Aust. J. Math. Anal. Appl. Vol. **19** (2022), No. 2, Art. 13, 9 pp. AJMAA



BOUNDS FOR THE EXTREMAL EIGENVALUES OF POSITIVE DEFINITE MATRICES

SHIVANI SINGH AND PRAVIN SINGH

Received 28 February, 2022; accepted 4 October, 2022; published 11 November, 2022.

UNISA, DEPARTMENT OF DECISION SCIENCES PO BOX 392, PRETORIA, 0003, SOUTH AFRICA.

University of KwaZulu-Natal, School of Mathematics Statistics and Computer Sciences Private Bag X54001, Durban, 4000, South Africa.

singhs2@unisa.ac.za
singhprook@gmail.com

ABSTRACT. We use a projection to achieve bounds for a vector function of the eigenvalues of a positive definite matrix. For various choices of the monotonic function we are able to obtain bounds for the extremal eigenvalues in terms of the traces of the matrix and its powers. These bounds are relatively simple to compute.

Key words and phrases: Positive definite; Eigenvalues; Bounds.

2010 Mathematics Subject Classification. Primary: 15A18, 15A45, 65F35.

ISSN (electronic): 1449-5910

[©] 2022 Austral Internet Publishing. All rights reserved.

1. Introduction

The eigenvalues λ of a $n \times n$ matrix are difficult to evaluate as solving the n_{th} degree polynomial equation $\det(\lambda \mathbf{I} - \mathbf{A})$ is challenging. As positive definite matrices are Hermitian and have real eigenvalues, their location on the real line is important. In some cases only the extremal eigenvalues are required. For example to solve the linear system Ax = b one could use a matrix splitting approach [1] and hence invoke an iterative scheme of the form $\mathbf{x}_{n+1} = \mathbf{B}\mathbf{x}_n + \mathbf{c}$, where B is related to A and c is related to b. For convergence it is necessary that the spectral radius $\rho(\mathbf{B}) < 1$. In addition the conditioning of a linear system is related to the ratio of the largest to smallest eigenvalues for positive definite systems. In 1946 Brauer [4] proved that $|\lambda| \leq \min\{R,C\}$, where $R = \max_i \sum_{j=1}^n |a_{ij}|$ and $C = \max_j \sum_{i=1}^n |a_{ij}|$. Gerschgorin [9] proved the famous inequality that the eigenvalues are located in the union of the n discs $|z - a_{ii}| \le \sum_{j=1; j \ne i}^{n} |a_{ij}|, i = 1, 2, \cdots, n$ in the complex plane. Brauer in 1958 [3] showed that the ovals of Cassini given by $|z - a_{ii}||z - a_{jj}| \le (\sum_{k=1; k \ne i}^{n} |a_{ik}|)(\sum_{k=1; k \ne j}^{n} |a_{jk}|), i, j = 1, 2, \cdots, n : i \ne j$ are even better than Gerschgorin's theorem in providing inclusion sets for the the spectrum $\sigma(\mathbf{A})$. In 1959 Brauer [2] bounded the spread sp(A) of matrices with real eigenvalues. Rayleigh's theorem [9] may also be used to locate the extremal eigenvalues of real symmetric matrices. Indeed $\lambda_1 = \max_{\|\mathbf{x}\|_2=1} \mathbf{x}^t \mathbf{A} \mathbf{x}$ and $\lambda_n = \min_{\|\mathbf{x}\|_2=1} \mathbf{x}^t \mathbf{A} \mathbf{x}$, where λ_1 and λ_n denote the largest and smallest eigenvalues of A and $\|.\|_2$ denotes the Euclidean norm. In this regard the interlacing property [9] of the eigenvalues of A and its principal sub-matrices would be useful. For positive definite Toeplitz matrices Dembo [8] in 1988 provided useful bounds for the extremal eigenvalues. Recently an interval containing the eigenvalues of real symmetric matrices was provided by Huang and Xu [7] using the trace(A) and trace(A²).

