Aust. J. Math. Anal. Appl.
Vol. **21** (2024), No. 1, Art. 1, 9 pp.
AJMAA

EIGENVALUE BOUNDS BASED ON PROJECTIONS

PRAVIN SINGH, SHIVANI SINGH, VIRATH SINGH

Received 9 February, 2023; accepted 15 September, 2023; published 16 February, 2024.

UNIVERSITY OF KWAZULU-NATAL, PRIVATE BAG X54001, DURBAN, 4001, SOUTH AFRICA.
UNIVERSITY OF SOUTH AFRICA, DEPARTMENT OF DECISION SCIENCES, PO BOX 392, PRETORIA,0003, SOUTH AFRICA.
UNIVERSITY OF KWAZULU-NATAL, PRIVATE BAG X54001, DURBAN, 4001, SOUTH AFRICA.
singhp@ukzn.ac.za, singhs2@unisa.ac.za, singhv@ukzn.ac.za

ABSTRACT. In this paper, we derive expressions for the bounds of the extremal eigenvalues of positive definite matrices. Our approach is to use a symmetric projection operator onto an n-2 dimensional subspace of the real space of n tuples. These bounds are based on traces of the matrix and its powers. They are relatively easy and inexpensive to compute.

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

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

ISSN (electronic): 1449-5910

1. INTRODUCTION

Let $\mathbf{A} \in \mathbf{R}^{n\times n}$ be a positive definite matrix, $\mathbf{x}, \mathbf{b} \in \mathbf{R}^n$ and consider the linear system $\mathbf{Ax} = \mathbf{b}$. The condition number cond$(\mathbf{A})$, the spectral radius $\rho(\mathbf{A})$ and the spread sp$(\mathbf{A})$ are important to the qualitative analysis of the behaviour of techniques aimed at solving the system. Accurate bounds on the spectrum $\sigma(\mathbf{A})$ are vital to the polynomial approximations of functions of $\mathbf{A}$, for example the inverse $\mathbf{A}^{-1}$ [7], using the spectral mapping theorem [3]. In almost all spheres of engineering and science, knowledge of these eigenvalues are crucial. While there exists numerical techniques of approximating these extremal values, usually an initial approximation is required. Eigenvalue location by Gerschgorin disks, ovals of Cassini [2], Rayleigh quotient, power method [3], bounds based on traces [8],[6], are some tools that are inexpensive, yet effective. In addition for symmetric tridiagonal matrices several bounds have been advocated [4]. Here we also improve the bounds of [8]. The bounds that we derive here are applicable to symmetric matrices (for a special class of functions), however we choose to concentrate on positive definite matrices (using a larger class of functions), as we shall describe below.

2. THEORY

Lemma 2.1. *Define* $\mathbf{P} \in \mathbf{R}^{n\times n}$ *by*

$$\mathbf{P} = \mathbf{I} - \mathbf{e_j}\mathbf{e_j^t} - \frac{(\mathbf{e} - \mathbf{e_j})(\mathbf{e} - \mathbf{e_j})^t}{n-1}, \tag{2.1}$$

where $\mathbf{e} \in \mathbf{R}^n$ *is the vector with all elements unity and* $\mathbf{e_j}$ *is the standard unit vector in* $\mathbf{R}^n$ *with unity in the* j^{th} *position. Then the following is true:*

(1) $\mathbf{P}$ *is idempotent and symmetric*

(2) *rank* $(\mathbf{P}) = n - 2$

(3) *an orthonormal basis for the nullspace* $N(\mathbf{P}) = \left\{\mathbf{e_j}, \frac{\mathbf{e} - \mathbf{e_j}}{\sqrt{n-1}}\right\}$

(4) $\mathbf{R}^n = R(\mathbf{P}) \oplus N(\mathbf{P})$ *is an orthogonal decomposition of* $\mathbf{R}^n$, *where* $R(\mathbf{P})$ *denotes the range of* $\mathbf{P}$.

Proof.

(1) By direct calculation it follows that $\mathbf{P} = \mathbf{P}^2$ and $\mathbf{P} = \mathbf{P}^t$.

