CONVEX MATRIX FUNCTIONS
WILLIAM WATKINS

Abstract. The purpose of this paper is to prove convexity properties for the tensor product, determinant, and permanent of hermitian matrices.

Let $C^n$ be the vector space of all complex $n$-tuples with the usual inner product $\langle \ , \ \rangle$ and let $H_n$ be the set of all $n$ by $n$ hermitian matrices. A matrix $A$ in $H_n$ is nonnegative if $\langle Ax, x \rangle \geq 0$ for all $x$ in $C^n$. If $A$ and $B$ are in $H_n$, we write $A \geq B$ if $A - B$ is nonnegative. A function $f$ from $H_n$ to $H_m$ is monotone if $A \geq B$ implies $f(A) \geq f(B)$, and convex if $f(\lambda A + (1 - \lambda)B) \leq \lambda f(A) + (1 - \lambda)f(B)$, for all $0 \leq \lambda \leq 1$.

Löwner [6] introduced the case where $f$ is induced by a real valued function and $m = n$. Other authors [2], [4], [5] have analysed this case further.

Example [9]. The inverse function is convex on the set of all invertible, nonnegative matrices in $H_n$.

Example [4]. The square root function is monotone on the set of all nonnegative matrices in $H_n$.

Some work has been done on the case where $m = 1$. That is, $f$ is a function from $H_n$ to the real numbers. For example, Marcus and Nikolai [8] have shown that each member of a class of generalized matrix functions is monotone. This class of functions contains the determinant and permanent. For other results of this type see [1].

In order to state the convexity property for the tensor product, let $m_1, \cdots, m_r$ be $r$ positive integers. It is well known [10, p. 268] that, for $x_i, y_i$ in $C^{m_i}$, $i = 1, \cdots, r$, the decomposable tensors $x_1 \otimes \cdots \otimes x_r$ and $y_1 \otimes \cdots \otimes y_r$ in $C^N$, $N = m_1 \cdots m_r$, satisfy

$\langle x_1 \otimes \cdots \otimes x_r, y_1 \otimes \cdots \otimes y_r \rangle = \langle x_1, y_1 \rangle \cdots \langle x_r, y_r \rangle$.

If $A_i$ is an $m_i$ by $m_i$ matrix $(i = 1, \cdots, r)$, then the tensor product $\otimes^r A_i$ is an $N$ by $N$ matrix satisfying

$\otimes^r A_i(x_1 \otimes \cdots \otimes x_r) = A_1 x_1 \otimes \cdots \otimes A_r x_r$,

for $x_i$ in $C^{m_i}$ $(i = 1, \cdots, r)$.

Received by the editors June 15, 1973.

Theorem 1. If $A_i$ and $B_i$ are matrices in $H_{m_i}$ with $0 \leq B_i \leq A_i$, $i = 1, \cdots, r$, and $0 \leq \lambda \leq 1$, then
\[
\otimes^r (\lambda A_i + (1 - \lambda)B_i) \leq \lambda \otimes^r A_i + (1 - \lambda) \otimes^r B_i.
\]

Definition (Generalized matrix function). Let $S_n$ denote the permutation group on $n$ letters and let $G$ be a subgroup of $S_n$ with irreducible character $\chi: G \to \mathbb{C}$. For each $n$ by $n$ complex matrix $A = (a_{ij})$, define
\[
d(A) = \sum_{\sigma} \chi(\sigma) \prod_{i=1}^{n} a_{\sigma i, i} \text{ (sum } \sigma \text{ in } G).
\]

The function $d$ depends on both the subgroup $G$ and its character $\chi$. If $G = S_n$ and $\chi(\sigma)$ is the sign of $\sigma$, then $d$ is the determinant function. If $G = S_n$ and $\chi \equiv 1$, then $d$ is the permanent function. For a fuller explanation see [7].

