Estimates of eigenspaces and eigenvalues of a matrix

Łukasz Struski[*], Jacek Tabor, and Piotr Zgliczyński [†]

Jagiellonian University,
Faculty of Mathematics and Computer Science,
Łojasiewicza 6, 30-348 Kraków, Poland

e-mail: struski@im.uj.edu.pl,
tabor@ii.uj.edu.pl,
umzglicz@cyf-kr.edu.pl

December 23, 2014

Abstract

We present a new tool based on cones for rigorous estimations of eigenvectors, eigenspaces and eigenvalues of a matrix. We introduce the notion of dominated matrix and present a theorem which shows that our method is a generalization of the Gerschgorin theorem in the case isolated Gerschgorin disk. Our approach is based on ideas from dynamical systems which allow us also to locate eigenspaces of the composition of matrices.

Key words and phrases: eigenvectors, eigenvalues, Gerschgorin theorem, cone condition, spectrum of the matrix

2010 Mathematics Subject Classification: 65F15, 37D30.

1 Introduction

Assume that our task is to find rigorous bounds for eigenvalues (all or some of them) and their corresponding eigenspaces of a matrix $M \in \mathbb{R}^{n\times n}$. First, one usually applies some iterative scheme, for example QR-algorithm, to obtain matrix A (almost diagonal) which is similar to M. Then one typically applies some abstract theorem to $\tilde{A}$ to infer the rigorous bounds on the eigenvalues and the eigenspaces. In this paper we present a method in this direction.

The main question we try to address can be stated as follows. Assume that we have a square matrix A which the entries (or blocks) on the diagonal 'dominate' the off-diagonal entries (blocks) and we want to obtain efficient computable bounds (a formula) for the spectrum and eigenspaces of A. Regarding the bounds on the spectrum almost all of the known methods are given by the Gerschgorin theorems and its modifications, for example the Brauer ovals [1, 7] or the generalization of the Gerschgorin theorem to the case of multi-dimensional blocks by Feingold and Varga [2]. Estimation of isolated eigenvectors from Gerschgorin's results are due to Wilkinson [8]. However, Wilkinson's result does not give the whole eigenspace in the case of not simple eigenvalue or a cluster of close eigenvalues. In [9], T. Yamamoto showed how find rigorous error bounds for computed single eigenvalues and eigenvectors of real matrices on the

[*] Corresponding author

[†] Research has been supported by Polish National Science Centre grant 2011/03B/ST1/04780

basis of an existence theorem for solutions of nonlinear systems using iteration Newton's method. However, the Yamamoto's approach gives no theoretical estimates for the bounds for computed eigenvalues and eigenvectors.

In this article we propose a new method for the estimates of eigenvalues and eigenspaces. Our approach is based on the ideas coming from the hyperbolic dynamics [6] and can be illustrated by the following simple two-dimensional example.

Example 1.1. Consider the matrix A is defined by the formula $A = \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix}$. Note that if we take the gray cone (see Figure 1) and we start to iterate points of this cone by matrix A then range of our cone will be reduced to the eigenspace corresponding to one of the eigenvectors of A.

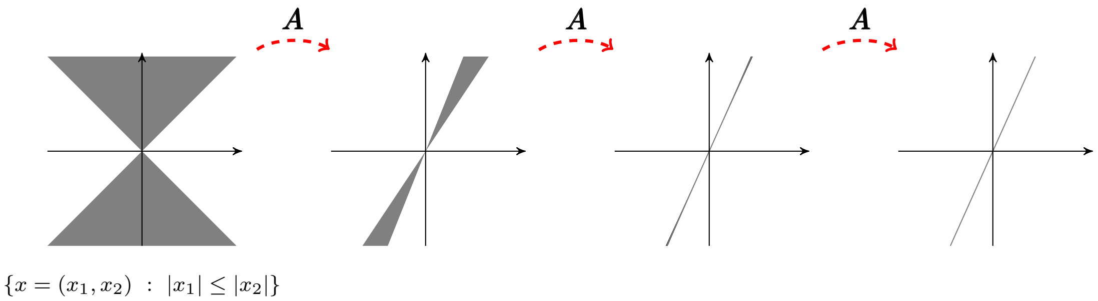


Figure 1: Transformation of the cone by the matrix A.

Iterating backward (A^{-1}) the cone $\{x = (x_1, x_2) \ : \ |x_1| \geq |x_2|\}$ we obtain second eigenspace.

This example illustrates that we can estimate eigenspace by using an invariant cone. We started to study this problem and it turned out that using forward and backward invariant cones we were able to give not only good bounds for eigenvectors but also for eigenvalues. In addition, by means of this tool we can locate the eigenspaces and eigenvalues of products of many matrices.

To explain our main results we introduce some basic notations. Let $\|\mathsf{x}\| := \max\limits_j |x_j|$. For $\mathsf{x} = (x_1, \dots, x_k, \dots, x_n) \in \mathbb{R}^n$ we set

$$\|\mathsf{x}\|_{\leq k} = \max_{i \leq k} |x_i| \quad \text{and} \quad \|\mathsf{x}\|_{>k} = \max_{i>k} |x_i|.$$

For linear map $A \colon \mathbb{R}^k \times \mathbb{R}^{n-k} \to \mathbb{R}^k \times \mathbb{R}^{n-k}$ we define the extension and contraction constants:

$$\begin{aligned} \rangle A \langle &= \ \inf\{R \in \mathbb{R}_+ \mid \|A\mathsf{x}\| \leqslant R \cdot \|x\| \text{ for all } \mathsf{x} \in \mathbb{R}^n : \|A\mathsf{x}\|_{\leq k} \geq \|A\mathsf{x}\|_{>k}\}, \\ \langle A \rangle &= \sup\{R \in \mathbb{R}_+ \mid \|A\mathsf{x}\| \geqslant R \cdot \|x\| \text{ for all } \mathsf{x} \in \mathbb{R}^n : \|\mathsf{x}\|_{\leq k} \leq \|\mathsf{x}\|_{>k}\}. \end{aligned}$$

We say that A is dominating if $\rangle A \langle < \langle A \rangle$. It turns out that composition of dominating maps is dominating, see Proposition 2.9.

We show two main results in our paper

Main Result I [Theorem 5.2]. *Let $A \in \mathbb{C}^{n \times n}$ be a matrix with an isolated Gerschgorin disk. Then A is dominating.*

Together with the following result we get that our method is generalization of the Gerschgorin theorem in the case of the isolated Gerschgorin disk of multiplicity one.

Main Result II [simplified version of Theorem 3.3]. *Let $A \colon \mathbb{R}^k \times \mathbb{R}^{n-k} \to \mathbb{R}^k \times \mathbb{R}^{n-k}$ be dominating. Then there exists a unique direct sum decomposition $F_1 \oplus F_2 = \mathbb{R}^n$ into A-invariant subspaces F_1, F_2 such that*

$$\sigma(A|_{F_1}) \subset \overline{B}(0, \rangle A \langle), \quad \sigma(A|_{F_2}) \subset \mathbb{C} \setminus B(0, \langle A \rangle).$$

Moreover, we have:

1) $\dim F_1 = k, \ \dim F_2 = n - k,$

2) $F_1 \subset \{\mathsf{x} \in \mathbb{R}^n : \|\mathsf{x}\|_{\leq k} \geq \|\mathsf{x}\|_{>k}\}$ *and* $F_2 \subset \{\mathsf{x} \in \mathbb{R}^n : \|\mathsf{x}\|_{\leq k} \leq \|\mathsf{x}\|_{>k}\}$,

3) $\|A|_{F_1}\| \leq \rangle A \langle$ *and* $\|(A|_{F_2})^{-1}\| \leq \langle A \rangle^{-1}$.

In comparison with the Gerschgorin's theorems our method has the following advantages:

- locate spectrum and eigenspaces of a matrix when multiple eigenvalues or clusters of very close eigenvalues are present,
- gives better estimation for isolated eigenvalues,
- allow to deal with composition of matrices.

As an illustration of the quality of the estimates produced by our method we considered the case of the isolated eigenvalue. Assume that

$$A = \begin{bmatrix} a_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $a_{11} \in \mathbb{C}$, $A_{12} \in \mathbb{C}^{1\times(n-1)}$, $A_{21} \in \mathbb{C}^{(n-1)\times 1}$ and $A_{22} \in \mathbb{C}^{(n-1)\times(n-1)}$ are such that a_{11} does not belong to the spectrum of A_{22}. From the implicit function theorem it follows that, if $\|A_{12}\|$ and $\|A_{21}\|$ are sufficiently small, then A has an eigenvalue close to a_{11} and this eigenvalue and the corresponding eigenvector depend analytically on A. Hence there exist analytic functions $\lambda(A) \in \mathbb{C}$ and $v(A) \in \mathbb{C}^n$, such that

$$Av(A) = \lambda(A)v(A).$$

Observe that if $A_{21} = 0$ or $A_{12} = 0$, then $\lambda(A) = a_{11}$. Therefore we expect the following behavior

$$\lambda(A) = a_{11} + O(\|A_{12}\| \cdot \|A_{21}\|) \tag{1}$$

$$v(A) = (1, 0, \dots)^T + O(\|A_{21}\|). \tag{2}$$

Using our approach we obtain the following bounds

$$\begin{aligned} |\lambda(A) - a_{11}| &\leq 2\|A_{12}\| \cdot \|A_{21}\| \cdot \|(A_{22} - a_{11} \cdot I_{\mathbb{C}^{n-1}})^{-1}\|, \\ \|v(A) - (1, 0, \dots, 0)^T\| &\leq 2\|A_{21}\| \cdot \|(A_{22} - a_{11} I_{\mathbb{C}^{n-1}}))^{-1}\| \cdot \|(1, 0, \dots, 0)^T\| \end{aligned}$$

provided $A_{22} - a_{11} I_{\mathbb{C}^{n-1}}$ is invertible and $\|(A_{22} - a_{11} \cdot I_{\mathbb{C}^{n-1}})^{-1}\|^{-2} - 4\|A_{12}\|\|A_{21}\| > 0$. This is the content of Theorem 4.3. Observe that our bounds satisfy (1) and (2).

The content of this paper can be briefly described as follows: in Section 2 we introduce notion of cones and build the concept of dominating matrix. In Section 3 we establish the main result: Theorem 3.3 which allow us to rigorously estimate eigenspaces and eigenvalues. In Section 4 we develop computable estimates for the eigenvalues and eigenspaces based on the results from the Section 3. In Section 5 we compare the proposed method with the Gerschgorin theorem in the case of the isolated Gerschgorin disk. We show that all matrices which have an isolated Gerschgorin disk, are dominating and if the radius of this disk nonzero, we obtain sharper bounds. This means that our approach can be used whenever the Gerschgorin disk is isolated. We also show examples of matrices for which we can not use the Gerschgorin theorem since the Gerschgorin disks cannot be separated, but our method still works, see Example 5.6.

1.1 Notation

By $\mathbb{R}$ and $\mathbb{C}$ we denote the sets of real, and complex numbers. The spectrum $\sigma(A)$ of a square matrix $A = [a_{ij}] \in \mathbb{C}^{n\times n}$ we define the collection of all eigenvalues of A, i.e.

$$\sigma(A) := \{\lambda \in \mathbb{C} \ : \ A - \lambda I \text{ is singular}\}.$$

By $I_{\mathbb{C}^n}$ we mean the identity matrix of size n, while $\mathbb{I}$ denotes the interval $[\![-1, 1]\!]$. For $\varepsilon > 0$ we put $B_{\mathbb{C}}(0, \varepsilon) := \{z \in \mathbb{C} \ : \ |z| < \varepsilon\}$.

2 Cones and dominating maps

In this section we introduce the basic concepts and tools of our method of invariant cones to locate the eigenspaces and bound the spectrum for matrices. For this end we modify the concept of cones from [5]. Our approach is strongly motivated by the methods from the theory of hyperbolic dynamical systems, in particular by the results of Newhouse [6], who obtained conditions for hyperbolic splitting on compact invariant set for a diffeomorphism in terms of its induced action on a cone-field and its complement.

