NORTH-HOLLAND

# Singular Values, Doubly Stochastic Matrices, and Applications

L. Elsner
*Fakultät für Mathematik*
*Universität Bielefeld*
*33501 Bielefeld, Germany*

and

S. Friedland
*Department of Mathematics, Statistics, and Computer Science*
*University of Illinois at Chicago*
*851 South Morgan Street*
*Chicago, Illinois 60607-7045*

Submitted by Raphael Loewy

ABSTRACT

The Hadamard square of any square matrix $A$ is bounded above and below by some doubly stochastic matrices times the square of the largest and the smallest singular values of $A$. Applications to graphs, permanents, and eigenvalue perturbations are discussed.

## 0. INTRODUCTION

For a set $D \subset \mathbf{C}$ let $M_n(D)$ be the set of $n \times n$ complex valued matrices with entries in $D$. Denote by $\Omega_n \subset M_n(\mathbf{R}_+)$ the convex set of doubly stochastic matrices. Let $A = (a_{ij})_1^n \in M_n(\mathbf{C})$. Set $A^{\bigcirc 2} = A \bigcirc A = (a_{ij}\bar{a}_{ij})_1^n \in M_n(\mathbf{R}_+)$. Here $\bigcirc$ denotes the Hadamard product of matrices. Denote by $\sigma_1(A) \geqslant \cdots \geqslant \sigma_n(A) \geqslant 0$ the singular values of $A$. If $\sigma_1(A) = \cdots = \sigma_n(A) = s > 0$, then $A$ is an $s$-multiple of a unitary matrix and $A^{\bigcirc 2}$ is an $s^2$-multiple of a doubly stochastic matrix. In this paper we generalize the above fact to arbitrary complex valued square matrices.

*LINEAR ALGEBRA AND ITS APPLICATIONS* 220:161–169 (1995)


0024-3795/95/$9.50
SSDI 0024-3795(95)00111-P

MAIN THEOREM. *Let $A \in M_n(\mathbf{C})$. Then there exist $A_1$, $A_2 \in \Omega_n$ such that*

$$\sigma_n(A)^2 A_1 \leqslant A^{\circ 2} \leqslant \sigma_1(A)^2 A_2.$$

Here the inequalities are taken in the entrywise sense. The lower estimate for $A^{\circ 2}$ seems to be much deeper and more useful than the upper estimate. We present applications to graph theory, permanents, and to perturbation estimates on the eigenvalues of matrices.

## 1. PROOF OF THE MAIN THEOREM

Let $B = (b_{ij})_1^n \in M_n(\mathbf{R}_+)$. For $I, J \subset \{1,\ldots,n\}$ let

$$B_{IJ} = \sum_{i \in I, j \in J} b_{ij}.$$

Denote by $|I|$, $|J|$ the cardinalities of $I$, $J$ respectively. $B$ is called $(0 \leqslant)$ $t$-doubly superstochastic (d.sps.) if $B \geqslant tB_1$ for some $B_1 \in \Omega_n$. In [1] Cruse showed that $B$ is $t$-d.sps. iff

$$B_{IJ} \geqslant t(|I| + |J| - n) \qquad \forall I, J \subset \{1,\ldots,n\}.$$

Set

$$\mu(B) = \min_{I,J \subset \{1,\ldots,n\},\, |I|+|J|>n} \frac{B_{IJ}}{|I| + |J| - n}.$$

Cruse's theorem yields that $B$ is $\mu(B)$ d.sps. The first part of the Main Theorem follows from the following lemma.

LEMMA 1.1. *Let $A \in M_n(\mathbf{C})$. Then*

$$\mu(A^{\circ 2}) \geqslant \sigma_n^2(A).$$

*Proof.* It is enough to assume that $A$ is nonsingular. Let $P, Q \in \Omega_n$ be permutation matrices. Clearly, $PAQ$ has the same singular values as $A$, and

$\mu((PAQ)^{\circ 2}) = \mu(A^{\circ 2})$. Thus, it suffices to show that

$$\sum_{i=1,j=1}^{i=p,j=q} |a_{ij}|^2 \geq (p+q-n)\sigma_n^2(A), \qquad 1 \leq p \leq q, \quad n < p+q.$$