2. THEORY

Lemma 2.1. Define $P \in \mathbb{R}^{n \times n}$ by $P = I - \frac{e e^t}{n}$, where $e \in \mathbb{R}^n$ is the vector with elements all unity. Then the following is true

- (1) P is idempotent
- (2) $rank(\mathbf{P}) = n 1$
- (3) a basis for the nullspace $N(\mathbf{P}) = \{\mathbf{e}\}$
- (4) $\mathbf{R}^n = R(\mathbf{P}) \oplus N(\mathbf{P})$ is an orthogonal decomposition of \mathbf{R}^n

Proof. (1) By direct calculation it follows that $P = P^2$.

(2)

$$rank(\mathbf{I}) = rank\left(\mathbf{I} - \frac{\mathbf{e}\,\mathbf{e}^t}{n} + \frac{\mathbf{e}\,\mathbf{e}^t}{n}\right)$$
$$\leq rank(\mathbf{P}) + rank\left(\frac{\mathbf{e}\,\mathbf{e}^t}{n}\right)$$
$$= rank(\mathbf{P}) + 1$$

Hence $rank(\mathbf{P}) \ge n-1$, but as \mathbf{P} is rank deficient since it is a projector, it follows that $rank(\mathbf{P}) = n-1$.

- (3) It follows from (2) that $N(\mathbf{P})$ is one dimensional and since $\mathbf{Pe} = \mathbf{0}$, $\{\mathbf{e}\}$ can be taken as a basis for $N(\mathbf{P})$.
- (4) It follows from the elementary theory of projectors that $\mathbf{R}^n = R(\mathbf{P}) \oplus N(\mathbf{P})$. If $\mathbf{x} \in R(\mathbf{P})$ and $\mathbf{y} \in N(\mathbf{P})$ then $\mathbf{x} = \mathbf{P}\mathbf{z}$ for some $\mathbf{z} \in \mathbf{R}^n$ and

$$\langle \mathbf{x}, \, \mathbf{y} \rangle = \langle \mathbf{Pz}, \, \mathbf{y} \rangle$$

= $\langle \mathbf{z}, \, \mathbf{Py} \rangle$
= 0.

Here $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^t \mathbf{x}$ denotes the standard inner product on \mathbf{R}^n .

Lemma 2.2. Let $\lambda = (\lambda_i) \in \mathbf{R}^n$ be the vector of eigenvalues of a positive definite matrix $\mathbf{A} \in \mathbf{R}^{n \times n}$ and $f : (0, \infty) \longrightarrow (0, \infty)$ be an increasing function. Order the eigenvalues such that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$$
.

Define $\mathbf{f}(\lambda) = [f(\lambda_1), f(\lambda_2), \dots, f(\lambda_n)]^t$. Choose $\mathbf{v} \notin N(\mathbf{P})$, define

$$m = \frac{\langle \mathbf{f}(\boldsymbol{\lambda}), \, \mathbf{e} \rangle}{n}$$

and

$$S^2 = \frac{\langle \mathbf{Pf}(\lambda), \, \mathbf{f}(\lambda) \rangle}{n}.$$

Then

(2.1)
$$|\langle \mathbf{f}(\lambda) - m\mathbf{e}, \mathbf{v} \rangle| \le S\sqrt{n\langle \mathbf{P}\mathbf{v}, \mathbf{v} \rangle}.$$

Proof. From the definition of **P** it follows that $\mathbf{Pf}(\lambda) = \mathbf{f}(\lambda) - m\mathbf{e}$. Also

$$\begin{aligned} |\langle \mathbf{Pf}(\boldsymbol{\lambda}), \, \mathbf{v} \rangle| &= |\langle \mathbf{P^2f}(\boldsymbol{\lambda}), \, \mathbf{v} \rangle| \\ &= |\langle \mathbf{Pf}(\boldsymbol{\lambda}), \, \mathbf{Pv} \rangle| \\ &\leq \sqrt{\langle \mathbf{Pf}(\boldsymbol{\lambda}), \, \mathbf{Pf}(\boldsymbol{\lambda}) \rangle} \sqrt{\langle \mathbf{Pv}, \, \mathbf{Pv} \rangle} \quad \text{(by Cauchy Schwarz)} \\ &= \sqrt{\langle \mathbf{Pf}(\boldsymbol{\lambda}), \, \mathbf{f}(\boldsymbol{\lambda}) \rangle} \sqrt{\langle \mathbf{Pv}, \, \mathbf{v} \rangle}. \end{aligned}$$