Definition 2.1. By a *cone-space* we understand a finite dimensional Banach space E with semi-norms $\rangle \cdot \langle$ (we call it *contracting*), $\langle \cdot \rangle$ (which we call *expanding*) such that

$$|||\mathsf{x}||| := \max(\,\rangle \mathsf{x} \langle, \langle \mathsf{x} \rangle\,)$$

defines an equivalent norm on E. By the *r-norm* for $r > 0$ on the cone-space E we take

$$|||\mathsf{x}|||_r := \max(\,\rangle \mathsf{x} \langle, r \cdot \langle \mathsf{x} \rangle\,).$$

Definition 2.2. Let E be a cone-space. We define the *r-contracting cone* in E by

$$\rangle E \langle_r := \{\mathsf{x} \in E : \,\rangle \mathsf{x} \langle \geq r \langle \mathsf{x} \rangle\},$$

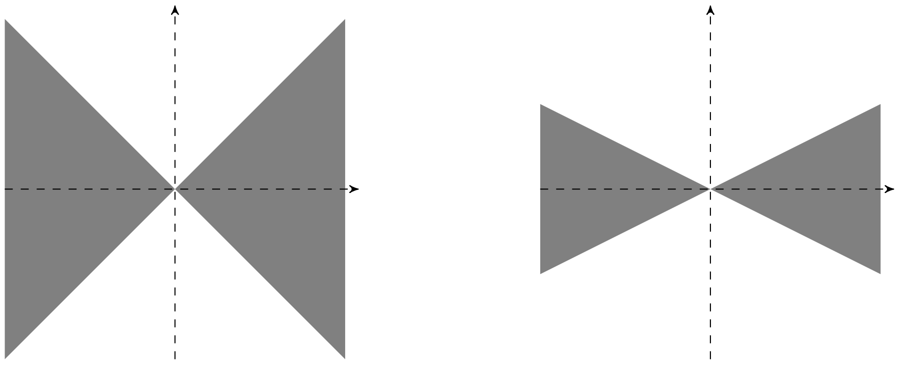
and the *r-expanding cone* in E by

$$\langle E \rangle_r := \{\mathsf{x} \in E : \,\rangle \mathsf{x} \langle \leq r \langle \mathsf{x} \rangle\}.$$

Note that

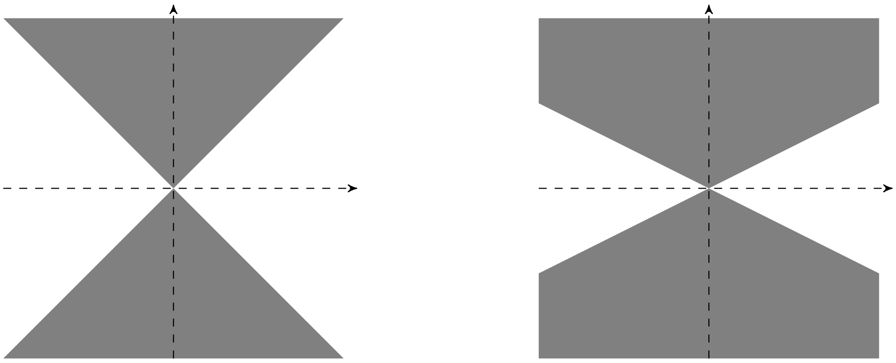
$$E = \rangle E \langle_r \cup \langle E \rangle_r. \tag{3}$$

In the same way we define r-contracting cone and r-expanding cone in subspace of E. If $r = 1$ we will omit the subscript r, in particular we speak of contracting cone. We introduce the scaling by r of semi-norms to have a better control over size of the cones (see Figure 2), which will consequently allow us to better locate the eigenvectors.



(a) The contracting cone in $\mathbb{R} \times \mathbb{R}$.

(b) The 2-contracting cone in $\mathbb{R}\times\mathbb{R}$.



(c) The expanding cone in $\mathbb{R}\times\mathbb{R}$.

(d) The 2-expanding cone in $\mathbb{R}\times\mathbb{R}$.

Figure 2: The cones in the cone-space $\mathbb{R}\times\mathbb{R}$.

If E has a fixed product structure $E = E_1 \times E_2$, we introduce a natural cone-space structure on E by defining seminorms

$$\rangle \mathsf{x} \langle := \|x_1\|, \quad \langle \mathsf{x}\rangle := \|x_2\| \quad \text{for } \mathsf{x} = (x_1, x_2) \in E_1 \times E_2.$$

In the proof of our main result, Theorem 3.3, the following proposition will play a crucial role.

Proposition 2.3. *Let $E = E_1 \times E_2$ be a cone-space and let $r > 0$ be given. Assume that we have direct sum decomposition $E = V_1 \oplus V_2$ such that*

$$V_1 \subset \rangle E \langle_r \quad \textit{and} \quad V_2 \subset \langle E\rangle_r.$$

Then $\dim V_1 = \dim E_1$ *and* $\dim V_2 = \dim E_2$.

Proof. Let $n := \dim E_1$ and $m := \dim E_2$. First we show that $\dim V_1 \leq n$. For an indirect proof, assume that $\dim V_1 > n$. Then there exist linearly independent vectors $\mathsf{v_1}, \ldots, \mathsf{v_{n+1}} \in V_1$. Obviously $\mathsf{v_i} = (w_i, z_i)$ for $i \in \{1, \ldots, n+1\}$ and unique $w_i \in E_1$, $z_i \in E_2$. Since $w_1, \ldots, w_{n+1} \in E_1$ and $\dim E_1 = n$ there exist a set of $n+1$ scalars, $\alpha_1, \ldots, \alpha_{n+1}$, not all zero, such that

$$\alpha_1 w_1 + \ldots + \alpha_{n+1} w_{n+1} = 0.$$

Note that

$$\mathsf{z} := \alpha_1 z_1 + \ldots + \alpha_{n+1} z_{n+1} \neq 0,$$

because otherwise the vectors $\mathsf{v_1}, \ldots, \mathsf{v_{n+1}}$ would not be linearly independent. Consequently we obtain

$$(0, \mathsf{z}) = \left(\sum_{i=1}^{n+1} \alpha_i w_i, \sum_{i=1}^{n+1} \alpha_i z_i \right) \in V_1 \subset \rangle E \langle_r,$$

and thus $r\|\mathsf{z}\| \leq \|0\|$, which implies that $\mathsf{z} = 0$. We get a contradiction with the fact the sequence of vectors $\mathsf{v_1}, \ldots, \mathsf{v_{n+1}}$ is linearly independent.

The proof that $\dim V_2 \leq m$ is analogous. Finally, since $\dim E = n + m$ and $\dim V_1 \leq n$, $\dim V_2 \leq m$ we obtain

$$\dim V_1 = n, \quad \text{and} \quad \dim V_2 = m.$$

□

By an *operator* we mean a linear mapping between cone-spaces E and F. We denote the space of all operators by $\mathcal{L}(E, F)$. If $F = E$, we denote $\mathcal{L}(E, E)$ by $\mathcal{L}(E)$.

Let $A \in \mathcal{L}(E, F)$. We define

$$\rangle A \langle_r := \inf\{R \in \mathbb{R}_+ \mid |||A\mathsf{x}|||_r \leqslant R\, |||\mathsf{x}|||_r \text{ for all } \mathsf{x} \in E : A\mathsf{x} \in \rangle F \langle_r\}, \tag{4}$$

$$\langle A\rangle_r := \sup\{R \in \mathbb{R}_+ \mid |||A\mathsf{x}|||_r \geqslant R\, |||\mathsf{x}|||_r \text{ for all } \mathsf{x} \in E : \mathsf{x} \in \langle E\rangle_r\}. \tag{5}$$

The following lemma is obvious.

Lemma 2.4. *Let $A \in \mathcal{L}(E,F)$.*

$$\rangle A\langle_r = (\inf\{|||\mathsf{x}\,|||_r \mid A\mathsf{x} \in \rangle F\langle_r,\ |||A\mathsf{x}|||_r = 1\})^{-1} \text{ when } A \text{ is invertible,} \tag{6}$$

$$\langle A\rangle_r = \inf\{|||A\mathsf{x}\,|||_r \mid \mathsf{x} \in \langle E\rangle_r,\ |||\mathsf{x}|||_r = 1\}. \tag{7}$$

Remark 2.5. *Observe, that*

$$|||A\mathsf{x}|||_r \leqslant \rangle A\langle_r |||\mathsf{x}|||_r \quad \text{for} \quad \mathsf{x} \in A^{-1}\rangle F\langle_r,$$

$$|||A\mathsf{x}|||_r \geqslant \langle A\rangle_r\, |||\,\mathsf{x}|||_r \quad \text{for} \quad \mathsf{x} \in \langle E\rangle_r.$$

The above definitions of $\rangle A\langle_r$ and $\langle A\rangle_r$ are modifications of analogous notions in [6], where $\langle A\rangle$ is called the expansion rate and $1/\rangle A\langle$ is the co-expansion rate. Using of those rates we can generalize the classical dominating maps which are relevant to our research.

Definition 2.6. We say that $A \in \mathcal{L}(E,F)$ is *r-dominating*, if

$$\rangle A\langle_r < \langle A\rangle_r.$$

By $\mathcal{D}_r(E,F)$ we denote the set of all $A \in \mathcal{L}(E,F)$ which are r-dominating. If $F = E$, we denote the space $\mathcal{D}_r(E,E)$ by $\mathcal{D}_r(E)$.

Observation 2.7. *Let $\tilde{E} \subset E$, $\tilde{F} \subset F$ be subspaces and let $A \in \mathcal{L}(E,F)$ be such that $A(\tilde{E}) \subset \tilde{F}$. Then $A|_{\tilde{E}} \in \mathcal{L}(\tilde{E},\tilde{F})$ and*

$$\rangle A|_{\tilde{E}}\langle_r \leq \rangle A\langle_r \quad \text{and} \quad \langle A\rangle_r \leq \langle A|_{\tilde{E}}\rangle_r.$$

Moreover, if $A \in \mathcal{D}_r(E,F)$ then $A \in \mathcal{D}_r(\tilde{E},\tilde{F})$.

Proof. It is a consequence of (4), (5) and Definition 2.6. □

It turns out that r-cones are invariant for r-dominant operators.

Theorem 2.8. *Let $A \in \mathcal{D}_r(E,F)$ and let $\mathsf{v} \in E$ be arbitrary. Then*

$$\mathsf{v} \in \langle E\rangle_r \implies A\mathsf{v} \in \langle F\rangle_r,$$

$$A\mathsf{v} \in \rangle F\langle_r \implies \mathsf{v} \in \rangle E\langle_r.$$

Proof. The proof is a simple modification of the proof of [5, Proposition 2.1]. □

As a consequence of the above theorem we obtain that composition of r-dominating maps is r-dominating. Moreover, we get estimate for expansion and contraction rates.

Proposition 2.9. *Let $A \in \mathcal{D}_r(F,G)$ and $B \in \mathcal{D}_r(E,F)$. Then $A \circ B \in \mathcal{D}_r(E,G)$ and*

$$\rangle A \circ B\langle_r \leq \rangle A\langle_r \cdot \rangle B\langle_r, \ \ \langle A \circ B\rangle_r \geq \langle A\rangle_r \cdot \langle B\rangle_r. \tag{8}$$

Proof. To prove the first inequality from (8), consider an $\mathsf{x} \in E$ such that $(A \circ B)(\mathsf{x}) \in \rangle G\langle_r$. From (4) and Theorem 2.8 we know that $B\mathsf{x} \in \rangle F\langle_r$, and thus we have

$$|||A \circ B(\mathsf{x})|||_r \leq \rangle A\langle_r \cdot |||\,B\mathsf{x}|||_r \leq \rangle A\langle_r \cdot \rangle B\langle_r \cdot |||\,\mathsf{x}\,|||_r\ .$$

Hence

$$\rangle A \circ B\langle_r \leq \rangle A\langle_r \cdot \rangle B\langle_r.$$

Using (5) and Theorem 2.8, we obtain the second inequality from (8).

