Filomat 35:4 (2021), 1205–1214
https://doi.org/10.2298/FIL2104205G

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

On Similarity of an Arbitrary Matrix to a Block Diagonal Matrix

Michael Gil'[a]

[a] *Department of Mathematics*
Ben Gurion University of the Negev
P.0. Box 653, Beer-Sheva 84105, Israel

Abstract. Let an $n \times n$ -matrix A have $m < n$ ($m \geq 2$) different eigenvalues λ_j of the algebraic multiplicity μ_j ($j = 1, ..., m$). It is proved that there are $\mu_j \times \mu_j$-matrices A_j, each of which has a unique eigenvalue λ_j, such that A is similar to the block-diagonal matrix $\hat{D} = \text{diag}\ (A_1, A_2, ..., A_m)$. I.e. there is an invertible matrix T, such that $T^{-1}AT = \hat{D}$. Besides, a sharp bound for the number $\kappa_T := \|T\|\|T^{-1}\|$ is derived. As applications of these results we obtain norm estimates for matrix functions non-regular on the convex hull of the spectra. These estimates generalize and refine the previously published results. In addition, a new bound for the spectral variation of matrices is derived. In the appropriate situations it refines the well known bounds.

1. Introduction

Let $\mathbb{C}^n$ be the n-dimensional complex Euclidean space with a scalar product $(.,.)$, the Euclidean norm $\|.\| = \sqrt{(.,.)}$ and unit matrix I. $\mathbb{C}^{n_1 \times n_2}$ denotes the set of all complex $n_1 \times n_2$-matrices.

For an $A \in \mathbb{C}^{n \times n}$, $\sigma(A)$ denotes the spectrum, $\|A\|$ is the spectral norm, i.e. the operator norm with respect to the Euclidean vector norm; A^* is the adjoint matrix; $\|A\|_F = (\text{trace}\ A^*A)^{1/2}$ is the Frobenius norm; λ_j ($j = 1, ..., m$) ($m \geq 2$) are the different eigenvalues of A enumerated in an arbitrary way; μ_j is the algebraic multiplicity of λ_j. So

$$\delta := \min_{j,k=1,...,m;\ k \neq j} |\lambda_j - \lambda_k| > 0 \tag{1.1}$$

and $\mu_1 + ... + \mu_m = n$. The aim of this paper is to show that there are matrices $A_j \in \mathbb{C}^{\mu_j \times \mu_j}$ ($j = 1, ..., m$) and an invertible matrix $T \in \mathbb{C}^{n \times n}$, such that

$$T^{-1}AT = \hat{D}, \text{ where } \hat{D} = \text{diag}\,(A_1, A_2, ..., A_m). \tag{1.2}$$

Besides, each block A_j has the unique eigenvalue λ_j of the algebraic multiplicity μ_j ($j = 1, ..., m$). In addition, we obtain an estimate for the (block-condition) number $\kappa_T := \|T\|\|T^{-1}\|$ and consider some applications of these results.

The paper consists of 7 sections. In Section 2, the preliminary results are presented. The main result of this paper-Theorem 3.1 is formulated in Section 3. The proof of Theorem 3.1 is divided into a series of lemmas which are presented in Sections 4 and 5. In Section 6 we discuss applications of Theorem 3.1. In

2010 *Mathematics Subject Classification*. Primary 15A04; Secondary 15A42, 15A18
Keywords. matrices; similarity; condition number; operator functions; matrix function; resolvent: spectrum perturbation
Received: 28 March 2020; Accepted: 06 July 2020
Communicated by Dragan S. Djordjević
Email address: gilmi@bezeqint.net (Michael Gil')

particular, we obtain norm estimates for matrix functions non-regular on the convex hull of the spectra and generalize the inequalities for functions of diagonalizable matrices. In addition, we obtain a bound for the spectral variation of two matrices, which refines the Elsner result, cf. [24, p. 168]. In Section 7 an illustrative example is given

2. Preliminary results

Let $\hat{\lambda}_k$ $(k = 1, ..., n)$ be all the eigenvalues of A taken with the multiplicities and enumerated in the following way:

$$\hat{\lambda}_1 = \hat{\lambda}_2 = ... = \hat{\lambda}_{\mu_1} = \lambda_1,$$

$$\hat{\lambda}_{\mu_1+1} = \hat{\lambda}_{\mu_1+2} = ... = \hat{\lambda}_{\mu_1+\mu_2} = \lambda_2, ...,$$

$$\hat{\lambda}_{\mu_1+\mu_2+...+\mu_{m-1}+1} = \hat{\lambda}_{\mu_1+\mu_2+...+\mu_{m-1}+2} = ... = \hat{\lambda}_{\mu_1+\mu_2+...+\mu_m} = \lambda_m.$$

