Distance bounds for prescribed multiple eigenvalues of matrix polynomials

Panayiotis J. Psarrakos

National Technical University of Athens

Introduction and preliminaries

Wilkinson's problem (1972) concerns computing the spectral norm distance from a matrix $A \in \mathbb{C}^{n\times n}$ with n distinct eigenvalues to the set of $n \times n$ matrices having multiple eigenvalues, and has a strong connection to ill-conditioning of eigenvalue problems.

Malyshev (1999) provided a solution to Wilkinson's problem by obtaining

$$\inf\{\|E\|_2 : \mu \text{ is a multiple eigenvalue of } A+E\}$$

$$= \sup_{\gamma>0} s_{2n-1}\left(\begin{bmatrix} I\mu - A & 0 \\ \gamma I & I\mu - A \end{bmatrix}\right),$$

where $s_1(\cdot) \geq s_2(\cdot) \geq s_3(\cdot) \geq \cdots$ denote the singular values of a matrix.

Mengi (2011) derived a characterization for the smallest perturbation to a matrix with an eigenvalue of specified algebraic multiplicity by proving

$$\inf \{ \|E\|_2 : \mu \text{ is an eigenvalue of } A+E \text{ of algebraic multiplicity } r \}$$

$$= \sup_{\gamma_{i,j} \in \mathbb{C} \setminus \{0\}} s_{rn-r+1} \left(\begin{bmatrix} I\mu - A & 0 & 0 & \cdots & 0 \\ \gamma_{2,1} I & I\mu - A & 0 & \cdots & 0 \\ \gamma_{3,1} I & \gamma_{3,2} I & I\mu - A & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{r,1} I & \gamma_{r,2} I & \gamma_{r,3} I & \cdots & I\mu - A \end{bmatrix} \right).$$

In this work, we study the case of polynomial eigenvalue problems, and estimate a weighted distance from a given $n \times n$ matrix polynomial to the $n \times n$ matrix polynomials that have a prescribed multiple eigenvalue.

Consider a **matrix polynomial (m.p.)**

$$P(\lambda) = A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_1\lambda + A_0,$$

where λ is a complex variable, $A_j \in \mathbb{C}^{n\times n}$, and $\det P(\lambda) \not\equiv 0$.

A scalar $\lambda_0 \in \mathbb{C}$ is an **eigenvalue** of $P(\lambda)$ if $P(\lambda_0)x_0 = 0$ for some $0 \neq x_0 \in \mathbb{C}^n$, known as a **(right) eigenvector** of $P(\lambda)$ corresponding to λ_0. The set of all eigenvalues of $P(\lambda)$,

$$\sigma(P) \;=\; \{\lambda \in \mathbb{C} : \det P(\lambda) = 0\}\,,$$

is the **spectrum** of $P(\lambda)$ and contains at most nm (finite) elements.

For a $\lambda_0 \in \sigma(P)$, its **algebraic multiplicity (a.m.)** is the multiplicity of λ_0 as a zero of $\det P(\lambda)$, and its **geometric multiplicity (g.m.)** is the dimension of the null space of $P(\lambda_0)$. It holds that a.m. $\geq$ g.m.

Suppose that for a $\lambda_0 \in \sigma(P)$, there are $x_0 (\neq 0), x_1, \dots, x_k \in \mathbb{C}^n$, s.t.

$$\sum_{j=0}^{\xi} \frac{1}{j!} P^{(j)}(\lambda_0)\, x_{\xi-j} = 0\,; \quad \xi = 0, 1, \dots, k.$$

Then x_0 is an eigenvector of λ_0, and $x_1, x_2, \dots, x_k$ are known as **generalized eigenvectors**. The set $\{x_0, x_1, \dots, x_k\}$ is said to be a **Jordan chain** of $P(\lambda)$ corresponding to $\lambda_0 \in \sigma(P)$.

