Lower Bounds for the Spread of a Matrix

Charles R. Johnson*
Institute for Physical Science and Technology
University of Maryland
College Park, Maryland 20742

and

Mathematical Sciences Dept.
Clemson University
Clemson, South Carolina 29631

and

Ravinder Kumar and Henry Wolkowicz†
Department of Mathematics
The University of Alberta
Edmonton, Alberta, Canada T6G 2G1

Dedicated to Helmut Wielandt on his 75th birthday

Submitted by Emeric Deutsch

ABSTRACT

A characterization of the spread of a normal matrix is used to derive several simple lower bounds for the spread. Comparisons are then made with several known bounds.

1. INTRODUCTION

We are interested in estimating the maximum distance between two eigenvalues of a given $n \times n$ matrix. For the matrix A with eigenvalues $\lambda_1, \ldots, \lambda_n$, we let

$$s(A) = \max_{i,j} |\lambda_i - \lambda_j| \tag{1.1}$$

denote the *spread of A*. Bounds for $s(A)$ have been given in [1] and [5–9]. In particular, Mirsky [6] has presented the following characterization for the spread.

*The work of this author was supported in part by National Science Foundation grant MCS 80-01611 and by Air Force Wright Aeronautical Laboratory Contract F 33615-81-K-3224.
†Research partially supported by NSERC A3388.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/85/$3.30

THEOREM 1.1. *If A is normal, then*

$$s(A) = \sup_{u,v} |(u, Au) - (v, Av)|$$
$$\geq \sqrt{3} \sup_{u,v} |(u, Av)|, \tag{1.2}$$

while if A is Hermitian, then

$$s(A) = 2 \sup_{u,v} |(u, Av)|, \tag{1.3}$$

where the upper bounds above are taken with respect to all orthonormal vectors u, v.

If we let

$$W(A) = \{(x, Ax) : (x, x) = 1\}$$

denote the *numerical range of* A (e.g. [3]), then it is well known that for normal matrices, $W(A)$ is the convex hull of the spectrum of A, denoted $\operatorname{co} \sigma(A)$. Thus the equality in (1.2) follows, and moreover the sup is attained if we choose u and v to be the eigenvectors corresponding to the eigenvalues for which the max is attained in (1.1). In fact, it is clear that this relation holds whenever $W(A) = \operatorname{co} \sigma(A)$. A characterization of such matrices A is given in [3].

By compactness and continuity, the sup is attained in (1.3) as well. In [2] it is shown that, for any A,

$$\sup_{\substack{\|u\| = \|v\| = 1 \\ (u,v) = 0}} |(u, Av)| = \min_{\alpha} \|A - \alpha I\|_{sp}, \tag{1.4}$$

where $\| \ \|_{sp}$ denotes spectral norm.

We derive lower bounds for $s(A)$ for A normal [or more generally for A satisfying $W(A) = \operatorname{co} \sigma(A)$] and for A Hermitian, from appropriate choices of u and v in (1.2) and (1.3). In particular, two bounds derived are

$$s(A) \geq \frac{1}{n-1} \left| \sum_{i \neq j} a_{ij} \right|, \tag{1.5}$$

and, if R_i denotes the ith row sum, ν denotes the standard deviation of the

row sums, and $R_{i_1} \geq R_{i_2} \geq \cdots \geq R_{i_n}$, then (for A symmetric)

$$s(A) \geq 2\nu \geq \left\{ \frac{2}{n} \sum_{j=1}^{[(n+1)/2]} \left(R_{i_j} - R_{i_{n-j+1}} \right)^2 \right\}^{1/2} \geq \frac{2}{n} \sum_{j=1}^{[(n+1)/2]} \left(R_{i_j} - R_{i_{n-j+1}} \right)$$

where $[\cdot]$ denotes greatest integer part. Comparisons of the bounds derived, with several known bounds, are made in Section 3.

In the remainder of this section we list some lower bound for $s(A)$, A normal or Hermitian, which have appeared in the literature.