By the Schur theorem [19] for any matrix $A \in \mathbb{C}^{n\times n}$ there is a non-unique unitary transform, such that A can be reduced to the triangular form:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & ... & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & a_{23} & ... & a_{2,n-1} & a_{2n} \\ . & . & . & ... & . & \\ 0 & 0 & 0 & ... & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & ... & 0 & a_{nn} \end{pmatrix}.$$

Besides, the diagonal entries are the eigenvalues enumerated as

$$a_{11} = a_{22} = ... = a_{\mu_1,\mu_1} = \lambda_1,$$

$$a_{\mu_1+1,\mu_1+1} = a_{\mu_1+2,\mu_1+2} = ... = a_{\mu_1+\mu_2,\mu_1+\mu_2} = \lambda_2, ...$$

$$a_{\mu_1+\mu_2+...+\mu_{m-1}+1,\mu_1+\mu_2+...+\mu_{m-1}+1} = a_{\mu_1+\mu_2+...+\mu_{m-1}+2,\mu_1+\mu_2+...+\mu_{m-1}+2}$$

$$= ... = a_{\mu_1+\mu_2+...+\mu_m,\mu_1+\mu_2+...+\mu_m} = \lambda_m.$$

Let $\{e_k\}_{k=1}^n$ be the corresponding orthonormal basis of the upper-triangular representation (the Schur basis). Denote

$$Q_i = \sum_{k=1}^{i} (., e_k)e_k \ (i = 1, ..., n); \Delta Q_k = (., e_k)e_k \ (k = 1, ..., n);$$

$$P_0 = 0, P_1 = \sum_{k=1}^{\mu_1} \Delta Q_k, P_2 = \sum_{k=1}^{\mu_1+\mu_2} \Delta Q_k, ..., P_j = \sum_{k=1}^{\mu_1+\mu_2+...+\mu_j} \Delta Q_k$$

and

$$\Delta P_j = P_j - P_{j-1} = \sum_{k=\nu_{j-1}+1}^{\nu_j} \Delta Q_k \text{ , where } \nu_0 = 0, \nu_j = \mu_1 + \mu_2 + ... + \mu_j \ \ (j = 1, ..., m).$$

In addition, put $A_{jk} = \Delta P_j A \Delta P_k$ $(j \neq k)$ and $A_j = \Delta P_j A \Delta P_j$ $(j, k = 1, ..., m)$. We can see that each P_j is an orthogonal invariant projection of A and

$$A = \begin{pmatrix} A_1 & A_{12} & A_{13} & ... & A_{1m} \\ 0 & A_2 & A_{23} & ... & A_{2m} \\ . & . & . & ... & . \\ 0 & 0 & 0 & ... & A_m \end{pmatrix}. \tag{2.1}$$

Besides, if $\mu_j = 1$, then $A_j = \lambda_j \Delta P_j$ and ΔP_j is one dimensional. If $\mu_j > 1$, then

$$A_j = \sum_{k=\nu_{j-1}+1}^{\nu_j} \Delta Q_k A \sum_{i=\nu_{j-1}}^{\nu_j} \Delta Q_i = \sum_{k=\nu_{j-1}+1}^{\nu_j} \Delta Q_k A \Delta Q_k + \sum_{i=\nu_{j-1}+1}^{\nu_j} \sum_{k=\nu_{j-1}+1}^{i-1} \Delta Q_k A \Delta Q_i$$

$$= \lambda_j \sum_{k=\nu_{j-1}+1}^{\nu_j} \Delta Q_k + V_j = \lambda_j \Delta P_j + V_j,$$

where

$$V_j = \sum_{i=\nu_{j-1}+1}^{\nu_j} \sum_{k=\nu_{j-1}+1}^{i-1} \Delta Q_k A Q_i.$$

In the matrix form the blocks A_j have the form

$$A_1 = \begin{pmatrix} \lambda_1 & a_{12} & a_{13} & \dots & a_{1,\mu_1-1} & a_{1\mu_1} \\ 0 & \lambda_1 & a_{23} & \dots & a_{2n-1} & a_{2n} \\ . & . & . & \dots & . & \\ 0 & 0 & 0 & \dots & \lambda_1 & a_{\mu_1-1,\mu_1} \\ 0 & 0 & 0 & \dots & 0 & \lambda_1 \end{pmatrix},$$

$$A_2 = \begin{pmatrix} \lambda_2 & a_{\mu_1+1,\mu_1+2} & a_{\mu_1+1,\mu_1+3} & \dots & a_{\mu_1+1,\mu_1+\mu_2-1} & a_{\mu_1+1,\mu_1+\mu_2} \\ 0 & \lambda_2 & a_{\mu_1+2,\mu_1+3} & \dots & a_{\mu_1+2,\mu_1+\mu_2-1} & a_{\mu_1+2,\mu_1+\mu_2} \\ . & . & . & \dots & . & \\ 0 & 0 & 0 & \dots & \lambda_2 & a_{\mu_1+\mu_2-1,\mu_1+\mu_2} \\ 0 & 0 & 0 & \dots & 0 & \lambda_2 \end{pmatrix},$$

etc. Besides, each V_j is a strictly upper-triangular (nilpotent) part of A_j. So A_j has the unique eigenvalue λ_j of the algebraic multiplicity μ_j: $\sigma(A_j) = \{\lambda_j\}$. We thus have proved the following result.