(2)

$$\begin{aligned}\text{rank}\,(\mathbf{I}) &= \text{rank}\,\left(\mathbf{P} + \mathbf{e_j}\mathbf{e_j^t} + \frac{(\mathbf{e} - \mathbf{e_j})(\mathbf{e} - \mathbf{e_j})^t}{n-1}\right) \\ &\leq \text{rank}\,(\mathbf{P}) + \text{rank}\,(\mathbf{e_j}\mathbf{e_j^t}) + \text{rank}\,\left(\frac{(\mathbf{e} - \mathbf{e_j})(\mathbf{e} - \mathbf{e_j})^t}{n-1}\right) \\ &= \text{rank}\,(\mathbf{P}) + 2\end{aligned}$$

Hence rank $(\mathbf{P}) \geq n-2$. As $\mathbf{P}\mathbf{e_j} = \mathbf{P}(\mathbf{e}-\mathbf{e_j}) = \mathbf{0}$ it follows that rank $(\mathbf{P}) = n-2$.

(3) This follows from (2).

(4) It follows from the elementary theory of projections that $\mathbf{R}^n = R(\mathbf{P}) \oplus N(\mathbf{P})$. This is an orthogonal decomposition as $\langle \mathbf{Pz}, \mathbf{y}\rangle = \langle \mathbf{z}, \mathbf{Py}\rangle = \mathbf{0}$, where $\mathbf{z} \in R(\mathbf{P})$ and $\mathbf{y} \in N(\mathbf{P})$. Here $\langle ., . \rangle$ denotes the standard inner product on $\mathbf{R}^n$.

∎

Definition 2.1. Let $\boldsymbol{\lambda} = (\lambda_i) \in \mathbf{R}^n$ 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 $\mathbf{f}(\boldsymbol{\lambda}) = [f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)]^t$.

Lemma 2.2. *Let*

$$m = \frac{\langle f(\boldsymbol{\lambda}), \mathbf{e}\rangle}{n} = \frac{trace\,(f(\mathbf{A}))}{n} \tag{2.2}$$

$$\mathbf{B} = f(\mathbf{A}) - m\mathbf{I} \tag{2.3}$$

$$S^2 = \frac{1}{n}\langle f(\boldsymbol{\lambda}) - m\mathbf{e}, f(\boldsymbol{\lambda}) - m\mathbf{e}\rangle \tag{2.4}$$

$$= \frac{trace\,(f(\mathbf{A})^2)}{n} - m^2 \tag{2.5}$$

$$= \frac{trace\,(\mathbf{B}^2)}{n} \tag{2.6}$$

then

$$|f(\lambda_j) - m)| \leq S\sqrt{n-1} \tag{2.7}$$

Proof. Write $f(\boldsymbol{\lambda})$ in terms of its orthogonal components as

$$f(\boldsymbol{\lambda}) = \mathbf{P}f(\boldsymbol{\lambda}) + \langle f(\boldsymbol{\lambda}), \mathbf{e_j}\rangle\mathbf{e_j} + \frac{\langle f(\boldsymbol{\lambda}), \mathbf{e}-\mathbf{e_j}\rangle}{\sqrt{n-1}}\frac{(\mathbf{e}-\mathbf{e_j})}{\sqrt{n-1}} \tag{2.8}$$

then it follows from the Pythagorean theorem [5] in an inner product space that

$$||f(\boldsymbol{\lambda})||^2 = \langle f(\boldsymbol{\lambda}),\, f(\boldsymbol{\lambda})\rangle \geq \langle f(\boldsymbol{\lambda}),\, \mathbf{e_j}\rangle^2 + \left(\frac{\langle f(\boldsymbol{\lambda}),\, \mathbf{e}-\mathbf{e_j}\rangle}{\sqrt{n-1}}\right)^2$$

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

$$= f(\lambda_j)^2 + \frac{[mn - f(\lambda_j)]^2}{n-1}$$

$$= \frac{n}{n-1} f(\lambda_j)^2 - \frac{2mn}{n-1} f(\lambda_j) + \frac{m^2n^2}{n-1}$$