Hence

$$|\langle \mathbf{f}(\boldsymbol{\lambda}) - m\mathbf{e}, \mathbf{v} \rangle| = |\langle \mathbf{Pf}(\boldsymbol{\lambda}), \mathbf{v} \rangle|$$

$$\leq \sqrt{\langle \mathbf{Pf}(\boldsymbol{\lambda}), \mathbf{f}(\boldsymbol{\lambda}) \rangle} \sqrt{\langle \mathbf{Pv}, \mathbf{v} \rangle}$$

$$= S\sqrt{n\langle \mathbf{Pv}, \mathbf{v} \rangle}$$

Theorem 2.3. Under the conditions of 2.1 and 2.2 we obtain bounds for $f(\lambda_i)$ given by

$$m - S\sqrt{n-1} \le f(\lambda_j) \le m + S\sqrt{n-1}.$$

Proof.

It suffices to choose $\mathbf{v} = \mathbf{e}_j$ (the standard basis vector in \mathbf{R}^n with unity in the j_{th} position) and substitute into (2.1).

Theorem 2.4. A lower bound for $f(\lambda_1)$ of the form

$$m + \frac{S}{\sqrt{n-1}} \le f(\lambda_1)$$

is satisfied.

Proof.

Consider

$$\langle f(\lambda_{1})\mathbf{e} - \mathbf{f}(\boldsymbol{\lambda}), f(\lambda_{1})\mathbf{e} - \mathbf{f}(\boldsymbol{\lambda}) \rangle$$

$$= \sum_{i=1}^{n} [f(\lambda_{1}) - f(\lambda_{i})]^{2}$$

$$\leq \sum_{i=1}^{n} [f(\lambda_{1}) - f(\lambda_{i})]^{2} + \sum_{i \neq j} [f(\lambda_{1}) - f(\lambda_{i})][f(\lambda_{1}) - f(\lambda_{j})]$$

$$= \left(\sum_{i=1}^{n} [f(\lambda_{1}) - f(\lambda_{i})]\right)^{2}$$

$$= \left(nf(\lambda_{1}) - \sum_{i=1}^{n} f(\lambda_{i})\right)^{2}$$

$$= n^{2} \left(f(\lambda_{1} - \sum_{i=1}^{n} \frac{f(\lambda_{i})}{n}\right)^{2}$$

$$= n^{2} \left(f(\lambda_{1}) - \frac{\langle \mathbf{f}(\boldsymbol{\lambda}), \mathbf{e} \rangle}{n}\right)^{2}$$

$$= n^{2} (f(\lambda_{1}) - m)^{2}$$

$$(2.3)$$

We also have that

$$\langle f(\lambda_{1})\mathbf{e} - \mathbf{f}(\boldsymbol{\lambda}), f(\lambda_{1})\mathbf{e} - \mathbf{f}(\boldsymbol{\lambda}) \rangle$$

$$= \langle f(\lambda_{1})\mathbf{e} - m\mathbf{e} + m\mathbf{e} - \mathbf{f}(\boldsymbol{\lambda}), f(\lambda_{1})\mathbf{e} - m\mathbf{e} + m\mathbf{e} - \mathbf{f}(\boldsymbol{\lambda}) \rangle$$

$$= (f(\lambda_{1}) - m)^{2} \langle \mathbf{e}, \mathbf{e} \rangle + \langle \mathbf{f}(\boldsymbol{\lambda}) - m\mathbf{e}, \mathbf{f}(\boldsymbol{\lambda}) - m\mathbf{e} \rangle$$

$$+ 2(m - f(\lambda_{1})) \langle \mathbf{f}(\boldsymbol{\lambda}) - m\mathbf{e}, \mathbf{e} \rangle$$

$$= n(f(\lambda_{1}) - m)^{2} + \langle \mathbf{Pf}(\boldsymbol{\lambda}), \mathbf{Pf}(\boldsymbol{\lambda}) \rangle + 2(m - f(\lambda_{1})) \langle \mathbf{Pf}(\boldsymbol{\lambda}), \mathbf{e} \rangle$$

$$= n(f(\lambda_{1}) - m)^{2} + \langle \mathbf{Pf}(\boldsymbol{\lambda}), \mathbf{f}(\boldsymbol{\lambda}) \rangle$$

$$(2.4)$$

Where the last term is zero due to orthogonality. Hence, we have from (2.3) and (2.4)

$$n(f(\lambda_1) - m)^2 + s^2 n \le n^2 (f(\lambda_1) - m)^2$$

from which the result follows.