Lemma 2.1. *An arbitrary matrix $A \in \mathbb{C}^{n\times n}$ can be reduced by a unitary transform to the block triangular form (2.1) with $A_j = \lambda_j \Delta P_j + V_j \in \mathbb{C}^{\mu_j\times\mu_j}$, where V_j is either a nilpotent operator, or $V_j = 0$. Besides, A_j has the unique eigenvalue λ_j of the algebraic multiplicity μ_j.*

3. Statement of the main result

The following quantity (the departure from normality) plays an essential role hereafter:

$$g(A) := [\|A\|_F^2 - \sum_{k=1}^{m} \mu_k |\lambda_k|^2]^{1/2}.$$

$g(A)$ enjoys the following properties:

$$g^2(A) \le 2\|A_I\|_F^2 \ (A_I = (A - A^*)/2i) \text{ and } g^2(A) \le \|A\|_F^2 - |\text{trace } A^2|,$$

cf. [15, Section 3.1]. If A is normal, then $g(A) = 0$. Introduce also the notations

$$d_j := \sum_{k=0}^{j} \frac{j!}{((j-k)!k!)^{3/2}} \quad (j = 0, ..., n-2), \quad \theta(A) := \sum_{k=0}^{n-2} \frac{d_k g^k(A)}{\delta^{k+1}}$$

and

$$\gamma(A) := \left(1 + \frac{g(A)\theta(A)}{\sqrt{m-1}}\right)^{2(m-1)}.$$

It is not hard to check that $d_j \le 2^j$. Now we are in a position to formulate the main result of this paper.

Theorem 3.1. *Let an $n \times n$-matrix A have $m \leq n$ ($m \geq 2$) different eigenvalues λ_j of the algebraic multiplicity μ_j ($j = 1, ..., m$). Then there are $\mu_j \times \mu_j$-matrices A_j each of which has a unique eigenvalue λ_j and such that (1.2) holds with the block-diagonal matrix $\hat{D} = \text{diag}\ (A_1, A_2, ..., A_m)$. Moreover,*

$$\kappa_T = \|T\|\|T^{-1}\| \leq \gamma(A). \tag{3.1}$$

As it was above mentioned, the proof of this theorem is presented in the next two sections. Theorem 3.1 is sharp: if A is normal, then $g(A) = 0$ and $\gamma(A) = 1$. Thus we obtain the equality $\kappa_T = 1$.

If all the eigenvalues are different: $m = n$, then Theorem 3.1 coincides with Theorem 6.1 from [15] (see also [13]). Besides, κ_T is the condition number. About the recent interesting investigations of the similarity of matrices see the papers [6, 7, 11, 17] and references therein.

4. An inequality for the norm of *T*

Recall that P_j are the orthogonal invariant projections defined in Section 2 and $\Delta P_j = P_j - P_{j-1}$; A_{jk} and A_j are also defined in Section 2. Put

$$\overline{P}_k = I - P_k, B_k = \overline{P}_k A \overline{P}_k \text{ and } C_k = \Delta P_k A \overline{P}_k \ \ (k = 1, ..., m-1).$$

By Lemma 2.1 A_j has the unique eigenvalue λ_j and A is represented by (2.1). Represent B_j and C_j in the block form:

$$B_j = \overline{P}_j A \overline{P}_j = \begin{pmatrix} A_{j+1} & A_{j+1,j+2} & ... & A_{j+1,m} \\ 0 & A_{j+2} & ... & A_{j+2,m} \\ . & . & ... & . \\ 0 & 0 & . & A_m \end{pmatrix}$$

and

$$C_j = \Delta P_j A \overline{P}_j = \begin{pmatrix} A_{j,j+1} & A_{j,j+2} & ... & A_{j,m} \end{pmatrix} \ \ (j = 1, ..., m-1).$$

Since B_j is a block triangular matrix, it is not hard to see that

$$\sigma(B_j) = \cup_{k=j+1}^m \sigma(A_k) = \cup_{k=j+1}^m \lambda_k \ \ (j = 1, ..., m-1),$$

cf. [15, Lemma 6.2]. So due to Lemma 2.1

$$\sigma(B_j) \cap \sigma(A_j) = \emptyset \ \ (j = 1, ..., m-1). \tag{4.1}$$

Under this condition, the equation

$$A_j X_j - X_j B_j = -C_j \ \ (j = 1, ..., m-1) \tag{4.2}$$

has a unique solution

$$X_j : \overline{P}_j \mathbb{C}^n \to \Delta P_j \mathbb{C}^n, \tag{4.3}$$

e.g. [2, Section VII.2] or [3].