As a simple consequence of (8) we obtain $A \circ B \in \mathcal{D}_r(E,G)$. □

In the remainder of this section we show how to estimate $\rangle A\langle_r$, $\langle A\rangle_r$. Consider two cone-spaces $E = E_1 \times E_2$ and $F = F_1 \times F_2$. Let $A\colon E \to F$ be an operator given in the matrix form by

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

By

$$|||A|||_r := \max\left(\|A_{11}\| + \frac{1}{r}\|A_{12}\|, r\|A_{21}\| + \|A_{22}\|\right)$$

we define the *r-norm* of operator A, where $\|.\|$ is an operator norm. Observe that it satisfies

$$|||A\mathsf{x}|||_r \leq |||A\,|||_r \cdot |||\,\mathsf{x}\,|||_r \qquad \text{for } \mathsf{x} \in E.$$

Note that in general it is not (except for the case when E_1 is one dimensional) the operator norm for $|||\cdot|||_r$.

Theorem 2.10. *Let $A = [A_{ij}]_{1\leq i,j\leq 2} \in \mathcal{L}(E_1 \times E_2, F_1 \times F_2)$ and $r \in (0,\infty)$ be given.*

1) We have

$$\rangle A\langle_r \leq \|A_{11}\| + \frac{1}{r}\|A_{12}\|.$$

2) Additionally, if A_{22} is invertible, then

$$\langle A\rangle_r \geq \|A_{22}^{-1}\|^{-1} - r\|A_{21}\|.$$

Proof. For the proof of the first inequality, we take $\mathsf{x} = (x_1, x_2) \in E_1 \times E_2$ such that $A\mathsf{x} \in \rangle F\langle_r$. From Definition 2.2 we have

$$\|A_{11}x_1 + A_{12}x_2\| \geq r\|A_{21}x_1 + A_{22}x_2\|, \tag{9}$$

and therefore

$$\begin{aligned} |||A\mathsf{x}|||_r &= \max(\|A_{11}x_1 + A_{12}x_2\|, r\|A_{21}x_1 + A_{22}x_2\|) \\ &\overset{by\,(9)}{=} \|A_{11}x_1 + A_{12}x_2\| \leq \|A_{11}\| \cdot \|x_1\| + \frac{1}{r}\|A_{12}\| \cdot r\|x_2\| \\ &\leq (\|A_{11}\| + \frac{1}{r}\|A_{12}\|) \cdot |||\mathsf{x}\,|||_r\,. \end{aligned}$$

For the proof of the second inequality, suppose that $\mathsf{x} = (x_1, x_2) \in \langle E\rangle_r$, where $x_1 \in E_1$, $x_2 \in E_2$. Then

$$\|x_1\| \leq r\|x_2\| = |||\mathsf{x}\,|||_r\,. \tag{10}$$

We know that

$$\|A_{22}x_2\| \geq \|A_{22}^{-1}\|^{-1}\|x_2\| \geq 0. \tag{11}$$

Finally, we obtain

$$\begin{aligned} |||A\mathsf{x}|||_r &\geq r\|A_{21}x_1 + A_{22}x_2\| \geq r\|A_{22}x_2\| - r\|A_{21}x_1\| \\ &\overset{by\,(11)}{\geq} r\|A_{22}^{-1}\|^{-1}\|x_2\| - r\|A_{21}\|\|x_1\| \overset{by\,(10)}{\geq} \left(\|A_{22}^{-1}\|^{-1} - r\|A_{21}\|\right) \cdot |||\mathsf{x}\,|||_r\,. \end{aligned}$$

□

Example 2.11. Let us verify that the matrix $A \in \mathcal{L}(\mathbb{C}\times\mathbb{C}, \mathbb{C}\times\mathbb{C})$, $A = \begin{bmatrix} 2 & 1.5 \\ 1 & 5 \end{bmatrix}$ is dominating. By Theorem 2.10 we have $\rangle A\langle \leq 3.5 < 4 \leq \langle A\rangle$, and therefore A is dominating.

Let us stress that the estimates from Theorem 2.10 are sharp, but there are cases when we do not have equalities in them.

Example 2.12. Let $A \in \mathcal{L}(\mathbb{C} \times \mathbb{C}, \mathbb{C} \times \mathbb{C})$ be given by the formula $A = \begin{bmatrix} 2 & 3 \\ 2 & 5 \end{bmatrix}$. We show that A is dominating. Observe that Theorem 2.10 does not allow us to decide whether this matrix A is dominating, since

$$\rangle A \langle \leq \|A_{11}\| + \|A_{12}\| = 5 \text{ and } \langle A \rangle \geq \|A_{22}^{-1}\|^{-1} - \|A_{21}\| = 3.$$

We calculate exactly $\rangle A \langle$ and $\langle A \rangle$ (we take the norm $\| \cdot \|_\infty$) from the formulas (6) and (7). The minimum of (7) is realized in points $(1, -1)^T$ and $(-1, 1)^T$. It is easy to see that the matrix A is invertible, so minimum of (6) is realized in points $(\frac{1}{2}, 0)^T$ and $(-\frac{1}{2}, 0)^T$. Hence

$$\rangle A \langle = 2 \quad \text{and} \quad \langle A \rangle = 3.$$

Finally, we obtain that A is dominating. Observe that for this example Gerschgorin theorem does not hold (it is impossible to separate Gerschgorin disks).

3 Localization of eigenspaces based on cones and dominating maps

In this section we show that the eigenspaces of the r-dominating operator A lie in the corresponding r-cones. Moreover, we can estimate $\sigma(A)$ with the help of $\rangle A \langle_r$, $\langle A \rangle_r$.

Lemma 3.1. *Let $A \in \mathcal{D}_r(E)$. Then*

$$\lambda \in \sigma(A) \implies |\lambda| \in [0, \rangle A \langle_r] \cup [\langle A \rangle_r, \infty). \tag{12}$$

Moreover $[0, \rangle A \langle_r] \cap [\langle A \rangle_r, \infty) = \emptyset$.

Proof. Since $A \in \mathcal{D}_r(E)$ we get $[0, \rangle A \langle_r] \cap [\langle A \rangle_r, \infty) = \emptyset$.

Now we show implication (12). Let λ be an eigenvalue of A and let $\mathsf{x} \in E$ be a corresponding eigenvector. By (3) we know that $\mathsf{x} \in \rangle E \langle_r \cup \langle E \rangle_r$. We consider two cases. First suppose that $\mathsf{x} \in \rangle E \langle_r$. Since x is an eigenvector, $A\mathsf{x} = \lambda \mathsf{x}$, and thus $A\mathsf{x} \in \rangle E \langle_r$. By (4) we get

$$|\lambda| \leq \rangle A \langle_r.$$

Now suppose that $\mathsf{x} \in \langle E \rangle_r$. By (5) we get

$$|\lambda| \geq \langle A \rangle_r,$$

which completes the proof. □

Let E be a finite dimensional vector space over the field $\mathbb{C}$ and let operator $A \colon E \to E$ be given. One can easily deduce from the Jordan theorem (see also [4, Appendix to Chapter 4] for the general case) that if $\sigma(A) = \sigma_1 \cup \sigma_2$ then there is a unique direct sum decomposition $E = E_{\sigma_1} \oplus E_{\sigma_2}$ such that $A(E_{\sigma_1}) \subset E_{\sigma_1}$, $A(E_{\sigma_2}) \subset E_{\sigma_2}$ and $\sigma(A|_{E_{\sigma_1}}) = \sigma_1$, $\sigma(A|_{E_{\sigma_2}}) = \sigma_2$. For $0 < c < d$ we define

$$E_{\leq c} := E_{\{\lambda \,:\, |\lambda| \leq c\}} \quad \text{and} \quad E_{\geq d} := E_{\{\lambda \,:\, |\lambda| \geq d\}}.$$

Theorem 3.2. *Let E be a finite dimensional cone-space and let $A \in \mathcal{D}_r(E)$. Then there is a direct sum decomposition $E = E_{\leq \rangle A \langle_r} \oplus E_{\geq \langle A \rangle_r}$ which satisfies*

$$E_{\leq \rangle A \langle_r} \subset \rangle E \langle_r, \; E_{\geq \langle A \rangle_r} \subset \langle E \rangle_r.$$

Proof. From Lemma 3.1 and the comments preceding our theorem we obtain a decomposition of E into A-invariant subspaces

$$E = E_{\leq \rangle A \langle_r} \oplus E_{\geq \langle A \rangle_r},$$

such that

$$\sigma(A|_{E_{\leq \rangle A \langle_r}}) = \{\lambda : |\lambda| \in [0, \rangle A \langle_r]\} \quad \text{and} \quad \sigma(A|_{E_{\geq \langle A \rangle_r}}) = \{\lambda : |\lambda| \in [\langle A \rangle_r, \infty)\}.$$

Now we show $E_{\leq \rangle A \langle_r} \subset \rangle E \langle_r$. Consider an arbitrary $\mathsf{x} \in E_{\leq \rangle A \langle_r}$. The case when $\mathsf{x} = 0$ is obvious. Assume that $\mathsf{x} \neq 0$. Without any loss of the generality we can assume that $\|\mathsf{x}\| = 1$. For an indirect proof, assume that $\mathsf{x} \notin \rangle E \langle_r$. Then by (3) we get $\mathsf{x} \in \langle E \rangle_r$. Let $\varepsilon > 0$ be arbitrary. From the fact that $\mathsf{x} \in E_{\leq \rangle A \langle_r}$, we know that

$$\limsup_{m\to+\infty} \sqrt[m]{|||A|^m_{E_{\leq \rangle A \langle_r}}|||} = \sup \sigma(A|_{E_{\leq \rangle A \langle_r}}) \leq \rangle A \langle_r. \tag{13}$$

Note that inequality (13) holds for all norms. For all $x \in E_{\leq \rangle A \langle_r}$ we obtain

$$\limsup_{m\to+\infty} \sqrt[m]{|||A^m \mathsf{x}|||} \leq \rangle A \langle_r,$$

and thus there exists an $M \in \mathbb{N}$ such that for all $m \in \mathbb{N}$

$$m \geq M \Rightarrow \sqrt[m]{|||A^m \mathsf{x}|||} \leq \rangle A \langle_r + \varepsilon.$$

Since $\mathsf{x} \in \langle E \rangle_r$ and from Theorem 2.8 we obtain

$$\mathsf{x} \in \langle E \rangle_r \Rightarrow A\mathsf{x} \in \langle E \rangle_r \Rightarrow \cdots \Rightarrow A^m \mathsf{x} \in \langle E \rangle_r.$$

Using (5) and Remark 2.5 we get

$$\begin{gathered} |||A\mathsf{x}||| \geq \langle A \rangle_r \; |||\, \mathsf{x}|||, \\ |||A^2\mathsf{x}||| = |||A(A\mathsf{x})||| \geq \langle A \rangle_r \; |||\, A\mathsf{x}||| \geq \langle A \rangle_r^2 \; |||\, \mathsf{x}|||, \\ \vdots \\ |||A^m\mathsf{x}||| \geq \langle A \rangle_r^m \; |||\, \mathsf{x} \,||| \, . \end{gathered}$$

Finally we have

$$\langle A \rangle_r = \sqrt[m]{\langle A \rangle_r^m} \leq \sqrt[m]{|||A^m \mathsf{x}|||} \leq \rangle A \langle_r + \varepsilon.$$

Since ε was arbitrary, we get a contradiction with the fact that A is r-dominating.

