MATHEMATICS MAGAZINE

Supermultiplicative Inequalities for the Permanent of Nonnegative Matrices

JOEL E. COHEN
Rockefeller University
New York, NY 10021

A fact about the determinant (abbreviated det) that is usually taught early and is very useful later is that $\det(AB) = \det(A)\det(B)$, where A and B are $n \times n$ matrices and n is a finite positive integer. This multiplicative identity for the determinant is related to a host of generalizations. In some generalizations, det is replaced by another real-valued or matrix-valued function f with matrix argument, where f is related in some way to the determinant. In some generalizations, the equality is replaced by an inequality. When the function f that replaces det satisfies $f(AB) \geqslant f(A)f(B)$, f is said to be supermultiplicative; when $f(AB) \leqslant f(A)f(B)$, f is said to be submultiplicative. Clearly, det is both supermultiplicative and submultiplicative.

The purpose of this note is to consider two well-known, apparently unrelated supermultiplicative functions of nonnegative matrices and to show that they are special cases of a natural, more general supermultiplicative function. All of these functions may be viewed as relatives of the determinant. Further, these supermultiplicative functions are surprisingly useful in the theory of products of random matrices. An application to products of random matrices is sketched at the end of this note.

For a fixed finite positive integer n, an $n \times n$ matrix with all nonnegative real elements will be called a nonnegative matrix. The well-known supermultiplicative functions of a nonnegative matrix to be considered here are the diagonal elements and the permanent (Theorems A and B).

THEOREM A. *If $A = (a_{ij})$ and $B = (b_{ij})$ are nonnegative matrices, then the ith diagonal element $(AB)_{ii}$ of AB is related to the ith diagonal elements a_{ii} of A and b_{ii} of B by $(AB)_{ii} \geqslant a_{ii}b_{ii}$, for $i = 1, \ldots, n$.*

Proof. $(AB)_{ii} = \sum_{j=1}^{n} a_{ij}b_{ji} \geqslant a_{ii}b_{ii}$.

Recall that $\operatorname{per}(A)$, the permanent of A, is a determinant that thinks positively, i.e., if $\sigma = (\sigma(1), \ldots, \sigma(n))$ is a permutation of $(1, \ldots, n)$, then $\operatorname{per}(A) = \sum_\sigma a_{1,\sigma(1)}a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$ where the summation runs over all permutations σ. Minc gives an encyclopedic account of permanents [6].

THEOREM B (Brualdi 1966). *If A and B are nonnegative, then* $\operatorname{per}(AB) \geqslant \operatorname{per}(A)\operatorname{per}(B)$.

Brualdi's (1966) proof of Theorem B is elementary. In outline, every term in $\operatorname{per}(A)\operatorname{per}(B)$ appears as a term of $\operatorname{per}(AB)$, and the other terms of $\operatorname{per}(AB)$ are nonnegative.

The previously known Theorems A and B are both special cases of a more general set of inequalities involving the permanent.

For $k = 1, \ldots, n$, define the kth *permanent-compound* of A (any matrix over a field will do, not necessarily a nonnegative matrix) as an $\binom{n}{k} \times \binom{n}{k}$ matrix $A_{[k]}$ with elements constructed as follows.

Let $Q_{k,n}=\{(i_1,i_2,\ldots,i_k)|1\leqslant i_1<i_2<\cdots<i_k\leqslant n\}$ and choose an ordering, say lexicographic, of the k-tuples in $Q_{k,n}$. By a slight abuse of notation, the elements of $A_{[k]}$ will be indexed by pairs $(\underline{i},\underline{j})\in Q_{k,n}\times Q_{k,n}$ rather than by pairs of integers. For $\underline{i}=(i_1,\ldots,i_k)\in Q_{k,n}$, $\underline{j}=(j_1,\ldots,j_k)\in Q_{k,n}$, define $A[\underline{i};\underline{j}]$ to be the $k\times k$ matrix that contains the elements in the intersections of rows $i_1,\ldots,i_k$ and columns $j_1,\ldots,j_k$. Then $A_{[k]}$ is the $\binom{n}{k}\times\binom{n}{k}$ matrix with $(\underline{i},\underline{j})$ element $(A_{[k]})_{\underline{i}\underline{j}}=\operatorname{per}(A[\underline{i};\underline{j}])$.