Lemma 4.1. *Let X_j be a solution to (4.2). Then*

$$(I - X_{m-1})(I - X_{m-2}) \cdots (I - X_1)\, A\, (I + X_1)(I + X_2) \cdots (I + X_{m-1}) = \hat{D}. \tag{4.4}$$

Proof. Due to (4.3) we can write $X_j = \Delta P_j X_j \overline{P}_j$. But $\Delta P_j \overline{P}_j = \overline{P}_j \Delta P_j = 0$. Therefore $X_j A_j = B_j X_j = X_j C_j = C_j X_j = 0$ and

$$X_j^2 = 0. \tag{4.5}$$

Since P_j is a projection invariant to A: $P_j A P_j = A P_j$, we can write $\overline{P}_j A P_j = 0$. Thus, $A = A_1 + B_1 + C_1$ and consequently,

$$(I - X_1) A (I + X_1) = (I - X_1)(A_1 + B_1 + C_1)(I + X_1) =$$

$$A_1 + B_1 + C_1 - X_1 B_1 + A_1 X_1 = A_1 + B_1.$$

Furthermore, $B_1 = A_2 + B_2 + C_2$. Hence,

$$(\overline{P}_1 - X_2)B_1(\overline{P}_1 + X_2) = (\overline{P}_1 - X_1)(A_2 + B_2 + C_2)(\overline{P}_1 + X_1) =$$

$$A_2 + B_2 + C_2 - X_2 B_2 + A_2 X_2 = A_2 + B_2.$$

Therefore,

$$(I - X_2)(A_1 + B_1)(I + X_2) = (P_1 + \overline{P}_1 - X_2)(A_1 + B_1)(P_1 + \overline{P}_1 + X_2) =$$
$$A_1 + (\overline{P}_1 - X_2)(A_1 + B_1)(\overline{P}_1 + X_2) = A_1 + A_2 + B_2.$$

Consequently,

$$(I - X_2)(A_1 + B_1)(I + X_2) = (I - X_2)(I - X_1)A(I + X_1)(I + X_2) = A_1 + A_2 + B_2.$$

Continuing this process and taking into account that $B_{m-1} = A_m$, we obtain

$$(I - X_{m-1})(I - X_{m-2}) \cdots (I - X_1)\, A\, (I + X_1)(I + X_2) \cdots (I + X_{m-1}) = A_1 + ... + A_m = \hat{D},$$

as claimed. □

Take

$$T = (I + X_1)(I + X_2) \cdots (I + X_{m-1}). \tag{4.6}$$

According to (4.5)

$$(I + X_j)(I - X_j) = (I - X_j)(I + X_j) = I.$$

So the matrix $I - X_j$ is inverse to $I + X_j$. Thus,

$$T^{-1} = (I - X_{m-1})(I - X_{m-2}) \cdots (I - X_1) \tag{4.7}$$

and *(4.4) can be written as (1.2)*. We thus arrive at

Corollary 4.2. *Let an $n \times n$-matrix A have $m \le n$ ($m \ge 2$) different eigenvalues λ_j of the algebraic multiplicity μ_j ($j = 1, ..., m$). Then there are $\mu_j \times \mu_j$-matrices A_j each of which has a unique eigenvalue λ_j and such that (1.2) holds with T defined by (4.6).*

By the inequalities between the arithmetic and geometric means from (4.6) and (4.7) we get

$$\|T\| \le \prod_{j=1}^{m-1}(1 + \|X_j\|) \le \left(1 + \frac{1}{m-1}\sum_{j=1}^{m-1}\|X_j\|\right)^{m-1} \tag{4.8}$$

and

$$\|T^{-1}\| \le \left(1 + \frac{1}{m-1}\sum_{k=1}^{m-1}\|X_k\|\right)^{m-1}. \tag{4.9}$$

5. Proof of Theorem 3.1

Consider the Sylvester equation

$$BX - X\tilde{B} = C, \tag{5.1}$$

where $B \in \mathbb{C}^{n_1 \times n_1}, \tilde{B} \in \mathbb{C}^{n_2 \times n_2}$ and $C \in \mathbb{C}^{n_1 \times n_2}$ are given; $X \in \mathbb{C}^{n_1 \times n_2}$ should be found. Assume that the eigenvalues $\lambda_k(B)$ and $\lambda_j(\tilde{B})$ of B and $\tilde{B}$, respectively, satisfy the condition.