Let

$$D_p = \operatorname{diag}(d_1, \dots, d_n), \qquad d_1 = \cdots = d_p = 1, \quad d_{p+1} = \cdots = d_n = 0.$$

Set $A' = D_p A D_q$. Let

$$V = \left\{x : x \in \mathbf{C}^n, (I - D_q)x = (I - D_p)Ax = 0\right\}.$$

It then follows that $m = \dim V \geq p + q - n$. Moreover, $(A - A')V = 0$. Set $W = AV = A'V$. We can view $A$ and $A'$ as representing the same linear transformation $T : V \to W$. Furthermore, $V$, $W$ are inner product spaces induced by the standard inner product in $\mathbf{C}^n$. Observe next

$$\sigma_m(T) = \min_{x \in V\setminus\{0\}} \frac{|Ax|}{|x|} \geq \min_{x \in \mathbf{C}^n\setminus\{0\}} \frac{|Ax|}{|x|} = \sigma_n(A).$$

Hence,

$$\sum_{i=1,j=1}^{i=p,j=q} |a_{ij}|^2 = \operatorname{trace}\left[(A')^* A'\right] = \operatorname{trace}(T^*T)$$

$$= \sum_{1}^{m} \sigma_k(T)^2 \geq m\sigma_m(T)^2 \geq (p+q-n)\sigma_n(A)^2.$$

The proof of the lemma is completed. ∎

For $B \in M_n(\mathbf{R}_+)$ let $r(B) = (r_1(B), \dots, r_n(B))$, $c(B) = (c_1(B), \dots, c_n(B))$ be the vectors corresponding to the row and column sums of $B$. Set

$$\nu(B) = \max_{1 \leq i \leq n} \max\left(r_i(B), c_i(B)\right).$$

LEMMA 1.2. *Let $B \in M_n(\mathbf{R}_+)$. Then there exists a doubly stochastic matrix $B_2 \in \Omega_n$ such that $B \leqslant \nu(B)B_2$.*

*Proof.* It is enough to assume that $B \neq 0$. Let $0 \leqslant k \leqslant 2n - 1$ be the number of rows and columns of $B$ whose sum is less than $\nu(B)$. If $k = 0$ then $B_2 = \nu(B)^{-1}B$ and the lemma holds. Assume that $k \geqslant 1$. Since

$$\sum_1^n r_i(B) = \sum_1^n c_i(B),$$

it follows that there exist $1 \leqslant p, q \leqslant n$ so that $r_p(B), c_q(B) < \nu(B)$. Let $B'$ be the matrix obtained from $B = (b_{ij})_1^n$ by replacing the entry $b_{pq}$ by the entry $b_{pq} + \nu(B) - \max(r_p(B), c_q(B))$. Thus, $B \leqslant B'$, $\nu(B) = \nu(B')$. Moreover, $k' < k$, where $k'$ is the number of rows and columns in $B'$ which are less than $\nu(B') = \nu(B)$. Continue this process until it stops and the lemma is verified. ■

As $\sigma_1(A) = \|A\| = \|A^*\|$, we easily deduce that $\nu(A^{\circ 2}) \leqslant \sigma_1(A)^2$. Lemma 1.2 yields the upper bound of the Main Theorem. We now note that the lower bound in the Main Theorem is sharp. Let $A \in M_3(\mathbf{R}_+)$ be a symmetric matrix whose entries on the first row and column are equal to 1 while all other entries are equal to 0. Then the eigenvalues of $A$ are 2, $-1$, 0 while its singular values are 2, 1, 0. Clearly, if $A \geqslant \rho A_1$ for some $A_1 \in \Omega_3$, then $\rho \leqslant 0$. In particular, there is no analog for Lemma 1.2 for the lower bound of $A^{\circ 2}$.

## 2. APPLICATIONS TO GRAPHS AND PERMANENTS

Let $G$ be a simple digraph on $n$ vertices. That is, $G$ is represented by a $0 - 1$ matrix $A(G) \in M_n(\{0, 1\})$. Vice versa, any $A \in M_n(\{0, 1\})$ represents a simple digraph $G$. $G$ is called strongly connected if any two distinct vertices are connected by a path in $G$. We say that $G$ is without loops if all the diagonal elements of $A(G)$ are equal to zero. $G$ is called $k$-regular if the out- and indegree of every vertex $i$ – $(r_i(A(G)), c_i(A(G)))$ are equal to $k$. $G$ has a spanning $k$-regular subgraph if there exists a $k$-regular graph $H$ on $n$ vertices such that $A(G) \geqslant A(H)$.

