Improving Eigenvalue Bounds Using Extra Bounds

Jorma Kaarlo Merikoski
Department of Mathematical Sciences
University of Tampere
P.O. Box 607
SF-33101 Tampere 10, Finland

and

Henry Wolkowicz
Department of Mathematics
The University of Alberta
Edmonton, Alberta, Canada T6G 2G1

Submitted by Emeric Deutsch

ABSTRACT

Bounds for various functions of the eigenvalues of a Hermitian matrix A, based on the traces of A and A^2, are improved. A technique is presented whereby these bounds can be improved by combining them with other bounds. In particular, the diagonal of A, in conjunction with majorization, is used to improve the bounds. These bounds all require $O(n^2)$ multiplications.

1. INTRODUCTION

Consider an $n \times n$ complex matrix A with real eigenvalues $\lambda_1 \geqslant \cdots \geqslant \lambda_n$, $n \geqslant 2$. Bounds for various functions of the eigenvalues were given in [7], [8], [9]. These bounds used the two traces, $\operatorname{tr} A$ and $\operatorname{tr} A^2$. The bounds were initially obtained using the Karush-Kuhn-Tucker conditions from optimization and were the best bounds obtainable given only $\operatorname{tr} A$, $\operatorname{tr} A^2$, and n; see [8]. The purpose of this paper is to show how one can use other information from the matrix to improve the bounds. All the proofs are elementary.

If A is Hermitian, then the eigenvalues majorize the diagonal elements. The use of this information can provide a large improvement in the bounds. The bounds referred to above generally became poorer as n became larger. Bringing in the diagonal elements overcomes this problem, in a large number

© Elsevier Science Publishing Co., Inc., 1985
52 Vanderbilt Ave., New York, NY 10017 0024-3795/85/$3.30

of cases. These improved bounds are the best possible bounds given only n, $\operatorname{tr} A$, $\operatorname{tr} A^2$, and the ordered diagonal elements of A. They all require $O(n^2)$ multiplications.

2. PRELIMINARIES

Let

$$m = \frac{\operatorname{tr} A}{n}, \qquad s^2 = \frac{\operatorname{tr} A^2}{n} - m^2.$$

Our bounds actually deal with the ordered vector of real numbers $\lambda = (\lambda_i)$ and the given first two moments

$$\sum_{i=1}^{n} \lambda_i = K, \qquad \sum_{i=1}^{n} \lambda_i^2 = L,$$

i.e., $K = \operatorname{tr} A$ and $L = \operatorname{tr} A^2$ are fixed. Thus we would like to get as much information as possible about the numbers λ_i, given the first two moments K and L. (Equivalently, we would like to get as much information as possible about the eigenvalues λ_i, given $\operatorname{tr} A$ and $\operatorname{tr} A^2$.) Bounds for various functions of λ have been presented in [7], [8], [9]. (Upper and lower bounds for the product $\lambda_1, \lambda_2, \ldots, \lambda_n$ are given in [3].) For example, for $1 \leqslant k \leqslant n$, let

$$\lambda_k^l := m - s\left(\frac{k-1}{n-k+1}\right)^{1/2}, \qquad \lambda_k^u := m + s\left(\frac{n-k}{k}\right)^{1/2},$$

$$\bar{\lambda}_1^l := m + \frac{s}{(n-1)^{1/2}}, \qquad \bar{\lambda}_n^u := m - \frac{s}{(n-1)^{1/2}}. \tag{2.1}$$

Then

$$\lambda_k^l \leqslant \lambda_k \leqslant \lambda_k^u, \tag{2.2a}$$

$$\bar{\lambda}_1^l \leqslant \lambda_1, \tag{2.2b}$$

$$\lambda_n \leqslant \bar{\lambda}_n^u. \tag{2.2c}$$

Equality holds on the left in (2.2a) if and only if

$$\lambda_1 = \cdots = \lambda_{k-1} \left(= \lambda_{k-1}^u\right) \quad \text{and} \quad \lambda_k = \cdots = \lambda_n;$$

on the right in (2.2a) if and only if

$$\lambda_1 = \cdots = \lambda_k \quad \text{and} \quad \lambda_{k+1} = \cdots = \lambda_n \left(= \lambda_{k+1}^l\right);$$

in (2.2b) if and only if

$$\lambda_1 = \cdots = \lambda_{n-1} \quad \left(\text{and} \quad \lambda_n = \lambda_n^l\right);$$

in (2.2c) if and only if

$$\left(\lambda_1 = \lambda_1^u \quad \text{and}\right) \quad \lambda_2 = \cdots = \lambda_n.$$

The following lemmas show some relationships between the numbers λ_i when their first two moments are fixed.

Lemma 2.1. *Given K, L fixed, let m_i elements of λ be equal to λ_i, m_j equal to λ_j, and m_k to λ_k, i.e.*

$$m_i\lambda_i + m_j\lambda_j + m_k\lambda_k = K,$$

$$m_i\lambda_i^2 + m_j\lambda_j^2 + m_k\lambda_k^2 = L.$$

Then

$$\frac{\partial\lambda_i}{\partial\lambda_j} = \frac{m_j(\lambda_j - \lambda_k)}{m_i(\lambda_k - \lambda_i)}.$$

Proof. Differentiating the two equations with respect to λ_j yields

$$m_j + m_i\frac{\partial\lambda_i}{\partial\lambda_j} + m_k\frac{\partial\lambda_k}{\partial\lambda_j} = 0,$$

$$m_j\lambda_j + m_i\lambda_i\frac{\partial\lambda_i}{\partial\lambda_j} + m_k\lambda_k\frac{\partial\lambda_k}{\partial\lambda_j} = 0.$$

