2010

Perturbations of functions of diagonalizable matrices

Michael I. Gil
gilmi@bezeqint.net

Follow this and additional works at: http://repository.uwyo.edu/ela

Recommended Citation
DOI: https://doi.org/10.13001/1081-3810.1375

This Article is brought to you for free and open access by Wyoming Scholars Repository. It has been accepted for inclusion in Electronic Journal of Linear Algebra by an authorized editor of Wyoming Scholars Repository. For more information, please contact scholcom@uwyo.edu.
Abstract. Let $A$ and $\tilde{A}$ be $n \times n$ diagonalizable matrices and $f$ be a function defined on their spectra. In the present paper, bounds for the norm of $f(A) - f(\tilde{A})$ are established. Applications to differential equations are also discussed.

Key words. Matrix valued functions, Perturbations, Similarity of matrices, Diagonalizable matrices.

AMS subject classifications. 15A54, 15A45, 15A60.

1. Introduction and statement of the main result. Let $\mathbb{C}^n$ be a Euclidean space with the scalar product $(\cdot, \cdot)$, Euclidean norm $\|\cdot\| = \sqrt{(\cdot, \cdot)}$ and identity operator $I$. $A$ and $\tilde{A}$ are $n \times n$ matrices with eigenvalues $\lambda_j$ and $\tilde{\lambda}_j$ ($j = 1, \ldots, n$), respectively. $\sigma(A)$ denotes the spectrum of $A$, $A^*$ is the adjoint to $A$, and $N_2^2(A)$ is the Hilbert-Schmidt (Frobenius) norm of $A$: $N_2^2(A) = \text{Trace}(A^*A)$.

In the sequel, it is assumed that each of the matrices $A$ and $\tilde{A}$ has $n$ linearly independent eigenvectors, and therefore, these matrices are diagonalizable. In other words, the eigenvalues of these matrices are semi-simple.

Let $f$ be a scalar function defined on $\sigma(A) \cup \sigma(\tilde{A})$. The aim of this paper is to establish inequalities for the norm of $f(A) - f(\tilde{A})$. The literature on perturbations of matrix valued functions is very rich but mainly, perturbations of matrix functions of a complex argument and matrix functions of Hermitian matrices were considered, cf. [1, 11, 13, 14, 16, 18]. The matrix valued functions of a non-Hermitian argument have been investigated essentially less, although they are very important for various applications; see the book [10].

The following quantity plays an essential role hereafter:

$$g(A) := \left[ N_2^2(A) - \sum_{k=1}^{n} |\lambda_k|^2 \right]^{1/2}.$$
Let \( \delta_j \) be the half-distance from \( \mu_j \) to the other eigenvalues of \( A \), namely,
\[
\delta_j := \min_{k=1, \ldots, m; k \neq j} |\mu_j - \mu_k|/2 > 0.
\]
Similarly,
\[
\tilde{\delta}_j := \min_{k=1, \ldots, \tilde{m}; k \neq j} |\tilde{\mu}_j - \tilde{\mu}_k|/2 > 0.
\]
Put
\[
\beta(A) := \sum_{j=1}^{m} p_j \sum_{k=0}^{n-1} \frac{g^k(A)}{\delta_j^k k!}
\]
and
\[
\beta(\tilde{A}) := \sum_{j=1}^{\tilde{m}} p_j \sum_{k=0}^{n-1} \frac{g^k(\tilde{A})}{\delta_j^k k!}.
\]
Here \( g^0(A) = \delta_j^0 = 1 \) and \( \tilde{g}^0(A) = \tilde{\delta}_j^0 = 1 \). According to (1.1),
\[
\beta(A) \leq \sum_{j=1}^{m} p_j \sum_{k=0}^{n-1} \frac{\sqrt{2N_2(A_j)}}{\delta_j^k k!}.
\]
In what follows, we put
\[
\frac{f(\lambda_k) - f(\tilde{\lambda}_j)}{\lambda_k - \tilde{\lambda}_j} = 0 \quad \text{if} \quad \lambda_k = \tilde{\lambda}_j.
\]
Now we are in a position to formulate our main result.

**Theorem 1.1.** Let \( A \) and \( \tilde{A} \) be \( n \times n \) diagonalizable matrices and \( f \) be a function defined on \( \sigma(A) \cup \sigma(\tilde{A}) \). Then the inequalities
\[
N_2(f(A) - f(\tilde{A})) \leq \beta(A)\beta(\tilde{A}) \max_{j,k} \left| \frac{f(\lambda_k) - f(\tilde{\lambda}_j)}{\lambda_k - \tilde{\lambda}_j} \right| N_2(A - \tilde{A}) \tag{1.2}
\]
and
\[
N_2(f(A) - f(\tilde{A})) \leq \beta(A)\beta(\tilde{A}) \max_{j,k} |f(\lambda_k) - f(\tilde{\lambda}_j)| \tag{1.3}
\]
are valid.