Mirsky [6,7] presents the following lower bounds for $s(A)$, A normal:

$$s(A) \geq \sqrt{3} \max_{i \neq j} |a_{ij}|,$$

$$s(A) \geq \max_{i \neq j} \left\{ \left(\operatorname{Re} a_{ii} - \operatorname{Re} a_{jj} \right)^2 + |a_{ij} + \bar{a}_{ji}|^2 \right\}^{1/2},$$

$$s(A) \geq \max_{i \neq j} \left\{ |a_{ii} - a_{jj}|^2 + \left(|a_{ij}| - |a_{ji}| \right)^2 \right\}^{1/2},$$

$$s(A) \geq \max_{i \neq j} \left(|a_{ij}| + |a_{ji}| \right),$$

$$s(A) \geq \max_{i \neq j} \left(\tfrac{1}{2} c_{ij} \right)^{1/2}, \tag{1.6}$$

where

$$c_{ij} = |a_{ii} - a_{jj}|^2 + \left| (a_{ii} - a_{jj})^2 + 4a_{ij}a_{ji} \right| + 2|a_{ij}|^2 + 2|a_{ji}|^2.$$

While if A is Hermitian, then:

$$s(A) \geq 2 \max_{i \neq j} |a_{ij}|,$$

$$s(A) \geq \max_{i \neq j} \left\{ (a_{ii} - a_{jj})^2 + 4|a_{ij}|^2 \right\}^{1/2}. \tag{1.7}$$

Brauer and Mewborn [1] present the following lower bounds for $s(A)$, A normal. Let $n \geq 2$, s_1 be the trace of any principal minor matrix B of order

$k \geq 3$, and s_2 the sum of the principal minors of order 2 of B. Then

$$s(A) \geq \begin{cases} \frac{2}{k}\left|(k-1)s_1^2 - 2ks_2\right|^{1/2}, & k \text{ even}, \\ \left(\frac{4}{k^2-1}\right)^{1/2}\left|(k-1)s_1^2 - 2ks_2\right|^{1/2}, & k \text{ odd}. \end{cases} \tag{1.8}$$

Let c_1 and c_2 be the first and second symmetric functions of the eigenvalues of A respectively. Set $K_2 = \{2(1-1/n)c_1^2 - 4c_2\}^{1/2}$. If A has real roots, then

$$s(A) \geq \begin{cases} \sqrt{\frac{2}{n}}\, K_2 & n \text{ even}, \\ \left(\frac{2n}{n^2-1}\right)^{1/2} K_2 & n \text{ odd}. \end{cases} \tag{1.9}$$

With A Hermitian,

$$s(A) \geq \frac{1}{2}\max_{i \neq j}\left\{a_{ii} + a_{jj} + \left[(a_{ii} - a_{jj})^2 + 4|a_{ij}|^2\right]^{1/2}\right\} - \frac{1}{2}\min_{i \neq j}\left\{a_{ii} + a_{jj} - \left[(a_{ii} - a_{jj})^2 + 4|a_{ij}|^2\right]^{1/2}\right\}, \tag{1.10}$$

and if in addition, $n \geq 3$, s_1 is the trace of any principal minor matrix B of order $k \geq 3$, and s_2 is the sum of the principal minors of order 2 of B, then

$$s(A) \geq \begin{cases} \frac{2}{k}\left\{(k-1)s_1^2 - 2ks_2\right\}^{1/2}, & k \text{ even}, \\ \left(\frac{4}{k^2-1}\right)^{1/2}\left\{(k-1)s_1^2 - 2ks_2\right\}^{1/2}, & k \text{ odd}. \end{cases} \tag{1.11}$$

Wolkowicz and Styan [9] present lower bounds for matrices A with real eigenvalues. Let

$$s^2 = \frac{\operatorname{tr} A^2}{n} - \left(\frac{\operatorname{tr} A}{n}\right)^2;$$

then

$$s(A) \geq \begin{cases} 2s, & n \text{ even}, \\ 2sn/(n^2-1)^{1/2}, & n \text{ odd}. \end{cases}$$

Lower bounds for general matrices are given in [10].

