{\sigma(i)\tau(i)}\geqslant\prod_{k=1}^{n}|A^k|^{1/n},\tag{i}$$

$$\frac{1}{n!}\sum_{\sigma,\tau\in S_n}\epsilon(\sigma\tau)\prod_{i=1}^{n}a^i_{\sigma(i)\tau(i)}\geqslant\prod_{k=1}^{n}|A^k|^{1/n}.\tag{ii}$$

The expression on the left-hand side of (ii) has been termed "mixed discriminant" in the literature [1, 8]. Analogously, the expression appearing in (i) may be thought of as the "mixed permanent" of the matrices $A^1,\dots,A^n$. When $A^1,\dots,A^n$ are all equal to A, the inequality in (i) specializes to the well-known result of Schur that if $A\geqslant 0$, then $\operatorname{per}A\geqslant|A|$.

We now introduce some notation. It will be convenient for us to assume throughout that the elements of S_n have been ordered in the following way: if $\sigma,\tau\in S_n$, then σ precedes τ if σ^{-1} precedes τ^{-1} in the lexicographic ordering, or equivalently, if the first nonzero difference $\sigma^{-1}(i)-\tau^{-1}(i)$, $i=1,2,\dots,n$, is negative. Thus, the elements of S_3 are ordered as follows:

$$123, 132, 213, 312, 231, 321$$

Let $A^k=(a^k_{ij})$, $k=1,2,\dots,n$, be $n\times n$ matrices, and let $A^1\times A^2\times\cdots\times A^n$ be their Kronecker product. Let $\Pi(A^1,\dots,A^n)$ be the $n!\times n!$ matrix defined as follows. Index the rows as well as the columns of $\Pi(A^1,\dots,A^n)$ by S_n. If $\sigma,\tau\in S_n$, then the (σ,τ) entry of $\Pi(A^1,\dots,A^n)$ is equal to

$$a^1_{\sigma(1)\tau(1)}a^2_{\sigma(2)\tau(2)}\cdots a^n_{\sigma(n)\tau(n)}.$$

It may be verified that $\Pi(A^1,\dots,A^n)$ is a principal submatrix of $A^1\times A^2\times\cdots\times A^n$. The following result is immediate from this observation.

LEMMA 2. *If $A^1,\dots,A^n$ are hermitian positive semidefinite, then so is $\Pi(A^1,\dots,A^n)$.*

We now give an example. If A^1, A^2, A^3 are given by

$$A^1=\begin{bmatrix}2&0&0\\0&2&1\\0&1&3\end{bmatrix},\qquad A^2=\begin{bmatrix}1&0&0\\0&3&2\\0&2&3\end{bmatrix},\qquad A^3=\begin{bmatrix}2&0&0\\0&3&-1\\0&-1&2\end{bmatrix},$$

then

$$\Pi(A^1, A^2, A^3) = \begin{bmatrix} 12 & -4 & 0 & 0 & 0 & 0 \\ -4 & 18 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & -1 & 0 & 0 \\ 0 & 0 & -1 & 9 & 0 & 0 \\ 0 & 0 & 0 & 0 & 12 & 4 \\ 0 & 0 & 0 & 0 & 4 & 18 \end{bmatrix}.$$

Note that since each A^i is a direct sum of a 1×1 matrix and a 2×2 matrix, $\Pi(A^1, A^2, A^3)$ turns out to be a direct sum of three 2×2 matrices due to the ordering of S_3 that we adopt. This observation is used in the proof of Theorem 6.

When $A^k = A$, $k = 1,2,\dots,n$; the matrix $\Pi(A^1,\dots,A^n) = \Pi(A,\dots,A)$ will be denoted simply by $\Pi(A)$, and this agrees with the notation first introduced by Soules [10]. The matrix $\Pi(A)$ has been denoted by $\tilde{A}$ in [2].

In a very important paper, Schur [9] generalized the Hadamard determinant inequality in a substantial way. We now describe two results from that paper. Let A be an $n\times n$ matrix, let G be a subgroup of S_n, and let λ be a character of G. Define the function

$$d_G^\lambda(A) = \sum_{\sigma\in G} \lambda(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}.$$

The following result from [9] is commonly known as Schur's inequality, and it has inspired a great amount of research (see, for example, [4], [5], [6], [7, Chapter VI]).

