Revisiting Schur’s bound on the largest singular value

Vladimir Nikiforov
Department of Mathematical Sciences, University of Memphis, Memphis TN 38152, USA, email: vnikifrv@memphis.edu

March 30, 2022

Abstract

We give upper and lower bounds on the largest singular value of a matrix using analogues to walks in graphs. For nonnegative matrices these bounds are asymptotically tight.

In particular, the following result improves a bound due to Schur. If $A = (a_{ij})$ is an $m \times n$ complex matrix, its largest singular value satisfies

$$\sigma^2 (A) \leq \max_{i \in [m]} \sum_{j \in [n]} |a_{ij}| c_j \leq \max_{a_{ij} \neq 0} r_i c_j,$$

where $r_i = \sum_{k \in [n]} |a_{ik}|$, $c_j = \sum_{k \in [m]} |a_{kj}|$.

Keywords: largest singular value, Schur's bound, singular values, walks.

AMS classification: 15A42

1 Introduction

How large the largest singular value $\sigma (A)$ of an $m \times n$ matrix $A = (a_{ij})$ can be? In 1911 Schur [4], p. 6, gave the bound

$$\sigma^2 (A) \leq \max_{i \in [m], j \in [n]} r_i c_j,$$

(1)

where $r_i = \sum_{k \in [n]} |a_{ik}|$, $c_j = \sum_{k \in [m]} |a_{kj}|$.

The aim of this note to strengthen this bound and give similar lower bounds on $\sigma (A)$.

In particular, our results imply that if $A$ is nonzero, then

$$\sigma^2 (A) \leq \max_{i \in [m]} \sum_{j \in [n]} |a_{ij}| c_j \leq \max_{a_{ij} \neq 0} r_i c_j.$$

(2)

Note that sometimes (2) is much stronger than (1). Indeed, letting $A$ be the adjacency matrix of the star $K_{1,n}$, inequality (1) gives $\sigma^2 (A) \leq n^2$, while (2) gives $\sigma^2 (A) \leq n$, which is best possible, in view of $\sigma^2 (A) = n$. 

For basic notation and definitions see [2]. In particular, $j_m$ denotes the vector of $m$ ones.

Given an $m \times n$ matrix $A = (a_{ij})$, for all $r \geq 0$ and $i, j \in [m]$, let $w^r_A (i, j)$ be the $(i, j)$th entry of $(AA^*)^r$. Set $w^r_A (k) = \sum_{i\in[m]} w^r_A (k, i)$ and $w^r_A = \sum_{i\in[m]} w^r_A (i)$.

Note that if $A$ is the adjacency matrix of a graph, then $w^r_A$ is the number of walks on $2r + 1$ vertices.

The following theorem generalizes inequality (2) and thus, inequality (1).

**Theorem 1** For every nonzero $m \times n$ matrix $A = (a_{ij})$ and all $r \geq 0$, $p \geq 1$,

$$
\sigma^{2p} (A) \leq \max_{k\in[m]}, w^r_{[A](k)\neq 0} \frac{w^{r+p}_{\{A\}} (k)}{w^r_{\{A\}} (k)}
$$

where $|A| = (|a_{ij}|)$.

The values $w^r_A$ can be used for lower bounds on $\sigma (A)$ as well.

**Theorem 2** For every matrix $A$ and all $r \geq 0$, $p \geq 1$,

$$
\sigma^{2p} (A) \geq \frac{w^{r+p}_{\{A\}}}{w^r_{\{A\}}}
$$

unless $\Sigma (AA^*) = 0$.

On the other hand, for almost all matrices $A$ and $r$ large, Theorems 1 and 2 are nearly optimal.

**Theorem 3** For every $m \times n$ matrix $A$ and all $p \geq 1$,

$$
\sigma^{2p} (A) = \lim_{r\to\infty} \frac{w^{r+p}_{\{A\}}}{w^r_{\{A\}}} = \lim_{r\to\infty} \max_{k\in[m], w^r_{[A](k)\neq 0}} \frac{w^{r+p}_{\{A\}} (k)}{w^r_{\{A\}} (k)}
$$

unless the eigenspace of $AA^*$ corresponding to $\sigma^2 (A)$ is orthogonal to $j_m$.

The following proposition sheds some light on Theorems 2 and 3.

**Proposition 4** For every $m \times n$ matrix $A$, the equality $w^1_{\{A\}} = 0$ holds if and only if $w^r_{\{A\}} = 0$ holds for all $r \geq 1$. If $w^1_{\{A\}} = 0$, then $j_m$ is an eigenvector to $AA^*$ corresponding to 0; consequently all eigenvectors of $AA^*$ to nonzero eigenvalues are orthogonal to $j_m$.