One of the unsolved problems in complexity theory is to decide if a simple digraph without loops contains an even cycle. See for example [6] and [12] and the references therein. Let $G$ be a $k$-regular simple digraph without loops. It was shown by the second author in [6] that $G$ contains an even cycle if $k \geqslant 7$. It was conjectured in [6] that the above result is valid for $k \geqslant 3$. This conjecture was proved by Thomassen in [12]. Moreover, Thomassen showed that if $G$ is a strongly connected simple digraph with no loops such that each in- and outdegree is at least 3, then $G$ contains an even cycle.

THEOREM 2.1. *Let $G$ be a simple graph on $n$ vertices. Then $G$ has a $\lfloor \sigma_n(A(G))^2 \rfloor$-regular spanning subgraph. If $G$ has no loops and $\sigma_n(A(G))^2 \geqslant 3$, then $G$ has an even cycle.*

*Proof.* Let $A = A(G)$. Note that $A^{\bigcirc 2} = A$. By Lemma 1.1, $\mu(A) \geqslant \sigma_n(A)^2$. As $A \in M_n(\{0,1\})$, Fulkerson's result [8] implies that $G$ has a $\lfloor \sigma_n(A(G))^2 \rfloor$-regular spanning graph. Assume that $G$ has no loops and $\sigma_n(A)^2 \geqslant 3$. Then Thomassen's result [12] implies that $G$ has an even cycle. ■

Let $B \in M_n(\mathbf{C})$. Let $\det B$, $\operatorname{per} B$ denote the determinant and the permanent of $B$ respectively. While the calculation of determinants is easy, the calculation of permanents is a well-known difficult problem even for $B \in M_n(\{0,1\})$. See for example [13]. It is not difficult to give upper estimates for $\operatorname{per} B$. For example,

$$|\operatorname{per}(b_{ij})_1^n| \leqslant \operatorname{per}(|b_{ij}|)_1^n \leqslant \prod_{i=1}^{n} \sum_{j=1}^{n} |b_{ij}|.$$

It is much harder to give lower estimates on $|p(B)|$. Assume that $B$ is a hermitian nonnegative definite matrix. Then Schur's inequality states $0 \leqslant \det B \leqslant \operatorname{per} B$; e.g. [10, 4.4.4]. The (now proved) van der Waerden conjecture claims that $\operatorname{per} B \geqslant n!/n^n$, $B \in \Omega_n$. We refer to [2] and [4] for the proofs. In [5] generalized van der Waerden type inequalities are treated. Let $B \in M_n(\mathbf{R}_+)$. Then Cruse's inequality states that $B \geqslant \mu(B)B_1$ for some $B_1 \in \Omega_n$. We combine this inequality with the van der Waerden inequality to deduce $\mu(B)^n n!/n^n \leqslant \operatorname{per} B$. Reference [7] contains more discussion on this approach to lower estimates of $\operatorname{per} B$.

Let $B = (b_{ij})_1^n \in M_n(\mathbf{R}_+)$. Set $B^{\bigcirc 1/2} = (\sqrt{b_{ij}})_1^n$. Let $E = (e_{ij})_1^n \in M_n(\mathbf{C})$, $|e_{ij}| = 1$, $i, j = 1, \ldots, n$,. Then the matrix $C = B^{\bigcirc 1/2} \bigcirc E = (\sqrt{b_{ij}}\, e_{ij})_1^n$ satisfies the equality $C^{\bigcirc 2} = B$. Moreover, any $C \in M_n(\mathbf{C})$, $C^{\bigcirc 2} = B$, is of the above form. The Main Theorem yields that $\mu(B) \geqslant \sigma_n(C)^2$. We have thus shown

THEOREM 2.2. *Let $B \in M_n(\mathbf{R}_+)$. Assume that $C^{\bigcirc 2} = B$, $C \in M_n(\mathbf{C})$.*

*Then*

$$\operatorname{per} B \geqslant \frac{\sigma_n(C)^{2n} n!}{n^n}.$$

*In particular, the above inequality holds for* $C = B^{\circ 1/2}$.