$$\rho_0(B, \tilde{B}) := \text{distance } (\sigma(B), \sigma(\tilde{B})) = \min_{j,k} |\lambda_k(B) - \lambda_j(\tilde{B})| > 0. \tag{5.2}$$

Then equation (5.1) has a unique solution X [3]. Due to [15, Corollary 5.8] (see also Corollary 6.2 from [14]) the inequality

$$\|X\|_F \le \|C\|_F \sum_{p=0}^{n_1+n_2-2} \frac{1}{\rho_0^{p+1}(B,\tilde{B})} \sum_{k=0}^{p} \binom{p}{k} \frac{g^k(\tilde{B}) g^{p-k}(B)}{\sqrt{(p-k)!k!}} \tag{5.3}$$

is valid and therefore

$$\|X\|_F \le \|C\|_F \sum_{p=0}^{n_1+n_2-2} \frac{d_p \hat{g}^p}{\rho_0^{p+1}(B,\tilde{B})}, \tag{5.4}$$

where $\hat{g} = \max\{g(B), g(\tilde{B})\}$.

Let us go back to equation (4.2). In this case $B = A_j$, $\tilde{B} = B_j$, $C = C_j$, $n_1 = \mu_j$, $n_2 = \hat{n}_j := \dim\ \overline{P}_j \mathbb{C}^n$, and due to (1.1), $\rho_0(A_j, B_j) \ge \delta$ $(j = 1, ..., n)$. In addition, $\mu_j + \hat{n}_j \le n$. Now (5.4) implies

$$\|X_j\|_F \le \|C_j\|_F \sum_{k=0}^{n-2} \frac{d_k \hat{g}_j^k}{\delta^{k+1}}, \tag{5.5}$$

where $\hat{g}_j = \max\{g(B_j), g(A_j)\}$.

Recall that $\{e_k\}_{k=1}^n$ denotes the Schur basis. So

$$Ae_k = \sum_{j=1}^{k} a_{jk} e_j \text{ with } a_{jk} = (Ae_k, e_j) \ \ (j = 1, ..., n).$$

We can write $A = D_A + V_A$ $(\sigma(A) = \sigma(D_A))$ with a normal (diagonal) matrix D_A defined by $D_A e_j = a_{kk} e_k = \hat{\lambda}_j e_k$ $(k = 1, ..., n)$ and a nilpotent (strictly upper-triangular) matrix V_A defined by $V_A e_k = a_{1k} e_1 + ... + a_{k-1,k} e_{k-1}$ $(k = 2, ..., n)$, $V_A e_1 = 0$. D_A and V_A will be called *the diagonal part and nilpotent part* of A, respectively. It can be $V_A = 0$, i.e. A is normal.

Besides, $g(A) = \|V_A\|_F$. In addition, the nilpotent part V_j of A_j is $\Delta P_j V_A \Delta P_j$ and the nilpotent part W_j of B_j is $\overline{P}_j V_A \overline{P}_j$. So V_j and W_j are orthogonal, and

$$g(A_j) = \|V_j\|_F \le \|V_A\|_F = g(A), g(B_j) = \|W_j\|_F \le \|V_A\|_F^2 = g(A).$$

Thus, from (5.5) it follows

$$\|X_j\|_F \le \|C_j\|_F \sum_{k=0}^{n-2} \frac{d_k g^k(A)}{\delta^{k+1}} = \|C_j\|_F \theta(A). \tag{5.6}$$

It can be directly checked that

$$\|C_j\|_F^2 = \sum_{k=j+1}^{m} \|A_{jk}\|_F^2$$

and

$$\sum_{j=1}^{m-1} \|C_j\|_F^2 = \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} \|A_{jk}\|_F^2 \le \sum_{j=1}^{m} \sum_{k=j}^{m} \|A_{jk}\|_F^2 - \sum_{j=1}^{m} \|A_{jj}\|_F^2 = \|A\|_F^2 - \sum_{j=1}^{m} \|A_{jj}\|_F^2.$$

Since $\|A_{kk}\|_F \ge \mu_k |\lambda_k|$, we have

$$\sum_{j=1}^{m-1} \sum_{k=j+1}^{m} \|A_{jk}\|_F^2 \le g^2(A),$$

and consequently,

$$\sum_{j=1}^{m-1} \|C_j\|_F^2 \le g^2(A). \tag{5.7}$$

Take T as is in (4.6). Then (4.8), (4.9) and (5.6) imply

$$\|T\| \le \left(1 + \frac{1}{m-1}\sum_{k=1}^{m-1}\|X_k\|_F\right)^{m-1} \le \left(1 + \frac{\theta(A)}{m-1}\sum_{k=1}^{m-1}\|C_k\|_F\right)^{m-1}$$

and

$$\|T^{-1}\| \le \left(1 + \frac{\theta(A)}{m-1}\sum_{k=1}^{m-1}\|C_k\|_F\right)^{m-1}.$$