Theorem 2. If $A$ and $B$ are matrices in $H_n$ with $0 \leq B \leq A$ and $0 \leq \lambda \leq 1$, then
\[
d(\lambda A + (1 - \lambda)B) \leq \lambda d(A) + (1 - \lambda)d(B).
\]

Corollary. If $A$ and $B$ are matrices in $H_n$ with $0 \leq B \leq A$ and $0 \leq \lambda \leq 1$, then
\[
det(\lambda A + (1 - \lambda)B) \leq \lambda \det A + (1 - \lambda)\det B
\]
and
\[
\text{per}(\lambda A + (1 - \lambda)B) \leq \lambda \text{ per } A + (1 - \lambda)\text{ per } B.
\]

Proofs.
Proof of Theorem 1. It is shown in [8] that if $A_1, B_1$ are in $H_{m_1}$ and $A_2, B_2$ are in $H_{m_2}$ with $0 \leq B_1 \leq A_1$ and $0 \leq B_2 \leq A_2$, then $A_1 \otimes A_2 \geq B_1 \otimes B_2$. Thus the right side of the identity
\[
\lambda(A_1 \otimes A_2) + (1 - \lambda)(B_1 \otimes B_2) - (\lambda A_1 + (1 - \lambda)B_1) \otimes (\lambda A_2 + (1 - \lambda)B_2)
\]
is nonnegative. Theorem 1 follows by induction.

In order to prove Theorem 2, we develop ideas relating the tensor product to the generalized matrix function $d$.

For each $\sigma$ in $S_n$, define an $N$ by $N$ ($N = n^n$) permutation matrix $P(\sigma)$ by $P(\sigma^{-1})x_1 \otimes \cdots \otimes x_n = x_{\sigma_1} \otimes \cdots \otimes x_{\sigma_n}$ for all $x_i$ in $\mathbb{C}^n$. Notice that $P(\sigma \mu) = P(\sigma)P(\mu)$. Define an $N$ by $N$ matrix $T$ by
\[
T = \frac{\chi(1)}{|G|} \sum_{\sigma} \chi(\sigma) P(\sigma) \text{ (sum } \sigma \text{ in } G).
\]
It follows from the orthogonality relations for irreducible characters [3, p. 219] that $T$ is an idempotent. The matrix $T$ is hermitian since the complex conjugate of $\chi(\sigma)$ is $\chi(\sigma^{-1})$ and $P(\sigma)^* = P(\sigma^{-1})$. If $A = (a_{ij})$ is an $n \times n$ matrix, then $\otimes^n A$ commutes with each $P(\sigma)$ and so it commutes with $T$.

Let $e_1, \ldots, e_n$ be the usual basis for $\mathbb{C}^n$. Then,

$$
(\otimes^n A)Te_1 \otimes \cdots \otimes e_n, Te_1 \otimes \cdots \otimes e_n) \\
= (T^*(\otimes^n A)Te_1 \otimes \cdots \otimes e_n, e_1 \otimes \cdots \otimes e_n) \\
= (T(\otimes^n A)e_1 \otimes \cdots \otimes e_n, e_1 \otimes \cdots \otimes e_n) \\
= (TAe_1 \otimes \cdots \otimes Ae_n, e_1 \otimes \cdots \otimes e_n) \\
= \frac{\chi(1)}{|G|} \sum_{\sigma} \chi(\sigma)(Ae_{\sigma 1} \otimes \cdots \otimes Ae_{\sigma n}, e_1 \otimes \cdots \otimes e_n) \\
= \frac{\chi(1)}{|G|} \sum_{\sigma} \chi(\sigma) \prod_{i} (Ae_{\sigma i}, e_i) \\
= \frac{\chi(1)}{|G|} d(A).
$$

In the second inequality, notice that $T^*(\otimes^n A)T = T(\otimes^n A)$, since $T$ and $\otimes^n A$ commute and $T$ is a hermitian idempotent. If $A$ and $B$ are in $H_n$ and $0 \leq A \leq B$ and $0 \leq \lambda \leq 1$, then by Theorem 1 we have

$$
\otimes^n (\lambda A + (1 - \lambda)B) \leq \lambda \otimes^n A + (1 - \lambda) \otimes^n B.
$$

By comparing inner products

$$
(\otimes^n (\lambda A + (1 - \lambda)B)Te_1 \otimes \cdots \otimes e_n, Te_1 \otimes \cdots \otimes e_n)
$$

and

$$
((\lambda \otimes^n A + (1 - \lambda) \otimes^n B)Te_1 \otimes \cdots \otimes e_n, Te_1 \otimes \cdots \otimes e_n),
$$

we get $d(\lambda A + (1 - \lambda)B) \leq \lambda d(A) + (1 - \lambda) d(B)$. The corollary consists of special cases.

**References**


5. F. Kraus, Über konvexe Matrixfunktionen, Math. Z. 41 (1936), 18–42.

DEPARTMENT OF MATHEMATICS, CALIFORNIA STATE UNIVERSITY, NORTH RIDGE, CALIFORNIA 91324