The above theorem can be considered as a generalization of Schur's inequality for nonnegative matrices. Given $B \in M_n(\mathbf{R}_+)$, it is of interest to find

$$\max_{C, C^{\circ 2} = B} \sigma_n(C)$$

and to characterize the extremal matrices.

## 3. PERTURBATIONS OF EIGENVALUES

For $A \in M_n(\mathbf{C})$ let $\|A\|_F = [\operatorname{trace}(AA^*)]^{1/2}$ be the Frobenius norm. Recall that $\sigma_1(A) = \|A\|_2$ is the spectral norm of $A$. Moreover,

$$\|AB\|_F \leqslant \min(\|A\|_F \|B\|_2, \|A\|_2 \|B\|_F).$$

Let $\mathrm{GL}_n(\mathbf{C}) \subset M_n(\mathbf{C})$ be the group of invertible matrices. Assume that $A \in \mathrm{GL}_n(\mathbf{C})$. Then $\sigma_n(A) = \sigma_1(A^{-1})^{-1}$. Furthermore $\operatorname{cond}(A) = \|A\|_2 \|A^{-1}\|_2 = \sigma_1(A)/\sigma_n(A)$ is the condition number of $A$. We prove the following generalization of the Hoffman-Wielandt inequality following the approach in [3].

THEOREM 3.1. *Let* $A, B \in M_n(\mathbf{C})$ *be diagonalizable, i.e.*

$$S^{-1}AS = \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n), \quad T^{-1}BT = M = \operatorname{diag}(\mu_1, \dots, \mu_n),$$
$$S, T \in \mathrm{GL}_n(\mathbf{C}).$$

*Then there exist permutations* $\pi$ *and* $\tilde{\pi}$ *of* $\{1, \dots, n\}$ *such that*

$$\sum_{k=1}^{n} |\lambda_k - \mu_{\pi_k}|^2 \leqslant \|A - B\|_F^2 \|S^{-1}\|_2^2 \|T\|_2^2 \|T^{-1}S\|_2^2$$
$$\leqslant \|A - B\|_F^2 \operatorname{cond}(T)^2 \operatorname{cond}(S)^2,$$

$$\sum_{k=1}^{n} |\lambda_k - \mu_{\tilde{\pi}_k}|^2 \geqslant \frac{\|A - B\|_F^2}{\|S\|_2^2 \|T\|_2^2 \|S^{-1}T\|_2^2}$$
$$\geqslant \|A - B\|_F^2 \operatorname{cond}(T)^{-2} \operatorname{cond}(S)^{-2}.$$

*Proof.* Observe that $A - B = S(\Lambda G - GM)T^{-1}$, where $G = S^{-1}T = (g_{ik})_1^n$. Thus, $S^{-1}(A - B)T = \Lambda G - GM$. Use the norm inequalities to obtain

$$\|A - B\|_F^2 \geqslant \|S^{-1}\|_2^{-2}\|T\|_2^{-2} \sum_{i,k=1}^{n} |g_{ik}|^2|\lambda_i - \mu_k|^2.$$

By the Main Theorem the last sum is greater than or equal to $\sigma_n(G)^2\sum_{i,k=1}^n d_{ik}|\lambda_i - \mu_k|^2$ for a suitable doubly stochastic matrix $D = (d_{ik})_1^n$. As the doubly stochastic matrices are convex combinations of permutation matrices (Birkhoff theorem), we have by the usual reasoning (see e.g. [3]) that there is a permutation $\pi$ such that

$$\sum_{i,k=1}^{n} d_{ik}|\lambda_i - \mu_k|^2 \geqslant \sum_{i=1}^{n} |\lambda_i - \mu_{\pi_i}|^2.$$

Use the relations

$$\sigma_n(G) = \sigma_n(S^{-1}T) = \|T^{-1}S\|_2^{-1} \geqslant \|S\|_2^{-1}\|T^{-1}\|_2^{-1}$$

to get the first inequality in Theorem 3.1. Apply the norm inequalities to $A - B = S(\Lambda G - MG)T^{-1}$ to deduce

$$\|A - B\|_F^2 \leqslant \|S\|_2^2\|T^{-1}\|_2^2 \sum_{i,k=1}^{n} |g_{ik}|^2|\lambda_i - \mu_k|^2.$$

Use the upper bound of the Main Theorem and the Birkhoff theorem to deduce the second inequality of the theorem. ∎

We remark that the first inequality of Theorem 3.1 has been proved in [11, p. 216] in a very different way. Also, for unitary $A$, $B$ the above theorem is well known. The methods of the proof of Theorem 3.1 yield a different proof of the following result due to Li [9].