But by the Schwarz inequality and (5.7),

$$(\sum_{j=1}^{m-1}\|C_j\|_F)^2 \le (m-1)\sum_{j=1}^{m-1}\|C_j\|_F^2 \le (m-1)g^2(A).$$

Thus,

$$\|T\|^2 \le \left(1 + \frac{\theta(A)}{\sqrt{m-1}}g(A)\right)^{2(m-1)} = \gamma(A)$$

and $\|T^{-1}\|^2 \le \gamma(A)$. Now (4.4) proves the theorem. □

6. Applications of Theorem 3.1

Let $f(z)$ be a scalar function, regular on $\sigma(A)$. Define $f(A)$ by the usual way via the Cauchy integral [2]. Since A_j are mutually orthogonal, we have

$$f(\hat{D}) = \text{diag }\ (f(A_1,...,f(A_m)) \text{ and } \|f(\hat{D})\| = \max_j \|\Delta P_j f(A_j)\|. \tag{6.1}$$

Let

$$r(z) = \sum_{k=0}^{n} c_k z^{n-k}$$

be the interpolation Lagrange-Silvester polynomial such that $r(\hat{\lambda}_j) = f(\hat{\lambda}_j)$ $(\hat{\lambda}_j \in \sigma(A), j = 1,...,n)$ and $r(A) = f(A)$, cf. [10, Section V.1].

Now (1.2) implies

$$f(A) = \sum_{k=0}^{n} c_k A^{n-k} = T^{-1}\sum_{k=0}^{n} c_k \hat{D}^{n-k} T = T^{-1} r(\hat{D}) T = T^{-1} f(\hat{D}) T.$$

Hence, (6.1) and (3.1) imply

Corollary 6.1. *Let $A \in \mathbb{C}^{n\times n}$. Then there is an invertible matrix T, such that*

$$\|f(A)\| \le \kappa_T \max_j \|\Delta P_j f(A_j)\| \le \gamma(A) \max_j \|\Delta P_j f(A_j)\|.$$

Due to Theorem 3.5 from the book [15] we have

$$\|f(A_j)\| \le \sum_{k=0}^{\mu_j - 1} |f^{(k)}(\lambda_j)| \frac{g^k(A_j)}{\sqrt{k!}}.$$

Take into account that $g(A_j) \le g(A)$ (see Section 5). Now Theorem 3.1 immediately implies.

Corollary 6.2. *Let $A \in \mathbb{C}^{n\times n}$. Then*

$$\|f(A)\| \le \gamma(A) \max_j \sum_{k=0}^{\mu_j-1} |f^{(k)}(\lambda_j)| \frac{g^k(A)}{(k!)^{3/2}}.$$

This corollary generalizes Corollary 6.1 from [15]. Moreover, in contrast to [15, Theorem 3.5] it can be applied to matrix functions non-regular on the convex hull of the spectra. For example, we have

$$\|e^{tA}\| \le \gamma(A) e^{\alpha(A)t} \sum_{k=0}^{\hat{\mu}-1} t^k \frac{g^k(A)}{(k!)^{3/2}} \quad (t \ge 0),$$

where $\alpha(A) = \max_k \operatorname{Re} \lambda_k$ and $\hat{\mu} = \max_j \mu_j$.

About the recent interesting results devoted to matrix-valued functions see the papers [9, 18] and references therein.

Now consider the resolvent. Then by (1.2) for $|z| > \max\{\|A\|, \|\hat{D}\|\}$ we have

$$R_z(A) = (A - zI)^{-1} = -\sum_{k=0}^{\infty} \frac{A^k}{z^{k+1}} = -T^{-1} \sum_{k=0}^{\infty} \frac{\hat{D}^k}{z^{k+1}} T = T^{-1} R_z(\hat{D}) T.$$

Extending this relation analytically to all regular z and taking into account that

$$R_z(\hat{D}) = \sum_{k=1}^{m} R_z(A_j) \text{ and } \|R_z(\hat{D})\| = \max_j \|\Delta P_j R_z(A_j)\| \quad (z \in \sigma(A)), \tag{6.2}$$

we get

Corollary 6.3. *Let $A \in \mathbb{C}^{n\times n}$. Then there is an invertible matrix T, such that*

$$\|R_z(A)\| \le \kappa_T \max_j \|\Delta P_j R_z(A_j)\| \le \gamma(A) \max_j \|\Delta P_j R_z(A_j)\|$$

for any regular z of A.