These can be solved for the partial derivatives. ∎

We can now state necessary and sufficient conditions for the existence of a perturbation which maintains the first two moments and the ordering of the λ_i's. We call such a perturbation *consistent*. We see that we need only worry about maintaining the ordering among the λ_i's. Note that when λ_i, λ_k, λ_j vary and the other λ_ts are constant, then by Lemma 2.1,

$$\lambda_j > \lambda_k > \lambda_i \quad \text{implies} \quad \frac{\partial \lambda_i}{\partial \lambda_j} > 0 \text{ and } \frac{\partial \lambda_i}{\partial \lambda_k} < 0. \tag{2.3}$$

Let $\lambda_i \uparrow$ denote a positive perturbation of λ_i and $\lambda_i \downarrow$ a negative one. Then (2.3) says that the only consistent perturbations are necessarily alternating:

$$\lambda_i \uparrow, \lambda_k \downarrow, \lambda_j \uparrow \quad \text{or} \quad \lambda_i \downarrow, \lambda_k \uparrow, \lambda_j \downarrow. \tag{2.4}$$

Thus if $\lambda_i \uparrow$ to $\lambda_i + \varepsilon$, with $\varepsilon > 0$, and $m_i = m_j = m_k = 1$, then $\lambda_k \downarrow$ to $\lambda_k - \delta$ and $\lambda_j \uparrow$ to $\lambda_j + \varepsilon - \delta$, where $\delta > 0$ and $\varepsilon - \delta > 0$. Similarly, we can define the perturbations for λ_k and λ_j if $\lambda_i \downarrow$ to $\lambda_i - \varepsilon$. Multiple λ_i, λ_k, λ_j's can be treated analogously. This restriction on the perturbations is, of course, due to the fact that the first two moments are fixed.

LEMMA 2.2. *Given the ordered vector $\lambda = (\lambda_i)$, then one of the perturbations* (2.4) *is consistent if and only if it does not contradict the ordering of the λ_i's.*

Proof. Necessity is clear. Now, suppose that the perturbation does not contradict the ordering. Let $\varepsilon > 0$, $i < k < j$, and

$$\lambda_i \to \lambda_i + \varepsilon, \quad \lambda_k \to \lambda_k - \delta, \quad \lambda_j \to \lambda_j - \varepsilon + \delta, \qquad \delta > \varepsilon.$$

Then we need only satisfy

$$(\lambda_i + \varepsilon)^2 + (\lambda_k - \delta)^2 + (\lambda_j - \varepsilon + \delta)^2 = \lambda_i^2 + \lambda_k^2 + \lambda_j^2.$$

If ε is small enough, we show that this quadratic has a solution $\delta > \varepsilon$. By assumption, the $\lambda_i, \lambda_k, \lambda_j$ must satisfy $\lambda_i \geq \lambda_k > \lambda_j$. Thus we obtain for

$\varepsilon < \lambda_k - \lambda_j$

$$2\delta = \lambda_k - \lambda_j + \varepsilon + \left\{(\lambda_k - \lambda_j + \varepsilon)^2 - 4\varepsilon(\lambda_i - \lambda_j + \varepsilon)\right\}^{1/2} > 2\varepsilon,$$

since the discriminant is positive for small enough ε. The perturbation of the second type in (2.4) can be treated similary. ∎

This lemma will enable us to give very simple proofs for bounds of certain functions $f(\lambda)$. The key is isolating the configuration of the λ_i's at which the particular bound is attained. The proof entails showing that there exists a consistent perturbation which will increase the value of $f(\lambda)$ if this configuration is not chosen.

3. IMPROVING A BOUND

The bounds (2.1) require $O(n^2)$ multiplications. The main work is in calculating the second moment, $\operatorname{tr} A^2 = \sum_{i,j}|a_{ij}|^2$. Now suppose that we have obtained, independently, some other information about the eigenvalues. This new information can sometimes be used to improve our bounds. The strategy involved is that we get an improvement in a bound if the new information contradicts the conditions for attainment of that bound. For example, the upper bound for the kth largest eigenvalue λ_k is $\lambda_k \leqslant \lambda_k^u$ with equality if and only if

$$\lambda_1 = \cdots = \lambda_k$$

and

$$\lambda_{k+1} = \cdots = \lambda_n = \lambda_{k+1}^l \left[= m - s\left(\frac{k}{n-k}\right)^{1/2}\right].$$

Thus, new information such as

$$\lambda_n \leqslant a_n < \lambda_{k+1}^l \left[= m - s\left(\frac{k}{n-k}\right)^{1/2}\right]$$

or

$$\lambda_1 \geqslant a_1 > \lambda_k^u \left[= m + s\left(\frac{n-k}{k}\right)^{1/2}\right],$$

for some given a_1, a_n, does not allow attainment of the bound and can be used to improve the bound. This new information might come from applying the Gershgorin's discs, for example.

Recall that

$$K = \operatorname{tr} A, \qquad L = \operatorname{tr} A^2,$$
$$m = \frac{K}{n}, \qquad s^2 = \frac{L}{n} - m^2. \tag{3.1}$$

Now let

$$a_1 \geqslant \cdots \geqslant a_n \tag{3.2}$$

be given, and for $1 \leqslant t \leqslant n$ define

$$K_t = K - \sum_{i=t+1}^{n} a_i, \qquad L_t = L - \sum_{i=t+1}^{n} a_i^2,$$
$$m_t = \frac{K_t}{t}, \qquad s_t^2 = \frac{L_t}{t} - m_t^2. \tag{3.3}$$

For $-t$, $1 \leqslant t \leqslant n$, define

$$K_{-t} = K - \sum_{i=1}^{n-t} a_i, \qquad L_{-t} = L - \sum_{i=1}^{n-t} a_i^2,$$
$$m_{-t} = \frac{K_{-t}}{t}, \qquad s_{-t}^2 = \frac{L_{-t}}{t} - m_{-t}^2. \tag{3.4}$$

We also define $\lambda_{k,t}^l$ and $\lambda_{k,-t}^l$ as λ_k^l in (2.1) with m, s replaced by m_t, s_t and m_{-t}, s_{-t}, respectively. We similarly redefine λ_k^u, $\bar{\lambda}_1^l$, and $\bar{\lambda}_n^u$.