Any eigenvalue of $P(\lambda)$ of g.m. p has p maximal Jordan chains associated with p linearly independent eigenvectors, with total number of vectors equal to its a.m. The largest length of Jordan chains of $P(\lambda)$ corresponding to a $\lambda_0 \in \sigma(P)$ is called the **index** of λ_0, and it is the size of the largest Jordan blocks of the Jordan canonical form of $P(\lambda)$ associated with λ_0. The index is equal to 1 **iff** a.m. = g.m.

We are interested in the spectra of perturbations of $P(\lambda)$ of the form

$$Q(\lambda) = P(\lambda) + \Delta(\lambda) = \sum_{j=0}^{m} (A_j + \Delta_j)\lambda^j$$

for arbitrary $\Delta_j \in \mathbb{C}^{n \times n}$. For $\mathrm{w} = \{w_0, w_1, \ldots, w_m\}$, $w_j \geq 0 \ (j > 0)$, $w_0 > 0$, and $\varepsilon > 0$, we define the set of perturbations of $P(\lambda)$,

$$\mathcal{B}(P, \varepsilon, \mathrm{w}) = \left\{ Q(\lambda) = \sum_{j=0}^{m} (A_j + \Delta_j)\lambda^j : \|\Delta_j\|_2 \leq \varepsilon\, w_j,\ j = 0, 1, \ldots, m \right\}.$$

The weights w_j allow freedom in how perturbations are measured.

The ε-**pseudospectrum** of $P(\lambda)$ [Tisseur-Higham, 2001] is defined as

$$\sigma_\varepsilon(P) = \{\mu \in \sigma(Q) : Q(\lambda) \in \mathcal{B}(P, \varepsilon, \mathrm{w})\},$$

and allows a “visualization” of the sensitivity of eigenvalues.

For the m.p. $P(\lambda)$ and a given $\mu \in \mathbb{C}$, we define the **distance from $P(\lambda)$ to μ as an eigenvalue of g.m. at least** r,

$$\mathcal{G}_r(\mu) = \inf\{\varepsilon \geq 0 : \exists\, Q(\lambda) \in \mathcal{B}(P, \varepsilon, \mathrm{w}) \text{ with } \mu \in \sigma(Q) \text{ of g.m.} \geq r\},$$

the **distance from $P(\lambda)$ to μ as an eigenvalue of a.m. at least** r,

$$\mathcal{E}_r(\mu) = \inf\{\varepsilon \geq 0 : \exists\, Q(\lambda) \in \mathcal{B}(P, \varepsilon, \mathrm{w}) \text{ with } \mu \in \sigma(Q) \text{ of a.m.} \geq r\},$$

and the **distance from $P(\lambda)$ to μ as an eigenvalue of a.m. at least r and index (exactly)** k,

$$\mathcal{E}_{r,k}(\mu) = \inf\{\varepsilon \geq 0 : \exists\, Q(\lambda) \in \mathcal{B}(P, \varepsilon, \mathrm{w}) \text{ with } \mu \in \sigma(Q) \text{ of a.m.} \geq r \text{ and index } k\}.$$

Computation of $\mathcal{G}_r(\mu)$

(generalizing [Schmidt, 1907] & [Eckart-Young, 1936])

For a given $\mu \in \mathbb{C}$, consider the **SVD** of the (constant) matrix $P(\mu)$

$$P(\mu) = U \,\text{diag}\,\{s_1(P(\mu)), s_2(P(\mu)), \dots, s_n(P(\mu))\}\, V^*,$$

where $U = [u_1 \; u_2 \; \cdots \; u_n]\,,\, V = [v_1 \; v_2 \; \cdots \; v_n] \in \mathbb{C}^{n\times n}$ are unitary.

Then define the $n \times n$ matrices

$$E = -U\,\text{diag}\{0, \dots, 0, s_{n-r+1}(P(\mu)), \dots, s_n(P(\mu))\}\, V^*$$

and

$$\Delta_j = \frac{w_j}{w(|\mu|)} \left(\frac{\overline{\mu}}{|\mu|} \right)^j E\,; \quad j = 0, 1, \dots, m,$$

assuming that $\overline{\mu}/|\mu| = 0$ whenever $\mu = 0$.

