PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 44, Number 1, May 1974

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 $(\ ,\ )$ and let H_n be the set of all n by n hermitian matrices. A matrix A in H_n is *nonnegative* if $(Ax, x) \geqq 0$ for all x in C^n. If A and B are in H_n, we write $A \geqq B$ if $A-B$ is nonnegative. A function f from H_n to H_m is *monotone* if $A \geqq B$ implies $f(A) \geqq f(B)$, and *convex* if $f(\lambda A+(1-\lambda)B) \leqq \lambda f(A)+(1-\lambda)f(B)$, for all $0 \leqq \lambda \leqq 1$.

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

$$(x_1 \otimes \cdots \otimes x_r, y_1 \otimes \cdots \otimes y_r) = (x_1, y_1) \cdots (x_r, y_r).$$

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.

AMS (MOS) subject classifications (1970). Primary 15A45; Secondary 15A15, 15A69.

THEOREM 1. *If A_i and B_i are matrices in H_{m_i} with $0 \leqq B_i \leqq A_i$, $i=1, \cdots, r$, and $0 \leqq \lambda \leqq 1$, then*

$$\otimes^r (\lambda A_i + (1-\lambda)B_i) \leqq \lambda \otimes^r A_i + (1-\lambda) \otimes^r B_i.$$

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 $\chi: G \to C$. For each n by n complex matrix $A=(a_{ij})$, define

$$d(A) = \sum \chi(\sigma) \prod_{i=1}^{n} a_{\sigma i, i} \quad (\text{sum } \sigma \text{ in } G).$$

The function d depends on both the subgroup G and its character χ. If $G=S_n$ and $\chi(\sigma)$ is the sign of σ, 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 $0 \leqq B \leqq A$ and $0 \leqq \lambda \leqq 1$, then*

$$d(\lambda A + (1-\lambda)B) \leqq \lambda d(A) + (1-\lambda) d(B).$$

COROLLARY. *If A and B are matrices in H_n with $0 \leqq B \leqq A$ and $0 \leqq \lambda \leqq 1$, then*

$$\det(\lambda A + (1-\lambda)B) \leqq \lambda \det A + (1-\lambda) \det B$$

and

$$\operatorname{per}(\lambda A + (1-\lambda)B) \leqq \lambda \operatorname{per} A + (1-\lambda) \operatorname{per} B.$$

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 $0 \leqq B_1 \leqq A_1$ and $0 \leqq B_2 \leqq A_2$, then $A_1 \otimes A_2 \geqq B_1 \otimes B_2$. Thus the right side of the identity

$$\begin{aligned} \lambda(A_1 \otimes A_2) + (1-\lambda)(B_1 \otimes B_2) - (\lambda A_1 + (1-\lambda)B_1) \\ \otimes (\lambda A_2 + (1-\lambda)B_2) = \lambda(1-\lambda)(A_1 - B_1) \otimes (A_2 - B_2) \end{aligned}$$

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 σ 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 C^n. Notice that $P(\sigma\mu) = P(\sigma)P(\mu)$. Define an N by N matrix T by

$$T = \frac{\chi(1)}{|G|} \sum \chi(\sigma) P(\sigma) \quad (\text{sum } \sigma \text{ in } G).$$

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 by n matrix, then $\otimes^n A$ commutes with each $P(\sigma)$ and so it commutes with T.

Let $e_1, \cdots, e_n$ be the usual basis for C^n. Then,

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

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 $0 \leqq A \leqq B$ and $0 \leqq \lambda \leqq 1$, then by Theorem 1 we have

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

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

References

1. R. Bellman, *Introduction to matrix analysis*, McGraw-Hill, New York, 1960, Chap. 8. MR **23** #A153.

2. J. Bendat and S. Sherman, *Monotone and convex operator functions*, Trans. Amer. Math. Soc. **79** (1955), 58–71. MR **18,** 588.

3. C. W. Curtis and I. Reiner, *Representation theory of finite groups and associative algebras*, Pure and Appl. Math., vol. 11, Interscience, New York, 1962. MR **26** #2519.

4. C. Davis, *Notions generalizing convexity for functions defined on spaces of matrices*, Proc. Sympos. Pure Math., vol. 7, Amer. Math. Soc., Providence, R.I., 1963, pp. 187–201. MR **27** #5771.

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

6. K. Löwner, *Über monotone Matrixfunktionen*, Math. Z. **38** (1934), 177–216.

7. M. Marcus and H. Minc, *Generalized matrix functions*, Trans. Amer. Math. Soc. **116** (1965), 316–329. MR **33** #2655.

8. M. Marcus and P. J. Nikolai, *Inequalities for some monotone matrix functions*, Canad. J. Math. **21** (1969), 485–494. MR **38** #5815.

9. M. H. Moore, *A convex matrix function*, Amer. Math. Monthly **80** (1973), 408–409.

10. G. D. Mostow and J. H. Sampson, *Linear algebra*, McGraw-Hill, New York, 1969. MR **42** #7673.

DEPARTMENT OF MATHEMATICS, CALIFORNIA STATE UNIVERSITY, NORTHRIDGE, CALIFORNIA 91324