For matrices with real eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$, we get that $s(A) = \lambda_1 - \lambda_n$. Thus lower bounds for λ_1 and upper bounds for λ_n provide lower bounds for $s(A)$. For example, in [4] it is shown that, for A Hermitian, $\lambda_1 \geq (1/n)\Sigma_{i,j} a_{ij}$. Since $\lambda_n \leq (1/n)\Sigma_i a_{ii} = (1/n)\Sigma_i \lambda_i$, we conclude that

$$s(A) \geq \frac{1}{n} \sum_{i \neq j} a_{ij}.$$

This lower bound is improved in Theorem 2.1.

2. LOWER BOUNDS FOR THE SPREAD

Given a Hermitian matrix A, the Rayleigh principle states that the spectral radius

$$\rho(A) = \max_i |\lambda_i| = \max_{u \neq 0} \frac{|u^* A u|}{u^* u}. \tag{2.1}$$

Thus, evaluating the Rayleigh quotient u^*Au/u^*u, for any choice of u, yields a lower bound for $\rho(A)$. In particular, choosing $u = e = (1, 1, \ldots, 1)'$ yields the lower bound

$$\rho(A) \geq \left| \sum_{i,j} \frac{a_{ij}}{n} \right|. \tag{2.2}$$

Merikoski [4] shows that (2.2) is a particularly good estimate when A is real, nonnegative elementwise, and symmetric.

From Theorem 1.1, we see that any choice of u, v orthonormal yields a lower bound for the spread $s(A)$. For example, Mirsky [6] notes that choosing $u = e_i$, and $v = e_j$, the ith and jth unit vectors respectively, with $i \neq j$, yields (for A Hermitian)

$$s(A) \geq 2 \max_{i \neq j} |a_{ij}|. \tag{2.3}$$

We now consider different choices for the orthonormal vectors u and v in (1.2) and (1.3). The first choice yields the following easily computable lower bound for $s(A)$.

THEOREM 2.1. *Suppose that A is real and normal (or more generally $W(A) = co\,\sigma(A)$). Then*

$$s(A) \geq \frac{1}{n-1}\left|\sum_{i\neq j} a_{ij}\right|. \tag{2.4}$$

Proof. Choose $u = (1/\sqrt{n})e = (1/\sqrt{n})(1,1,\ldots,1)'$ and $v = (1/\sqrt{2})e_k - (1/\sqrt{2})e_l$, where e_k and e_l are the kth and lth unit vectors respectively. Then (1.2) implies that

$$\begin{aligned} s(A) &\geq \frac{1}{n}\sum_{i,j} a_{ij} - \min_{k\neq l}\left(\frac{a_{kk}+a_{ll}-2a_{kl}}{2}\right) \\ &\geq \frac{1}{n}\sum_{i,j} a_{ij} - \frac{1}{n(n-1)}\frac{1}{2}\sum_{k\neq l}(a_{kk}+a_{ll}) + \frac{1}{n(n-1)}\sum_{k\neq l} a_{kl} \\ &= \frac{1}{n}\sum_{i,j} a_{ij} - \frac{1}{2n(n-1)}\sum_{k} 2(n-1)a_{kk} + \frac{1}{n(n-1)}\sum_{k\neq l} a_{kl} \\ &= \frac{1}{n-1}\sum_{i\neq j} a_{ij}. \end{aligned} \tag{2.5}$$

The result now follows by noting that $s(-A) = s(A)$. ∎

The above lower bound for $s(A)$ is extremely easy to calculate. Note that it differs from the lower bound for the spectral radius in (2.2) only in that the diagonal elements are ignored and $1/n$ is replaced by the larger $1/(n-1)$. Moreover, the bound is attained when $A = J$, the matrix whose elements are all equal to 1. In Section 3 we will see that the bound (2.4) is particularly good when A is nonnegative elementwise and symmetric.

