Singular values and doubly stochastic matrices

L. Elsner
Fakultät für Mathematik
Universität Bielefeld
33501 Bielefeld
Germany

S. Friedland
Department of Mathematics, Statistics, and Computer Science
University of Illinois at Chicago
Chicago, Illinois 60680
USA

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.

§1. 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 $A^{\circ 2} = A \circ \bar{A} = (a_{ij}\bar{a}_{ij})_1^n \in M_n(\mathbf{R}_+)$. Denote by $\sigma_1(A) \geq \cdots \geq \sigma_n(A) \geq 0$ the singular values of A. If $\sigma_1(A) = ... = \sigma_n(A) = s > 0$, then A is an s-multiple of a unitary matrix and $A^{\circ 2}$ is an s^2-multiple of a doubly stochastic matrix. In this paper we generalize the above fact to arbitrary complex valued square matrices.

Main Theorem. *Let $A \in M_n(\mathbf{C})$. Then there exist $A_1, A_2 \in \Omega_n$ so that*

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

Here the inequalities are taken in the entrywise sense. The lower estimate for $A^{\circ 2}$ seems to be much deeper than the upper estimate. We bring 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 $I, J \subset \{1, ..., n\}$ 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 $(0 \leq) t$ doubly super stochastic (d.sp.s.) if $B \geq tB_1$ for some $B_1 \in \Omega_n$. In [**Cru**] Cruse showed that B is t d.sp.s. iff

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

Set

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

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

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

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

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

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

Let

$$D_p = diag(d_1, ..., d_n), d_1 = ... = d_p = 1, d_{p+1} = ... = d_n = 0.$$

Set $A' = D_p A D_q$. Let

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

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,

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

$$\sum_1^m \sigma_k(T)^2 \geq m\sigma_m(T)^2 \geq (p+q-n)\sigma_n(A)^2.$$

The proof of the lemma is completed. $\diamond$

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

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

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

Proof. It is enough to assume that $B \neq 0$. Let $0 \leq k \leq 2n-1$ 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 $k \geq 1$. Since

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

it follows that there exists $1 \leq p, q \leq n$ 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, $B \leq B', \nu(B) = \nu(B')$. Moreover, $k' < k$, where k' is the number of rows and columns in B' which are less then $\nu(B') = \nu(B)$. Continue this process until it stops and the lemma is verified. $\diamond$

As $\sigma_1(A) = \|A\| = \|A^*\|$ we easily deduce that $\nu(A^{\circ 2}) \leq \sigma_1(A)^2$. Lemma (1.2) yields the upper bound of the Main Theorem.

§2. Applications to graphs and permanents

Let G be a simple digraph on n vertices. That is, G is represented by $0-1$ 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 in degree of every vertex i $(r_i(A(G)), c_i(A(G)))$ is equal to k. G has a spanning k-regular subgraph if there exists a k-regular graph H on n vertices so that $A(G) \geq A(H)$.

One of the unsolved problems in complexity theory is to decide if a simple digraph without loops contains an even cycle. See for example [**Fri2**] and [**Tho**] and the references therein. Let G be a k-regular simple digraph without loops. It was shown by the second author in [**Fri2**] that G contains an even cycle if $k \geq 7$. It was conjectured in [**Fri2**] that the above result is valid for $k \geq 3$. This conjecture was proved by Thomassen in [**Tho**]. Moreover, Thomassen showed that if G is a strongly connected simple digraph with no loops such each in and out degree is at least 3 then G contains an even cycle.

(2.1) Theorem *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 $\sigma_n(A(G))^2 \geq 3$ then G has an even cycle.*

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

Let $B \in M_n(\mathbf{C})$. Let $det(B), per(B)$ 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 [**Val**]. It is not difficult to give upper estimates for $per(B)$. For example,

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

It is much harder to give lower estimates on $|p(B)|$. Assume that B is a hermitian nonnegative definite matrix. Then Schur's inequality claims $0 \leq det(B) \leq per(B)$, e.g. [**M-M**, 4.4.4]. The (now proved) van der Waerden conjecture claims that $per(B) \geq \frac{n!}{n^n}, B \in \Omega_n$. We refer to [**Ego**] and [**Fal**] for the proofs. In [**Fri1**] generalized van der Waerden type inequalities are treated. Let $B \in M_n(\mathbf{R}_+)$. Then Cruse's inequality states that $B \geq \mu(B)B_1$ for some $B_1 \in \Omega_n$. We combine this inequality with the van der Waerden inequality to deduce $\frac{\mu(B)^n n!}{n^n} \leq per(B)$. [**F-L-S**] contains more discussion on this approach to lower estimates of $per(B)$.

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

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

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

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

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 $||A||_F = (trace(AA^*))^{\frac{1}{2}}$ be the Frobenius norm. Recall that $\sigma_1(A) = ||A||_2$ is the spectral norm of A. Moreover,

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

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

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

$$S^{-1}AS = \Lambda = diag(\lambda_1, \ldots, \lambda_n), \; T^{-1}BT = M = diag(\mu_1, \ldots, \mu_n), \; S, T \in GL_n(\mathbf{C}).$$

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

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

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

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

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.**[Els]**) that there is a permutation π such that

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

Use the relations

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

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

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

Use the upper bound of the Main Theorem and the Birkhoff theorem to deduce the second inequality of the theorem. $\diamond$

We remark that the first inequality of Theorem 3.1 has been proved in (**[St-S]**, p.216) in a very different way. The methods of the proof of Theorem 3.1 yield a different proof of the following result due to R. Li [**Li**].

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

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

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

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

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

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

References

[**Cru**] A.B. Cruse, A proof of Fulkerson's characterization of permutation matrices, *Linear Algebra Appl.* 12 (1975), 21-28.

[**Ego**] G.P. Egorichev, The solution of the van der Waerden problem for permanents, *Dokl. Akad. Nauk SSSR* 258 (1981), 1041-1044.

[**Els**] L. Elsner, A Note on the Hoffman - Wielandt Theorem. *Linear Algebra Appl.* 182 (1993), 235-237.

[**Fal**] D.I. Falikman, A proof of van der Waerden's conjecture on the permanents of a doubly stochastic matrix, *Mat. Zametki* 29 (1981), 931-938.

[**Fri1**] S. Friedland, A proof of a generalized van der Waerden conjecture on permanents, *Linear Multilin. Algebra* 11 (1982), 107-120.

[**Fri2**] S. Friedland, Every 7-regular digraph contains an even cycle, *J. Combin. Theory Ser. B* 46 (1989), 249-252.

[**F-L-S**] S. Friedland, C.K. Li and H. Schneider, Additive decompositions of nonnegative matrices with applications to permanents and scaling, *Linear Multilin. Algebra* 23 (1988), 63-78.

[**Ful**] D.R. Fulkerson, The maximum number of disjoint permutations contained in a matrix of zeros and ones, *Canad. J. Math.* 16 (1964), 729-735.

[**Li**] R.-C. Li, On perturbations of matrix pencils with real spectra, *Math. Comp.*, (1994), to appear.

[**M-M**] M. Marcus and H. Minc, *A survey of matrix theory and matrix inequalities*, Prindle, Weber & Schmidt, Boston 1964.

[**St-S**] G.W.Stewart- J.Sun, *Matrix perturbation theory*, Academic Press (1990)

[**Tho**] C. Thomassen, The even cycle problem for directed graphs, *Journal AMS* 5 (1992), 217-229.

[**Val**] L.G. Valiant, The complexity of computing the permanent, *Theoretical Computer Science* 8 (1979), 189-201.