Linear and Multilinear Algebra, 1990, Vol. 27, pp. 105–110
Reprints available directly from the publisher
Photocopying permitted by license only

Printed in the United States of America

On the Variation of Permanents

RAJENDRA BHATIA
Indian Statistical Institute, Delhi Center, 7, S.J.S. Sansanwal Marg, New Delhi, 110016, India

and

LUDWIG ELSNER
Fakultät für Mathematik, Universität Bielefeld, Postfach 8640, 4800 Bielefeld 1, F.R.G.

(*Received June 21, 1989; in final form November 1, 1989*)

It is shown that for any two n-by-n matrices A, B the inequality

$$|\text{per}(A)-\text{per}(B)| \leqslant n\|A-B\| \max(\|A\|, \|B\|)^{n-1}$$

holds for $\| \ \| = \| \ \|_p$, the l_p-operator norm $(1 \leqslant p \leqslant \infty)$. There are operator norms for which this inequality is invalid for some A, B.

1. INTRODUCTION

It is well known that for any operator norm $\| \ \|$ on the set $\mathbb{C}^{n,n}$ of n-by-n complex matrices and for any $A, B \in \mathbb{C}^{n,n}$ the inequality

$$|\det(A)-\det(B)| \leqslant n\|A-B\| \max(\|A\|, \|B\|)^{n-1}$$

holds, see [3].

Influenced by the general recent interest in permanents, the question of the variation of permanents also has been studied. In [1] and [2] we proved for the spectral norm and the l_1 and l_∞ norm resp. the inequality above with "det" replaced by "per". In this note we show that this holds for all l_p-norms $(1 \leqslant p \leqslant \infty)$, but not for all operator norms.

The main results is:

THEOREM *Let $A, B \in \mathbb{C}^{n,n}$. Then*

$$|\text{per}(A)-\text{per}(B)| \leqslant n\|A-B\| \max(\|A\|, \|B\|)^{n-1} \tag{1}$$

for $\| \ \| = \| \ \|_p$, $1 \leqslant p \leqslant \infty$.

Here per(A) denotes the permanent of $A=(a_{ik})$, i.e.,

$$\text{per}(A) = \sum_\sigma a_{1\sigma_1} a_{2\sigma_2} \cdots a_{n\sigma_n}$$

where σ runs through all permutations of $\{1, \ldots, n\}$. For $x=(x_1, \ldots, x_n) \in \mathbb{C}^n$ and

$1 \leqslant p \leqslant \infty$ we define as usual the l_p-norm

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

and the associated operator norm

$$\|A\|_p = \max\{\|Ax\|_p : \|x\|_p \leqslant 1\}.$$

The proof of the theorem is given in Section 2. In the concluding remarks we mention that (1) holds also for the weighted l_p-norms $\left(\sum_{i=1}^{n} |d_i x_i|^p \right)^{1/p}$, $d_i > 0$ and that (1) fails to hold for certain operator norms.

For any rectangular matrices F, G we denote by $F \otimes G$ the usual tensor product, as, e.g., defined in [5].

2. PROOF OF THE THEOREM

We start with three preparatory lemmata.

LEMMA 1 *If* $A \in \mathbb{C}^{m,m}$, $B \in \mathbb{C}^{n,n}$ *then*

$$\|A \otimes B\|_p = \|A\|_p \|B\|_p. \tag{2}$$

Proof We make use of the obvious relation

$$\|u \otimes v\|_p = \|u\|_p \|v\|_p, \qquad u \in \mathbb{C}^m, \quad v \in \mathbb{C}^n.$$

There exist $x \in \mathbb{C}^m$, $\|x\|_p = 1$, $y \in \mathbb{C}^n$, $\|y\|_p = 1$ such that $\|Ax\|_p = \|A\|_p$, $\|By\|_p = \|B\|_p$. As $\|x \otimes y\|_p = 1$, we have

$$\|A \otimes B\|_p \geqslant \|(A \otimes B)(x \otimes y)\|_p = \|Ax \otimes By\|_p = \|Ax\|_p \|By\|_p = \|A\|_p \|B\|_p. \tag{3}$$