The m.p. $\Delta(\lambda) = \sum_{j=0}^{m} \Delta_j \lambda^j$ satisfies $\Delta(\mu) = E$, and the perturbation

$$Q(\lambda) = P(\lambda) + \Delta(\lambda) = \sum_{j=0}^{m} (A_j + \Delta_j)\lambda^j$$

lies on $\partial \mathcal{B}\left(P, \frac{s_{n-r+1}(P(\mu))}{w(|\mu|)}, \text{w}\right)$. Moreover, for every $j = n-r+1, \ldots, n-1, n,$ we have $Q(\mu)v_j = 0$ and $u_j^* Q(\mu) = 0$.

Theorem 1 *The distance from* $P(\lambda)$ *to* μ *as an eigenvalue of geometric multiplicity* r *is*

$$\mathcal{G}_r(\mu) \; = \; \frac{s_{n-r+1}(P(\mu))}{w(|\mu|)},$$

and $Q(\lambda)$ *above is an optimal perturbation.*

A lower bound for $\mathcal{E}_{r,k}(\mu)$ (generalizing [Malyshev, 1999])

By our definitions and the previous theorem,

$$\mathcal{E}_1(\mu) = \frac{s_n(P(\mu))}{w(|\mu|)} \leq \mathcal{E}_r(\mu) \leq \frac{s_{n-r+1}(P(\mu))}{w(|\mu|)} = \mathcal{G}_r(\mu) \leq \mathcal{E}_{r,1}(\mu).$$

For $0 \neq \gamma \in \mathbb{C}$, we define

$$F_k[P(\lambda);\gamma] = \begin{bmatrix} P(\lambda) & 0 & \cdots & 0 \\ \gamma\, P^{(1)}(\lambda) & P(\lambda) & \cdots & 0 \\ \frac{\gamma^2}{2!}\, P^{(2)}(\lambda) & \gamma\, P^{(1)}(\lambda) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\gamma^{k-1}}{(k-1)!}\, P^{(k-1)}(\lambda) & \frac{\gamma^{k-2}}{(k-2)!}\, P^{(k-2)}(\lambda) & \cdots & P(\lambda) \end{bmatrix}.$$

Lemma 2 *Suppose* $\lambda_0 \in \mathbb{C}$ *is an eigenvalue of* $P(\lambda)$ *of a.m. at least* r *and index* k*. Then for any scalar* $\gamma \neq 0$,

$$s_{kn-r+1}(F_k[P(\lambda_0);\gamma]) = 0.$$

Theorem 3 *For any* $\gamma > 0$,

$$\mathcal{E}_{r,k}(\mu) \geq \frac{s_{kn-r+1}(F_k[P(\mu);\gamma])}{\|F_k[w(|\mu|);\gamma]\|}\,; \quad k = 1, 2, \ldots, r$$

and

$$\mathcal{E}_r(\mu) \geq \min_{k=1,2,\ldots,nm} \frac{s_{kn-r+1}(F_k[P(\mu);\gamma])}{\|F_k[w(|\mu|);\gamma]\|}.$$

An upper bound for $\mathcal{E}_r(\mu)$ (generalizing [Malyshev, 1999])

Without loss of generality, we may assume that $\gamma > 0$.

For $r \in \{2, 3, \dots, n\}$, let $\begin{bmatrix} u_1(\gamma) \\ u_2(\gamma) \\ \vdots \\ u_r(\gamma) \end{bmatrix}, \begin{bmatrix} v_1(\gamma) \\ v_2(\gamma) \\ \vdots \\ v_r(\gamma) \end{bmatrix} \in \mathbb{C}^{rn}$, with

$u_j(\gamma), v_j(\gamma) \in \mathbb{C}^n$ $(j = 1, 2, \dots, r)$, be left and right singular vectors of $s_{rn-r+1}(F_r[P(\mu); \gamma])$, respectively. We define the $n \times r$ matrices $U(\gamma) = [\, u_1(\gamma)\ u_2(\gamma)\ \cdots\ u_r(\gamma) \,]$ and $V(\gamma) = [\, v_1(\gamma)\ v_2(\gamma)\ \cdots\ v_r(\gamma) \,]$.