But due to Theorem 3.2 from [15] we have

$$\|R_z(A_j)\| \le \sum_{k=0}^{\mu_j-1} \frac{g^k(A_j)}{\rho^{k+1}(A_j, z)\sqrt{k!}} \quad (z \notin \sigma(A_j)),$$

where $\rho(A, z)$ is the distance between z and the spectrum of A. Clearly, $\rho(A_j, z) \ge \rho(A, z)$ $(j = 1, ..., m)$. Now Theorem 3.1 and (6.2) imply

Corollary 6.4. *Let $A \in \mathbb{C}^{n\times n}$. Then*

$$\|R_z(A)\| \le \gamma(A) \sum_{k=0}^{\hat{\mu}-1} \frac{g^k(A)}{\rho^{k+1}(A, z)\sqrt{k!}} \quad (\lambda \notin \sigma(A)).$$

Furthermore, let A and $\tilde{A}$ be complex $n \times n$-matrices. Recall that

$$sv_A(\tilde{A}) := \max_{t \in \sigma(\tilde{A})} \min_{s \in \sigma(A)} |t - s|$$

is the spectral variation of $\tilde{A}$ with respect to A, cf. [24]. We need the following technical lemma.

Lemma 6.5. *Let A and $\tilde{A}$ be linear operators in $\mathbb{C}^n$ and $q := \|A - \tilde{A}\|$. In addition, let*

$$\|R_\lambda(A)\| \leq F\left(\frac{1}{\rho(A,\lambda)}\right) \quad (\lambda \notin \sigma(A)),$$

where $F(x)$ is a monotonically increasing continuous function of a non-negative variable x, such that $F(0) = 0$ and $F(\infty) = \infty$. Then $sv_A(\tilde{A}) \leq z(F,q)$, where $z(F,q)$ is the unique positive root of the equation $qF(1/z) = 1$.

For the proof see [15, Lemma 1.10]. Now Corollary 6.4 implies $sv_A(\tilde{A}) \leq z(A,q)$, where $z(A,q)$ is the unique positive root of the equation

$$q\gamma(A)\sum_{k=0}^{\hat{\mu}-1} \frac{g^k(A)}{z^{k+1}\sqrt{k!}} = 1.$$

This equation is equivalent to the algebraic one

$$z^{\hat{\mu}} = q\gamma(A)\sum_{k=0}^{\hat{\mu}-1} \frac{g^k(A)z^{\hat{\mu}-k-1}}{\sqrt{k!}}. \tag{6.3}$$

Various estimates for the roots of algebraic equations, can be found for instance, in [4, 20] and references therein. For example, if

$$\zeta(A,q) := q\gamma(A)\sum_{k=0}^{\hat{\mu}-1} \frac{g^k(A)}{\sqrt{k!}} < 1, \tag{6.4}$$

then due to Lemma 3.17 from [15], we have $z^{\hat{\mu}}(A,q) \leq \zeta(A,q)$. So we arrive at

Corollary 6.6. *Let A and $\tilde{A}$ be $n \times n$-matrices. Then $sv_A(\tilde{A}) \leq z(A,q)$. If, in addition, condition (6.4) holds, then $sv_A^{\hat{\mu}}(\tilde{A}) \leq \zeta(A,q)$.*

In the next section we compare our results with the Elsner inequality:

$$sv_A(\tilde{A}) \leq q^{1/n}(\|A\| + \|\tilde{A}\|)^{1-1/n}, \tag{6.5}$$

cf. [24, p. 168].

7. Example

To illustrate Corollary 6.6 consider the matrices

$$A = \begin{pmatrix} -1 & a_{12} & a_{13} & a_{14} \\ 0 & -1 & a_{23} & a_{24} \\ 0 & 0 & 1 & a_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ and } \tilde{A} = \begin{pmatrix} -1 & a_{12} & a_{13} & a_{14} \\ a_{21} & -1 & a_{23} & a_{24} \\ a_{31} & a_{32} & 1 & a_{34} \\ a_{41} & a_{42} & a_{43} & 1 \end{pmatrix}$$

The eigenvalues of A are $\lambda_1 = \lambda_2 = -1, \lambda_3 = \lambda_4 = 1$. So $m = 2, \mu_1 = \mu_2 = 2, \delta = 2$,

$$g^2(A) = \sum_{k=1}^{4}\sum_{j=1}^{k-1} |a_{jk}|^2,$$

$d_0 = 1, d_1 = 1$, and $d_2 \leq 4$. Hence,

$$\theta(A) \leq \theta_1(A) := \frac{1}{2}(1 + \frac{g(A)}{2} + g^2(A)) \text{ and } \gamma(A) \leq \gamma_1(A),$$

where $\gamma_1(A) := (1 + g(A)\theta_1(A))^2$. According to (6.3) consider the equation $z^2 = q\gamma_1(A)(z + g(A))$. So one can take $z(A, q) = z_1(A, q)$, where

$$z_1(A, q) := \frac{1}{2}q\gamma_1(A) + \sqrt{\frac{1}{4}q^2\gamma_1^2(A) + q\gamma_1(A)g(A)}.$$