As examples, if we assume the lexicographic ordering of the elements of $Q_{k,n}$ is chosen, $A_{[1]}=A$ and if $n=3$, then

$$A_{[2]}=\begin{pmatrix} a_{11}a_{22}+a_{21}a_{12} & a_{11}a_{23}+a_{21}a_{13} & a_{12}a_{23}+a_{22}a_{13}\\ a_{11}a_{32}+a_{31}a_{12} & a_{11}a_{33}+a_{31}a_{13} & a_{12}a_{33}+a_{32}a_{13}\\ a_{21}a_{32}+a_{31}a_{22} & a_{21}a_{33}+a_{31}a_{23} & a_{22}a_{33}+a_{32}a_{23}\end{pmatrix}$$

Regardless of the ordering of $Q_{k,n}$, $A_{[n]}=\operatorname{per}(A)$. Changing the ordering of the elements of $Q_{k,n}$ simultaneously permutes the rows and columns of the permanent-compound matrix, leaving the diagonal elements $(A_{[k]})_{\underline{i}\underline{i}}$ on the diagonal. The kth permanent-compound of the $n\times n$ identity matrix is the $\binom{n}{k}\times\binom{n}{k}$ identity matrix. For any scalar c, $(cA)_{[k]}=c^kA_{[k]}$. If A^* denotes the conjugate transpose of A, then $(A_{[k]})^*=(A^*)_{[k]}$ because $\operatorname{per}(A)=\operatorname{per}(A^T)$, where A^T is the transpose of A ([6], p. 16) and the conjugate of a product of two complex numbers is the product of their conjugates.

Theorem A is the special case of the following Theorem 1 when $k=1$ and Theorem B is the special case when $k=n$. Thus Theorem 1 is a natural generalization of Theorems A and B.

THEOREM 1. *If A and B are nonnegative matrices, then for $k=1,\ldots,n$, and all $\underline{i}\in Q_{k,n}$, $((AB)_{[k]})_{\underline{i}\underline{i}}\geqslant(A_{[k]})_{\underline{i}\underline{i}}(B_{[k]})_{\underline{i}\underline{i}}$.*

Before proving Theorem 1, note that Theorem 1 is unaffected by the ordering chosen for $Q_{k,n}$ because the theorem deals only with diagonal elements of the permanent-compound.

To prove Theorem 1, two easy lemmas are needed. Let $(AB)[\underline{i};\underline{j}]$ denote the $k\times k$ matrix that contains the elements in the intersections of rows $i_1,\ldots,i_k$ and columns $j_1,\ldots,j_k$ of the product matrix AB.

LEMMA 1. *For any $\underline{i}\in Q_{k,n}$ and any nonnegative A,B, $(AB)[\underline{i};\underline{i}]\geqslant A[\underline{i};\underline{i}]B[\underline{i};\underline{i}]$, where the inequality applies elementwise.*

Proof. For $1\leqslant g,h\leqslant k$, let i_g and i_h denote any two elements of $\underline{i}$. Then

$$(AB)_{i_gi_h}=\sum_{m=1}^{n}a_{i_gm}b_{mi_h}\geqslant\sum_{p=1}^{k}a_{i_gi_p}b_{i_pi_h}=(A[\underline{i};\underline{i}]B[\underline{i};\underline{i}])_{i_gi_h}.$$

LEMMA 2. *For any nonnegative A, B, if $A\geqslant B$ (elementwise), then* $\operatorname{per}(A)\geqslant\operatorname{per}(B)$.

This is obvious.

Proof of Theorem 1. Let $\underline{i}$ denote the ith element of $Q_{k,n}$ in the chosen ordering. Then for any nonnegative A,B,

$$\begin{aligned}((AB)_{[k]})_{ii}&=\operatorname{per}\{(AB)[\underline{i};\underline{i}]\}\\ &\geqslant\operatorname{per}\{A[\underline{i};\underline{i}]B[\underline{i};\underline{i}]\}\quad\text{(by Lemmas 1 and 2)}\\ &\geqslant\operatorname{per}\{A[\underline{i};\underline{i}]\}\operatorname{per}\{B[\underline{i};\underline{i}]\}\quad\text{(by Theorem B)}\end{aligned}$$

$$= (A_{[k]})_{ii}(B_{[k]})_{ii}.$$

THEOREM 2. *If A and B are nonnegative, then for $k = 1,\ldots,n$, and $(\underline{i},\underline{j}) \in Q_{k,n} \times Q_{k,n}$,*

$$((AB)_{[k]})_{\underline{i}\underline{j}} \geqslant \max\left\{(A_{[k]})_{\underline{i}\underline{h}}(B_{[k]})_{\underline{h}\underline{j}} | \underline{h} \in Q_{k,n}\right\}.$$

The proof parallels the proof of Theorem 1 exactly. The supermultiplicative inequality in Theorem 1 describes the special case of Theorem 2 when $\underline{i} = \underline{h} = \underline{j}$.