Hence

$$f(\lambda_j)^2 - 2mf(\lambda_j) + m^2 + m^2(n-1) \leq \frac{n-1}{n}\langle f(\boldsymbol{\lambda}),\, f(\boldsymbol{\lambda})\rangle$$

$$(f(\lambda_j) - m)^2 \leq (n-1)\left(\frac{\langle f(\boldsymbol{\lambda}),\, f(\boldsymbol{\lambda})\rangle}{n} - m^2\right)$$

$$= (n-1)\left(\frac{trace\,(f(\mathbf{A})^2)}{n} - m^2\right)$$

$$= (n-1)S^2.$$

The result follows by taking the square root.
∎

Theorem 2.3. *Upper and lower bounds for $f(\lambda_1)$ and $f(\lambda_n)$ are given by*

$$f(\lambda_1) \leq m + S\sqrt{n-1} \tag{2.9}$$

$$f(\lambda_n) \geq m - S\sqrt{n-1} \tag{2.10}$$

Proof. Let $j = 1$ and $j = n$ in 2.2. ∎

Theorem 2.4. *Lower and upper bounds for $f(\lambda_1)$ and $f(\lambda_n)$ are given by*

$$f(\lambda_1) \geq m + \frac{S}{\sqrt{n-1}} \tag{2.11}$$

$$f(\lambda_n) \leq m - \frac{S}{\sqrt{n-1}} \tag{2.12}$$

Proof. We use the fact that for real numbers $f(\lambda_i),\ i = 1,\ 2,\ \cdots,\ n$ the variance S satisfies the inequality [1].

$$S^2 \leq [f(\lambda_1) - m][m - f(\lambda_n)]. \tag{2.13}$$

We prove only (2.11) as (2.12) is proved similarly. From (2.13) and (2.10) we have

$$\begin{aligned} f(\lambda_1) &\geq m + \frac{S^2}{m - f(\lambda_n)} \\ &\geq m + \frac{S}{\sqrt{n-1}} \end{aligned}$$

∎

Theorem 2.5. *All* $f(\lambda_j)$ *are bounded below by*

$$f(\lambda_j) \geq trace\,(f(\mathbf{A})) - \sqrt{(n-1)\,trace\,(f(\mathbf{A})^2)} \tag{2.14}$$

Proof. From (1) and the Pythagorean theorem we have

$$\begin{aligned} ||f(\boldsymbol{\lambda})||^2 = \langle f(\boldsymbol{\lambda}),\ f(\boldsymbol{\lambda})\rangle &\geq \left(\frac{\langle f(\boldsymbol{\lambda}),\ \mathbf{e} - \mathbf{e_j}\rangle}{\sqrt{n-1}}\right)^2 \\ &= \frac{(\langle f(\boldsymbol{\lambda}),\ \mathbf{e}\rangle - f(\lambda_j))^2}{n-1} \end{aligned}$$

and hence

$$[trace\,(f(\mathbf{A})) - f(\lambda_j)]^2 \leq (n-1)trace\,(f(\mathbf{A})^2)$$

from which the result follows. This result is particularly useful for $j = n$ if the right hand side in (2.14) is positive. ∎

Theorem 2.6. $f(\lambda_1)$ *is bounded below by*

$$f(\lambda_1) \geq m + \frac{S^2}{\sqrt{(n-1)trace\,(f(\mathbf{A})^2)} - m(n-1)} \tag{2.15}$$

Proof. Use inequality (2.14) with $j = n$ together with (2.13) and solve for $f(\lambda_1)$ ∎

Lemma 2.7. *Let* $\mathbf{v} \notin N(\mathbf{P})$ *then*

$$|\langle \mathbf{P}f(\boldsymbol{\lambda}),\ \mathbf{v}\rangle| \leq \langle \mathbf{P}f(\boldsymbol{\lambda}),\ f(\boldsymbol{\lambda})\rangle^{\frac{1}{2}} \langle \mathbf{P}\mathbf{v},\ \mathbf{v}\rangle^{\frac{1}{2}}$$