Now we show the reverse inequality. Let I_m denote the m-by-m unit matrix. Partition a given vector $x \in \mathbb{C}^{mn}$ in the form

$$x = (x^1, \ldots, x^m), \qquad x^i \in \mathbb{C}^n, \quad i = 1, \ldots, m.$$

Then, as

$$I_m \otimes B = \begin{bmatrix} B & & & 0 \\ & B & & \\ & & \ddots & \\ & 0 & & B \end{bmatrix}$$

we get

$$\|(I_m \otimes B)x\|_p^p = \|(Bx^1, \ldots, Bx^m)\|_p^p = \sum_{i=1}^{m} \|Bx^i\|_p^p \leqslant \|B\|_p^p \sum_{i=1}^{m} \|x^i\|_p^p = \|B\|_p^p \|x\|_p^p$$

and we have

$$\|I_m \otimes B\|_p \leqslant \|B\|_p. \tag{4}$$

It is well known that there exists a permutation P such that $F \otimes G = P(G \otimes F)P^T$ for square F, G. In particular, $A \otimes I_n = P(I_n \otimes A)P^T$, and as the l_p-norm is invariant under P, we get from (4)

$$\|A \otimes I_n\|_p \leqslant \|A\|_p. \tag{5}$$

Using $A \otimes B = (A \otimes I_n)(I_m \otimes B)$ and the submultiplicativity of the l_p-norm we get from (4) and (5)

$$\|A \otimes B\|_p \leqslant \|A\|_p \|B\|_p \tag{6}$$

and by (3) and (6) we have (2). ■

An obvious consequence is

LEMMA 2 *For* $A_i \in \mathbb{C}^{n_i, n_i}$, $i = 1, \ldots, s$

$$\|A_1 \otimes \cdots \otimes A_s\|_p = \prod_{i=1}^{s} \|A_i\|_p. \tag{7}$$

Let $A \in \mathbb{C}^{n,n}$, and denote the nth tensor power of A by $\bigotimes^n A$, i.e.,

$$\bigotimes^n A = \underbrace{A \otimes A \otimes \cdots \otimes A}_{n\text{-times}}.$$

LEMMA 3 *For* $A, B \in \mathbb{C}^{n,n}$

$$\| \bigotimes^n A - \bigotimes^n B\|_p \leqslant n \|A - B\|_p \max(\|A\|_p, \|B\|_p)^{n-1}. \tag{8}$$

Proof Let

$$C_i = \underbrace{A \otimes \cdots \otimes A}_{n-i} \otimes (A - B) \otimes \underbrace{B \otimes \cdots \otimes B}_{i-1}.$$

Then

$$\bigotimes^n A - \bigotimes^n B = \sum_{i=1}^{n} C_i.$$

Now (8) follows from (7). ■

Proof of the Theorem The usual way to index the coefficients of vectors in $\bigotimes^n \mathbb{C}^n$ is by using multiindices

$$\alpha = (\alpha_1, \ldots, \alpha_n) \in \{1, \ldots, n\}^n.$$

Let $M \subset \{1, \ldots, n\}^n$ be defined by

$$M = \{\alpha : \alpha_i \neq \alpha_j \text{ for } i \neq j\}$$

and introduce the projection P_M by

$$P_M x = \xi = (\xi_\alpha), \qquad x = (x_\alpha), \qquad \xi_\alpha = \begin{cases} x_\alpha & \alpha \in M \\ 0 & \alpha \notin M \end{cases}.$$

Let $e = (1, \ldots, 1)$. It is easy to see that

$$P_M(\bigotimes^n A)P_M e = \operatorname{per}(A) P_M e. \tag{9}$$

In fact, for $\alpha \in M$

$$((\bigotimes^n A)P_M e)_\alpha = \sum_{\beta \in M} (\bigotimes^n A)_{\alpha\beta} = \sum_{\beta \in M} a_{\alpha_1\beta_1} a_{\alpha_2\beta_2} \cdots a_{\alpha_n\beta_n} = \mathrm{per}(A).$$

Using $\|P_M \xi\|_p \leqslant \|\xi\|_p$ we get

$$|\mathrm{per}(A) - \mathrm{per}(B)| \, \|P_M e\|_p = \|P_M(\bigotimes^n A - \bigotimes^n B)P_M e\|_p \leqslant \|(\bigotimes^n A - \bigotimes^n B)P_M e\|_p$$
$$\leqslant \|(\bigotimes^n A - \bigotimes^n B\|_p \|P_M e\|_p.$$

This together with (8) gives (1). ■

3. CONCLUDING REMARKS

Remark 1 Lemma 1 (and 2) may be there in the literature; we were, however, unable to find it. We are thankful to K. B. Sinha for an argument leading to the proof given here. It should be mentioned that it is easy to adapt the proof so as to show the result (2) for nonsquare matrices. The same result also holds for operators on the Lebesgue spaces L_p, $1 \leqslant p \leqslant \infty$.