If $rank(V(\gamma)) = r$, then we can construct a m.p. $\Delta_\gamma(\lambda)$ s.t. $Q_\gamma(\lambda) = P(\lambda) + \Delta_\gamma(\lambda)$ has μ as an eigenvalue of a.m. at least r.

We define the quantities $\phi_i = \frac{w^{(i)}(|\mu|)}{(i!)\, w(|\mu|)} \left(\frac{\overline{\mu}}{|\mu|}\right)^i$, $i = 1, 2, \ldots, r$, setting $\overline{\mu}/|\mu| = 0$ whenever $\mu = 0$. Then we define the $r \times r$ upper triangular Toeplitz matrix

$$\Theta_\gamma = [\theta_{i,j}] = \begin{bmatrix} 1 & -\gamma\phi_1 & \gamma^2(\phi_1^2 - \phi_2) & \gamma^3(2\phi_1\phi_2 - \phi_3 - \phi_1^3) & \cdots \\ 0 & 1 & -\gamma\phi_1 & \gamma^2(\phi_1^2 - \phi_2) & \cdots \\ 0 & 0 & 1 & -\gamma\phi_1 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

whose entries above the diagonal are given by the recursive formulae

$$\theta_{i,j} = -\theta_{i,i}\gamma^{j-i}\phi_{j-i} - \theta_{i,i+1}\gamma^{j-(i+1)}\phi_{j-(i+1)} - \cdots - \theta_{i,j-1}\gamma\phi_1\,;\ 1 \le i < j \le r.$$

Denoting by $V(\gamma)^\dagger$ the Moore-Penrose pseudoinverse of $V(\gamma)$, we consider the $n \times n$ matrix

$$\Delta_\gamma = -s_{rn-r+1}(F_r[P(\mu);\gamma])\,U(\gamma)\,\Theta_\gamma V(\gamma)^\dagger,$$

and we define the matrices

$$\Delta_{\gamma,j} = \frac{w_j}{w(|\mu|)}\left(\frac{\overline{\mu}}{|\mu|}\right)^j \Delta_\gamma\,;\quad j = 0, 1, \ldots, m$$

and the matrix polynomial

$$\Delta_\gamma(\lambda) = \sum_{j=0}^{m} \Delta_{\gamma,j}\lambda^j.$$

Then, the scalar μ is a multiple eigenvalue of the m.p.

$$Q_\gamma(\lambda) \;=\; P(\lambda) + \Delta_\gamma(\lambda) \;=\; \sum_{j=0}^{m} (A_j + \Delta_{\gamma,j})\lambda^j$$

with $\left\{v_1(\gamma), \gamma^{-1}v_2(\gamma), \gamma^{-2}v_3(\gamma), \ldots, \gamma^{-(r-1)}v_r(\gamma)\right\}$ as an associated Jordan chain of length r.

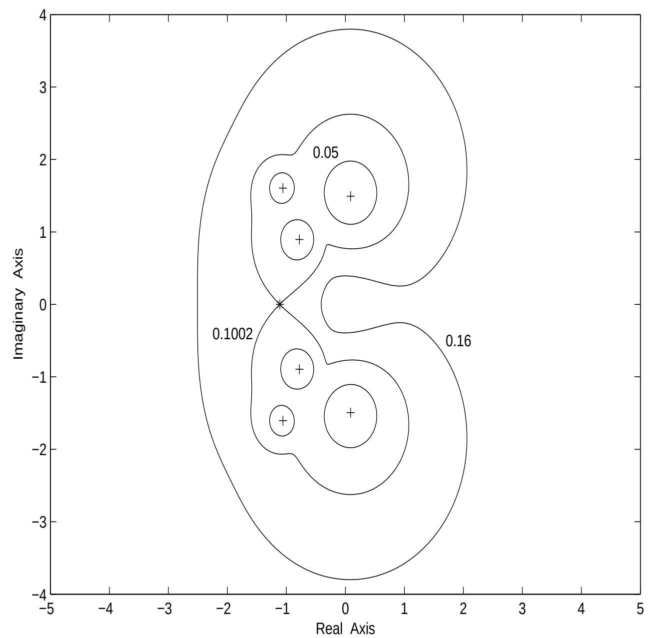
Theorem 4 *For any real $\gamma > 0$ s.t. $rank(V(\gamma)) = r \le n$, it holds that*