The permanent-compound is closely related to similarly defined objects, some of which have similar properties, e.g., the determinant-compound or adjugate matrix ([5], pp. 86–87), the induced matrix ([6], p. 87) and the combinatorial compound matrix [2]. The determinant-compound matrix is defined in the same way as the permanent-compound matrix except that per is replaced by det. The induced matrix is defined in terms of permanents of submatrices of a given matrix, but the definition is a bit more elaborate than that of the permanent-compound matrix given above. Like the simple determinant, both the determinant-compound matrix and the induced matrix preserve the product of matrices, i.e., the determinant-compound matrix of a product of matrices is the product of the determinant-compound matrices, and similarly for the induced matrices ([5], pp. 86–87; [6], p. 87). Although I know of no earlier definition of the permanent-compound matrix, I make no claim to be the first to consider it.

To conclude, I sketch the application of supermultiplicative functions to the theory of products of random matrices [4]. To avoid complications, I will describe only a special case of available results. Even so, this sketch presumes some familiarity with probability theory and does not pretend to be self-contained. Details appear in Key's paper [4]. Mathematical background and scientific applications are given in Cohen, Kesten and Newman [3].

Suppose $\{A_j\colon j = 1,2,\ldots\}$ is a sequence of matrices chosen independently and identically distributed from a finite set of positive $n \times n$ matrices. Suppose $\|A\|$ is any fixed norm of a matrix A. For any positive integer t, let $M_t = A_1A_2\ldots A_t$ be the product of the first t matrices from the sequence. Denoting the mathematical expectation or average by the symbol $E(\cdot)$, as is customary in probability theory, the limiting rate of growth of the norm

$$\log\lambda = \lim_{t\to\infty} t^{-1}E(\log\|M_t\|)$$

exists in this example (as well as for many other random sequences $\{A_j\colon j = 1,2,\ldots\}$). In the degenerate case when all the $A_j = A$ for some fixed positive matrix A, $\log\lambda$ is just the logarithm of the largest eigenvalue of A. So the limiting growth rate denoted by $\log\lambda$ may be thought of as the analog, for products of random matrices, of the logarithm of the largest eigenvalue of a fixed positive matrix.

Let f be a continuous, homogeneous, supermultiplicative function of a positive matrix argument. Here homogeneous means that for $c \geqslant 0$, $f(cA) = cf(A)$. Key [4] proved that for such a function f,

$$\log\lambda = \lim_{t\to\infty} t^{-1}E(\log f(M_t)) \qquad \text{and} \qquad \log\lambda = \lim_{t\to\infty} t^{-1}\log f(M_t)$$

with probability 1, and, moreover, the function f_t defined by $f_t = u^{-1}E\log f(M_u)$, where $u = 2^t$, increases monotonically to $\log \lambda$ with increasing t. Key's result provides a means of bounding $\log \lambda$ from below, namely, by computing f_t for finite values of t.

For sequences of positive matrices, Key cited two nontrivial functions of a positive matrix argument that satisfy the requirements of being continuous, homogeneous and supermultiplicative: the ith diagonal element, and the permanent raised to the power $1/n$. Theorem 1 above expands the set of nontrivial functions that can be used to bound $\log \lambda$ from below, for it implies easily that, for $k = 1, \ldots, n$, the functions $\{(A_{[k]})_{ii}\}^{1/k}$ are continuous, homogeneous, and supermultiplicative.

This work was supported in part by U.S. National Science Foundation grant BSR 87-05047 and the hospitality of Mr. and Mrs. William T. Golden.

REFERENCES

1. Richard A. Brualdi, Permanent of the direct product of matrices, *Pacific Journal of Mathematics* 16(3) (1966), 471–482.
2. Richard A. Brualdi and Li Qiao, On the combinatorial compound matrix, *Journal of Mathematical Research and Exposition* (China) 1 (1988), 153–162.
3. Joel E. Cohen, Harry Kesten, and Charles M. Newman, eds., *Random Matrices and Their Applications*, Contemporary Mathematics, Vol. 50, American Mathematical Society, Providence, RI, 1986.
4. Eric S. Key, Lower bounds for the maximal Lyapunov exponent, *Journal of Theoretical Probability* 3(3) (1990), 477–488.
5. C. C. MacDuffee, *The Theory of Matrices*, Chelsea Publishing, New York, 1946.
6. Henryk Minc, *Permanents*, Encyclopedia of Mathematics and Its Applications, Vol. 6, Addison-Wesley Publishing Co., Reading, MA, 1978.