Now consider the *partitions* defined by the disjoint sets I, J, K

$$\begin{aligned} &\varnothing \neq I \subset \{1,\ldots,n\}, \\ &\varnothing \neq J \subset \{1,\ldots,n\}\setminus I, \\ &K = \{1,\ldots,n\}\setminus(I\cup J) \end{aligned}$$

Let s and t denote the cardinality of I and J respectively. We then get:

THEOREM 2.2. *Suppose that I, J define a partition as above.*

(i) *If A is normal (or more generally $W(A) = \operatorname{co}\sigma(A)$), then*

$$s(A) \geq \left| \frac{1}{t} \sum_{i,j \in I} a_{ij} - \frac{1}{s} \sum_{i,j \in J} a_{ij} \right|. \tag{2.6}$$

(ii) *If A is Hermitian, then*

$$s(A) \geq \frac{2}{\sqrt{st}} \left| \sum_{\substack{i \in I \\ j \in J}} a_{ij} \right|. \tag{2.7}$$

Proof. Let $u = (1/\sqrt{t})\Sigma_{i \in I} e_i$ and $v = (1/\sqrt{s})\Sigma_{j \in J} e_j$. Then (2.6) follows from (1.2), while (2.7) follows from (1.3). ∎

The above bounds depend on the choice of the partition defined by I and J. In Section 3 we provide simple strategies for choosing a partition.

Now let

$$R_i = \sum_{j=1}^{n} a_{ij}$$

denote the *ith row sum* of A. The following two theorems use the row sums to obtain lower bounds for $s(A)$.

THEOREM 2.3. *Suppose that*

$$R = (R_i)$$

is the vector of row sums of A and that R is real. Let

$$m = \frac{1}{n} \sum_{i=1}^{n} R_i$$

and

$$v^2 = \frac{1}{n} \sum_{i=1}^{n} R_i^2 - m^2 \tag{2.8}$$

be the mean and variance of R respectively. Then

$$s(A) \geq \sqrt{3}\, v \qquad \text{if } A \text{ is normal}, \tag{2.9}$$

$$s(A) \geq 2v \qquad \text{if } A \text{ is Hermitian}. \tag{2.10}$$

Proof. First suppose A is Hermitian. Let $v = e = (1,1,\ldots,1)^t$ and $u = (u_j)$ with

$$u_j = \frac{R_j - m}{\sqrt{n}\, v}. \tag{2.11}$$

Then (1.3) implies that

$$s(A) \geq \frac{2}{\sqrt{n}} u^t A v = 2v.$$

Note that the vector u is the solution of the optimization problem

$$\text{maximize} \quad \{ u^t R : u^t e = 0,\ \|u\| = 1 \}, \tag{2.12}$$

which can be verified using Lagrange multipliers. The normal case follows similarly if we use (1.2) rather than (1.3). ∎

Now let $[\cdot]$ denote greatest integer part. If we order the row sums by magnitude we obtain the following.

THEOREM 2.4. *Suppose that A has real row sums and that the row sums of A are ordered*

$$R_{i_1} \geq R_{i_2} \geq \cdots \geq R_{i_n}.$$

Then

$$s(A) \geq \left\{ \frac{K}{n} \sum_{j=1}^{[(n+1)/2]} \left(R_{i_j} - R_{n-j+1} \right)^2 \right\}^{1/2}, \tag{2.13}$$

$$s(A) \geq \frac{K}{n} \sum_{j=1}^{[(n+1)/2]} \left(R_{i_j} - R_{i_{n-j+1}} \right), \tag{2.14}$$

where $K = \frac{3}{2}$ if A is normal, while $K = 2$ if A is Hermitian.

Proof. Let $r = (r_j)$ with

$$r_j = R_{i_j} - R_{i_{n-j+1}}$$

for $j = 1, \ldots, [(n+1)/2]$, and let $v = e = (1, 1, \ldots, 1)'$ and $u = (u_j)$ with

$$u_{i_j} = \begin{cases} \alpha_j & \text{if} \quad j < \left[\frac{n+1}{2}\right], \\ -\alpha_{n-j+1} & \text{if} \quad j > \left[\frac{n+1}{2}\right], \\ 0 & \text{if} \quad j = \left[\frac{n+1}{2}\right] \end{cases} \tag{2.15}$$

for some $\alpha = (\alpha_j)$ to be determined. Then (1.3) implies that, in the Hermitian case,

$$\begin{aligned} s(A) &\geq \frac{2}{\sqrt{2\sum_j \alpha_j^2 \sqrt{n}}} |u'Av| \\ &= \sqrt{\frac{2}{n}} \frac{1}{\|\alpha\|} |u'(R_i)| \\ &= \sqrt{\frac{2}{n}} \frac{1}{\|\alpha\|} \sum_{j=1}^{[n+1/2]} \alpha_j (R_{i_j} - R_{i_{n-j+1}}) \\ &= \sqrt{\frac{2}{n}} \left(\frac{1}{\|\alpha\|}\alpha, r\right). \end{aligned} \tag{2.16}$$