$$\mathcal{E}_r(\mu) \;\le\; \frac{s_{rn-r+1}(F_r[P(\mu);\gamma])}{w(|\mu|)} \left\| U(\gamma)\,\Theta_\gamma V(\gamma)^\dagger \right\|,$$

and $Q_\gamma(\lambda) \,\in\, \partial\mathcal{B}\left(P, \frac{s_{rn-r+1}(F_r[P(\mu);\gamma])}{w(|\mu|)} \left\| U(\gamma)\,\Theta_\gamma V(\gamma)^\dagger \right\|, \mathrm{w}\right)$ *above has μ as an eigenvalue with a Jordan chain of length at least r.*

A numerical example

Let $P(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & -1 & 6 \end{bmatrix} \lambda + \begin{bmatrix} 2 & 1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 10 \end{bmatrix}$ and w $= \{10, 6.11, 3\}$. The ε-pseudospectra of $P(\lambda)$ for $\varepsilon = 0.05, 0.1002,$ $0.16,$ are:



For the scalar $\mu = 3 + \mathrm{i}$, we get

$$\mathcal{G}_2(3+\mathrm{i}) = \frac{s_2(P(3+\mathrm{i}))}{w(|3+\mathrm{i}|)} = \frac{32.1524}{59.3240} = 0.5420,$$

