COMPLETELY STRONG SUPERADDITIVITY OF GENERALIZED MATRIX FUNCTIONS

MINGHUA LIN AND SUVRIT SRA

Abstract. We prove that generalized matrix functions satisfy a block-matrix strong superadditivity inequality over the cone of positive semidefinite matrices. Our result extends a recent result of Paksoy-Turkmen-Zhang [6]. As an application, we obtain a short proof of a classical inequality of Thompson (1961) on block matrix determinants.

1. Introduction

Let $\mathbb{M}_n$ denote the algebra of all $n \times n$ complex matrices. Let $A \subset \mathbb{M}_n$. A functional $f : A \to \mathbb{R}$ is called superadditive if for all $A, B \in A$

$$f(A + B) \geq f(A) + f(B),$$

and it is called strongly superadditive if for all $A, B, C \in A$

$$f(A + B + C) + f(C) \geq f(A + C) + f(B + C).$$

It is known (e.g., [8, Eq.(5)]) that the determinant is strongly superadditive (and so superadditive) over the cone of positive semidefinite matrices. That is,

$$\det(A + B + C) + \det(C) \geq \det(A + C) + \det(B + C)$$

for $A, B, C \succeq 0$.

Definition 1.1. Let $\chi$ be a character of the subgroup $G$ of the symmetric group $S_n$. The generalized matrix function $d^G_\chi : \mathbb{M}_n \to \mathbb{C}$ is defined by

$$d^G_\chi(A) := \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^n a_{i \sigma(i)},$$

where $A = [a_{ij}]$.

When $G = S_n$ and $\chi(\sigma) = \text{sgn}(\sigma)$ then $d^G_\chi(A)$ reduces to the determinant $\det(A)$, while for $\chi(\sigma) \equiv 1$ we obtain $d^G_\chi(A) = \text{per}(A)$, the permanent of $A$.

Recently, Paksoy, Turkmen and Zhang [6] presented a natural extension of (1.1) via an embedding approach and through tensor products. More precisely, for $A, B, C \succeq 0$ they proved

$$d^G_\chi(A + B + C) + d^G_\chi(C) \geq d^G_\chi(A + C) + d^G_\chi(B + C).$$

This paper extends the above-cited strong superadditivity results to block matrices, thereby obtaining “completely strong superadditivity” for generalized matrix functions.

Before stating our problem formally, let us fix some notation. The conjugate transpose of $X \in \mathbb{M}_n$ is denoted by $X^*$. For Hermitian matrices $X, Y \in \mathbb{M}_n$, the inequality $X \succeq Y$ means

---

2010 Mathematics Subject Classification. 15A45, 15A69.

Key words and phrases. Generalized matrix function, strongly superadditive, determinantal inequality, positive definite matrix.

1
$X - Y$ is positive semidefinite. Let $\mathbb{M}_m(\mathbb{M}_n)$ be the algebra of $m \times m$ block matrices with each block in $\mathbb{M}_n$. We will denote members of $\mathbb{M}_m(\mathbb{M}_n)$ via bold letters such as $A$. A map (not necessarily linear) $\phi : \mathbb{M}_n \to \mathbb{M}_k$ is positive if it maps positive semidefinite matrices to positive semidefinite matrices. This map is completely positive if for each positive integer $m$, the blockwise map $\Phi : \mathbb{M}_m(\mathbb{M}_n) \to \mathbb{M}_m(\mathbb{M}_k)$ defined by

$$
\Phi \left( \begin{bmatrix} A_{i,j} \end{bmatrix}_{i,j=1}^m \right) = \begin{bmatrix} \phi(A_{i,j}) \end{bmatrix}_{i,j=1}^m
$$

is positive. The determinant is well-known to be completely positive [2]. More generally, it is known that the generalized matrix functions are completely positive (e.g., [9, Theorem 3.1]).

The following definition extends the notion of strong superadditivity.