Analogously, to prove inclusion $E_{\geq \langle A \rangle_r} \subset \langle E \rangle_r$, assume that $\mathsf{x} \in E_{\geq \langle A \rangle_r}$ and $\mathsf{x} \notin \langle E \rangle_r$. Then $\mathsf{x} \in \rangle E \langle_r$. Since $\sigma(A|_{E_{\geq \langle A \rangle_r}}) = \sigma_{\geq \langle A \rangle_r} := \{\lambda \;:\; |\lambda| \geq \langle A \rangle_r\}$ and $0 \notin \sigma_{\geq \langle A \rangle_r}$ we know that $A|_{E_{\geq \langle A \rangle_r}} : E_{\geq \langle A \rangle_r} \to E_{\geq \langle A \rangle_r}$ is invertible. Let $\varepsilon > 0$ be arbitrary. Using the fact that $\mathsf{x} \in E_{\geq \langle A \rangle_r}$, by dual result (13), we know that

$$\limsup_{m\to+\infty} \sqrt[m]{|||A|^{-m}_{E_{\geq \langle A \rangle_r}} \mathsf{x}|||} \leq \langle A \rangle_r^{-1},$$

and thus there exists an $M \in \mathbb{N}$ such that for all $m \in \mathbb{N}$

$$m \geq M \Rightarrow \sqrt[m]{|||A|^{-m}_{E_{\geq \langle A \rangle_r}} \mathsf{x}|||} \leq \langle A \rangle_r^{-1} + \varepsilon. \tag{14}$$

From the Observation 2.7 and Theorem 2.8 we get

$$\mathsf{x} \in \rangle E_{\geq \langle A \rangle_r} \langle_r \Rightarrow A|^{-1}_{E_{\geq \langle A \rangle_r}} \mathsf{x} \in \rangle E_{\geq \langle A \rangle_r} \langle_r \Rightarrow \cdots \Rightarrow A|^{-m}_{E_{\geq \langle A \rangle_r}} \mathsf{x} \in \rangle E_{\geq \langle A \rangle_r} \langle_r,$$

and from (4) and Remark 2.5 we have

$$\begin{gathered} |||\mathsf{x}||| \leq \rangle A|_{E_{\geq \langle A \rangle_r}} \langle_r |||A|^{-1}_{E_{\geq \langle A \rangle_r}} \mathsf{x}|||, \\ |||A|^{-1}_{E_{\geq \langle A \rangle_r}} \mathsf{x}||| \leq \rangle A|_{E_{\geq \langle A \rangle_r}} \langle_r |||A|^{-2}_{E_{\geq \langle A \rangle_r}} \mathsf{x}|||, \\ \vdots \\ |||A|^{-m+1}_{E_{\geq \langle A \rangle_r}} \mathsf{x}||| \leq \rangle A|_{E_{\geq \langle A \rangle_r}} \langle_r |||A|^{-m}_{E_{\geq \langle A \rangle_r}} \mathsf{x} \,||| \, . \end{gathered}$$

Hence

$$|||\, \mathsf{x}||| \leq (\, \rangle A|_{E_{\geq \langle A \rangle_r}} \langle_r)^m \; |||\, A|^{-m}_{E_{\geq \langle A \rangle_r}} \mathsf{x} \,||| \, . \tag{15}$$

Finally from the Observation 2.7 and (14), (15) we obtain

$$\rangle A \langle_r \geq \rangle A|_{E_{\geq \langle A \rangle_r}} \langle_r = \sqrt[m]{(\, \rangle A|_{E_{\geq \langle A \rangle_r}} \langle_r)^m} \geq \sqrt[m]{\frac{1}{|||A|^{-m}_{E_{\geq \langle A \rangle_r}} \mathsf{x}|||}} \geq \frac{1}{\langle A \rangle_r^{-1} + \varepsilon} = \langle A \rangle_r \cdot \frac{1}{1 + \varepsilon \cdot \langle A \rangle_r},$$

which gives a contradiction with the fact that A is r-dominating. □

Now we are ready to state the main result on the eigenspaces and eigenvalue location using our method of cones and dominating maps.

Theorem 3.3. *Let $E = E_1 \times E_2$ be a finite dimensional cone-space and let $A \in \mathcal{D}_r(E)$. Then there exists a unique direct sum decomposition $E = F_1 \oplus F_2$ of A-invariant subspaces F_1, F_2 such that*

$$\sigma(A|_{F_1}) \subset \overline{B}(0, \rangle A\langle_r), \quad \sigma(A|_{F_2}) \subset \mathbb{C} \setminus B(0, \langle A\rangle_r).$$

Moreover, we have:

1) $\mathrm{dim}F_1 = \mathrm{dim}E_1,\ \mathrm{dim}F_2 = \mathrm{dim}E_2$*,*

2) $F_1 \subset \rangle E\langle_r$ *and* $F_2 \subset \langle E\rangle_r$*,*

3) $\|A|_{F_1}\| \leq \rangle A\langle_r$ *and* $\|(A|_{F_2})^{-1}\| \leq \langle A\rangle_r^{-1}$*.*

Proof. From Theorem 3.2 we know that exists a unique direct sum decomposition $E = E_{\leq \rangle A\langle_r} \oplus E_{\geq \langle A\rangle_r}$ which satisfies

$$E_{\leq \rangle A\langle_r} \subset \rangle E\langle_r,\ E_{\geq \langle A\rangle_r} \subset \langle E\rangle_r.$$

We take $F_1 = E_{\leq \rangle A\langle_r}$ and $F_2 = E_{\geq \langle A\rangle_r}$. By Proposition 2.3 we obtain $\mathrm{dim}F_1 = \mathrm{dim}E_1$ and $\mathrm{dim}F_2 = \mathrm{dim}E_2$.

Now we show that $\sigma(A|_{F_1}) \subset \overline{B}(0, \rangle A\langle_r)$. Let $\mathsf{x} \in F_1$ be an eigenvector of A and let λ be the eigenvalue of A corresponding to x. Since x is an eigenvector ($A\mathsf{x} = \lambda\mathsf{x}$) and $F_1 \subset \rangle E\langle_r$ therefore $A\mathsf{x} \in \rangle E\langle_r$. By (4) we obtain that $|\lambda| \leq \rangle A\langle_r$, so we get $\sigma(A|_{F_1}) \subset \overline{B}(0, \rangle A\langle_r)$.

Now suppose that $\mathsf{x} \in F_2$. Since $F_2 \subset \langle E\rangle_r$ and by (5) we get $|\lambda| \geq \langle A\rangle_r$. Hence $\sigma(A|_{F_2}) \subset \mathbb{C} \setminus B(0, \langle A\rangle_r)$.

The inequalities of item 3) we obtain from (4) and (5). □

As a direct consequence of the above theorem we obtain the following conclusion.

Corollary 3.4. *Let $r \in (0, \infty)$ and $n \in \mathbb{N}$. Assume that an operator $A \in \mathcal{D}_r(\mathbb{C} \times \mathbb{C}^{n-1})$ is given. Then there exists unique eigenvalue λ of A such that $|\lambda| \leq \rangle A\langle_r$ and the eigenspace corresponding to λ is one-dimensional. The unique (after rescaling) eigenvector x corresponding to the eigenvalue λ satisfies*

$$\mathsf{x} \in (1, 0, \ldots, 0)^T + \{0\} \times \overline{B}_{\mathbb{C}}(0, 1/r)^{n-1} \subset (1, 0, \ldots, 0)^T + \frac{1}{r} \cdot (0, \mathbb{I}, \ldots, \mathbb{I})^T + \frac{1}{r} \cdot (0, \mathbb{I}, \ldots, \mathbb{I})^T i.$$

Proof. It is a direct consequence of Theorem 3.3 and Definition 2.2. □

Because at the origin of our approach based on cones and dominating maps is the theory of hyperbolic dynamical systems, so our method should be well suited to locate the eigenspaces and eigenvalues of products of many matrices. In the example below we contrast our approach with a naive approach, which tries to diagonalize a matrix obtained as a product of many matrices. The essential feature of this example is that the matrices we multiply are known with some accuracy only.

Example 3.5. Let the matrices $A_i \in \mathcal{L}(\mathbb{R} \times \mathbb{R})$, $i \in \{1, \ldots, 15\}$ be such that

$$A_i \in \begin{bmatrix} \llbracket 0, 0.5 \rrbracket & \varepsilon\mathbb{I} \\ \varepsilon\mathbb{I} & \llbracket 1.5, 2 \rrbracket \end{bmatrix},$$

where $\varepsilon = 0.01$ and $\mathbb{I} = \llbracket -1, 1 \rrbracket$. Consider the matrix $B := A_{15} \cdot \ldots \cdot A_1$.

From Theorem 2.10 we obtain that $A_i \in \mathcal{D}(\mathbb{R} \times \mathbb{R})$ and

$$\rangle A_i\langle \leq 0.5 + \varepsilon,\ \langle A_i\rangle \geq 1.5 - \varepsilon.$$

From Theorem 2.8 and Proposition 2.9 we conclude that $B \in \mathcal{D}(\mathbb{R} \times \mathbb{R})$ and

$$\rangle B\langle \leq \rangle A_{15}\langle \cdot \ldots \cdot \rangle A_1\langle, \quad \langle B\rangle \geq \langle A_{15}\rangle \cdot \ldots \cdot \langle A_1\rangle.$$

From Theorem 3.3 we obtain that eigenvalues λ_1 and λ_2 of B such that

$$|\lambda_1| \leq (0.5 + \varepsilon)^{15} \text{ and } |\lambda_2| \geq (1.5 - \varepsilon)^{15}.$$

Now, a naive method will ask first for a computation of B. Using interval arithmetic we obtained

$$B \in \left[\begin{array}{cc} [\![-1.45687, 1.45693]\!] & [\![-218.543, 218.544]\!] \\ [\![-218.543, 218.544]\!] & [\![433.611, 32782.94]\!] \end{array} \right].$$

However, there exists matrix B_1 within the bounds given above, which has both eigenvalues larger than 1. For example, let us consider

$$B_1 = \left[\begin{array}{cc} 1 & 100 \\ -100 & 521 \end{array} \right].$$

This matrix have the eigenvalues $\lambda_1 = 21$ and $\lambda_2 = 501$. Consequently, this means that none of the methods applied to the product matrix will not give us the expected estimation $|\lambda_1| < 1$ and $|\lambda_2| > 1$.

4 Estimations of the eigenvalues and eigenvectors

In this section we develop computable estimates for the eigenvalues and eigenspaces based on the results from the previous section.

Lemma 4.1. *Let $A \in \mathcal{L}(E_1 \times E_2)$ be given such that*

$$A := \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

If A_{22} is invertible, $d = \|A_{22}^{-1}\|^{-1} - \|A_{11}\| > 0$ and $\Delta := d^2 - 4\|A_{12}\|\|A_{21}\| > 0$ then

$$A \in \mathcal{D}_r(E_1 \times E_2) \quad \textit{for} \quad \begin{cases} r \in \left(\dfrac{d - \sqrt{\Delta}}{2\|A_{21}\|}, \dfrac{d + \sqrt{\Delta}}{2\|A_{21}\|} \right) & \textit{if } \|A_{21}\| \neq 0 \\ r \in \left(\dfrac{\|A_{12}\|}{d}, \infty \right) & \textit{if } \|A_{21}\| = 0 \end{cases}.$$

Proof. Let $a := \|A_{12}\|$, $b := \|A_{11}\|$ and $c := \|A_{21}\|$. Making use of Theorem 2.10 it suffices to show that

$$b + \frac{a}{r} < (d + b) - cr.$$

Multiplying both sides of the above inequality by the positive number r we get the inequality

$$cr^2 - dr + a < 0. \tag{16}$$

If $c = 0$ then we get $r > \frac{a}{d}$. Suppose now that $c \neq 0$. Since from our assumption follows that $\Delta > 0$ we see inequality (16) is satisfied for

$$r \in \left(\frac{d - \sqrt{\Delta}}{2c}, \frac{d + \sqrt{\Delta}}{2c} \right).$$

□

Remark 4.2. *Let A be an operator, which satisfies the assumptions of Lemma 4.1 (in particular $\Delta > 0$). Let $a := \|A_{12}\|$, $b := \|A_{11}\|$ and $c := \|A_{21}\| \neq 0$. It is easy to see, that*

$$\frac{d - \sqrt{\Delta}}{2c} < \frac{d}{2c} < \frac{d + \sqrt{\Delta}}{2c} < \frac{d}{c}.$$

Therefore, if A satisfies the assumptions of Lemma 4.1 and $\|A_{21}\| \neq 0$ and we want to find possibly largest r for which A is r-dominating, then we can take $r = \frac{d}{2\|A_{21}\|}$. With this choice we have $r < r_{max} < 2r$, where r_{max} is the supremum the set of r's obtained in the above lemma, therefore we might not be optimal, but we obtain easily manageable expression.

We present now our main result on the location of an isolated eigenvalue and its eigenspace.