Theorem 2.5. An upper bound for $f(\lambda_n)$ of the form

$$f(\lambda_n) \le m - \frac{S}{\sqrt{n-1}}$$

is satisfied.

Proof.

Consider

$$\langle \mathbf{f}(\boldsymbol{\lambda}) - f(\lambda_n)\mathbf{e}, \, \mathbf{f}(\boldsymbol{\lambda}) - f(\lambda_n)\mathbf{e} \rangle$$

$$= \sum_{i=1}^{n} [f(\lambda_i) - f(\lambda_n)]^2$$

$$\leq \sum_{i=1}^{n} [f(\lambda_i) - f(\lambda_n)]^2 + \sum_{\substack{i,j=1\\i\neq j}}^{n} [f(\lambda_i) - f(\lambda_n)][f(\lambda_j) - f(\lambda_n)]$$

$$= \left(\sum_{i=1}^{n} [f(\lambda_i) - f(\lambda_n)]\right)^2$$

$$= \left(\sum_{i=1}^{n} [f(\lambda_i) - nf(\lambda_n)]\right)^2$$

$$= n^2 \left(\frac{\sum_{i=1}^{n} f(\lambda_i)}{n} - f(\lambda_n)\right)^2$$

$$= n^2 \left(\frac{\langle \mathbf{f}(\boldsymbol{\lambda}), \, \mathbf{e} \rangle}{n} - f(\lambda_n)\right)^2$$

$$= n^2 (m - f(\lambda_n))^2$$

$$(2.6)$$

We also have that

$$\langle \mathbf{f}(\boldsymbol{\lambda}) - f(\lambda_n)\mathbf{e}, \, \mathbf{f}(\boldsymbol{\lambda}) - f(\lambda_n)\mathbf{e} \rangle$$

$$= \langle \mathbf{f}(\boldsymbol{\lambda}) - m\mathbf{e} + m\mathbf{e} - f(\lambda_n)\mathbf{e}, \, \mathbf{f}(\boldsymbol{\lambda}) - m\mathbf{e} + m\mathbf{e} - f(\lambda_n)\mathbf{e} \rangle$$

$$= (m - f(\lambda_n))^2 \langle \mathbf{e}, \, \mathbf{e} \rangle + \langle \mathbf{f}(\boldsymbol{\lambda}) - m\mathbf{e}, \, \mathbf{f}(\boldsymbol{\lambda}) - m\mathbf{e} \rangle$$

$$+ 2(m - f(\lambda_n)) \langle \mathbf{f}(\boldsymbol{\lambda}) - m\mathbf{e}, \, \mathbf{e} \rangle$$

$$= n(m - f(\lambda_n))^2 + \langle \mathbf{Pf}(\boldsymbol{\lambda}), \, \mathbf{Pf}(\boldsymbol{\lambda}) \rangle + 2(m - f(\lambda_n)) \langle \mathbf{Pf}(\boldsymbol{\lambda}), \, \mathbf{e} \rangle$$

$$= n(m - f(\lambda_n))^2 + \langle \mathbf{Pf}(\boldsymbol{\lambda})\mathbf{f}(\boldsymbol{\lambda}) \rangle$$
(2.7)

Where the last term is zero due to orthogonality. Hence we have from (2.6) and (2.7)

$$nS^{2} + n(m - f(\lambda_{n}))^{2} \le n^{2}(m - f(\lambda_{n}))^{2}$$

from which the result follows.

Theorem 2.6. Under the conditions of 2.1 and 2.2, we have the optimal bounds for $f(\lambda_1)$ and $f(\lambda_n)$ given by

$$m + \frac{S}{\sqrt{n-1}} \le f(\lambda_1) \le m + S\sqrt{n-1}$$
$$\max \{0, m - S\sqrt{n-1}\} \le f(\lambda_n) \le m - \frac{S}{\sqrt{n-1}}$$

Proof.