**Definition 1.2.** Let $A = [A_{i,j}]_{i,j=1}^m, B = [B_{i,j}]_{i,j=1}^m, C = [C_{i,j}]_{i,j=1}^m \in \mathbb{M}_m(\mathbb{M}_n)$ be Hermitian. A map $\phi : \mathbb{M}_n \to \mathbb{M}_k$ is said to be completely strongly superadditive (CSS) if for each positive integer $m$, the map $\Phi$ defined in (1.4) satisfies

$$
\Phi(A + B + C) + \Phi(C) \geq \Phi(A + C) + \Phi(B + C).
$$

Our main assertion in this paper is as follows.

**Theorem 1.3.** Generalized matrix functions are CSS over the cone of positive semidefinite matrices. In particular, the determinant and permanent are CSS.

We slightly overload the notation and extract a special case for later use. For any $A = [A_{i,j}]_{i,j=1}^m \in \mathbb{M}_m(\mathbb{M}_n)$, define $\det_m(A) := [\det A_{i,j}]_{i,j=1}^m$.

**Corollary 1.4.** Let $A, B \in \mathbb{M}_m(\mathbb{M}_n)$ be positive semidefinite. Then

$$
\det_m(A + B) \geq \det_m(A) + \det_m(B).
$$

In particular,

$$
\det \left( \det_m(A + B) \right) \geq \det \left( \det_m(A) \right) + \det \left( \det_m(B) \right).
$$

The proof of Theorem 1.3 is given in Section 2. In Section 3, we apply Corollary 1.4 to obtain a new proof of a determinantal inequality due to Thompson (1961).

2. Auxiliary Results and Proof of Theorem 1.3

We start by recalling standard notation from multilinear algebra [4, 5]. Let $\mathcal{Y}$ be an $n$-dimensional Hilbert space, and let $\chi$ be a character of degree 1 on a subgroup $G$ of $S_m$, the symmetric group on $m$ elements. The symmetrizer induced by $\chi$ on the tensor product space $\otimes^m \mathcal{Y}$ is defined by its action

$$
S(v_1 \otimes \cdots \otimes v_m) := \frac{1}{|G|} \sum_{\sigma \in G} \chi(\sigma) v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(m)}.
$$

Elements of the form (2.1) span a vector space that is denoted as

$$
\mathcal{V}_\chi^m(G) := S(\otimes^m \mathcal{Y}) \subset \otimes^m \mathcal{Y}.
$$

This vector space is the space of the symmetry class of tensors associated with $G$ and $\chi$. It can be verified that $\mathcal{V}_\chi^m(G)$ is an invariant subspace of $\otimes^m \mathcal{Y}$. The elements of $\mathcal{V}_\chi^m(G)$ are denoted by the following “star-product”:

$$
v_1 \star \cdots \star v_m := S(v_1 \otimes \cdots \otimes v_m).
$$
For any linear operator $T$ on $V$ there is a unique induced operator $K(T) : \mathcal{V}_\chi^m(G) \to \mathcal{V}_\chi^m(G)$ which satisfies (see also [3] for related material):

\begin{equation}
K(T)(v_1 \star \cdots \star v_m) = Tv_1 \star \cdots \star Tv_m.
\end{equation}

This operation is usually written as $K(T)v^* = T v^*$, where $v^* \equiv v_1 \star \cdots \star v_m$.

From an orthonormal basis for $V$ we can induce an orthonormal basis for $\mathcal{V}_\chi^m(G)$, which will allow us to write down a matrix representation of the operator $K(T)$. To define such a matrix we need some more notation from [4].

Let $\Gamma_{m,n}$ denote the totality of sequences $\alpha = (\alpha_1, \ldots, \alpha_m)$ such that $1 \leq \alpha_i \leq n$ for $1 \leq i \leq m$. Thus, $|\Gamma_{m,n}| = n^m$. Two sequences $\alpha$ and $\beta$ in $\Gamma_{m,n}$ are said to be $G$-equivalent, denoted $\alpha \sim_G \beta$, if there exists a permutation $\sigma \in G$ such that $\alpha = (\beta_{\sigma(1)}, \ldots, \beta_{\sigma(n)})$. This equivalence partitions $\Gamma_{m,n}$ into equivalence classes; let $\Delta$ be a system of distinct representatives for these equivalence classes; we order sequences in $\Delta$ using lexicographic order.