Theorem 4.3. *Let $A = [a_{ij}]_{1\leq i,j\leq n} \in \mathcal{L}(\mathbb{C}\times\mathbb{C}^{n-1})$ be given in the block from by*

$$A := \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{11} = a_{11}$. Assume that $A_{22} - a_{11}\cdot I_{\mathbb{C}^{n-1}}$ is invertible and $\|(A_{22} - a_{11}\cdot I_{\mathbb{C}^{n-1}})^{-1}\|^{-2} - 4\|A_{12}\|\|A_{21}\| > 0$. Then

1) there exists a unique eigenvalue λ of A which satisfies

$$|\lambda - a_{11}| \leq 2\|A_{12}\|\cdot\|A_{21}\|\cdot\|(A_{22} - a_{11}\cdot I_{\mathbb{C}^{n-1}})^{-1}\|,$$

2) the eigenspace corresponding to λ is one-dimensional and there exist unique δ_2, …, $\delta_n \in \mathbb{C}$,

$$\|(0,\delta_2,\dots,\delta_n)^T\| \leq 2\|A_{21}\|\cdot\|(A_{22} - a_{11}\cdot I_{\mathbb{C}^{n-1}})^{-1}\|\cdot\|(1,0,\dots,0)^T\|$$

such that $(1,\delta_2,\dots,\delta_n)^T$ is the eigenvector corresponding to λ.

Proof. It is easy to see that if $A_{21} = 0$, then theorem holds. Therefore we will assume that $\|A_{21}\| > 0$.

In order to apply Lemma 4.1 to matrix $A - a_{11}I_{\mathbb{C}^n}$ we set $a := \|A_{12}\|$, $c := \|A_{21}\|$ and $d = \|(A_{22} - a_{11}\cdot I_{\mathbb{C}^{n-1}})^{-1}\|^{-1}$. By Lemma 4.1 and Remark 4.2 we get $A - a_{11}\cdot I_{\mathbb{C}^n} \in \mathcal{D}_{d/(2c)}(\mathbb{C}\times\mathbb{C}^{n-1})$, and from Corollary 3.4 and Theorem 2.10 we conclude that there exists a unique eigenvalue λ of A which satisfies

$$|\lambda - a_{11}| < \rangle A - a_{11}I\langle_{\frac{d}{2c}} \leq \frac{1}{\frac{d}{2c}}\|A_{12}\| = \frac{2ac}{d}.$$

From Theorem 3.3 (second point) we know that eigenspace, which contains eigenvector corresponding to the λ, lies in $\rangle\mathbb{C}\times\mathbb{C}^{n-1}\langle_{\frac{d}{2c}}$. Hence (see Definition 2.2) we obtain unique δ_2, …, $\delta_n \in \mathbb{C}$, $\|(0,\delta_2,\dots,\delta_n)^T\| \leq 2\|A_{21}\|\cdot\|(A_{22} - a_{11}\cdot I_{\mathbb{C}^{n-1}})^{-1}\|\cdot\|(1,0,\dots,0)^T\|$ such that $(1,\delta_2,\dots,\delta_n)^T$ is the eigenvector corresponding to λ. □

Let us stress here that in the proof Theorem 4.3 through Lemma 4.1 we used estimates for $\rangle A\langle$ and $\langle A\rangle$ provided by Theorem 2.10, which may fail establish that a matrix is dominating for a dominating matrix. If this is the case we will use Theorem 3.3. This happens in Examples 5.4 and 5.5.

The following lemma shows how $\|(A - zI)^{-1}\|^{-1}$ can be computed in arbitrary norm, when A is close to the diagonal matrix.

Lemma 4.4. *Let $n \in \mathbb{N}$, $z \in \mathbb{C}$ and $A \in \mathbb{C}^{n\times n}$ be given. Let A be decomposed into $A = J + E$ where J is a diagonal matrix and E equals zero on the diagonal. Assume that $J - z\cdot I_{\mathbb{C}^n}$ is invertible and $\|(J - z\cdot I_{\mathbb{C}^n})^{-1}\|^{-1} - \|E\| > 0$. Then*

$$\|(A - z\cdot I_{\mathbb{C}^n})^{-1}\|^{-1} \geq \|(J - z\cdot I_{\mathbb{C}^n})^{-1}\|^{-1} - \|E\|.$$

Proof. It is well-known that for an invertible operator B we have

$$(B - C)^{-1} = \sum_{n=0}^{\infty}(B^{-1}C)^n B^{-1} \quad \text{for } C \in \mathbb{C}^{n\times n} \ : \ \|C\| < 1/\|B^{-1}\|.$$

Hence, if $\|C\| < 1/\|B^{-1}\|$, then

$$\|(B - C)^{-1}\| \leq \frac{\|B^{-1}\|}{1 - \|B^{-1}\|\cdot\|C\|},$$

so we obtain

$$\|(B - C)^{-1}\|^{-1} \geq \frac{1}{\|B^{-1}\|}(1 - \|B^{-1}\|\cdot\|C\|) = \frac{1}{\|B^{-1}\|} - \|C\|. \tag{17}$$

From (17) applied to $B = J - zI_{\mathbb{C}^n}$ and $C = -E$ we get assertion of the lemma. □

Now we present results about the location of the eigenspaces.

Theorem 4.5. *Let $k, n \in \mathbb{N}$ such that $0 \le k \le n$ and $A \in \mathcal{L}(\mathbb{C}^k \times \mathbb{C}^{n-k})$ be given in the block from by*

$$A := \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{11} \in \mathcal{L}(\mathbb{C}^k)$, $A_{12} \in \mathcal{L}(\mathbb{C}^k, \mathbb{C}^{n-k})$, $A_{21} \in \mathcal{L}(\mathbb{C}^{n-k}, \mathbb{C}^k)$ and $A_{22} \in \mathcal{L}(\mathbb{C}^{n-k})$. Assume that A_{22} is invertible, $d := \|A_{22}^{-1}\|^{-1} - \|A_{11}\| > 0$ and $d^2 - 4\|A_{12}\|\|A_{21}\| > 0$. Then there exists a unique direct sum decomposition $\mathbb{C}^k \times \mathbb{C}^{n-k} = F_1 \oplus F_2$, such that F_1 and F_2 are A-invariant subspaces F_1, F_2, $\dim F_1 = k$, $\dim F_2 = n-k$ and

$$\begin{aligned} F_1 &\subset \left\{ (x_1, x_2) \in \mathbb{C}^k \times \mathbb{C}^{n-k} : \|x_2\| \le \frac{2\|A_{21}\|}{d}\|x_1\| \right\}, \\ F_2 &\subset \left\{ (x_1, x_2) \in \mathbb{C}^k \times \mathbb{C}^{n-k} : \frac{2\|A_{21}\|}{d}\|x_1\| \le \|x_2\| \right\}. \end{aligned} \tag{18}$$

Moreover, we have

$$\sigma(A|_{F_1}) \subset \overline{B}\left(0, \|A_{11}\| + \frac{2\|A_{12}\| \cdot \|A_{21}\|}{d}\right), \quad \sigma(A|_{F_2}) \subset \mathbb{C} \setminus B\left(0, \|A_{22}^{-1}\|^{-1} - \frac{d}{2}\right). \tag{19}$$

Proof. Let $c := \|A_{21}\|$. If $c = 0$, the assertion holds. Assume that $c \neq 0$. By Lemma 4.1 we get $A \in \mathcal{D}_{d/(2c)}(\mathbb{C}^k \times \mathbb{C}^{n-k})$, and from Theorem 3.3 we know that exists a direct sum decomposition $\mathbb{C}^k \times \mathbb{C}^{n-k} = F_1 \oplus F_2$ such that $\dim F_1 = k$, $\dim F_2 = n-k$ and F_1, F_2 are invariant. The properties (18) and (19) are consequences of Theorem 3.3 and Theorem 2.10 and Definition 2.2, respectively. □

Corollary 4.6. *We use the same notation and decomposition of the matrix A as in Theorem 4.5. Assume that for some $z \in \mathbb{C}$ matrices $A_{11} - zI_{\mathbb{C}^k}$, $A_{22} - zI_{\mathbb{C}^{n-k}}$ are invertible and $d := \|(A_{22} - zI_{\mathbb{C}^{n-k}})^{-1}\|^{-1} - \|A_{11} - zI_{\mathbb{C}^k}\| > 0$, $d^2 - 4\|A_{12}\|\|A_{21}\| > 0$. Then there exists a unique direct sum decomposition $\mathbb{C}^k \times \mathbb{C}^{n-k} = F_1 \oplus F_2$ into A-invariant subspaces F_1, F_2 such that* $\dim F_1 = k$, $\dim F_2 = n-k$ *and*

$$\begin{aligned} F_1 &\subset \left\{ (x_1, x_2) \in \mathbb{C}^k \times \mathbb{C}^{n-k} : \|x_2\| \le \frac{2\|A_{21}\|}{d}\|x_1\| \right\}, \\ F_2 &\subset \left\{ (x_1, x_2) \in \mathbb{C}^k \times \mathbb{C}^{n-k} : \frac{2\|A_{21}\|}{d}\|x_1\| \le \|x_2\| \right\}. \end{aligned}$$

Moreover, we have

$$\begin{aligned} \sigma(A|_{F_1}) &\subset \overline{B}\left(z, \|A_{11} - zI_{\mathbb{C}^k}\| + \frac{2\|A_{12}\| \cdot \|A_{21}\|}{d}\right), \\ \sigma(A|_{F_2}) &\subset \mathbb{C} \setminus B\left(z, \|(A_{22} - zI_{\mathbb{C}^{n-k}})^{-1}\|^{-1} - \frac{d}{2}\right). \end{aligned}$$

4.1 Gerschgorin theorems

For to the convenience of the reader, in this section we recall the Gerschgorin theorem and its modifications.

We have a matrix A which has a block structure

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \cdots & \dots & \dots & \dots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix},$$

where A_{ij} are matrices and A_{ii} are square matrices.

Let $V = \bigoplus_{i=1}^n V_i$, where V_i are finite dimensional vector spaces over $\mathbb{C}$, and $A : V \to V$ be decomposed into blocks $A_{ij} : V_j \to V_i$ $i, j = 1, 2, \dots, n$, so that for $v = v_1 + \dots + v_n$, where $v_i \in V_i$ holds

$$A(v_1 + \dots + v_n) = \sum_i \sum_j A_{ij} v_j. \tag{20}$$

We define Gerschgorin disks $G_i(A)$ for the block matrix A by

$$\begin{aligned} R_i(A) &= \sum_{j,j\neq i} \|A_{ij}\|, \\ G_i(A) &= \{\lambda \in \mathbb{C} \ : \ A_{ii} - \lambda I_i \text{ exists and } \|(A_{ii} - \lambda I_i)^{-1}\|^{-1} \leq R_i(A)\}, \quad i = 1, \dots, n, \end{aligned}$$

where I_{V_i} is an identity map on V_i. If A is known from the context, then we will usually drop A and write just R_i and G_i. Similarly, we write I instead of I_{V_i}.

Theorem below we present the generalizations of Gershgorin Theorems due to Feingold and Varga [2].

Theorem 4.7. *[2, Theorem 2]*

$$\sigma(A) \subset \bigcup_{i=1}^{n} G_i.$$

Theorem 4.8. *[2, Theorem 4] Assume that $J \subset \{1, \dots, n\}$ is such that*

$$\left(\bigcup_{j\in J} G_j\right) \cap \left(\bigcup_{j\notin J} G_j\right) = \emptyset.$$

Then the number of eigenvalues of A (counting with multiplicities) contained in $\left(\bigcup_{j\in J} G_j\right)$ is equal to $\sum_{j\in J} dim V_j$.

Now we give a theorem about the location of the eigenvectors based on the Wilkinson argument [8].

Theorem 4.9. *Assume that for some $j \in \{1, \dots, n\}$*

$$G_j \cap G_k = \emptyset, \quad \textit{for } k = 1, 2, \dots, n, \ k \neq j.$$

Then if $v = (v_1 + \dots + v_n)$ is an eigenvector corresponding to $\lambda \in G_j$, then $\|v_k\| \leq \|v_j\|$ for $k = 1, \dots, n$.