Choose j=1 and j=n in Theorem 2.3 and compare with the bounds from Theorem 2.5 and Theorem 2.6. Use the fact that $f(\lambda_n)>0$ and that $m-\frac{S}{\sqrt{n-1}}$ could be negative.

Lemma 2.7. Consider

$$(2.8) f(\lambda_n) \le m - \frac{S}{\sqrt{n-1}} \le m + \frac{S}{\sqrt{n-1}} \le f(\lambda_1).$$

Equality holds on the left $\iff \lambda_2 = \lambda_3 = \cdots = \lambda_n$, on the right $\iff \lambda_1 = \lambda_2 = \cdots = \lambda_{n-1}$ and in the centre $\iff \lambda_1 = \lambda_2 = \cdots = \lambda_n$

Proof. Equality holds on the left \iff equality holds in (2.5), hence

$$\sum_{\substack{i,j=1\\i\neq j}}^{n} [f(\lambda_i) - f(\lambda_n)][f(\lambda_j) - f(\lambda_n)] = 0.$$

Since this is the sum of positive terms we have $f(\lambda_2) = f(\lambda_3) = \cdots = f(\lambda_n) \iff \lambda_2 = \lambda_3 = \cdots = \lambda_n$. Equality holds on the right \iff equality holds in (2.2), hence

$$\sum_{\substack{i,j=1\\i\neq j}}^{n} [f(\lambda_1) - f(\lambda_i)][f(\lambda_1) - f(\lambda_j)] = 0.$$

Since this is the sum of positive terms we have $f(\lambda_1) = f(\lambda_2) = \cdots = f(\lambda_{n-1}) \iff \lambda_1 = \lambda_2 = \cdots = \lambda_{n-1}$. Equality holds in the centre $\iff S = 0$. Hence $\langle \mathbf{Pf}(\lambda), \mathbf{f}(\lambda) \rangle = 0$, which implies that $\mathbf{f}(\lambda) \in N(\mathbf{P})$, so $\mathbf{f}(\lambda) = c\mathbf{e}$ for some constant c. This implies that $f(\lambda_1) = f(\lambda_2) = \cdots = f(\lambda_n)$ and hence $\lambda_1 = \lambda_2 = \cdots = \lambda_n$.

Theorem 2.8. Consider

$$(2.9) m - S\sqrt{n-1} \le f(\lambda_n) \le m - \frac{S}{\sqrt{n-1}}$$

(2.10)
$$m + \frac{S}{\sqrt{n-1}} \le f(\lambda_1) \le m + S\sqrt{n-1}$$

(2.11)

Equality holds on the left of $(2.9) \iff$ equality holds on the left of $(2.10) \iff$ the n-1 largest eigenvalues are equal. Equality holds on the right of $(2.9) \iff$ equality holds on the right of $(2.10) \iff$ the n-1 smallest eigenvalues are equal.

Proof. Equality holds on the left of (2.9)

$$\iff f(\lambda_n) = m - S\sqrt{n-1}$$

$$\iff S\sqrt{n-1} = m - f(\lambda_n)$$

$$\iff \frac{S}{\sqrt{n-1}} = \frac{m - f(\lambda_n)}{n-1}$$

$$\iff m + \frac{S}{\sqrt{n-1}} = m + \frac{m - f(\lambda_n)}{n-1}$$

$$= \frac{mn - f(\lambda_n)}{n-1}$$

$$= \frac{f(\lambda_1) + f(\lambda_2) + \dots + f(\lambda_{n-1})}{n-1}$$

$$\geq f(\lambda_1).$$

Hence equality holds on the left of (2.10) and by Lemma 2.7 the largest n-1 eigenvalues are equal. Equality holds on the right of (2.9)

$$\iff f(\lambda_n) = m - \frac{S}{\sqrt{n-1}}$$

$$\iff m - f(\lambda_n) = \frac{S}{\sqrt{n-1}}$$

$$\iff (n-1)(m-f(\lambda_n)) = S\sqrt{n-1}$$

$$\iff mn - (n-1)f(\lambda_n) = m + S\sqrt{n-1}$$

$$\iff f(\lambda_1) + f(\lambda_2) + \dots + f(\lambda_n) - (n-1)f(\lambda_n) = m + S\sqrt{n-1}$$

$$\iff f(\lambda_1) = m + S\sqrt{n-1} \quad (\lambda_2 = \lambda_3 = \dots = \lambda_n \text{ by Lemma 2.7})$$