For all $\alpha \in \Gamma_{m,n}$ the set of all permutations $\sigma \in G$ for which $\alpha \sigma = \alpha$ is called the stabilizer of $\alpha$ and is denoted by $G_\alpha$. Clearly, it is a subgroup of $G$; we denote its order by $\nu(\alpha)$. We define the set $\bar{\Delta} \subset \Delta$ consisting of those $\alpha \in \Delta$ for which $G_\alpha \subset \ker \chi$. Since $\chi$ was assumed to be a character of degree 1, $\ker \chi$ is the set of permutations $\sigma$ for which $\chi(\sigma) = 1$. Thus, $\alpha \in \bar{\Delta}$ if and only if $\chi(\sigma) = 1$ for all $\sigma \in G_\alpha$. Therefore,

\begin{equation}
\sum_{\sigma \in G_\alpha} \nu(\alpha), \quad \text{if } \alpha \in \bar{\Delta},
\end{equation}

\begin{equation}
0, \quad \text{if } \alpha \notin \bar{\Delta}.
\end{equation}

Now suppose $B = \{e_1, \ldots, e_n\}$ is an orthonormal basis for $V$. Then,

\begin{equation}
B^* := \{e_{\alpha_1} \star \cdots \star e_{\alpha_m} \mid \alpha \in \bar{\Delta}\},
\end{equation}

is an orthogonal basis for $\mathcal{V}_\chi^m(G)$, which can be normalized to obtain an orthonormal basis—see e.g., [4, Theorem 3.2], which proves that

\begin{equation}
\bar{B}^* = \{(\sqrt{|G|/\nu(\alpha)})(e_{\alpha_1} \star \cdots \star e_{\alpha_m} \mid \alpha \in \bar{\Delta})
\end{equation}

is an orthonormal basis for $\mathcal{V}_\chi^m(G)$ with respect to the induced inner product on $\otimes^m V$. Moreover, $\dim \mathcal{V}_\chi^m(G) = \bar{|\Delta|}$.

Let $T \in \mathcal{L}(V, V)$. From [4, Theorem 4.1] we know that $K(T) = \otimes^m T \mid \mathcal{V}_\chi^m(G)$, the restriction of the tensor space $\otimes^m T$ to the symmetry class $\mathcal{V}_\chi^m(G)$. Thus, $K(T)v^* = (\otimes^m T)v^*$. Finally, it can be shown that [4, p. 126] that for multi-indices $\alpha, \beta \in \bar{\Delta}$, the $(\alpha, \beta)$ entry of $K(A)$ is given by

\begin{equation}
[K(A)]_{\alpha, \beta} = \frac{1}{\sqrt{\nu(\alpha)\nu(\beta)}} d^G_\chi(A^*[\beta|\alpha]),
\end{equation}

where $A^*[\beta|\alpha]$ is the $(\beta, \alpha)$ submatrix of $A^*$. For self-adjoint $A$, we see that we can recover $d^G_\chi(A)$ picking out a diagonal entry of $K(A)$ corresponding to $\beta = \alpha = (1, \ldots, m)$.

With this notation in hand we can state the following easy but key lemma.

**Lemma 2.1.** Let $T \in \mathcal{L}(V, V)$ be a self-adjoint operator with $A$ as its matrix representation. Let $K(T)$ be the induced operator corresponding to the symmetry class described by $\chi$ and subgroup $G \subset S_m$, and let $K(A)$ be the matrix representation of $K(T)$. Then, there exists a matrix $Z$ (of suitable size) such that

\begin{equation}
K(A) = Z^*(\otimes^m A)Z.
\end{equation}
Proof. From the discussion above it follows that $[K(A)]_{\alpha,\beta} = \langle K(T)e_\alpha^*, e_\beta^* \rangle$. Since $K(T)v^* = (\otimes^m T)v^*$, we obtain $[K(A)]_{\alpha,\beta} = \langle (\otimes^m A)e_\alpha^*, e_\beta^* \rangle$. Collecting the vectors $e_\alpha^*$ into a suitable matrix $Z$ (note $ZZ^* = I$), we therefore immediately obtain

$$K(A) = Z^* (\otimes^m A) Z. \quad \Box$$

Observe that Lemma 2.1 easily yields the well-known multiplicativity of $K$, i.e.,

$$(2.7) \quad K(AB) = K(A)K(B),$$

since $\otimes^k(AB) = (\otimes^k A)(\otimes^k B)$ and $ZZ^* = I$.