We now show how to improve the bounds in [9] if we are given other information about the eigenvalues λ_i. First suppose we have an improved upper bound for λ_n. This allows us to simultaneously improve the upper bound of λ_k and lower bound of λ_{k+1}.

THEOREM 3.1. *Suppose that* $1 \leqslant k \leqslant n-2$ *and that* a_n *satisfies*

$$\lambda_n \leqslant a_n \leqslant m - s\left(\frac{k}{n-k}\right)^{1/2} \left(= \lambda_{k+1}^l\right). \tag{3.5}$$

Then

$$\lambda_{k+1}^{l} \leqslant \lambda_{k+1,n-1}^{l} \leqslant \lambda_{k+1} \leqslant \lambda_{k} \leqslant \lambda_{k,n-1}^{u} \leqslant \lambda_{k}^{u}. \tag{3.6}$$

Equality holds in the leftmost and rightmost inequalities in (3.6) *if and only if it holds on the right in* (3.5). *Equality holds in the second inequality in* (3.6) *if and only if it holds in the fourth if and only if*

$$\lambda_n = a_n, \qquad \lambda_{k+1} = \cdots = \lambda_{n-1}, \qquad \lambda_1 = \cdots = \lambda_k. \tag{3.7}$$

Proof. The improved bounds in (3.6) and the conditions for equality are obtained by fixing $\lambda_n = a_n$ and applying the bounds in (2.2a) to the remaining $n-1$ eigenvalues $\lambda_1, \ldots, \lambda_{n-1}$. These now satisfy

$$\sum_{i=1}^{n-1} \lambda_i = K_{n-1}, \qquad \sum_{i=1}^{n-1} \lambda_i^2 = L_{n-1}. \tag{3.8}$$

We need only show that we must set $\lambda_n = a_n$. Suppose not; consider an ordered vector (λ_i) with $\lambda_n < a_n$. Then the conditions for equality in (2.2a) are violated; hence we cannot have both $\lambda_1 = \cdots = \lambda_k$ and $\lambda_{k+1} = \cdots = \lambda_n$. First suppose that the latter does not hold. Then for some i, j $(k+1 \leqslant i < j \leqslant n)$ we have $\lambda_{k+1} = \cdots = \lambda_i > \lambda_{i+1}$, $\lambda_{j-1} > \lambda_j = \cdots = \lambda_n$. Now the perturbation

$$\lambda_k \uparrow, \lambda_i \downarrow, \lambda_j \uparrow \tag{3.9}$$

is consistent. (We perturb each $\lambda_l = \lambda_k \uparrow$, $l \leqslant k$, if these exist). Similarly, if $\lambda_1 = \cdots = \lambda_i > \lambda_{i+1}$, $\lambda_{j-1} \geqslant \lambda_j = \cdots = \lambda_k$, then the perturbation $\lambda_i \downarrow, \lambda_j \uparrow, \lambda_n \downarrow$ is consistent. Since we can do this for each $\lambda_j = \lambda_k$ including λ_k, we can always increase λ_k by a perturbation if $\lambda_n < a_n$. Thus λ_k is maximized if $\lambda_k = a_n$, which proves the fourth inequality in (3.6). The conditions for equality are obtained from the conditions for equality in (2.2a). The second inequality is proved similarly. ∎

Note that in (2.2), the upper bound for λ_{n-1}, the lower bound for λ_n, and the lower bound for λ_1 all hold together, i.e., they hold if and only if $\lambda_1 = \cdots = \lambda_{n-1} = \bar{\lambda}_1^l = \lambda_{n-1}^u$ and $\lambda_n = \lambda_n^l$. Thus (3.5) cannot be used to improve these bounds. This is the reason k is restricted to be $\leqslant n-2$ in the theorem.

An improved lower bound for λ_1 allows similar improvements.

THEOREM 3.2. *Suppose that* $2 \leqslant k \leqslant n-1$ *and that* a_1 *satisfies*

$$\lambda_1 \geqslant a_1 \geqslant m + s\left(\frac{n-k}{k}\right)^{1/2} \left(=\lambda_k^u\right). \tag{3.10}$$

Then

$$\lambda_k^u \geqslant \lambda_{k-1,-(n-1)}^u \geqslant \lambda_k \geqslant \lambda_{k+1} \geqslant \lambda_{k,-(n-1)}^l \geqslant \lambda_{k+1}^l. \tag{3.11}$$

Equality holds in the leftmost and rightmost inequalities in (3.11) *if and only if it holds on the right in* (3.10). *Equality holds in the second inequality in* (3.11) *if and only if it holds in the fourth if and only if*

$$\lambda_1 = a_1, \qquad \lambda_2 = \cdots = \lambda_k, \qquad \lambda_{k+1} = \cdots = \lambda_n. \tag{3.12}$$