THEOREM 3.2. *Let $T \in M_n(C)$, and assume that $\Lambda^{(i)} = \mathrm{diag}(\lambda_j^{(i)}) \in M_n(\mathbf{C})$, $i = 1, \dots, 4$, are diagonal matrices. Then there is a permutation $\pi$ such that*

$$\|\Lambda^{(1)}T\Lambda^{(2)} - \Lambda^{(3)}T\Lambda^{(4)}\|_F^2 \geqslant \sigma_n(T)^2 \sum_{i=1}^{n} |\lambda_i^{(1)}\lambda_{\pi_i}^{(2)} - \lambda_i^{(3)}\lambda_{\pi_i}^{(4)}|^2.$$

*Proof*. Set $T = (t_{ij})_1^n$. From the Main Theorem, $(|t_{ij}|^2)_1^n \geqslant \sigma_n(T)^2 D$, where $D = (d_{ij})_1^n$ is a doubly stochastic matrix. Then

$$\|\Lambda^{(1)}T\Lambda^{(2)} - \Lambda^{(3)}T\Lambda^{(4)}\|_F^2 = \sum_{i,j=1}^{n} |t_{ij}|^2 |\lambda_i^{(1)}\lambda_j^{(2)} - \lambda_i^{(3)}\lambda_j^{(4)}|^2$$

$$\geqslant \sigma_n(T)^2 \sum_{i,j=1}^{n} d_{ij} |\lambda_i^{(1)}\lambda_j^{(2)} - \lambda_i^{(3)}\lambda_j^{(4)}|^2.$$

Use the Birkhoff theorem to deduce the existence of a permutation for which the claimed inequality holds. ∎

It was pointed out to us by R.-C. Li that our proof of Theorem 3.2 yields the following inequality for any $A = (a_{ij})_1^n$, $T \in M_n(\mathbf{C})$:

$$\|A \circ T\|_F^2 \geqslant \sigma_n^2(T) \min_{\pi} \sum_{i=1}^{n} |a_{i\pi_i}|^2.$$

## REFERENCES

1 A. B. Cruse, A proof of Fulkerson's characterization of permutation matrices, *Linear Algebra Appl.* 12:21–28 (1975).
2 G. P. Egorichev, The solution of the van der Waerden problem for permanents, *Dokl. Akad. Nauk SSSR* 258:1041–1044 (1981).
3 L. Elsner, A note on the Hoffman-Wielandt theorem, *Linear Algebra Appl.* 182:235–237 (1993).
4 D. I. Falikman, A proof of van der Waerden's conjecture on the permanents of a doubly stochastic matrix, *Mat. Zametki* 29:931–938 (1981).
5 S. Friedland, A proof of a generalized van der Waerden conjecture on permanents, *Linear and Multilinear Algebra* 11:107–120 (1982).
6 S. Friedland, Every 7-regular digraph contains an even cycle, *J. Combin. Theory Ser. B* 46:249–252 (1989).
7 S. Friedland, C. K. Li, and H. Schneider, Additive decompositions of nonnegative matrices with applications to permanents and scaling, *Linear and Multilinear Algebra* 23:63–78 (1988).
8 D. R. Fulkerson, The maximum number of disjoint permutations contained in a matrix of zeros and ones, *Canad. J. Math.* 16:729–735 (1964).
9 R.-C. Li, On perturbations of matrix pencils with real spectra, *Math. Comp.*, to appear.

10 M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Prindle, Weber & Schmidt, Boston, 1964.
11 G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic, 1990.
12 C. Thomassen, The even cycle problem for directed graphs, *J. Amer. Math. Soc.* 5:217–229 (1992).
13 L. G. Valiant, The complexity of computing the permanent, *Theoret. Comput. Sci.* 8:189–201 (1979).

*Received 25 November 1993; final manuscript accepted 8 May 1994*