3. RESULTS

Consider $f(x) = x^k, \ k \in \mathbb{N}$, the set on natural numbers, then

$$m = \frac{\operatorname{trace}(\mathbf{A}^k)}{n}$$

$$S^2 = \frac{\operatorname{trace}(\mathbf{A}^{2k})}{n} - m^2$$

$$= \frac{||\mathbf{A}^k||_F^2}{n} - m^2,$$

k	λ_1	λ_n
1	[6.52673, 10.5622]	(0, 2.47927]
2	[7.45819, 10.0933]	(0, 3.06192]
3	[8.06212, 10.0168]	(0, 3.50284]

Table 3.1: Bounds

and we obtain the bounds

$$\sqrt[k]{m + \frac{S}{\sqrt{n-1}}} \le \lambda_1 \le \sqrt[k]{m + S\sqrt{n-1}}$$

$$\lambda_n \le \sqrt[k]{m - \frac{S}{\sqrt{n-1}}}$$

$$\lambda_n \ge \begin{cases} 0 & \text{if } m - S\sqrt{n-1} < 0 \\ \sqrt[k]{m - S\sqrt{n-1}} & \text{otherwise} \end{cases}$$

When k = 1 we obtain the bounds of [5]. Consider the test matrix [6]

$$\mathbf{A} = \left[\begin{array}{cccc} 5 & 4 & 1 & 1 \\ 4 & 5 & 1 & 1 \\ 1 & 1 & 4 & 2 \\ 1 & 1 & 2 & 4 \end{array} \right]$$

with spectrum $\sigma(\mathbf{A}) = \{1, 2, 5, 10\}$. We summarize results for k = 1, 2, 3 in Table 1. For this example it is clear that the upper bounds get better as k increases. However it may not be prudent to use large k as the computation of powers of \mathbf{A} may be too expensive. However for sparse matrices the bounds that we have provided could be useful. Also the usage of non polynomial functions is prohibited due to the complexity of evaluating m and S.

4. CONCLUSION

We have provided useful bounds for the extremal eigenvalues of positive definite matrices. These bounds are a useful addition to the arsenal of tools already available to locate the spectrum.

REFERENCES

- [1] VARGA, RICHARD S, Matrix Iterative Analysis, New Jersey: Prentice-Hall, (1962).
- [2] A. BRAUER and A. C. MEWBOM, The greatest distance between two characteristic roots of a matrix, *Duke Mathematical Journal* **26(4)** (1959), pp. 653–661,.
- [3] A. BRAUER, Limits for the characteristic roots of a matrix VII, *Duke Math. J.* **25** (1958), pp. 583–590
- [4] M. MARCUS and H. MINT, A Survey of Matrix Theory and Matrix Inequalities, Prindle, Weber and Schmidt, Boston (1964).
- [5] H. WOLKOWICZ, G.P.H. STYAN, Bounds for eigenvalues using traces, *Linear Algebra Appl.* **29** (1980), pp. 471–506.
- [6] R.T. GREGORY and D.L. KARNEY, A Collection of Matrices for Testing Computational Algorithms, New York, Robert E. Krieger Publishing Company, (1978).

- [7] TING-ZHU HUANG and CHANG-XIAN XU, Bounds for the extreme eigenvalues of symmetric matrices, *ZAMM* **83(3)** (2003), pp. 214–216.
- [8] DEMBO, A, Bounds on the extreme eigenvalues of positive definite Toeplitz matrices, *IEEE Trans. Inform. Theory*, MR 89b:15028 **34** (1988), pp. 352–355.
- [9] HORN, R. and JOHNSON, Matrix Analysis, Cambridge: Cambridge University Press, (2012).