Theorem 3. *Let $A \geqslant 0$ be an $n\times n$ matrix, let G be a subgroup of S_n, and let λ be a character of G. Then*

$$\frac{1}{\lambda(\mathrm{id})} d_G^\lambda(A) \geqslant |A|.$$

Schur derives Theorem 3 by first proving the following result and then making use of the fact that $[1/\lambda(\mathrm{id})]d_G^\lambda(A)$ is in the field of values of $\Pi(A)$. In fact, if λ is the vector of order $n!$ whose σth entry is $\lambda(\sigma)/\sqrt{o(G)}$, then

$$\frac{1}{\lambda(\mathrm{id})} d_G^\lambda(A) = \lambda^* \Pi(A) \lambda.$$

THEOREM 4. *If $A \geqslant 0$, then $|A|$ is the smallest eigenvalue of $\Pi(A)$.*

It is interesting to note that although Theorem 3 has received a good amount of attention, Theorem 4 has remained more or less unnoticed. This point was also made in [2]. In fact, in a 1983 paper, Soules [10] gives Theorem 4 as an open problem, along with the other (still unsolved) open conjecture that if $A \geqslant 0$, then $\operatorname{per} A$ is the largest eigenvalue of $\Pi(A)$. The result in Theorem 4 has, however, been noted in some works of Marcus [4, 6]. In [4], Marcus gives an alternative proof of Theorem 4 using the Cauchy-Schwartz inequality. The purpose of the present paper is to show that when $A^k \geqslant 0$, $k = 1, 2, \dots, n$, the matrix $\Pi(A^1, \dots, A^n)$ dominates a certain diagonal matrix, whose diagonal entries are in terms of the principal minors of $A^1, \dots, A^n$ (see Theorem 6). The result obtained is more general than Theorem 4.

The determinant and the permanent are two of the most extensively studied functions associated with a matrix. There are various generalizations of these two functions which appear in the literature. If λ is a character of S_n, the term "immanant" has been used by Littlewood [3] for the function $\sum_{\sigma \in S_n} \lambda(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}$. If G is a subgroup of S_n and if λ is a character of G, then the term "generalized matrix function" or "Schur function" has been used to denote the function

$$d_G^{\lambda}(A) = \sum_{\sigma \in G} \lambda(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}.$$

It must be remarked that quite frequently, further restrictions are placed on the character while defining a generalized matrix function. The typical restrictions are that the character is irreducible or that it is of degree 1.

If α is an arbitrary function from S_n to the complex numbers, then we use the term "mixed Schur function" to denote

$$d^{\alpha}(A^1, \dots, A^n) = \sum_{\sigma, \tau \in S_n} \alpha(\sigma) \overline{\alpha(\tau)} \prod_{i=1}^{n} a^{i}_{\sigma(i)\tau(i)}.$$

The main inequality that we obtain for mixed Schur functions is stated in Theorem 7.

2. RESULTS

If A is an $n \times n$ matrix, we denote by A_j, $j = 1, 2, \dots, n$, the principal submatrix of A formed by deleting the first j rows and the first j columns of A. Also, we make the convention that $A_0 = A$ and that $|A_n| = 1$.

THEOREM 5. *Let $A^k > 0$, $k = 1,2,\ldots,n$, be $n \times n$ matrices, and let $B^k > 0$ be obtained from A^k by replacing a_{11}^k with $|A^k|/|A_1^k|$ and by replacing all remaining entries in the first row and the first column of A^k with zeros, $k = 1,2,\ldots,n$. Then*

$$\Pi(A^1,\ldots,A^n) \geqslant \Pi(B^1,\ldots,B^n).$$

Proof. For $k = 1,2,\ldots,n$, let C^k be obtained from A^k by replacing a_{11}^k with $a_{11}^k - |A^k|/|A_1^k|$. It is easy to show (see, for example, [2, Theorem 1]) that $C^k \geqslant 0$. It follows from the definition that

$$\Pi(C^1,\ldots,C^n) = \Pi(A^1,\ldots,A^n) - \Pi(B^1,\ldots,B^n)$$

and the proof is complete, since $\Pi(C^1,\ldots,C^n) \geqslant 0$ by Lemma 2.

THEOREM 6. *Let $A^k > 0$, $k = 1,2,\ldots,n$, be $n \times n$ matrices, and let $Z = (z_{ij})$ be the $n \times n$ matrix defined by*

$$z_{ij} = \frac{|A_{i-1}^j|}{|A_i^j|}, \qquad i,j = 1,2,\ldots,n.$$

Furthermore, let $D(A^1,\ldots,A^n)$ be the diagonal matrix of order $n!$ with its σth diagonal entry equal to $\prod_{i=1}^n z_{i\sigma(i)}$. Then

$$\Pi(A^1,\ldots,A^n) \geqslant D(A^1,\ldots,A^n).$$

Proof. Using the notation and the conclusion of Theorem 5, we have

$$\Pi(A^1,\ldots,A^n) \geqslant \Pi(B^1,\ldots,B^n).$$

In view of the ordering of S_n that we employ, it can be seen that

$$\Pi(B^1,\ldots,B^n) = \bigoplus_{k=1}^{n} \frac{|A^k|}{|A_1^k|}\Pi(A_1^1, A_1^2,\ldots,A_1^{k-1}, A_1^{k+1},\ldots,A_1^n),$$

where $\oplus$ denotes direct sum. Now the result follows by induction.