Proof. Use the Cauchy Schwarz inequality and the fact that $\mathbf{P}$ is a symmetric projection to get

$$\begin{aligned} |\langle \mathbf{P}f(\boldsymbol{\lambda}),\ \mathbf{v}\rangle| &= |\langle \mathbf{P}^2 f(\boldsymbol{\lambda}),\ \mathbf{v}\rangle| \\ &= |\langle \mathbf{P}f(\boldsymbol{\lambda}),\ \mathbf{P}\mathbf{v}\rangle| \\ &\leq \langle \mathbf{P}f(\boldsymbol{\lambda}),\ \mathbf{P}f(\boldsymbol{\lambda})\rangle^{\frac{1}{2}} \langle \mathbf{P}\mathbf{v},\ \mathbf{P}\mathbf{v}\rangle^{\frac{1}{2}} \\ &= \langle \mathbf{P}f(\boldsymbol{\lambda}),\ f(\boldsymbol{\lambda})\rangle^{\frac{1}{2}} \langle \mathbf{P}\mathbf{v},\ \mathbf{v}\rangle^{\frac{1}{2}}. \end{aligned} \tag{2.16}$$

∎

Theorem 2.8. *Define*

$$m_j = \frac{\langle f(\boldsymbol{\lambda}), \mathbf{e} - \mathbf{e_j} \rangle}{n-1} \tag{2.17}$$

$$= \frac{trace\,(f(\mathbf{A})) - f(\lambda_j)}{n-1}, \tag{2.18}$$

then all the eigenvalues satisfy

$$|f(\lambda_k) - m_j| \le \sqrt{n-2}\left(\frac{trace\,(f(\mathbf{A})^2) - f(\lambda_j)^2}{n-1} - m_j^2\right)^{\frac{1}{2}}, \tag{2.19}$$

for $k \neq j$ *and*

$$|f(\lambda_j) - m| \le S\sqrt{n-1} \tag{2.20}$$

Proof. Use 2.7 with $\mathbf{v} = \mathbf{e_k},\ k \neq j$ to get

$$\langle \mathbf{P} f(\boldsymbol{\lambda}),\, \mathbf{e_k} \rangle = f(\lambda_k) - \frac{trace\,(f(\mathbf{A})) - f(\lambda_j)}{n-1}$$

$$= f(\lambda_k) - m_j$$

$$\langle \mathbf{P} f(\boldsymbol{\lambda}),\ f(\boldsymbol{\lambda}) \rangle = trace\,(f(\mathbf{A})^2) - f(\lambda_j)^2 - \frac{[trace\,(f(\mathbf{A})) - f(\lambda_j)]^2}{n-1}$$

$$= trace\,(f(\mathbf{A})^2) - f(\lambda_j)^2 - (n-1)m_j^2$$

$$\langle \mathbf{P}\mathbf{e_k},\, \mathbf{e_k} \rangle = \frac{n-2}{n-1}$$

and the result follows by substitution into (2.16). Result (2.20) has already been proved in 2.2, however it also follows by noting that the discriminant in (2.19) is positive which yields

$$f(\lambda_j)^2 - \frac{2\,trace\,(f(A))f(\lambda_j)}{n} + \frac{[trace\,(f(A))]^2}{n} - \frac{n-1}{n} trace\,(f(A)^2) \le 0 \tag{2.21}$$

which simplifies to

$$f(\lambda_j)^2 - 2mf(\lambda_j) + m^2 + (n-1)m^2 - (n-1)(m^2 + S^2) \le 0 \tag{2.22}$$

or

$$\left[f(\lambda_j)^2 - m\right]^2 \le (n-1)S^2 \tag{2.23}$$

from which the result follows. With $k = 1$ and $j = n$ in (2.19) we have the upper bound

$$f(\lambda_1) \le m_n + \sqrt{n-2}\left(\frac{trace\,(f(\mathbf{A})^2) - f(\lambda_n)^2}{n-1} - m_n^2\right)^{\frac{1}{2}} \tag{2.24}$$

and with $k = n$ and $j = 1$ in (2.19) we have the lower bound

$$f(\lambda_n) \ge m_1 - \sqrt{n-2}\left(\frac{trace\,(f(\mathbf{A})^2) - f(\lambda_1)^2}{n-1} - m_1^2\right)^{\frac{1}{2}} \tag{2.25}$$