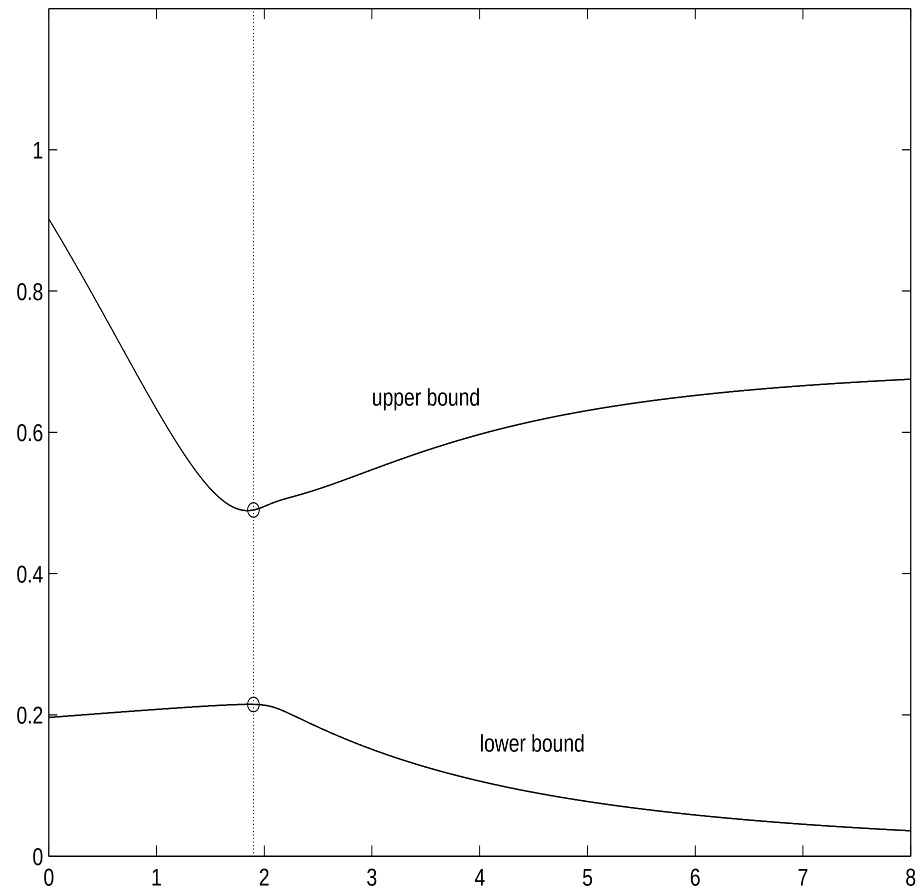
with an optimal perturbation

$$\hat{Q}(\lambda) = \begin{bmatrix} 0.4134 + \mathrm{i}\,0.0607 & -0.1918 + \mathrm{i}\,0.0808 & -0.0090 - \mathrm{i}\,0.0002 \\ 0.0403 - \mathrm{i}\,0.0302 & 0.4100 + \mathrm{i}\,0.2467 & -0.0611 - \mathrm{i}\,0.0173 \\ -0.0037 + \mathrm{i}\,0.0010 & 0.0599 + \mathrm{i}\,0.0176 & 3.0019 + \mathrm{i}\,0.0016 \end{bmatrix} \lambda^2$$
$$+ \begin{bmatrix} -1.1726 - \mathrm{i}\,0.2606 & 0.5773 + \mathrm{i}\,0.0325 & -0.0172 - \mathrm{i}\,0.0062 \\ 0.0974 - \mathrm{i}\,0.0325 & -0.2315 - \mathrm{i}\,0.5476 & 0.8931 - \mathrm{i}\,0.0727 \\ -0.0077 - \mathrm{i}\,0.0004 & -0.8955 + \mathrm{i}\,0.0727 & 6.0025 + \mathrm{i}\,0.0043 \end{bmatrix} \lambda$$
$$+ \begin{bmatrix} 0.3144 - \mathrm{i}\,1.0114 & 0.3269 - \mathrm{i}\,0.1683 & -0.0235 - \mathrm{i}\,0.0185 \\ -0.8319 + \mathrm{i}\,0.0000 & -1.7334 - \mathrm{i}\,2.5223 & -0.1283 - \mathrm{i}\,0.1682 \\ -0.0118 - \mathrm{i}\,0.0046 & 0.1245 + \mathrm{i}\,0.1669 & 10.0017 + \mathrm{i}\,0.0080 \end{bmatrix}$$

that has $\mu = 3 + \mathrm{i}$, as an eigenvalue with a.m. $=$ g.m. $= 2$.

The graphs of the lower bound for $\mathcal{E}_{2,2}(3+\mathrm{i})$ and the upper bound for $\mathcal{E}_2(3+\mathrm{i})$ are illustrated below, and for $\gamma = 1.9$,

$$0.2149 \leq \mathcal{E}_{2,2}(3+\mathrm{i}) = \mathcal{E}_2(3+\mathrm{i}) \leq 0.4901 < 0.5420 = \mathcal{G}_2(3+\mathrm{i}).$$