This is a maximum for $\alpha = r$. This proves (2.13); (2.14) follows by choosing $\alpha_j = 1$ for $j < [(n+1)/2]$ in (2.15) or by the Cauchy-Schwarz inequality. The normal case follows similarly using the inequality in (1.2). ■

Note that (2.10) in Theorem 2.3 resembles (1.12), but (2.10) uses the variance of the row sums whereas (1.12) uses the variance of the eigenvalues themselves, found from $\operatorname{tr} A$ and $\operatorname{tr} A^2$. Theorem 2.3 is proved using an optimization problem. We find the best possible choice in (1.2) and (1.3) for the vector u once the vector $v = e$ is chosen. Thus Theorem 2.3 provides better bounds than Theorem 2.4.

3. COMPARISONS OF BOUNDS

In this section we compare several of the bounds presented above. The comparisons are done using 50 10×10 real, symmetric matrices. The matrices are chosen using a uniform random number generator on $[0,1]$. The test is preformed twice, once for nonnegative (elementwise) matrices and once for general symmetric matrices.

The bounds compared are

$$s_1(A)=\frac{2}{\sqrt{st}}\left|\sum_{\substack{i\in I\\ j\in J}} a_{ij}\right|,$$

where I, J defines the partition [see (2.7)] found by the strategy given below;

$$s_2(A)=2\max_{i\neq j}|a_{ij}|;$$

$$s_3(A)=2s,$$

where $s^2=\operatorname{tr}A^2/n-(\operatorname{tr}A/n)^2$;

$$s_4(A)=\frac{2}{n-1}\left|\sum_{i<j} a_{ij}\right|;$$

$$s_5(A)=\left\{\frac{2}{n}\sum_{j=1}^{[(n+1)/2]}\left(R_{i_j}-R_{i_{n-j+1}}\right)^2\right\}^{1/2};$$

$$s_6(A)=\frac{2}{\sqrt{st}}\left|\sum_{\substack{i\in I\\ j\in J}} a_{ij}\right|,$$

where $s+t=n$ and I, J defines a random partition;

$$s_7(A)=2\nu,$$

where ν is the standard deviation of the row sums; and

$$s_8(A)=\frac{2}{n}\sum_{j=1}^{[(n+1)/2]}\left(R_{i_j}-R_{i_{n-j+1}}\right),$$

where R_{i_j} are the ordered row sums.

The strategy used to obtain the partition for $s_1(A)$, in pseudo programming language, is:

```
SET I = J = ∅, k = 1
WHILE ((k ≤ n(n − 1)/2) AND (I, J ≠ (1,2,..., n))) DO
  a_ij = kth largest off-diagonal element of A
  IF ((i, j) ∉ [(I × I) ∪ (J × J)]) THEN
    IF(Σ_{k∈J} a_ik + Σ_{k∈I} a_kj + a_ij ≥ 0) THEN
      IF (j ∈ I OR i ∈ J) THEN
        SET I = I ∪ {j} AND J = J ∪ {i}
      ELSE
        SET I = I ∪ {i} AND J = J ∪ {j}
      END IF
    END IF
  END IF
END WHILE
STOP
```

We first present the results for nonnegative (elementwise), symmetric matrices. The relative error for the ith bound is

$$\frac{s(A) - s_i(A)}{s(A)}.$$

The means and standard deviations for the relative errors for the eight bounds are:

Mean: 0.15326, 0.69717, 0.46146, 0.23500, 0.75504, 0.23377, 0.74966, 0.79687.

Standard deviation: 0.05521, 0.01992, 0.01199, 0.02854, 0.06194, 0.06580, 0.06347, 0.05553.