Note also that, using Proposition 5 below, Theorem 1 can be extended to partitioned matrices. In particular, if $A$ is an $m \times n$ matrix partitioned into $pq$ blocks $A_{ij}$, $i \in [p]$, $j \in [q]$, then

$$
\sigma^2 (A) \leq \max_i \sum_{k=1}^n \sigma (A_{ik}) \sum_{k=1}^m \sigma (A_{kj}) \leq \max_{A_{ij} \neq 0} \sum_{k=1}^n \sigma (A_{ik}) \sum_{k=1}^m \sigma (A_{kj}).
$$

**Proposition 5** Let the matrix $A$ be partitioned into $p \times q$ blocks $A_{ij}$, $i \in [p]$, $j \in [q]$. For all $i \in [p]$ and $j \in [q]$, let $b_{ij} = \sigma (A_{ij})$. Then the matrix $B = (b_{ij})$ satisfies $\sigma (A) \leq \sigma (B)$.
2 Proofs

Proof of Theorem 1 Since $\sigma (A) \leq \sigma (|A|)$, to simplify the presentation, we shall assume that $A$ is nonnegative. Likewise, dropping all zero rows, we may assume that $A$ has no zero rows, that is to say, $w_A^p (i) > 0$ for all $i \in [m]$. Set $b_{ii} = w_A^p (i)$ for $i \in [m]$ and let $B$ be the diagonal matrix with main diagonal $(b_{11}, \ldots, b_{mm})$. Since $B^{-1} (AA^*)^T B$ has the same spectrum as $(AA^*)^T$, the value $\sigma^{2r} (A)$ is bounded from above by the maximum row sum of $B^{-1} (AA^*)^T B$ - say the sum of the $q$th row - and so,

$$
\sigma^{2r} (A) \leq \sum_{i \in [m]} w_A^r (q, i) \frac{w_A^p (i)}{w_A^p (q)} = \frac{1}{w_A^p (q)} \sum_{i \in [m]} w_A^r (q, i) \sum_{j \in [m]} w_A^p (i, j)
= \frac{1}{w_A^p (q)} \sum_{j \in [m]} \sum_{i \in [m]} w_A^r (q, i) w_A^p (i, j) = \frac{1}{w_A^p (q)} \sum_{j \in [m]} w_A^{r+p} (q, j)
= \frac{w_A^{r+p} (q)}{w_A^p (q)} \leq \max_{k \in [m]} \frac{w_A^{r+p} (k)}{w_A^p (k)},
$$

completing the proof. \qed

Proof of inequalities (2) Theorem 1 with $r = 0$ and $p = 1$ implies that

$$
\sigma^2 (A) \leq \max_{i \in [m]} w_A^1 (i) = \max_{i \in [m]} \sum_{k \in [m], j \in [n]} |a_{ij}| |a_{kj}| = \max_{i \in [m]} \sum_{j \in [n]} |a_i| \sum_{k \in [m]} |a_{kj}| = \max_{i \in [m]} \sum_{a_{ij} \neq 0} |a_{ij}| c_j.
$$

Suppose the maximum in the right hand side is attained for $i = k$. Then,

$$
\sum_{j \in [n]} |a_{kj}| c_j = \sum_{j \in [n]} \frac{|a_{kj}|}{r_k} r_k c_j \leq \sum_{j \in [n]} \frac{|a_{kj}|}{a_{kj} \neq 0} r_k c_j = \max_{a_{ij} \neq 0} r_i c_j,
$$

completing the proof in this case. \qed

In the proofs below we shall assume that $\sigma = \sigma_1 \geq \cdots \geq \sigma_m$ are the singular values of $A$. Let $AA^* = VDV^*$ be the unitary decomposition of $AA^*$; thus, the columns of $V$ are the unit eigenvectors to $\sigma_1^2, \ldots, \sigma_m^2$ and $D$ is the diagonal matrix with $\sigma_i^2, \ldots, \sigma_m^2$ along its main diagonal. Writing $\Sigma (B)$ for the sum of the entries of a matrix $B$, note that for every $l \geq 0$,

$$
w_A^l = \Sigma ((AA^*)^l) = \Sigma (VD^l V^*) = \sum_{i \in [m]} c_i \sigma_i^{2l},
$$

where $c_i = \left| \sum_{j \in [m]} v_{ji} \right|^2 \geq 0$ is independent of $l$.