THEOREM 7. *Let* $A^i > 0$, $i = 1,2,\ldots,n$, *be* $n\times n$ *matrices, and let* $\alpha: S_n \to \mathbf{C}$. *Then*

$$\sum_{\sigma,\tau\in S_n} \alpha(\sigma)\overline{\alpha(\tau)} \prod_{i=1}^{n} a^i_{\sigma(i)\tau(i)} \geqslant \sum_{\sigma\in S_n} |\alpha_\sigma|^2 \prod_{i=1}^{n} z_{i\sigma(i)}, \tag{i}$$

$$\sum_{\sigma,\tau\in S_n} \alpha(\sigma)\overline{\alpha(\tau)} \prod_{i=1}^{n} a^i_{\sigma(i)\tau(i)} \geqslant \operatorname{per} Z \geqslant n!\left(\prod_{k=1}^{n} |A^k|\right)^{1/n}, \tag{ii}$$

if $|\alpha(\sigma)| = 1$ *for all* $\sigma \in S_n$.

Proof. Let α be the vector of order $n!$ whose σth entry is $\alpha(\sigma)$. Then, by Theorem 6,

$$\alpha^*\Pi(A^1,\ldots,A^n)\alpha \geqslant \alpha^* D(A^1,\ldots,A^n)\alpha,$$

and that is (i).

The first inequality in (ii) follows immediately from (i), whereas the second inequality follows by an application of the arithmetic-mean–geometric-mean inequality, since

$$\begin{aligned}\operatorname{per} Z &\geqslant n!\left(\prod_{i=1}^{n} z_{i\sigma(i)}\right)^{1/n!}\\ &= n!\left(\prod_{i,j=1}^{n} z_{ij}\right)^{1/n}\end{aligned}$$

and

$$\begin{aligned}\prod_{i,j=1}^{n} z_{ij} &= \prod_{i=1}^{n}\left(\prod_{j=1}^{n} z_{ij}\right)\\ &= \prod_{i=1}^{n} |A^i|.\end{aligned}$$

That completes the proof of the theorem.

Theorem 1 is clearly a simple consequence of Theorem 7. Setting $A^1 = A$ and $A^2 = \cdots = A^n = B$ in Theorem 1, we get the following.

COROLLARY 8. *Let $A \geqslant 0$, $B \geqslant 0$ be $n \times n$ matrices. Then*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \operatorname{per} B(i,j) \geqslant n|A|^{1/n}|B|^{1-1/n}, \tag{i}$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}|B(i,j)| \geqslant n|A|^{1/n}|B|^{1-1/n}, \tag{ii}$$

where $B(i,j)$ is the submatrix obtained by deleting the ith row and the jth column of B.

It is a pleasure to thank Professor Russ Merris for taking a careful look at the first draft of the paper and for his hospitality during my stay in Hayward.

REFERENCES

1 A. D. Alexandrov, On the theory of mixed volumes of convex bodies (in Russian), *Mat. Sb.* 3(45):227–251 (1938).
2 R. B. Bapat and V. S. Sunder, An extremal property of the permanent and the determinant, *Linear Algebra Appl.* 76:153–163 (1986).
3 D. E. Littlewood, *The Theory of Group Characters*, Oxford U.P., London, 1940.
4 M. Marcus, On two classical results of I. Schur, *Bull. Amer. Math. Soc.* 70:685–688 (1964).
5 M. Marcus and G. W. Soules, Some inequalities for combinatorial matrix functions, *J. Combin. Theory* 2(2):145–163 (1967).
6 M. Marcus, Inequalities for matrix functions of combinatorial interest, *SIAM J. Appl. Math.* 17(6):1023–1031 (1969).
7 R. Merris, *Multilinear Algebra*, Monograph Ser., Inst. for the Interdisciplinary Applications of Algebra and Combinatorics, Univ. of California, Santa Barbara, 1975.
8 H. Minc, Theory of permanents 1978–1981, *Linear and Multilinear Algebra* 12:227–263 (1983).
9 I. Schur, Uber endliche Gruppen und Hermitesche Formen, *Math Z.* 1:184–207 (1918).
10 G. W. Soules, Constructing symmetric nonnegative matrices, *Linear and Multilinear Algebra* 13:241–251 (1983).

Received 14 June 1985; revised 13 November 1985