Proof. To show that $\|v_k\| \leq \|v_j\|$ we will reason by the contradiction. Assume that for some $i \neq 0$ holds $\|v_i\| \geq \|v_k\|$ for $k = 1, \dots, n$ and $\|v_i\| > \|v_j\|$. We will apply the basic argument from the generalized Gerschgorin theorem (Theorem 4.7) to prove that $\lambda \in G_i$. This will lead to a contradiction, because $\lambda \in G_j$, hence $\lambda \in G_j \cap G_i \neq \emptyset$.

We have

$$\begin{aligned} \lambda v_i &= A_{ii} v_i + \sum_{k\neq i} A_{ik} v_k \\ (\lambda I - A_{ii}) v_i &= \sum_{k\neq i} A_{ik} v_k \\ \|(\lambda I - A_{ii})^{-1}\|^{-1} \|v_i\| &\leq \sum_{k\neq i} \|A_{ik}\| \|v_k\| \\ \|(\lambda I - A_{ii})^{-1}\|^{-1} &\leq \sum_{k\neq i} \|A_{ik}\| \frac{\|v_k\|}{\|v_i\|} \leq \sum_{k\neq i} \|A_{ik}\| \end{aligned}$$

hence $\lambda \in G_i$. We obtained the contradiction. This finishes the proof. □

One of the easiest ways to improve the estimation of the eigenvalues from the Gerschgorin theorem is through the scaling the basis of our domain. This approach is well known and can be found in the original article of Gerschgorin [3].

Assume, that we have matrix $A \in \mathbb{C}^{n\times n}$ and let $\mathsf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ such that $x_i > 0$ for all $i \in \{1, \dots, n\}$. With this vector x we define the matrix $X \in \mathbb{R}^{n\times n}$ with the elements of x on the leading diagonal, and 0 elsewhere. Note, that the matrix X is nonsingular and matrix $X^{-1}AX$ is similar to A therefore $\sigma(X^{-1}AX) = \sigma(A)$. If $A = [a_{ij}]_{1\leq i,j\leq n}$, then

$$X^{-1}AX = \left[\frac{a_{ij}x_j}{x_i}\right]_{1\leq i,j\leq n}$$

and

$$G_i = \overline{B}\Big(a_{ii}, \sum_{j\neq i} \frac{|a_{ij}|x_j}{x_i}\Big) \qquad \text{for } i = 1,\dots,n.$$

4.2 Example

In the following example we consider a matrix with multi-dimensional block for which we estimate eigenspaces.

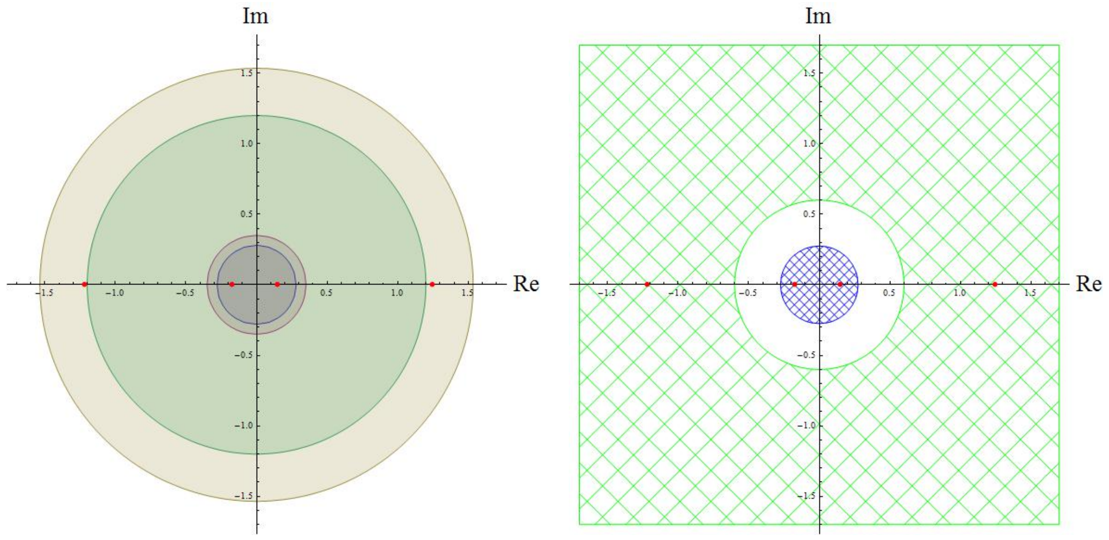
Example 4.10. Consider the matrix $A \in \mathcal{L}(\mathbb{C}^2 \times \mathbb{C}^2)$ be given by

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 0. & 0.15 & 0.11 & 0.02 \\ 0.2 & 0. & 0.1 & 0.05 \\ \hline 0.01 & 0.025 & 0. & 1.5 \\ 0.15 & 0.05 & 1. & 0. \end{array}\right].$$

We have $\|A_{11}\|_\infty = 0.2$, $\|A_{12}\|_\infty = 0.15$, $\|A_{21}\|_\infty = 0.2$. From Theorem 4.5 ($d = \|(A_{22}^{-1}\|_\infty^{-1} - \|A_{11}\|_\infty = 1 - 0.2 = 0.8 > 0$ and $d^2 - 4\|A_{12}\|_\infty\|A_{21}\|_\infty = 0.52 > 0$) we know that there exist eigenspaces F_1 and F_2, which satisfy

$$\begin{aligned} F_1 &\subset \big\{(x_1,x_2) \in \mathbb{C}^2 \times \mathbb{C}^2 : \|x_2\| \leq 0.5\|x_1\|\big\}, \\ F_2 &\subset \big\{(x_1,x_2) \in \mathbb{C}^2 \times \mathbb{C}^2 : \|x_1\| \leq 2\|x_2\|\big\}. \end{aligned}$$

and $\sigma(A_{F_1}) \subset \overline{B}(0, 0.275)$, $\sigma(A_{F_2}) \subset \mathbb{C} \setminus B(0, 0.6)$ (see Figure 3b).



(a) Gerschgorin circles.

(b) Estimates based on Theorem 4.5. The white annulus in does not contain any eigenvalue.

Figure 3: Gerschgorin and our circles with approximate eigenvalues in Example 4.10.

Observe that when using the Gerschgorin theorem with one-dimensional blocks with scalings, as described at the end of Section 4.1, we will not be able to separate the spectrum of A, because the centers of Gerschgorin circles are located at zero.

Now we discuss what happens when we use the generalized Gerschgorin theorems from [2]. First rescale the matrix A by $X = \begin{bmatrix} 2 & 0 \\ 0 & I \end{bmatrix}$ (we take the same rescaling as in our method, see Remark 4.2) to get

$$\tilde{A} = X^{-1}AX = \begin{bmatrix} A_{11} & \frac{1}{2}A_{12} \\ 2A_{21} & A_{22} \end{bmatrix}.$$

We use the Theorems 4.7 and 4.8 applied to the above block decomposition, and obtain the generalized Gerschgorin disks:

$$\begin{aligned} G_1(\tilde{A}) &= \left\{ \lambda \in \mathbb{C} : \|(A_{11} - \lambda I)^{-1}\|_\infty^{-1} \leq \frac{1}{2}\|A_{12}\|_\infty \right\}, \\ G_2(\tilde{A}) &= \left\{ \lambda \in \mathbb{C} : \|(A_{22} - \lambda I)^{-1}\|_\infty^{-1} \leq 2\|A_{21}\|_\infty \right\}. \end{aligned}$$

We want to show that $G_1(\tilde{A}) \cap G_2(\tilde{A}) = \emptyset$. Let us we check that $G_1(\tilde{A}) \subset \overline{B}(0, 0.25)$. We have

$$(A_{11} - \lambda I)^{-1} = \frac{1}{\lambda^2 - 0.03} \begin{bmatrix} -\lambda & -0.15 \\ -0.2 & -\lambda \end{bmatrix},$$

so we get

$$\|(A_{11} - \lambda I)^{-1}\|_\infty^{-1} = \frac{|\lambda^2 - 0.03|}{0.2 + |\lambda|}.$$

For $\lambda \in G_1(\tilde{A}) \subset \mathbb{C}$ we have

$$\frac{|\lambda^2 - 0.03|}{0.2 + |\lambda|} \leq 0.075.$$

Performing simple mathematical operations and changing the coordinate system to the polar one we obtain

$$40000r^4 - 15r(160r\cos(2q) + 15r + 6) + 27 \leq 0, \quad r = |\lambda| \in [0, \infty),\ \varphi \in [0, 2\pi).$$

Solving the above inequality we get

$$\sup r = \frac{3}{80}\left(1 + \sqrt{33}\right) < \frac{21}{80}.$$

This means that $G_1(\tilde{A}) \subset \overline{B}(0, 21/80)$. Now we show that $\lambda \notin G_2(\tilde{A})$ for an arbitrary $\lambda \in \overline{B}(0, 21/80)$.

Indeed we have

$$(A_{22} - \lambda I)^{-1} = \frac{1}{\lambda^2 - 1.5} \begin{bmatrix} -\lambda & -1.5 \\ -1 & -\lambda \end{bmatrix}.$$

It is easy to see that for $\lambda \in \overline{B}(0, 21/80)$ we have

$$\|(A_{22} - \lambda I)^{-1}\|_\infty < \frac{\frac{3}{2} + \frac{21}{80}}{\frac{3}{2} - \left(\frac{21}{80}\right)^2} = \frac{3760}{3053}.$$

Hence

$$\|(A_{22} - \lambda I)^{-1}\|_\infty^{-1} > \frac{3053}{3760} > \frac{80}{100}, \qquad \lambda \in G_1(\tilde{A}) \subset \overline{B}(0, 21/80).$$

Finally, we get $G_1(\tilde{A}) \cap G_2(\tilde{A}) = \emptyset$ (see Figure 4) and therefore we obtain from Theorem 4.7 and 4.8 that two eigenvalues belong to $G_1(\tilde{A})$ while the remaining two eigenvalues are inside $G_2(\tilde{A})$. As one can see, we get better estimation for eigenvalues close to 0 from generalized Gerschgorin theorem with scaling $r = 2$, than from Theorem 4.5 but by generalized Gerschgorin theorem we can not get eigenspaces.

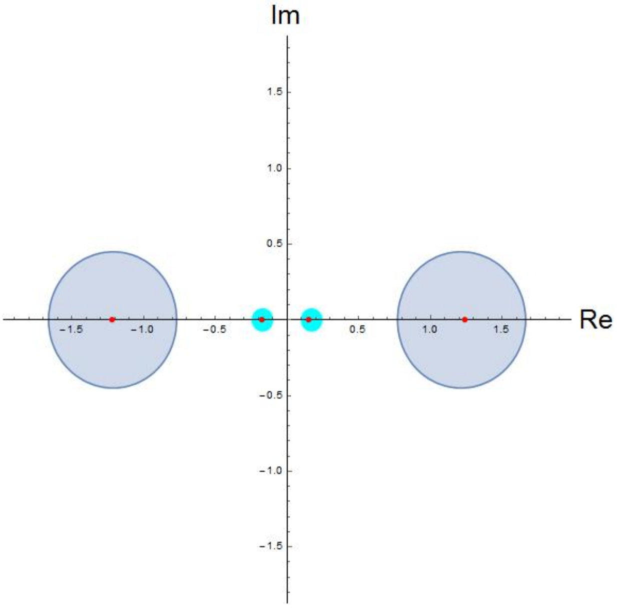


Figure 4: Generalized Gerschgorin circles: $G_1(\tilde{A})$ – greater circles and $G_2(\tilde{A})$ – smaller ones in Example 4.10 (compare Fig. 3b).

5 Comparisons in the case of the isolated Gerschgorin disk

In this section we compare our method of cones with the Gerschgorin theorem with rescaling of the basis, when trying to estimate an eigenvalue in an isolated Gerschgorin disk and corresponding eigenvector. Throughout this section we will use the $\|\cdot\|_\infty$ norm.

5.1 The isolation of first Gerschgorin disk implies that the matrix $A - a_{11}I$ is dominating

When applying Theorem 4.8 with the splitting $\mathbb{C}\oplus\mathbb{C}^{n-1}$ we will have two generalized Gerschgorin disks

$$\begin{aligned} G_1(A) &= \overline{B}(a_{11}, \|A_{12}\|_\infty) = \overline{B}(a_{11}, \sum_{j\neq 1}|a_{1j}|), \\ G_2(A) &= \{\lambda\in\mathbb{C} \ : \ \|(A_{22}-\lambda I)^{-1}\|_\infty^{-1} \leq \max_{j=2,\ldots,n}|a_{j1}|\}. \end{aligned}$$

Now we develop computable bounds for $G_2(A)$.