Proof. The improved bounds in (3.11) and the conditions for equality are obtained by fixing $\lambda_1 = a_1$, analogous to the proof of Theorem 3.1. We can again use the perturbation technique to show where the bounds are attained and that we must fix $\lambda_1 = a_1$. We then apply the bounds in (2.2a) to the remaining $n-1$ eigenvalues $\lambda_2, \dots, \lambda_n$. Since the first eigenvalue is λ_2 and not λ_1, we get a difference among the indices, e.g. $\lambda_2 \leqslant \lambda_{1,-(n-1)}^u$. ∎

Now suppose that we have improved the upper and/or lower bound for λ_k as above, e.g., given (3.5) and $1 \leqslant k \leqslant n-2$, we obtain (3.6) and (3.7). If we have additional information about λ_{n-1}, e.g.

$$\lambda_{n-1} \leqslant a_{n-1} \leqslant m_{n-1} - s_{n-1}\left(\frac{k}{n-k-1}\right)^{1/2} \left(=\lambda_{k+1,n-1}^l\right), \tag{3.13}$$

then we can perform the same procedure to get, for $1 \leqslant k \leqslant n-3$, that

$$\lambda_k^l \leqslant \lambda_{k+1,n-1}^l \leqslant \lambda_{k+1,n-2}^l \leqslant \lambda_{k+1} \leqslant \lambda_k \leqslant \lambda_{k,n-2}^u \leqslant \lambda_{k,n-1}^u \leqslant \lambda_k^u. \tag{3.14}$$

Equality holds in the leftmost and rightmost inequalities if and only if it holds on the right in (3.5). Equality holds in the second and sixth inequalities if and only if it holds on the right in (3.5) and (3.13). Equality holds in the third and fifth inequalities if and only if

$$\lambda_{n-1} = a_{n-1}, \quad \lambda_n = a_n, \quad \lambda_{k+1} = \cdots = \lambda_{n-2}, \quad \lambda_1 = \cdots \lambda_k. \tag{3.15}$$

We can continue improving the bounds using additional information on λ_{n-3} and/or on λ_2. In fact, the above provides a valid algorithm for improving any of the bounds in [7], [9]. For example, suppose we wish to improve the upper bound for λ_k given some additional information. Then we need only check if the conditions for attainment of the bound are violated. In the next section, we see how to use the diagonal elements to improve the bounds. We complete this section by deriving improvements for $\bar{\lambda}_1^l$ and $\bar{\lambda}_n^u$.

THEOREM 3.3. *Suppose that a_n satisfies*

$$\lambda_n \geq a_n \geq m - s(n-1)^{1/2} \left(= \lambda_n^l \right). \tag{3.16}$$

Then

$$\lambda_1 \geq \bar{\lambda}_{1,n-1}^l \geq \bar{\lambda}_1^l \tag{3.17}$$

and

$$\lambda_{n-1} \leq \bar{\lambda}_{n-1,n-1}^u \leq \lambda_{n-1}^u. \tag{3.18}$$

Equality holds on the right in (3.17) *and* (3.18) *if and only if it holds on the right in* (3.16). *Equality holds on the left in* (3.17) *if and only if*

$$\lambda_n = a_n, \qquad \lambda_1 = \cdots = \lambda_{n-2} \quad \left(\textit{and} \quad \lambda_{n-1} = \lambda_{n-1,n-1}^l\right). \tag{3.19}$$

Equality holds on the left in (3.18) *if and only if*

$$\lambda_n = a_n, \qquad \lambda_2 = \cdots = \lambda_{n-1} \quad \left(\textit{and} \quad \lambda_1 = \lambda_{1,n-1}^u\right). \tag{3.20}$$

Proof. Again we need only show that necessarily $\lambda_n = a_n$. Then, the improvements come from applying (2.2) to the remaining $n-1$ numbers $\lambda_1, \dots, \lambda_{n-1}$, which must now satisfy $\sum_{i=1}^{n-1} \lambda_i = K_{n-1}$ and $\sum_{i=1}^{n-1} \lambda_i^2 = L_{n-1}$. Suppose not, i.e. $\lambda_n > a_n$. From the conditions for equality for the attainment of λ_n^l, we see that $\lambda_1 > \lambda_{n-1}$. Let i be such that $\lambda_1 \geq \lambda_i > \lambda_{i+1} \geq \lambda_{n-1}$. Then the perturbation $\lambda_i \downarrow, \lambda_{i+1} \uparrow, \lambda_n \downarrow$ is consistent. Since we can do this for each j, $1 \leq j \leq i$, we see that we can decrease λ_1. Similarly, since we can do this for each j, $i+1 \leq j \leq n-1$, we see that we can increase λ_{n-1}. Thus in both cases we must have $\lambda_n = a_n$. ■

By applying the above theorem to $-A$ we get equivalent improvements to λ_2^l and $\bar{\lambda}_n^u$.

THEOREM 3.4. *Suppose that a_1 satisfies*

$$\lambda_1 \leq a_1 \leq m + s(n-1)^{1/2} \left(= \lambda_1^u \right). \tag{3.21}$$

Then

$$\lambda_n \leqslant \bar{\lambda}^u_{n-1,\,-(n-1)} \leqslant \bar{\lambda}^u_n \tag{3.22}$$

and

$$\lambda_2 \geqslant \bar{\lambda}^l_{1,\,-(n-1)} \geqslant \lambda^l_2. \tag{3.23}$$

Equality holds on the right (in both) if and only if it holds on the right in (3.21). *Equality holds on the left in* (3.22) *if and only if*

$$\lambda_1 = a_1, \qquad \lambda_3 = \cdots = \lambda_n \quad \left(\textit{and} \quad \lambda_2 = \lambda^u_{1,\,-(n-1)}\right). \tag{3.24}$$