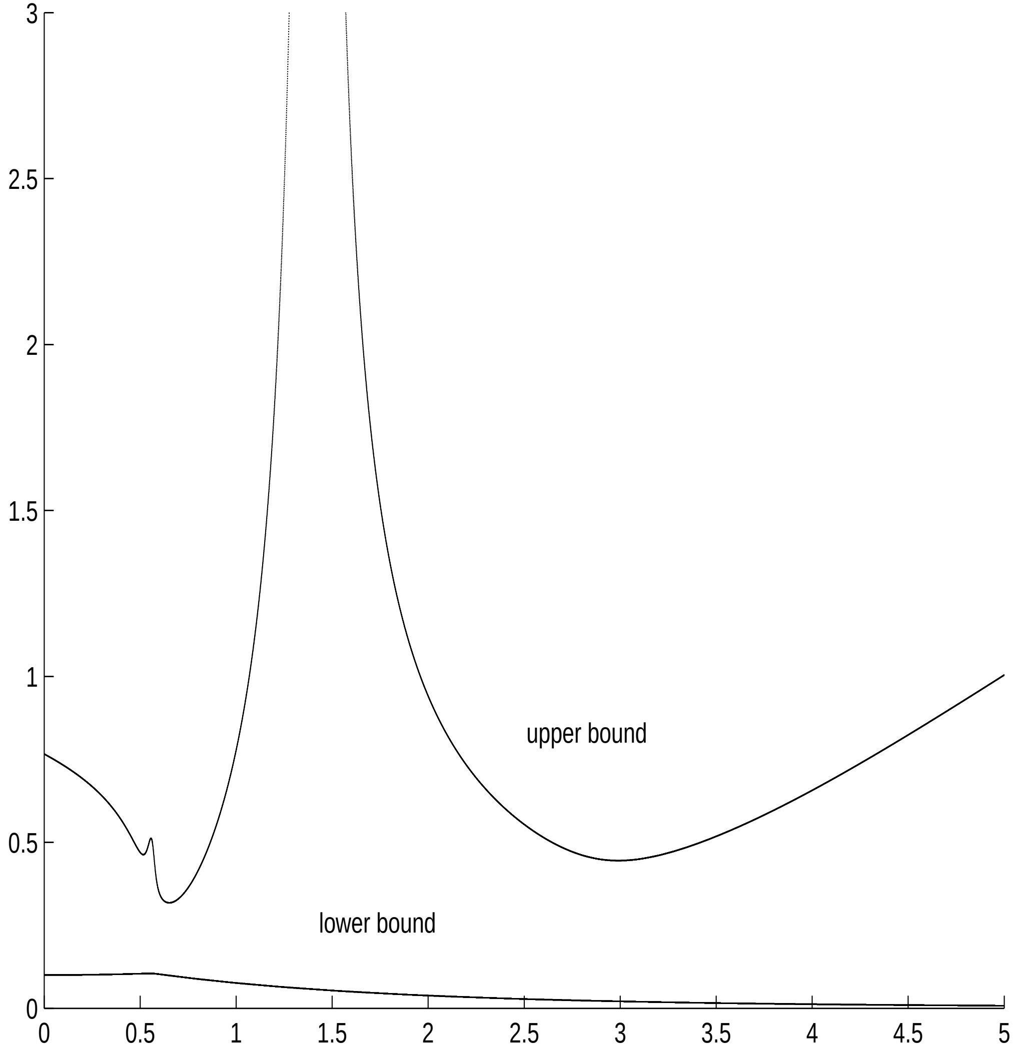


For $\mu = -1.1105$, a pseudospectra approach [Boulton-Lancaster-Ps., 2008] implies that

$$\mathcal{E}_2(-1.1105) = \mathcal{E}_{2,2}(-1.1105) = \mathcal{E}_1(-1.1105) = 0.1002.$$

The graphs of the lower bound for $\mathcal{E}_{3,3}(-1.1105)$ and the upper bound for $\mathcal{E}_3(-1.1105)$ are given below. For $\gamma = 0.5530$ and $\gamma = 0.6518$, we get, respectively,

$$\mathcal{E}_{3,3}(-1.1105) \geq 0.1048 \quad \text{and} \quad \mathcal{E}_3(-1.1105) \leq 0.3177.$$


upper bound
lower bound
0
0.5
1
1.5
2
2.5
3
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

References

L. Boulton, P. Lancaster and P. Psarrakos, *Math. Comp.*, 2008.

C. Eckart and G. Young, *Psychometrika*, 1936.

I. Gohberg, P. Lancaster and L. Rodman, *"Matrix Polynomials"*, 1982.

P. Lancaster and M. Tismenetsky, *"The Theory of Matrices"*, 1985.

A.N. Malyshev, *Numer. Math.*, 1999.

E. Mengi, *Numer. Math.*, 2011.

N. Papathanasiou and P. Psarrakos, *Linear Algebra Appl.*, 2008.

P. Psarrakos, *Linear Algebra Appl.*, 2012.

E. Schmidt, *Math. Anal.*, 1907.

F. Tisseur and N.J. Higham, *SIAM J. Matrix Anal. Appl.*, 2001.

J.H. Wilkinson, *Numer. Math.*, 1972.