Lemma 5.1. *Let $A = [a_{ij}] \in \mathbb{C}^{n\times n}$. Then*

$$\min_i(|a_{ii}| - \sum_{j\neq i}|a_{ij}|) \leq \sup\{\lambda\in\mathbb{R} \mid \forall x\in\mathbb{C}^n \quad \|Ax\|_\infty \geq \lambda\|x\|_\infty\}. \tag{21}$$

If A is invertible, then

$$\min_i(|a_{ii}| - \sum_{j\neq i}|a_{ij}|) \leq \frac{1}{\|A^{-1}\|_\infty}. \tag{22}$$

Proof. Let

$$S := \min_i(|a_{ii}| - \sum_{j\neq i}|a_{ij}|). \tag{23}$$

Let us take any $x\in\mathbb{C}^n$, such that $\|x\| = 1$. Let i be such that $|x_i| = 1$. We have

$$|(Ax)_i| \geq (|a_{ii}||x_i| - \sum_{j\neq i}|a_{ij}|\cdot|x_j|) \geq (|a_{ii}| - \sum_{j\neq i}|a_{ij}|) \geq S > 0.$$

Hence

$$\|Ax\| \geq S.$$

This establishes (21)

For the second part observe that, if A is invertible, then

$$\sup\{\lambda \in \mathbb{R} \mid \forall x \in \mathbb{C}^n \quad \|Ax\|_\infty \geq \lambda \|x\|_\infty\} = \frac{1}{\|A^{-1}\|_\infty}. \tag{24}$$

□

From Lemma 5.1 it follows that

$$\begin{aligned} G_2(A) \quad &\subset \quad \{\lambda \in \mathbb{C} \mid \min_{i=2,\dots,n} (|a_{ii} - \lambda| - \sum_{j \neq 1,i} |a_{ij}|) \leq \max_{j=2,\dots,n} |a_{j1}|\} \\ &= \quad \{\lambda \in \mathbb{C} \mid \exists i = 2, \dots, n \ |a_{ii} - \lambda| \leq \sum_{j \neq 1,i} |a_{ij}| + \max_{j=2,\dots,n} |a_{j1}|\}. \end{aligned}$$

So we see that $G_1(A) \cap G_2(A) = \emptyset$ if the following condition holds

$$|a_{11} - a_{ii}| > \sum_{j \neq 1} |a_{1j}| + \sum_{j \neq 1,i} |a_{ij}| + \max_{j=2,\dots,n} |a_{j1}| \quad \text{for all } i = 2, \dots, n. \tag{25}$$

If we will use the classical Gerschgorin theorem, i.e. blocks are one-dimensional, then to have $G_1 \cap G_i = \emptyset$ for

$$|a_{11} - a_{ii}| > R_1 + R_i = \sum_{j \neq 1} |a_{1j}| + \sum_{j \neq i} |a_{ij}| \quad \text{for all } i = 2, \dots, n. \tag{26}$$

Observe that in both cases we have the same Gerschgorin disk G_1, so the bound for the first eigenvalue will be the same, provided we have empty intersections with other disks. Observe that (25) implies (26).

Now we show one of the main results of this paper, which states that if a matrix $A = [a_{ij}]$ has an isolated Gerschgorin disk G_1, then $A - a_{11}I$ is dominating (relative to the splitting $\mathbb{C} \times \mathbb{C}^{n-1}$) and under very mild assumptions the bound obtained from the method of cones is better that the one from the Gershgorin theorem.

Theorem 5.2. *Let $A \in \mathbb{C}^{n \times n}$ be given by the formula*

$$A = \left[\begin{array}{c|ccc} a_{11} & a_{12} & \dots & a_{1n} \\ \hline a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

Assume that the matrix A satisfies inequality (26). *Then the matrix $A - a_{11}I$ is dominating. Moreover, we have*

$$\rangle A - a_{11}I \langle \ \leq \sum_{j \neq 1} |a_{1j}| < \langle A - a_{11}I \rangle.$$

and if $\sum_{j \neq 1} |a_{1j}| > 0$, then

$$\rangle A - a_{11}I \langle \ < \sum_{j \neq 1} |a_{1j}|.$$

Proof. Without any loss of the generality we can assume that $a_{11} = 0$. Let us denote $V = \mathbb{C} \oplus \mathbb{C}^{n-1}$

In order to estimate $\rangle A \langle$ we will first find a bound for the set Z of x, such that $A\mathsf{x} \in \rangle V \langle$. Then we will compute $\rangle A \langle$ on Z.

Let

$$\delta_k = |a_{kk}| - \sum_{j \neq k} |a_{kj}| - \sum_{j \neq 1} |a_{1j}|. \tag{27}$$

From our assumptions it follows that

$$\delta = \min_{k=2,\ldots,n} \delta_k > 0.$$

Let $\epsilon > 0$ be such that

$$\epsilon |a_{k1}| < \delta_k, \quad k = 2, \ldots, n. \tag{28}$$

Assume now that $\mathsf{x} = (x_1, \mathsf{x}_2)$ such that $|x_1| \leq (1+\epsilon)\|\mathsf{x}_2\|_\infty$. We will show that $Ax \notin \rangle V \langle$.

We can assume that $\|x_2\|_\infty = 1$ and $|x_k| = 1$. Then we have

$$|(Ax)_k| \geq |a_{kk}| - \sum_{j \notin \{1,k\}} |a_{kj}| - (1+\epsilon)|a_{k1}| = |a_{kk}| - \sum_{j \neq k} |a_{kj}| - \epsilon |a_{k1}|$$

$$\overset{by\,(27)}{=} \sum_{j \neq 1} |a_{1j}| + \delta_k - \epsilon |a_{k1}| \overset{by\,(28)}{>} \sum_{j \neq 1} |a_{1j}| \geq \|A_{12} x_2\| = \|(A\mathsf{x})_1\|_\infty.$$

Hence $A\mathsf{x} \notin \rangle V \langle$.

Therefore, if $A\mathsf{x} \in \rangle V \langle$, then $|x_1| > (1+\epsilon)\|\mathsf{x}_2\|_\infty$. In particular, we obtain

$$\text{if } A\mathsf{x} \in \rangle V \langle, \text{ then } \quad |||\, \mathsf{x} ||| = |x_1| > (1+\epsilon)\|\mathsf{x}_2\|_\infty. \tag{29}$$

Now we are ready to estimate $\rangle A \langle$. Let $\mathsf{x} = (x_1, \mathsf{x}_2)$ is such that $A\mathsf{x} \in \rangle V \langle$, then

$$|||A\mathsf{x}||| = \|A_{12}\mathsf{x}_2\|_\infty \leq \|A_{12}\|_\infty \cdot \|\mathsf{x}_2\|_\infty$$

$$\overset{by\,(29)}{\leq} \|A_{12}\|_\infty \frac{|||\mathsf{x}|||}{1+\epsilon} = \frac{1}{1+\epsilon} \sum_{j \neq 1} |a_{1j}| \; |||\, \mathsf{x} \,||| \; .$$

Hence $\rangle A \langle \; \leq \sum_{j \neq 1} |a_{1j}|$, but if $\sum_{j \neq 1} |a_{1j}| > 0$, then $\rangle A \langle \; < \sum_{j \neq 1} |a_{1j}|$.

Now we estimate $\langle A \rangle$. We will use Lemma 2.4. Let's take arbitrary $\mathsf{x} = (x_1, \mathsf{x}_2)$ such that $\|\mathsf{x}_2\|_\infty = 1$ and $|x_1| \leq 1$. Let $k = 2, \ldots, n$ be such that $|x_k| = 1$. From (26) we obtain

$$|(Ax)_k| \geq |a_{kk}| - \sum_{j \neq k} |a_{kj}| \overset{by\,(26)}{>} \sum_{j \neq 1} |a_{1j}|.$$

Hence $\|Ax\|_\infty > \sum_{j \neq 1} |a_{1j}|$. Therefore we have shown that

$$\langle A \rangle > \sum_{j \neq 1} |a_{1j}|.$$

□

Remark 5.3. *Observe that from Theorem 5.2 we know that our method of cones(i.e. Theorem 3.3) can be used for all matrices which have an isolated Gerschgorin disk. Moreover, we obtain*

$$|\lambda - a_{11}| \leq \; \rangle A - a_{11} I \langle \; \leq \frac{1}{1+\epsilon} R_1 = \frac{1}{1+\epsilon} \sum_{j \neq 1} |a_{1j}|.$$

This means that, if R_1 (the radius of the first Gerschogorin disk) is nonzero, then the estimate of the first eigenvalue from our method based on cones is better than the one obtained from the Gerschgorin theorem. This is also valid for all possible rescalings in the application of the Gerschogorin theorem, we should apply the same scaling in the method of cones.

5.2 Comparison of Theorem 4.3 with the Gerschgorin theorems

In the proof of Theorem 4.3 we applied Theorem 3.3 to the matrix $A - a_{11}I$ to estimate the size of the eigenvalue, λ_1, close to 0. We looked for possibly large parameter r, such that $A - a_{11}I$ is r-dominating and then we obtain

$$|a_{11} - \lambda_1| \leq \rangle A \langle_r \leq \frac{\|A_{12}\|}{r}.$$

This is exactly G_1 obtained from the Gerschgorin theorem for $\tilde{A}_r$.

The optimization with respect of r performed in the proof of the Theorem 4.3 to obtain the formula can be also repeated by suitable rescaling using the original Gerschgorin theorem as long $G_1(\tilde{A}_r)$ is disjoint from other Gerschgorin disks for $\tilde{A}_r$. Therefore both approaches differ only with the range of r's over which the optimization can be performed. In fact we are only interested in the upper bound for r in both methods.

Let $(1, \delta_2, \ldots, \delta_n)$ be the eigenvector corresponding to λ_1. We obtain from Theorem 3.2 the bound $1 \geq r\|(\delta_2, \ldots, \delta_n\|$, while from Theorem 4.9 applied to $\tilde{A}$ after returning to the original base we have $|\delta_i| \leq 1/r$. Hence the result is the same for the method based on cones and the Gerschgorin Theorem.

The example below demonstrates that it is possible to use the Gerschgorin theorem to isolate and estimate the eigenvector and eigenvalue, while assumptions of Theorem 4.3 and also assumptions of the generalized Gerschgorin Theorems 4.7 and 4.8 are not satisfied. This appears to contradict Theorem 5.2, but it does not, because in the proof Theorem 4.3 we used an expression for $\rangle A \langle$ from Lemma 2.10, which turns out to be an overestimation, see also Example 2.12. By Theorem 5.2 we know that the considered matrix is dominating, hence we can estimate the eigenpair using Theorem 3.3, see Example 5.5.

Example 5.4. Let $A \in \mathcal{L}(\mathbb{C} \times \mathbb{C}^2)$ be given by the formula

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{c|cc} 0 & 1 & 0 \\ \hline 0.5 & 2 & 0 \\ 50 & 0 & 100 \end{array}\right].$$

The classical Gerschgorin disks are

$$G_1 = \overline{B}(0,1), \quad G_2 = \overline{B}(2, 0.5), \quad G_3 = \overline{B}(100, 50).$$

It is clear that they are mutually disjoint, hence from the Gerschgorin theorem there exists an eigenvalue λ, $|\lambda| \leq 1$.

Now, we look at our Theorem 4.3 to estimate the eigenvalue close to 0. We have $A_{11} = 0$ and

$$\|A_{12}\|_\infty = 1, \quad \|A_{21}\|_\infty = 50, \quad \|(A_{22} - A_{11} \cdot I)^{-1}\|_\infty = 0.5,$$
$$\|(A_{22} - A_{11} \cdot I)^{-1}\|_\infty^{-2} - 4\|A_{12}\| \cdot \|A_{21}\|_\infty = 4 - 200 < 0.$$

Therefore assumptions of Theorem 4.3 are not satisfied.