Equality holds on the left in (3.23) *if and only if*

$$\lambda_1 = a_1, \qquad \lambda_2 = \cdots = \lambda_{n-1} \quad \left(\textit{and} \quad \lambda_n = \lambda^l_{n-1,\,-(n-1)}\right). \tag{3.25}$$

We can continue to get further improvements if we have an upper bound for λ_2 and/or a lower bound for λ_3, and so on. We can also improve the bounds for a general λ_k using modified forms of (3.16) and/or (3.21). For example, suppose $2 \leqslant k \leqslant n-1$ and we are given

$$\lambda_n \geqslant a_n \geqslant m - s\left(\frac{k}{n-k}\right)^{1/2} \left(= \lambda^l_{k+1}\right). \tag{3.26}$$

Then in fact $\lambda_{k+1} \geqslant \cdots \geqslant \lambda_n \geqslant a_n$, and the conditions for equality on the right in (2.2a) cannot be attained. Set $a_{k+1} = \cdots = a_n$. Then we get that

$$\lambda_k \leqslant \bar{\lambda}^u_{k,k} \leqslant \lambda^u_k \tag{3.27}$$

with equality on the right if and only if equality holds on the right in (3.26). Equality holds on the left if and only if

$$\lambda_{k+1} = \cdots = \lambda_n = a_n, \qquad \lambda_2 = \cdots = \lambda_k \quad (\text{and} \quad \lambda_1 = \lambda^u_{1,k}). \tag{3.28}$$

4. BEST BOUNDS WITH A FIXED DIAGONAL

The bounds in [7], [9] are the best possible given n, $\operatorname{tr} A$, and $\operatorname{tr} A^2$. We now use the results of the previous section to obtain the best possible bounds given n, $\operatorname{tr} A$, $\operatorname{tr} A^2$, and the diagonal elements $a_{ii} = a_i$, $i = 1, \dots, n$. Without loss of generality, we assume that $a_1 \geqslant \cdots \geqslant a_n$.

Equivalently, we obtain the best bounds for the ordered vector $\lambda = (\lambda_i)$ given n and the first two moments

$$\sum \lambda_i = K, \qquad \sum \lambda_i^2 = L$$

and that the vector $\lambda = (\lambda_i)$ majorizes the vector $a = (a_i)$:

$$\begin{gathered} \lambda_1 \geqslant a_1, \\ \lambda_1 + \lambda_2 \geqslant a_1 + a_2, \\ \vdots \\ \lambda_1 + \cdots + \lambda_{n-1} \geqslant a_1 + \cdots + a_{n-1}, \\ \lambda_1 + \cdots + \lambda_n = a_1 + \cdots + a_n \end{gathered}$$

(see [4], [5]).

The following program, written in a pseudo programming language, calculates the best possible upper and lower bounds for λ_k, $k = 1, \dots, n$. The bounds are the best possible using only $n, \operatorname{tr} A, \operatorname{tr} A^2, a$. The main work is still the calculation of $\operatorname{tr} A^2$, and thus the bounds require $O(n^2)$ multiplications. We improve the bounds (2.2) using the techniques of Theorems 3.1 and 3.2. Note that it might be possible to improve a bound for λ_{k-1} (or λ_{k+1}) when it could not be done for λ_k.

BEGIN PROGRAM:

COMMENT: Let $1 \leqslant k \leqslant n-1$. Calculate the best upper bound for λ_k and lower bound for λ_{k+1} given n, $\operatorname{tr} A$, $\operatorname{tr} A^2$, and the ordered diagonal $a_1 \geqslant \cdots \geqslant a_n$.

INITIALIZATION: Input A, k, and a_i, $i = 1, \dots, n$.
Set

$$K = \sum_{i=1}^{n} a_i, \qquad L = \sum_{i=1}^{n} a_i^2 + 2 \sum_{i<j} |a_{ij}|^2,$$

$$m = \frac{K}{n}, \qquad s^2 = \frac{L}{n} - m^2,$$

$$\lambda_{k+1}^{l} = m - s\left(\frac{k}{n-k}\right)^{1/2}, \qquad \lambda_k^{u} = m + s\left(\frac{n-k}{k}\right)^{1/2}.$$

WRITE: The first upper bound for λ_k and lower bound for λ_{k+1} are respectively: λ_k^u and λ_{k+1}^l.
Set $i=1$, $j=n$, $t=n$, $\bar{k}=k$, $\nu=\lambda_k^u$, $\mu=\lambda_{k+1}^l$.

LOOP: *While* ($\mu > a_j$ or $\nu < a_i$) *do*:
If ($\mu > a_j$) *do*:

$$K=K-a_j, \qquad L=L-a_j^2, \qquad j=j-1.$$

Else do:

$$K=K-a_i, \qquad L=L-a_i^2, \qquad i=i+1, \qquad \bar{k}=\bar{k}-1$$

End if

$$t=t-1, \qquad m=\frac{K}{t}, \qquad s^2=\frac{L}{t}-m^2,$$

$$\mu=m-s\left(\frac{\bar{k}}{t-\bar{k}}\right)^{1/2}, \qquad \nu=m+s\left(\frac{t-\bar{k}}{\bar{k}}\right)^{1/2}.$$

WRITE: The $(n-t)$th improved upper bound for λ_k and lower bound for λ_{k+1} are respectively:

$$\nu \text{ and } \mu.$$

ENDLOOP: *Endwhile*

WRITE: The last bounds are the best possible given n, $\operatorname{tr} A$, $\operatorname{tr} A^2$, and $a_1 \geqslant \cdots \geqslant a_n$. Equality holds if and only if

$$\lambda_1=a_1,\dots, \quad \lambda_{i-1}=a_{i-1}, \quad \lambda_{j+1}=a_{j+1},\dots, \quad \lambda_n=a_n,$$
$$\lambda_i=\cdots=\lambda_k, \qquad \lambda_{k+1}=\cdots=\lambda_j. \tag{4.2}$$