Remark 2 In the proof of the inequality for the variation of determinants [3] Friedland uses the fact that det A is an eigenvalue of $\bigotimes^n A$ with an eigenvector independent of A. The relation (9) can be interpreted that per(A) is an eigenvalue of a principal submatrix of A, where its position and the eigenvector are independent of A.

Remark 3 By the preceding remarks we have $|\mathrm{per}(A)| \leqslant \|\bigotimes^n A\|_p$, and hence by (7)

$$|\mathrm{per}(A)| \leqslant \|A\|^n \tag{10}$$

for $\| \; \| = \| \; \|_p$, $1 \leqslant p \leqslant n$. The special case $p = 2$ can be found in [4]. For another derivation of (10) see Remark 5.

Remark 4 Given a diagonal matrix $D \in \mathbb{C}^{n,n}$ with positive diagonal entries d_i consider the norm

$$\|x\|_{D,p} = \|Dx\|_p = \left(\sum_{i=1}^{n} |d_i x_i|^p \right)^{1/p}.$$

The operator norm corresponding to this vector norm is given by

$$\|A\|_{D,p} = \|DAD^{-1}\|_p.$$

As $\mathrm{per}(DXD^{-1}) = \mathrm{per}(X)$ for $X \in \mathbb{C}^{n,n}$ we obtain (1) also for $\| \; \| = \| \; \|_{D,p}$, D as above and $1 \leqslant p \leqslant \infty$.

Remark 5 It is, however, not true that (1) holds for all operator norms (contrary to a conjecture in [2]). This can be seen as follows:

Let $\| \; \|$ be an operator norm such that (1) holds for all A, B. If we set $B = tA$ in (1) we get for $0 < t < 1$

$$|\mathrm{per}(A)| \leqslant \frac{n(1-t)}{1-t^n} \|A\|^n.$$

Considering $t \to 1$ we get $|\mathrm{per}(A)| \leqslant \|A\|^n$. Hence, as the infimum of all operator norms of a fixed matrix A is given by its spectral radius $\rho(A)$, and as there are matrices A such that $\rho(A) < |\mathrm{per}(A)|$ (e.g., $A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$ where $\rho(A) = 0 < |\mathrm{per}(A)| = 2$), we find that (1) does not hold for all operator norms. We might as well present a counter-example: Take $A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$, $B = tA, 0 < t < 1$, $\|x\| = \max(\frac{1}{2}|x_1 + x_2|, |x_1 - x_2|)$. Then $\|A\| = 1$ for the associated operator norm and $|\mathrm{per}(A) - \mathrm{per}(B)| = 2(1 - t^2) > 2(1 - t) = 2\|A - B\| \max(\|A\|, \|B\|)$.

Remark 6 Observe that we have actually proved a result slightly stronger than (1), namely

$$|\mathrm{per}(A) - \mathrm{per}(B)| \leqslant \|A - B\|_p \cdot \frac{\|A\|_p^n - \|B\|_p^n}{\|A\|_p - \|B\|_p}.$$

Here it is agreed that $\dfrac{a^n - a^n}{a - a} = na^{n-1}$. This result follows if we use in the proof of Lemma 3 that $\|C_i\|_p = \|A - B\|_p \|A\|_p^{n-i} \|B\|_p^{i-1}$.

References

[1] R. Bhatia, Variation of symmetric tensor powers and permanents, *Lin. Alg. Appl.* **62** (1984), 269–276.
[2] L. Elsner, A note on the variation of permanents, *Lin. Alg. Appl.* **109** (1988), 37–39.
[3] S. Friedland, Variation of tensor powers and spectra, *Lin. Multilin. Alg.* **12** (1982), 81–98.
[4] M. Marcus and H. Minc, Generalized matrix functions, *Trans. Amer. Math. Soc.* **116** (1965), 316–329.
[5] M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Boston, 1964.