Due to Corollary 6.6 we have

$$sv_A(\tilde{A}) \leq z_1(A, q). \tag{7.1}$$

The Elsner inequality (6.5) gives us

$$sv_A(\tilde{A}) \leq q^{1/4}(\|A\| + \|\tilde{A}\|)^{3/4}. \tag{7.2}$$

We can see that under the condition

$$z_1(A, q) < q^{1/4}(\|A\| + \|\tilde{A}\|)^{3/4} \tag{7.3}$$

inequality (7.1) is sharper than (7.2). For example, if A is "close" to normal, then $g(A)$ is "small" and $\gamma_1(A)$ is "close" to one, and (7.3) is certainly holds. So our results can considerably improve (6.5) if we have an information about the multiplicities on the eigenvalues of A. About the recent perturbation results for matrices see the interesting papers [1, 5, 8, 16, 22, 23] and references given therein.

References

[1] J. Benasseni, Lower bounds for the largest eigenvalue of a symmetric matrix under perturbations of rank one. Linear Multilinear Algebra 59 (2011), no. 5, 565–569.
[2] R. Bhatia, Matrix Analysis, Springer, New York, 1997.
[3] R. Bhatia, and P. Rosenthal, How and why to solve the matrix equation $AX - XB = Y$, Bull. London Math. Soc., 29, (1997) 1–21.
[4] P. Borwein and T. Erdelyi, Polynomials and Polynomial Inequalities, Springer-Verlag, New York, 1995.
[5] J. Ccapa and R.L. Soto, On spectra perturbation and elementary divisors of positive matrices. Electron. J. Linear Algebra 18 (2009), 462–481.
[6] J.A. Dias da Silva and C.R. Johnson, Cospectrality and similarity for a pair of matrices under multiplicative and additive composition with diagonal matrices. Linear Algebra Appl. , 326, no. 1-3 (2001) 15–25.
[7] D. Djokovic, Universal zero patterns for simultaneous similarity of several matrices. Oper. Matrices, 1, no. 1, (2007) 113–119.
[8] R. Fernandes, Small perturbations and pairs of matrices that have the same immanent. Linear Multilinear Algebra 58, no. 7-8, (2010) 977–991.
[9] B. Fritzsche, B. Kirstein and A. Lasarow, Orthogonal rational matrix-valued functions on the unit circle: Recurrence relations and a Favard-type theorem. Math. Nachr., 279, no. 5–6 (2006) 513-542, .
[10] F.R. Gantmakher, Theory of Matrices, Nauka, Moscow 1967. (In Russian).
[11] A. George and K. Ikramov, Unitary similarity of matrices with quadratic minimal polynomials. Linear Algebra Appl., 349, no. 1-3, (2002) 11–16.
[12] M.I. Gil', Perturbations of functions of diagonalizable matrices, Electr. J. of Linear Algebra, 20, (2010) 303–313.
[13] M.I. Gil', A bound for condition numbers of matrices, Electr. J. of Linear Algebra, 27, (2014) 162–171.
[14] M.I. Gil', Resolvents of operators on tensor products of Euclidean spaces , Linear and Multilinear Algebra 64 (4), (2016) 699-716
[15] M.I. Gil', Operator Functions and Operator Equations. World Scientific, New Jersey, 2018.
[16] L. Glebsky and L.M. Rivera, On low rank perturbations of complex matrices and some discrete metric spaces. Electron. J. Linear Algebra 18 (2009), 302–316.
[17] T. Jiang, X. Cheng, and L. Chen, An algebraic relation between consimilarity and similarity of complex matrices and its applications. J. Phys. A, Math. Gen., 39, no. 29, (2006) 9215–9222.
[18] A. Lasarow, Dual Szegő pairs of sequences of rational matrix-valued functions. Int. J. Math. Math. Sci., 2006, no. 5 (2006) 1–37.
[19] M. Marcus and H. Minc, A Survey of Matrix Theory and Matrix Inequalities. Allyn and Bacon, Boston, 1964.
[20] G.V. Milovanović, D.S. Mitrinović, and Th. M. Rassias, Topics in Polynomials: Extremal Problems, Inequalities, Zeros, World Scientific, Singapore, 1996.
[21] C. Rajian and T. Chelvam, On similarity invariants of *EP* matrices. East Asian Math. J., 23, no. 2 (2007) 207–212.
[22] A.C.M. Ran and M. Wojtylak, Eigenvalues of rank one perturbations of unstructured matrices. Linear Algebra Appl. 437 (2012), no. 2, 589–600.
[23] L. Rodman, Lipschitz properties of structure preserving matrix perturbations. Linear Algebra Appl. 437 , no. 7 (2012), 1503–1537.
[24] G.W. Stewart and Sun Ji-guang. Matrix Perturbation Theory, Academic Press, New York, 1990.