END PROGRAM:

THEOREM 4.1. *Let $1 \leqslant k \leqslant n-1$. The above program finds the best possible upper bound for λ_k and the best lower bound for λ_{k+1} given n, $\operatorname{tr} A$, $\operatorname{tr} A^2$, and the ordered diagonal elements $a_1 \geqslant \cdots \geqslant a_n$.*

Proof. We need only apply Theorems 3.1 and 3.2 and check the conditions for equality to hold at each improvement step of the program. At the first step, the majorization yields that necessarily $a_n \geqslant \lambda_n$. Now if $\mu > a_n$, we cannot attain the bounds and still satisfy $a_n \geqslant \lambda_n$, since $\lambda_n \geqslant \mu$. Thus we fix $\lambda_n = a_n$. The majorization now yields that $a_{n-1} \geqslant \lambda_{n-1}$, and we can continue thus. Similarly, if $a_1 > \nu$, we must fix $a_1 = \lambda_1$. In this case we must change the index $\bar{k}$ to $\bar{k}-1$, since we have left the $n-1$ numbers $\lambda_2 \geqslant \cdots \geqslant \lambda_n$, and so λ_k becomes now the $(k-1)$th ordered number. Also, once we fix $a_1 = \lambda_1$, the majorization implies that $a_2 \leqslant \lambda_2$. The algorithm continues by applying Theorem 3.1 and 3.2 to the remaining t numbers $\lambda_i \geqslant \cdots \geqslant \lambda_k \geqslant \cdots \geqslant \lambda_j$. Note that when the algorithm stops, we have equality if and only if (4.2) holds, and then necessarily this satisfies the majorization. Note also that if $j = \bar{k}+1$ or $i = \bar{k}-1$, then necessarily $\mu \leqslant a_j$ and $\nu \geqslant a_i$, respectively. ∎

5. EXAMPLES

We now illustrate the above algorithm with several examples. We recall that the bounds (2.2) satisfy $\bar{\lambda}_1^l = \lambda_{n-1}^u$, $\lambda_n^u = \lambda_2^l$.

EXAMPLE 5.1. Let

$$A = \begin{bmatrix} 2 & 4 & 1 & 1 \\ 4 & -1 & 2 & -1 \\ 1 & 2 & -3 & 2 \\ 1 & -1 & 2 & 3 \end{bmatrix}.$$

The eigenvalues of A are

$$5.48, 3.44, -2.60, -5.32.$$

The ordered diagonal is

$$(a_i) = (3, 2, -1, -3).$$

Starting with $k = 1$, we find that

$$-2.28 \leqslant \lambda_2 \leqslant \lambda_1 \leqslant 7.84.$$

Since $-2.28 > a_4 = -3$, we can improve these bounds by fixing $\lambda_4 = a_4$.

We then get

$$-1.90 \leqslant \lambda_2 \leqslant \lambda_1 \leqslant 7.80.$$

The conditions for equality ($\lambda_1 = 7.80$, $\lambda_2 = \lambda_3 = -1.90$, and $\lambda_4 = -3$) satisfy the majorization constraint ($7.80 \geqslant a_1 = 3$ and $-1.90 \leqslant a_3 = -1$). Thus we cannot get a further improvement. For $k = 2$, we get

$$-4.13 \leqslant \lambda_3 \leqslant \lambda_2 \leqslant 4.63.$$

Since $4.63 \geqslant a_1 = 3$ and $-4.13 \leqslant a_4 = -3$, we cannot get a further improvement. For $k = 3$, we see that

$$-7.34 \leqslant \lambda_4 \leqslant \lambda_3 \leqslant 2.78.$$

Now $2.78 < a_1 = 3$, so that we can improve these bounds by fixing $\lambda_1 = a_1$. We then get

$$-7.33 \leqslant \lambda_4 \leqslant \lambda_3 = 2.67.$$

No further improvement is possible, since $2.67 \geqslant a_2 = 2$ and $-7.33 \leqslant a_4 = -3$. By the above, $\lambda_1 \geqslant 3$ and $\lambda_n \leqslant -3$ (better than $\lambda_1 \geqslant \lambda_3^u = 2.78$ and $\lambda_4 \leqslant \lambda_2 = -2.28$). In conclusion, we have inclusion regions,

$$3 \leqslant \lambda_1 \leqslant 7.80,$$

$$-1.90 \leqslant \lambda_2 \leqslant 4.63,$$

$$-4.13 \leqslant \lambda_3 \leqslant 2.67,$$

$$-7.33 \leqslant \lambda_4 \leqslant -3.$$

EXAMPLE 5.2. Let

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & -1 \end{bmatrix}.$$

The eigenvalues of A are

$$4.52, 1.39, -1.91.$$

The ordered diagonal is

$$(a_i) = (3, 2, -1).$$

Starting with $k = 1$, we find that

$$-0.52 \leqslant \lambda_2 \leqslant \lambda_1 \leqslant 5.05.$$

Since $-0.52 > -1$, we can improve this by fixing $\lambda_3 = a_3 = -1$. Then

$$0.0 \leqslant \lambda_2 \leqslant \lambda_1 \leqslant 5.0.$$

For $k = 2$,

$$-2.38 \leqslant \lambda_3 \leqslant \lambda_2 \leqslant 3.19.$$

This cannot be improved, since $3.19 \geqslant a_1 = 3$ and $-2.38 \leqslant a_3 = -1$. Because $\lambda_1^{-l} = \lambda_2^{u} = 3.19$ and $\lambda_3 \leqslant -1$ (better than $\lambda_3 \leqslant \lambda_2^{l} = 0$), we finally have

$$\begin{aligned} 3.19 &\leqslant \lambda_1 \leqslant 5.0, \\ 0.0 &\leqslant \lambda_2 \leqslant 3.19, \\ -2.38 &\leqslant \lambda_3 \leqslant 1. \end{aligned}$$