Proof of Proposition 4 In the notation above we see that $w_A^l = 0$ if and only if $c_i = 0$ for every nonzero $\sigma_i$, thus if and only if $w_A^1 = 0$. 

\[3\]
Note that \( w^1_A = \sum (AA^*) = \langle AA^*j_m, j_m \rangle \); hence, if \( w^1_A = 0 \), then \( j_m \) is an eigenvector of \( AA^* \) to 0. Indeed, since \( AA^* \) is positive semidefinite, by the Rayleigh principle, \( \langle AA^*x, x \rangle = 0 \) implies \( AA^*x = 0 \). The proof is completed.

Proof of Theorem 3

In the above notation we see that

\[
\sigma^{2p} w^{2p}_A = \sum_{i \in [m]} c_i \sigma_i^{2p} \sigma_i^{2p} \geq \sum_{i \in [m]} c_i \sigma_i^{2p+2r} = w_i^{p+r}.
\]

The proof is completed by Proposition 4.

Proof of Theorem 4

Assume that there is an eigenvector of \( AA^* \) to \( \sigma^2(A) \) that is not orthogonal to \( j_m \). Therefore, we may assume that \( c_1 > 0 \). Hence,

\[
\lim_{r \to \infty} \frac{\sum_{i \in [m]} c_i \sigma_i^{2p+2r}}{\sum_{i \in [m]} c_i \sigma_i^{2p}} = \sigma^{2p} \lim_{r \to \infty} \frac{\sum_{i=\sigma_1} c_i}{\sum_{i=\sigma_1} c_i} = \sigma^{2p},
\]

proving the first equality of the theorem.

For \( k \in [m] \) and every \( l \geq 0 \), the value \( w^l_A(k) \) is the \( k \)th row sum of the matrix \( VD^tV^* \); hence

\[
w^l_A(k) = \sum_{i \in [m]} \sum_{j \in [m]} v_{ki} \sigma_i^{2p} v_{ji} = \sum_{i \in [m]} \sigma_i^{2p} v_{ki} \sum_{j \in [m]} v_{ji} = \sum_{i \in [m]} b_i \sigma_i^{2p},
\]

where \( b_i = v_{ki} \sum_{j \in [m]} v_{ji} \) is independent of \( l \). Writing \( t \) for the largest number such that \( \sum_{i=\sigma_1} b_i \neq 0 \), we see that

\[
\lim_{r \to \infty} \frac{w^{r+p}_A(k)}{w^p_A(k)} = \sigma^t \leq \sigma^{2p}.
\]

On the other hand, since

\[
\frac{\max_{k \in [m]} w^{r+p}_A(k)}{w^p_A(k)} = \frac{\sum_{i \in [m]} w^{r+p}_A(i)}{\sum_{i \in [m]} w^p_A(i)}
\]

we obtain

\[
\liminf_{r \to \infty} \max_{k \in [m]} \frac{w^{r+p}_A(k)}{w^p_A(k)} \geq \sigma^{2p},
\]

completing the proof.

Proof of Proposition 5

Let \( A = (a_{ij}) \) be an \( m \times n \) matrix and \( [m] = \cup_{i=1}^p P_i \) and \( [n] = \cup_{i=1}^q Q_i \) be the partitions of its index sets. Select unit vectors \( x = (x_1, \ldots, x_n) \) and \( y = (y_1, \ldots, y_m) \) such that \( \sigma(A) = \langle Ax, y \rangle \). Then we have

\[
\sigma(A) = \langle Ax, y \rangle = \sum_{i \in [m], k \in [n]} a_{ik} x_k \bar{y}_i = \sum_{r \in [p], s \in [q]} \sum_{i \in P_r} \sum_{k \in Q_s} a_{ik} x_k \bar{y}_i \leq \sum_{r \in [p], s \in [q]} \sigma(A_{rs}) \sqrt{\sum_{i \in P_r} \left| x_i \right|^2} \sum_{k \in Q_s} \left| y_k \right|^2 \leq \sigma(B),
\]

where
completing the proof. □

Concluding remarks
Theorem 1 and 2 extend Theorems 5 and 16 of [3], that in turn generalize a number of results about the spectral radius of graphs - see, e.g., the references of [3].
Inequality (3) implies the essential result of the paper [1]; however, we admit that this paper triggered the present note.

References

[1] K.C. Das, R. Bapat, A sharp upper bound on the spectral radius of weighted graphs, preprint available at http://com2mac.postech.ac.kr/papers/2005/05-20.pdf

[2] R. Horn, C. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985. xiii+561 pp.

[3] V. Nikiforov, Walks and the spectral radius of graphs, Linear Algebra Appl. 418 (2006), 257-268.

[4] I. Schur, Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen, Journal für Reine und Angew. Mathematik, 140 (1911), 1–28.