Observe that also assumptions of the generalized Gerschgorin Theorems 4.7 and 4.8 for the decomposition given above are not satisfied. Our generalized Gerschgorin disks are

$$\begin{aligned} G_1(A) &= \overline{B}(0,1), \\ G_2(A) &= \{\lambda \,:\, \|(A_{22} - \lambda I)^{-1}\|_\infty^{-1} \leq 50\}. \end{aligned}$$

We have

$$(A_{22} - \lambda I)^{-1} = \frac{1}{(100-\lambda)(2-\lambda)} \begin{bmatrix} 100-\lambda & 0 \\ 0 & 2-\lambda \end{bmatrix},$$

hence

$$\|(A_{22} - \lambda I)^{-1}\|_\infty^{-1} = \min(|2-\lambda|, |100-\lambda|).$$

It is easy to see that $G_1(A) \cap G_2(A) \neq \emptyset$, therefore we cannot use these theorems.

Rescaling: When applying our method based on cones we should look for the largest r such that $\tilde{A}_r$ is 1-dominating and when using the Gerschgorin theorem we look for r such that $G_1(\tilde{A}_r)$ have empty intersection with others Gerschgorin circles for $\tilde{A}_r$.

For the Gerschorin disks we need to have the following inequalities

$$\frac{1}{r} < 2 - r/2, \quad \frac{1}{r} < 100 - 50r.$$

We obtain $\sup r = 1 + \sqrt{\frac{49}{50}} \approx 2$. Hence we obtain bound $|\lambda| \leq\approx 1/2$.

For the approach based on cones we need to find largest r, such that $\tilde{A}_r$ is 1-dominating. Using Theorem 2.10 we obtain the following condition

$$\frac{1}{r}\|A_{12}\| = \frac{1}{r} < \|A_{22}^{-1}\|^{-1} - r\|A_{21}\| = 2 - 50r.$$

Easy computations show that no such r exists in this case. Similar effect we get if we use the generalized Gerschgorin theorem.

In the following example we show that despite the fact that the matrix A from Example 5.4 does not satisfy the assumptions of Theorem 4.3, we can use our method of cones (we apply Theorem 3.3) to estimate the eigenvalue close to zero.

Example 5.5. Recall that $A \in \mathcal{L}(\mathbb{C} \times \mathbb{C}^2)$ of Example 5.4 was given by the formula

$$A = \left[\begin{array}{c|cc} 0 & 1 & 0 \\ \hline 0.5 & 2 & 0 \\ 50 & 0 & 100 \end{array}\right].$$

From Theorem 5.2 we know that the matrix A is dominating, so we can estimate the eigenvalue λ close to zero by $|\lambda| \le \rangle A\langle$ (see Theorem 3.3). From Lemma 2.4 we have

$$\rangle A\langle = \frac{1}{\min\left(\|x\|_\infty \text{ for } \mathsf{x} \in \mathbb{R}^3 \text{ such that } \|A\mathsf{x}\|_{>1} \le \|A\mathsf{x}\|_{\le 1} = 1\right)},$$

where $\|\mathsf{x}\|_{\le k} := \max_{i \le k} |x_i|$ and $\|\mathsf{x}\|_{>k} := \max_{i>k} |x_i|$ for $\mathsf{x} = (x_1, \ldots, x_k, \ldots, x_n) \in \mathbb{R}^n$, see (6).

The problem to calculate this constant comes down to solve simple optimization problem. We obtain

$$\min\big(\|x\|_\infty \text{ for } \mathsf{x} \in \mathbb{R}^3 \text{ such that } \|A\mathsf{x}\|_{>1} \le \|A\mathsf{x}\|_{\le 1} = 1\big) = 2.$$

This minimum is realized in the points $\big(-2, 1, \frac{99}{100}\big)^T$, $\big(-2, 1, \frac{101}{100}\big)^T$, $\big(2, -1, -\frac{101}{100}\big)^T$ and $\big(2, -1, -\frac{99}{100}\big)^T$. Hence $\rangle A\langle = \frac{1}{2}$. By Theorem 3.3 we get that eigenvalue close to zero satisfies

$$|\lambda| \le \rangle A\langle = \frac{1}{2}.$$

The bound $|\lambda| \le \frac{1}{2}$ can be obtained also from the Gerschgorin theorem, see Example 5.4 (*'Rescaling'*). Note that so far we did not improve the matrix A through the scaling $X = \begin{bmatrix} r & 0 \\ 0 & I \end{bmatrix}$ for $r \in (0, \infty)$. From Theorem 5.2 and again calculations from Example 5.4 (*'Rescaling'*) we know that our method work even if we rescale our matrix A by the matrix X for $r < 1 + \sqrt{\frac{49}{50}}$. For $r = \frac{9}{5}$ we obtain $|\lambda| \le \rangle A\langle = \frac{9}{26} < \frac{1}{2}$.

In the following two examples in view of the complicated mathematical calculations we will not try to apply the generalized Gerschgorin theorem (in both examples assumptions of Theorems 4.7 and 4.9 are satisfied). In the first example we construct a matrix such that the matrix $A - a_{11}I$ will be 1-dominating, while there will be no isolation of the first Gerschgorin disk.

Example 5.6. Let $A \in \mathcal{L}(\mathbb{C} \times \mathbb{C}^2)$ be given by the formula

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{c|cc} 0 & 0.75 & 0 \\ \hline \epsilon_1 & 1 & 0.5 \\ \epsilon_2 & 0.5 & 100 \end{array}\right],$$

where ϵ_1, ϵ_2 are sufficiently small. Observe that $G_1(A) \cap G_2(A) = \overline{B}(0, 0.75) \cap \overline{B}(1, 0.5 + \epsilon_1) \neq \emptyset$, hence the Gerschgorin theorem does not give us that $\lambda_1 \in G_1(A)$.

It is easy to see that $A - a_{11}I$ will be 1-dominating. Indeed from Theorem 2.10 we have

$$\rangle A\langle_1 \le \|A_{12}\| = 3/4, \quad \langle A\rangle_1 \ge \|A_{22}^{-1}\|^{-1} - \|A_{21}\| \approx 1.$$

Hence A is 1-dominating and Theorem 3.3 implies that $\lambda_1 \in G_1(A)$.

Rescaling: We set $\epsilon_1 = \epsilon_2 = 0.1$. We optimize by rescaling by r. The Gerschgorin disks approach leads to the following inequalities

$$\frac{3}{4r} < 0.5 - r/10.$$

There is no such r for which this holds.

The approach based on cones requires that

$$\frac{1}{r}\|A_{12}\| = \frac{3}{4r} < \|A_{22}^{-1}\|^{-1} - r\|A_{21}\| \approx 1 - \frac{r}{10}.$$

We obtain

$$\sup r \approx 5 + \sqrt{\frac{35}{2}}.$$

Hence we get $|\lambda| \leq 0.0817$.

The following example illustrates the case of the matrix A for which both methods discussed above can be applied.

Example 5.7. We put

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{c|cc} 0 & \frac{2}{5} & -\frac{1}{5} \\ \hline \frac{1}{5} & \frac{3}{2} & \frac{2}{5} \\ -\frac{1}{10} & \frac{3}{10} & 2 \end{array}\right].$$

First, by Theorem 4.3 we estimate the eigenvalue close to 0. We have $a_{11} = 0$ and

$$\|A_{12}\|_\infty = \frac{3}{5}, \quad \|A_{21}\|_\infty = \frac{1}{5}, \quad \|(A_{22} - a_{11}\cdot I)^{-1}\|_\infty = \frac{5}{6},$$
$$\|(A_{22} - a_{11}\cdot I)^{-1}\|_\infty^{-2} - 4\|A_{12}\|\cdot\|A_{21}\|_\infty = \frac{24}{25} > 0.$$

Therefore assumptions of Theorem 4.3 are satisfied and we obtain that the eigenvalue λ close to 0 satisfies $|\lambda| \leq \frac{1}{5}$.

Now we use the Gerschgorin theorems to estimate the eigenvalue close to 0. The Gerschgorin disks are

$$G_1(A) = \overline{B}\left(0, \frac{3}{5}\right), \quad G_2(A) = \overline{B}\left(\frac{3}{2}, \frac{3}{5}\right) \quad \text{and} \quad G_3(A) = \overline{B}\left(2, \frac{2}{5}\right).$$

It is easy to see that $G_1(A) \cap G_2(A) = \emptyset$, $G_1(A) \cap G_3(A) = \emptyset$ but we rescale the matrix A (with $r = 3$, which is the same rescaling as in our method), we obtain the matrix

$$\tilde{A}_r = \begin{bmatrix} 0 & \frac{2}{15} & -\frac{1}{15} \\ \frac{3}{5} & \frac{3}{2} & \frac{2}{5} \\ -\frac{3}{10} & \frac{3}{10} & 2 \end{bmatrix},$$

and consequently $G_1(\tilde{A}_r) \cap G_2(\tilde{A}_r) = \emptyset$ and $G_1(\tilde{A}_r) \cap G_3(\tilde{A}_r) = \emptyset$. Hence from the Gerschgorin theorem there exists an eigenvalue λ such that $|\lambda| \leq \frac{1}{5}$.

Rescaling: We look for the largest r for each method, which allows us to obtain the best estimation for the eigenvalue λ close to 0.

For Gerschgorin disks we need to solve the following inequalities

$$\frac{3}{2} > \frac{3}{5r} + \frac{r}{5} + \frac{2}{5}, \quad 2 > \frac{3}{5r} + \frac{r}{10} + \frac{3}{10}.$$

We obtain $\sup r = \frac{1}{4}\left(11 + \sqrt{73}\right)$. Hence we obtain bound $|\lambda| \leq\approx 0.1228$.

The cone based approach requires

$$\frac{1}{r}\|A_{12}\| = \frac{3}{5r} < \|A_{22}^{-1}\|^{-1} - r\|A_{21}\| = \frac{6}{5} - \frac{r}{5}.$$

We obtain $\sup r = 3 + \sqrt{6}$. Hence we obtain the bound $|\lambda| \leq\approx 0.110102$.

By doing the same calculations as above for the transpose of the matrix A we obtain

$$\sup r = \frac{1}{2}\left(3 + \sqrt{6}\right) \quad \text{and} \quad |\lambda| \leq\approx 0.110102,$$

from the classical Gerschgorin theorem, and for cone based we get

$$\sup r = \frac{1}{46}\left(72 + \sqrt{3597}\right),\ |\lambda| \leq\approx 0.104565.$$

As one can see the use of cone based approach gives us better estimation of the eigenvalue close to zero than the classical Gerschgorin theorem with rescaling.

Conclusions: As one can see from above examples and theorems our method is better than Gerschgorin theorem and its modifications. The main advantages of our method are:

- locate spectrum and eigenspaces of a matrix when multiple eigenvalues or clusters of very close eigenvalues are present,
- gives better estimation for isolated eigenvalues,
- allow to deal with composition of matrices.

References

[1] A. Brauer. Limits for the characteristic roots of a matrix. II. *Duke Mathematical Journal*, 14(1):21–26, 1947.

[2] D. G. Feingold and R. S. Varga. Block diagonally dominant matrices and generalizations of the gerschgorin circle theorem. *Pacific J. Math*, 12(4):1241–1250, 1962.

[3] S. Gerschgorin. Über die Abgrenzung der Eigenwerte einer Matrix. *Izv. Akad. Nauk. SSSR Ser. fiz-mat.*, 6:749–754, 1931.

[4] M. C. Irwin. *Smooth dynamical systems*, volume 94. Academic Press, 1980.

[5] T. Kułaga and J. Tabor. Hyperbolic dynamics in graph-directed IFS. *Journal of Differential Equations*, 251(12):3363–3380, 2011.

[6] Sh. Newhouse. Cone-fields, Domination, and Hyperbolicity. *Modern Dynamical Systems and Applications*, pages 419–433, 2004.

[7] R. S. Varga. *Geršgorin and his circles.* Springer, 2004.

[8] J.H. Wilkinson. Rigorous Error Bounds for Computer Eigensystems. *The Computer Journal*, 4(3):230–241, 1961.

[9] T. Yamamoto. Error bounds for computed eigenvalues and eigenvectors. *Numerische Mathematik*, 34(2):189–199, 1980.