EXAMPLE 5.3. Let

$$A = \begin{bmatrix} 2 & 4 & 1 & 1 & 3 & 1 \\ 4 & -1 & 2 & 3 & 1 & 2 \\ 1 & 2 & -3 & 2 & 1 & -1 \\ 1 & 3 & 2 & 3 & 1 & 1 \\ 3 & 1 & 1 & 1 & 1 & -2 \\ 1 & 2 & -1 & 1 & -2 & -2 \end{bmatrix},$$

which has eigenvalues

$$8.82, 2.74, 1.12, -2.91, -4.16, -5.62,$$

and

$$(a_i) = (3, 2, 1, -1, -2, -3).$$

Starting with $k=1$, we see that

$$-2.19 \leqslant \lambda_2 \leqslant \lambda_1 \leqslant 10.95.$$

Since $-2.19 > a_6 = -3$, we can obtain an improvement by setting $\lambda_6 = a_6$. Then

$$-1.98 \leqslant \lambda_2 \leqslant \lambda_1 \leqslant 10.92.$$

Now $-1.98 > a_5 = -2$, so we can improve again by fixing $\lambda_5 = a_5$. Then

$$-1.97 \leqslant \lambda_2 \leqslant \lambda_1 \leqslant 10.92.$$

This cannot be further improved, since $-1.97 \leqslant a_4 = -1$. For $k=2$,

$$-3.46 \leqslant \lambda_3 \leqslant \lambda_2 \leqslant 6.93;$$

for $k=3$,

$$-4.90 \leqslant \lambda_4 \leqslant \lambda_3 \leqslant 4.90;$$

and for $k=4$,

$$-6.93 \leqslant \lambda_5 \leqslant \lambda_4 \leqslant 3.46;$$

none of which can be improved. For $k=5$,

$$-10.95 \leqslant \lambda_6 \leqslant \lambda_5 \leqslant 2.19.$$

This can be improved, since $2.19 < a_1 = 3$. Fixing $\lambda_1 = a_1$ yields

$$-10.92 \leqslant \lambda_6 \leqslant \lambda_5 \leqslant 1.98.$$

This can be improved again, since $1.98 < a_2 = 2$. Fixing $\lambda_2 = a_2$ gives

$$-10.92 \leqslant \lambda_6 \leqslant \lambda_5 \leqslant 1.97,$$

which cannot be improved. Moreover, $\lambda_1 \geqslant 3$ (better than $\lambda_1 \geqslant \bar{\lambda}_5^{\mu} = 2.19$),

$\lambda_6 \leq -3$ (better than $\bar{\lambda}'_2 = -2.19$). In conclusion,

$$\begin{aligned} 3 &\leq \lambda_1 \leq 10.92, \\ -1.97 &\leq \lambda_2 \leq 6.93, \\ -3.46 &\leq \lambda_3 \leq 4.90, \\ -4.90 &\leq \lambda_4 \leq 3.46, \\ -6.93 &\leq \lambda_5 \leq 1.97, \\ -10.92 &\leq \lambda_6 \leq -3. \end{aligned}$$

Example 5.4. Let

$$A = \begin{bmatrix} 1 & 3 & 2 & 1 & 4 & 2 & 1 \\ 3 & -1 & 2 & -1 & -1 & 2 & 3 \\ 2 & 2 & -4 & -2 & -1 & -1 & 1 \\ 1 & -1 & -2 & 2 & -2 & 1 & 0 \\ 4 & -1 & -1 & -2 & -3 & 0 & 0 \\ 2 & 2 & -1 & 1 & 0 & 3 & 1 \\ 1 & 3 & 1 & 0 & 0 & 1 & 4 \end{bmatrix}.$$

The eigenvalues are

$$8.31, 4.12, 3.06, 0.92, -1.08, -5.26, -8.06.$$

The ordered diagonal is

$$(a_i) = (4,3,2,1,-1,-2,-3,-4)$$

For $k = 1$ we get three improvements. The first estimate and the improvements are

$$-1.84 \leq \lambda_2 \leq \lambda_1 \leq 13.03,$$

-1.37 ....12.83,

-0.93 ....12.71,

-0.90 ....12.71.

For $k = 2$, we have two improvements:

$$-3.00 \leqslant \lambda_3 \leqslant \lambda_4 \leqslant 8.51,$$

$$-2.74 \quad \ldots\ldots\ldots\ldots 8.48,$$

$$-2.65 \quad \ldots\ldots\ldots\ldots 8.48.$$

For $k = 3$, there are no improvements:

$$-4.22 \leqslant \lambda_4 \leqslant \lambda_3 \leqslant 6.29.$$

For $k = 4$, there are no improvements:

$$-5.72 \leqslant \lambda_5 \leqslant \lambda_4 \leqslant 4.79.$$

For $k = 5$, there is one improvement:

$$-7.94 \leqslant \lambda_6 \leqslant \lambda_5 \leqslant 3.58,$$

$$-7.93 \quad \ldots\ldots\ldots\ldots 3.47.$$

For $k = 6$, there are three improvements:

$$-12.46 \leqslant \lambda_7 \leqslant \lambda_6 \leqslant 2.41,$$

$$-12.35 \quad \ldots\ldots\ldots\ldots 2.07,$$

$$-12.31 \quad \ldots\ldots\ldots\ldots 1.83,$$

$$-12.31 \quad \ldots\ldots\ldots\ldots 1.77.$$

In conclusion,

$$\begin{aligned} 2.41 &\leqslant \lambda_1 \leqslant 12.71, \\ -0.90 &\leqslant \lambda_2 \leqslant 8.48, \\ -2.65 &\leqslant \lambda_3 \leqslant 6.29, \\ -4.22 &\leqslant \lambda_4 \leqslant 4.79, \\ -5.72 &\leqslant \lambda_5 \leqslant 3.47, \\ -7.93 &\leqslant \lambda_6 \leqslant -1.77, \\ -12.31 &\leqslant \lambda_7 \leqslant -1.84. \end{aligned}$$

6. USING SOME OTHER EXTRA BOUNDS

A general method to find upper bounds for the spectral radius (i.e., the largest absolute value of the eigenvalues) is to use (submultiplicative) matrix norms. It is well known that for any A,

$$\rho(A) \leqslant \mu(A), \tag{6.1}$$

where ρ is the spectral radius and μ is an arbitrary matrix norm. Better bounds, but more complicated, can be obtained if μ is a suitable *matricial norm* [1]; then

$$\rho(A) \leqslant \rho(\mu(A)). \tag{6.2}$$

If $\lambda_1 = \rho(A)$ (such happens, e.g., if A is (elementwise) nonnegative or symmetric nonnegative definite), we can use these results to overestimate λ_1.

If A is symmetric, every Rayleigh quotient is a well-known lower bound for λ_1. Especially, the bound

$$\frac{1}{n}\sum_i \sum_k a_{ik} \leqslant \lambda_1 \tag{6.3}$$

often seems to be good if A is also nonnegative [6]. Then a still better more complicated bound is

$$\prod_i R_i^{R_i/\sum_k R_k} \leqslant \lambda_1, \tag{6.4}$$

the R_is denoting the column sums, see [2]. Thus we can use (6.3) or (6.4) for symmetric nonnegative matrices.

For example, let

$$A = \begin{pmatrix} 4 & 0 & 2 & 3 \\ 0 & 5 & 0 & 1 \\ 2 & 0 & 6 & 0 \\ 3 & 1 & 0 & 7 \end{pmatrix}$$

with eigenvalues 9.376, 6.423, 4.775, and 1.426. The bounds (2.2) are [9]

$$\begin{aligned} 7.158 &\leqslant \lambda_1 \leqslant 10.475, \\ 3.842 &\leqslant \lambda_2 \leqslant 8.372, \\ 2.628 &\leqslant \lambda_3 \leqslant 7.158, \\ 0.525 &\leqslant \lambda_4 \leqslant 3.842. \end{aligned}$$

The bound (6.3) yields

$$\lambda_1 \geqslant 8.5.$$

Hence, by Theorem 3.2,

$$0.758 \leqslant \lambda_4 \leqslant \lambda_3 \leqslant 6.371,$$

$$2.629 \leqslant \lambda_3 \leqslant \lambda_2 \leqslant 8.242.$$

Slightly better results could be obtained by using (6.4), which gives $\lambda_1 > 8.696$.

To find an extra upper bound for λ_1, we use the matricial norm

$$\mu(A) = \left\| \begin{pmatrix} \|A_{11}\|_2 & \|A_{12}\|_2 \\ \|A_{21}\|_2 & \|A_{22}\|_2 \end{pmatrix} \right\|_2$$

where A is partitioned into 2×2 submatrices and $\|\cdot\|_2$ denotes the largest singular value. Then we obtain

$$\lambda_1 \leqslant 9.835,$$

which is better than the bound obtained using standard easily computable matrix norms ($\|A\|_1 = \|A\|_\infty = 11$, $\|A\|_F = 12.4$). Theorem 3.4 now implies

$$\lambda_2 \geqslant 5.206, \qquad \lambda_4 \leqslant 2.905.$$

In conclusion, we have

$$5.206 \leqslant \lambda_2 \leqslant 8.242,$$

$$2.629 \leqslant \lambda_3 \leqslant 6.371,$$

$$0.758 \leqslant \lambda_4 \leqslant 2.905.$$

We are grateful to Professor George P. H. Styan for valuable discussions.

REFERENCES

1 E. Deutsch, Matrical norms, *Numer. Math.* 16:73–84 (1970).
2 E. Deutsch, Lower bounds for the Perron root of a non-negative irreducible matrix, *Math. Proc. Cambr. Phil. Soc.* 92:49–54 (1982).

3 B. Grone, C. Johnson, E. Marques de Sa, and H. Wolkowicz, Improving Hadamard's inequality. Research Report, The University of Alberta, 1983.
4 A. Horn, Doubly stochastic matrices and the diagonal of a rotation matrix, *Amer. J. Math*. 76:620–630 (1959).
5 A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979.
6 J. K. Merikoski, On a lower bound for the Perron eigenvalue, *BIT* 19:39–42 (1979).
7 J. K. Merikoski, G. P. H. Styan, and H. Wolkowicz, Bounds for ratios of eigenvalues using traces, *Linear Algebra Appl*. 55:105–124 (1983).
8 H. Wolkowicz, Generating eigenvalue bounds using optimization. Research Report, The University of Alberta, 1983.
9 H. Wolkowicz and G. P. H. Styan, Bounds for eigenvalues using traces, *Linear Algebra Appl*. 29:471–506 (1980).

Received 27 September 1983; revised 16 February 1984