Next, we refer to the following result from [8, Lemma 2.2].

**Lemma 2.2.** Let $A,B,C \in \mathbb{M}_\ell$ be positive semidefinite. Then

$$\otimes^k(A + B + C) + \otimes^kC \geq \otimes^k(A + C) + \otimes^k(B + C)$$

for any positive integer $k$.

An immediate corollary of Lemmas 2.1 and 2.2 is the following.

**Corollary 2.3.** Let $A,B,C \in \mathbb{M}_\ell$ be positive semidefinite. Then

$$(2.8) \quad K(A + B + C) + K(C) \geq K(A + C) + K(B + C).$$

**Lemma 2.4.** Let $A = [A_{i,j}]_{i,j=1}^m \in \mathbb{M}_m(\mathbb{M}_n)$ be positive semidefinite. Then the matrix $[K(A_{i,j})]_{i,j=1}^m$ is a compression of the matrix $K(A)$.

Proof. We follow an approach similar to [9]. Since $A \succeq 0$, we can write it as $A = R^* R$. Now partition $R = [R_1, \ldots, R_m]$ where each $R_i$, $1 \leq i \leq m$, is an $mn \times n$ complex matrix. With this partitioning we see that $A_{i,j} = R_i^* R_j$, also, with this notation, we have $R_i = RE_i$, where $E_i$ is a suitable $mn \times n$ matrix that extracts the $i$th block from $R$.

The crucial property to exploit is the multiplicativity of $K$ and that $K(A^*) = K(A)^*$ [4, Theorem 4.2]. Consider, thus the block matrix $[K(A_{i,j})]_{i,j=1}^m$. We have

$$K(A_{i,j}) = K(R_i^* R_j) = K(E_i^* R^* RE_j) = K(E_i)^* K(R^* R)K(E_j) = P_i^* K(A) P_j.$$

In other words,

$$[K(A_{i,j})] = P^* K(A) P, \quad \text{where} \quad P = \begin{bmatrix} P_1 & & \\ & \ddots & \\ & & P_m \end{bmatrix}. \quad \Box$$

We are now in a position to present a proof of Theorem 1.3.

**Proof of Theorem 1.3.** Let $A = [A_{i,j}]_{i,j=1}^m$, $B = [B_{i,j}]_{i,j=1}^m$, $C = [C_{i,j}]_{i,j=1}^m \in \mathbb{M}_m(\mathbb{M}_n)$ be positive semidefinite. By Corollary 2.3,

$$(2.9) \quad K(A + B + C) + K(C) \geq K(A + C) + K(B + C).$$

By Lemma 2.4, $[K(A_{i,j})]_{i,j=1}^m$ is a compression of $K(A)$, which, combined with (2.9) yields the inequality

$$[K(A_{i,j} + B_{i,j} + C_{i,j})]_{i,j=1}^m + [K(C_{i,j})]_{i,j=1}^m \geq [K(A_{i,j} + C_{i,j})]_{i,j=1}^m + [K(B_{i,j} + C_{i,j})]_{i,j=1}^m.$$

Taking into account (2.6), it follows that

$$[d^G_{\chi}(A_{i,j} + B_{i,j} + C_{i,j})]_{i,j=1}^m + [d^G_{\chi}(C_{i,j})]_{i,j=1}^m \geq [d^G_{\chi}(A_{i,j} + C_{i,j})]_{i,j=1}^m + [d^G_{\chi}(B_{i,j} + C_{i,j})]_{i,j=1}^m,$$

therewith establishing the theorem. \Box
3. A proof of Thompson’s result

Thompson [7] proved the following elegant determinantal inequality.

**Theorem 3.1.** Let \( A \in \mathbb{M}_m(\mathbb{M}_n) \) be positive semidefinite. Then

\[
\det A \leq \det(\det_m(A)).
\]

As an application of our result, we present a new proof of Theorem 3.1.