Thus $s_1(A)$ appears to be the best bound, followed, by $s_6(A)$ and $s_4(A)$. The random partitions $s_6(A)$ do surprisingly well. The row sums do not do very well. The i, jth position of the following 8×8 matrix gives the number of times $s_i(A)$ was the jth best bound:

$$\begin{pmatrix} 41 & 8 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 39 & 1 & 9 & 1 \\ 0 & 0 & 0 & 50 & 0 & 0 & 0 & 0 \\ 3 & 23 & 24 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 10 & 40 & 0 \\ 6 & 19 & 25 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 11 & 39 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 49 \end{pmatrix}$$

Thus we see that $s_1(A)$ was the best bound 41 out of 50 times, while $s_6(A)$ and $s_4(A)$ again did very well. The i, jth position of the following upper triangular array gives the number of times $s_i(A) > s_j(A)$:

$$\begin{matrix} 0 & 50 & 50 & 46 & 50 & 44 & 50 & 50 \\ & 0 & 0 & 0 & 40 & 0 & 39 & 49 \\ & & 0 & 0 & 50 & 0 & 50 & 50 \\ & & & 0 & 50 & 25 & 50 & 50 \\ & & & & 0 & 0 & 0 & 50 \\ & & & & & 0 & 50 & 50 \\ & & & & & & 0 & 50 \\ & & & & & & & 0 \end{matrix}$$

We now present the above data for general, symmetric matrices.

Mean: 0.67687, 0.66358, 0.36987, 0.87802, 0.45679, 0.86221, 0.44488, 0.5503.

Standard deviation: 0.13219, 0.02669, 0.03295, 0.09104, 0.12825, 0.10370, 0.13184, 0.11397.

$$\begin{pmatrix} 0 & 3 & 3 & 6 & 11 & 22 & 2 & 3 \\ 0 & 1 & 0 & 5 & 24 & 19 & 1 & 0 \\ 36 & 0 & 12 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 2 & 23 & 24 \\ 0 & 14 & 30 & 5 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 3 & 4 & 20 & 22 \\ 14 & 32 & 3 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 31 & 12 & 2 & 3 & 1 \end{pmatrix}$$

$$\begin{matrix} 0 & 22 & 1 & 46 & 5 & 46 & 3 & 13 \\ & 0 & 0 & 48 & 2 & 46 & 2 & 11 \\ & & 0 & 50 & 36 & 50 & 36 & 49 \\ & & & 0 & 1 & 22 & 1 & 1 \\ & & & & 0 & 49 & 0 & 50 \\ & & & & & 0 & 1 & 4 \\ & & & & & & 0 & 50 \\ & & & & & & & 0 \end{matrix}$$

We see that the results are drastically different: $s_3(A)$ and the row sums $s_7(A)$ do quite well, while the sums $s_1(A)$, $s_4(A)$, and $s_6(A)$ no longer do as well.

REFERENCES

1 A. Brauer and A. C. Mewborn, The greatest distance between two characteristic roots of a matrix, *Duke Math. J.* 26:653–661 (1959).

2 R. Grone, C. R. Johnson, E. Marques de Sa, and H. Wolkowicz, Constrained ranges of sesquilinear forms, manuscript.

3 C. R. Johnson, Normality and the numerical range, *Linear Algebra Appl.* 15:89–94 (1976).

4 J. K. Merikoski, On a lower bound for the Perron eigenvalue, *BIT* 19:39–42 (1979).

5 H. Minc and M. Marcus, *A Survey of Matrix Theory and Matrix Inequalities*, Prindle, Weber and Schmidt, 1964.

6 L. Mirsky, Inequalities for normal and Hermitian matrices, *Duke Math. J.* 24:591–599 (1957).

7 L. Mirsky, The spread of a matrix, *Mathematics* 3:127–130 (1956).

8 W. V. Parker, The characteristic roots of matrices, *Duke Math. J.* 12:519–526 (1945).

9 H. Wolkowicz and G. P. H. Styan, Bounds for eigenvalues using traces, *Linear Algebra Appl.* 29:471–506 (1980).

10 H. Wolkowicz and G. P. H. Styan, More bounds for eigenvalues using traces, *Linear Algebra Appl.* 31:1–17 (1980).

Received 6 April 1984; revised 27 September 1984