∎

Theorem 2.9. *The spread* $sp(f(\mathbf{A}))$ *is bounded above by* $S\sqrt{2n}$

Proof. Consider the symmetric projector $\bar{\mathbf{P}}$ defined by

$$\bar{\mathbf{P}} = \mathbf{I} - \frac{(\mathbf{e_1} - \mathbf{e_n})(\mathbf{e_1} - \mathbf{e_n})^t}{2} - \frac{\mathbf{e}\mathbf{e}^t}{n}. \tag{2.26}$$

It is easily shown that $rank(\bar{\mathbf{P}}) = n - 2$ and that an orthonormal basis for $N(\bar{\mathbf{P}})$ is given by

$$N(\bar{\mathbf{P}}) = span\left\{\frac{\mathbf{e_1} - \mathbf{e_n}}{\sqrt{2}}, \frac{\mathbf{e}}{\sqrt{n}}\right\} \tag{2.27}$$

Write $f(\boldsymbol{\lambda})$ in terms of its orthogonal components as

$$f(\boldsymbol{\lambda}) = \bar{\mathbf{P}} f(\boldsymbol{\lambda}) + \frac{\langle f(\boldsymbol{\lambda}),\, \mathbf{e_1} - \mathbf{e_n}\rangle}{\sqrt{2}} \frac{(\mathbf{e_1} - \mathbf{e_n})}{\sqrt{2}} + \frac{\langle f(\boldsymbol{\lambda}),\, \mathbf{e}\rangle}{\sqrt{n}} \frac{\mathbf{e}}{\sqrt{n}} \tag{2.28}$$

then it follows from the Pythagorean theorem [5] in an inner product space that

$$||f(\boldsymbol{\lambda})||^2 = \langle f(\boldsymbol{\lambda}),\, f(\boldsymbol{\lambda})\rangle \geq \frac{\langle f(\boldsymbol{\lambda}),\, \mathbf{e_1} - \mathbf{e_n}\rangle^2}{2} + \frac{\langle f(\boldsymbol{\lambda}),\, \mathbf{e}\rangle^2}{n} \tag{2.29}$$

$$trace(f(\mathbf{A})^2) \geq \frac{(f(\lambda_1) - f(\lambda_n))^2}{2} + nm^2. \tag{2.30}$$

Recall that $trace(f(\mathbf{A})^2) = n(S^2 + m^2)$ so that (2.30) simplifies to

$$2nS^2 \geq (f(\lambda_1) - f(\lambda_n))^2 \tag{2.31}$$

$$f(\lambda_1) - f(\lambda_n) \leq S\sqrt{2n} \tag{2.32}$$

∎

When $f(x) = x^k$, we obtain the result

$$\begin{aligned} sp(\mathbf{A}^k) &\leq S\sqrt{2n} \\ &\leq \sqrt{2\left(trace(\mathbf{A}^{2k}) - \frac{(trace(\mathbf{A}^k))^2}{n}\right)} \end{aligned} \tag{2.33}$$

, which for $k = 1$ agrees with the result derived in [8].

Alternative Proofs

Here is an alternative proof of Theorem 2.3 and Theorem 2.4.

Proof. Let $z = f(\lambda_1) - m$ and use (2.13) and (2.25), also separately let $z = m - f(\lambda_n)$ and use (2.13) and (2.24). Both approaches lead to the quartic inequality in z given by

$$z^4 - \frac{n^2 - 2n + 2}{n - 1} S^2 z^2 + S^4 \leq 0., \tag{2.34}$$

the solution of which yields

$$\frac{S}{\sqrt{n-1}} \leq z \leq S\sqrt{n-1}.. \tag{2.35}$$

Substituting for $z = f(\lambda_1) - m$ and separately $z = m - f(\lambda_n)$ gives the required result. ∎

3. RESULTS

Consider the test matrix [8]