**Proof of Theorem 3.1.** As \( A \geq 0 \), we may write \( A = T^*T \) with \( T = [T_{i,j}]_{i,j=1}^m \) being block upper triangular. If \( A \) is singular, (3.1) is trivial. So we assume otherwise. We may further assume \( T_{i,i} = I_n \), the \( n \times n \) identity matrix, by pre- and post-multiplying both sides of (3.1) with \( \prod_{i=1}^m \det T_{i,i}^{-1} \) and \( \prod_{i=1}^m \det T_{i,i}^{-1} \), respectively. Thus, it suffices to show

\[
\det(\det_m(T^*T)) \geq 1.
\]

(3.2)

This reformulation is exactly what Thompson did in [7].

We prove (3.2) by induction. When \( m = 2 \),

\[
\det(\det_2(T^*T)) = \det \begin{bmatrix}
1 & \det T_{1,2} \\
\det T_{1,2} & \det(I_n + T_{1,2}^*T_{1,2})
\end{bmatrix}
= \det(I_n + T_{1,2}^*T_{1,2}) - \det(T_{1,2}^*T_{1,2}) \geq 1.
\]

Suppose (3.2) is true for \( m = k \), and then the case \( m = k + 1 \). For notational convenience, we denote \( T = \begin{bmatrix}
I_n & V \\
0 & \hat{T}
\end{bmatrix} \), where \( V = [T_{1,1} \cdots T_{1,m}] \) and \( \hat{T} = [T_{i+1,j+1}]_{i,j=1}^k \). Let \( D = [\det T_{1,1} \cdots \det T_{1,m}] \). Clearly, \( D^*D = \det_k(V^*V) \).

Now compute

\[
T^*T = \begin{bmatrix}
I_n & V^* \\
0 & \hat{T}
\end{bmatrix}^* \begin{bmatrix}
I_n & V \\
0 & \hat{T}
\end{bmatrix} = \begin{bmatrix}
I_n & V \\
V^* & \hat{T}^*\hat{T} + V^*V
\end{bmatrix}.
\]

Then

\[
\det(\det_m(T^*T)) = \det \begin{bmatrix}
D \\
D^* & \det_k(\hat{T}^*\hat{T} + V^*V)
\end{bmatrix}
= \det(\det_k(\hat{T}^*\hat{T} + V^*V) - D^*D)
\geq \det(\det_k(\hat{T}^*\hat{T}) + \det_k(V^*V) - D^*D)
= \det(\det_k(\hat{T}^*\hat{T})) \geq 1,
\]

in which the first inequality is by (1.5), while the second one is by the induction hypothesis. This completes the proof. \( \square \)

**Acknowledgements.** MathOverflow brought the authors to work on this topic together (see http://mathoverflow.net/q/173088/).

**References**

[1] R. Bhatia, Matrix Analysis, GTM 169, Springer-Verlag, New York, 1997.
[2] L.-K. Hua, Inequalities involving determinants (in Chinese), Acta Math. Sinica, 5 (1955) 463-470. [Translated into English: Transl. Amer. Math. Soc. Ser. II, 32, 265272 (1963)].
[3] C.-K. Li, A. Zaharia, Induced operators on symmetry classes of tensors, Trans. Amer. Math. Soc. 354 (2002) 807-836.
[4] M. Marcus. Finite dimensional multilinear algebra, Vol. 1, Marcel Dekker, 1973.
[5] R. Merris, Multilinear Algebra, Gordon & Breach, Amsterdam, 1997.
[6] V. Paksoy, R. Turkmen, F. Zhang, Inequalities of generalized matrix functions via tensor products, Electron. J. Linear Algebra 27 (2014) 332-341.
[7] R. C. Thompson, A determinantal inequality for positive definite matrices, Canad. Math. Bull. 4 (1961) 57-62.
[8] L. Tie, K.-Y. Cai, and Y. Lin. Rearrangement inequalities for Hermitian matrices. Linear Algebra Appl. 434 (2011) 443-456.
[9] F. Zhang, Positivity of matrices with generalized matrix functions, Acta. Math. Sinica, 28 (2012) 1779-1786.

Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada, V8W 3R4.
E-mail address: mlin87@ymail.com

MPI for Intelligent Systems, Tübingen, Germany
E-mail address: suvrit@gmail.com