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

with spectrum $\sigma(\mathbf{A}) = \{1.4257,\ 4.7754,\ 6.423,\ 9.3759\}$ accurate to four decimal places and choose. We choose $f(x) = x^k,\ \ k \in \mathbb{N}$ as polynomial functions of $\mathbf{A}$ are easy to evaluate. Our results are summarized for $k = 1, 2, 3, 4$ in Table 3.1 using equations (2.9)-(2.12). If lower bounds for $f(\lambda_n)$ are negative then we bound λ_n below by zero (positive definite) otherwise we take the k_{th} root to recover the bound. For $k = 1$, we obtain the bounds of [8] and [6]. From (14) we get $\lambda_n \geq 0.5058$ for $k = 1$ and zero for larger k (negative $f(\lambda_n)$). Inequalities (2.24) and (2.25) relate $f(\lambda_1)$ and $f(\lambda_n)$. If either λ_1 or λ_n are known fairly accurately then they may be used. With $\lambda_1 = 9.3759$ we get from (2.25) $\lambda_n \geq 1.2675$ for $k = 1$ and negative values for larger k. Results using (2.24) with $\lambda_n = 1.4257$ are summarized in Table 3.2. From 3.1 we find that the bounds get better for λ_1 for larger k, while worse for λ_n. However there is no need to use the same $f(x)$ for the upper and lower bounds here. Inequalities (2.24) and (2.25) are useful but we need to use the same $f(x)$ here. From (2.33) we obtain $sp(\mathbf{A}) = 8.1240$ and $sp(\mathbf{A}^2) = 89.7218$, which differs not much from the exact values of 7.9502 and 85.8749 respectively. In addition our work applies to symmetric matrices if the domain of $f(x)$ is $(-\infty, \infty)$ on which $f(x)$ increases. In this case we are restricted to $f(x) = x^{2k+1},\ k \in \mathbb{N}$.

k	λ_1	λ_n
1	$[7.1583, 10.4749]$	$[0.5251, 3.8417]$
2	$[7.5375, 9.6666]$	$(0, 4.4928]$
3	$[7.8461, 9.4672]$	$(0, 4.8978]$
4	$[8.0840, 9.4083]$	$(0, 5.1552]$

Table 3.1: Bounds, (2.9)-(2.12)

k	λ_1
1	9.5497
2	9.4552
3	9.4116
4	9.3918

Table 3.2: Bounds, (2.24)

4. CONCLUSION

We have illustrated by using a simple projection how effective formulae can be derived for the bounds of extremal eigenvalues of real positive definite matrices. We advocate using these formulae for relatively small k together with other known results to bound these extremal eigenvalues. In addition we have derived a useful result for the spread of functions of the matrix. For sparse matrices the larger values of k can be chosen as less numerical effort is required.

REFERENCES

[1] R. BHATIA and C. DAVIS, A better bound on the variance, *Amer. Math. Monthly*, **107** (2000), pp. 353–357.

[2] A. BRAUER, Limits for the characteristic roots of a matrix VII, *Duke Math. J.*, **25** (1958), pp. 583–590.

[3] R. HORN and C. A. JOHNSON, *Matrix Analysis*, Cambridge University Press, (2012).

[4] T. Z. HUANG and C. X. XU, Bounds for the extreme eigenvalues of symmetric matrices, *ZAMM* **83** (2003), No. 3, pp. 214–216.

[5] E. KREYSZIG, *Introductory Functional Analysis with Applications*, John Wiley, Hoboken, 1978.

[6] R. SHARMA, R. KUMAR and R. SAINI, Note on Bounds for Eigenvalues using Traces, *Functional Analysis*, arXiv:1409.0096v1, (2014).

[7] H. TAL-EZER, Polynomial approximation of functions of matrices and applications, *J. Sci. Comput.* (1989), No. 4, pp. 25–60. `https://doi.org/10.1007/BF01061265`.

[8] H. WOLKOWICZ and G. P. H. STYAN, Bounds for eigenvalues using traces, *Linear Algebra Appl*. **29** (1980), pp. 471–506.