The proof of this theorem is divided into lemmas which are presented in the next two sections. The importance of Theorem 1.1 lies in the fact that the right-hand sides of inequalities (1.2) and (1.3) only involve universal quantities calculated for \( A \) and \( \tilde{A} \), and the values of the function \( f \) on the spectra \( \sigma(A) \) and \( \sigma(\tilde{A}) \), but e.g. no matrices performing similarities of \( A \) and \( \tilde{A} \) to diagonal ones.
2. The basic lemma. Let $T$ and $\tilde{T}$ be the invertible matrices performing similarities of $A$ and $\tilde{A}$ to diagonal ones, that is,

$$T^{-1}AT = S,$$

(2.1)

and

$$\tilde{T}^{-1}\tilde{A}\tilde{T} = \tilde{S},$$

(2.2)

where $S = \text{diag}(\lambda_j)$ and $\tilde{S} = \text{diag}(\tilde{\lambda}_j)$. Everywhere below, $\{d_k\}_{k=1}^n$ is the standard orthonormal basis and $\|B\| = \sup \{\|Bh\|/\|h\| : h \in \mathbb{C}^n\}$ for an $n \times n$ matrix $B$. Recall that $N_2(AB) \leq \|B\|N_2(A)$.

We begin with the following lemma.

**Lemma 2.1.** Let the hypotheses of Theorem 1.1 hold. Then

$$N_2^2(f(A) - f(\tilde{A})) \leq \|\tilde{T}^{-1}\|T\|^2 \sum_{j,k=1}^n |(f(\lambda_k) - f(\tilde{\lambda}_j))(T^{-1}\tilde{T}d_j, d_k)|^2. \quad (2.3)$$

**Proof.** We have

$$N_2^2(A - \tilde{A}) = N_2^2(TST^{-1} - \tilde{T}\tilde{S}\tilde{T}^{-1}) = N_2^2(TST^{-1}\tilde{T}\tilde{T}^{-1} - TT^{-1}\tilde{T}\tilde{S}\tilde{T}^{-1})$$

$$= N_2^2(T(ST^{-1}\tilde{T} - T^{-1}\tilde{T}\tilde{S}\tilde{T}^{-1})) \leq \|T^{-1}\|T\|^2N_2^2(ST^{-1}\tilde{T} - T^{-1}\tilde{T}\tilde{S}).$$

But

$$N_2^2(ST^{-1}\tilde{T} - T^{-1}\tilde{T}\tilde{S}) = \sum_{j,k=1}^n |(ST^{-1}\tilde{T} - T^{-1}\tilde{T}\tilde{S})d_j, d_k|^2$$

$$= \sum_{j,k=1}^n |(T^{-1}\tilde{T}d_j, S^*d_k) - (T^{-1}\tilde{T}\tilde{S}d_j, d_k)|^2 = J,$$

(2.4)

where

$$J := \sum_{j,k=1}^n |(\lambda_k - \tilde{\lambda}_j)(T^{-1}\tilde{T}d_j, d_k)|^2.$$

So,

$$N_2^2(A - \tilde{A}) \leq \|\tilde{T}^{-1}\|T\|^2J.$$
Similarly, taking into account that
\[ T^{-1}f(A)T = f(S) = \text{diag} (f(\lambda_j)) \quad \text{and} \quad \tilde{T}^{-1}f(\tilde{A})\tilde{T} = f(\tilde{S}) = \text{diag} (f(\tilde{\lambda}_j)), \]
we get (2.3), as claimed. \( \square \)

¿From the previous lemma we obtain at once:

**Corollary 2.2.** Under the hypotheses of Theorem 1.1, we have
\[ N_2(f(A) - f(\tilde{A})) \leq \kappa_T \tilde{\kappa}_T \left[ \sum_{j,k=1}^{n} |f(\lambda_k) - f(\tilde{\lambda}_j)|^2 \right]^{1/2}, \]
where
\[ \kappa_T := \|T\|\|T^{-1}\| \quad \text{and} \quad \tilde{\kappa}_T := \|\tilde{T}\|\|\tilde{T}^{-1}\|. \]

Recall that
\[ \frac{f(\lambda_k) - f(\tilde{\lambda}_j)}{\lambda_k - \tilde{\lambda}_j} = 0 \quad \text{if} \quad \lambda_k = \tilde{\lambda}_j. \]

Note that (2.3) implies
\[ N_2^2(f(A) - f(\tilde{A})) \leq \|\tilde{T}^{-1}\|^2\|T\|^2 \sum_{j,k=1}^{n} \frac{|f(\lambda_k) - f(\tilde{\lambda}_j)|}{\lambda_k - \tilde{\lambda}_j}^2 (\lambda_k - \tilde{\lambda}_j)(T^{-1}\tilde{T}d_j, d_k)^2 \]
\[ \leq \|\tilde{T}^{-1}\|^2\|T\|^2 \max_{j,k} \left| \frac{f(\lambda_k) - f(\tilde{\lambda}_j)}{\lambda_k - \tilde{\lambda}_j} \right|^2 J. \quad (2.5) \]

It follows from (2.4) that
\[ J = N_2^2(T^{-1}AT - T^{-1}\tilde{A}\tilde{T}) \leq \|T^{-1}\|^2\|\tilde{T}\|^2 N_2^2(A - \tilde{A}). \]

Thus, (2.5) yields:

**Corollary 2.3.** Under the hypotheses of Theorem 1.1, the inequality
\[ N_2(f(A) - f(\tilde{A})) \leq \kappa_T \tilde{\kappa}_T N_2(A - \tilde{A}) \max_{j,k} \left| \frac{f(\lambda_k) - f(\tilde{\lambda}_j)}{\lambda_k - \tilde{\lambda}_j} \right| \quad (2.6) \]
is true.

Since
\[ \left| \frac{f(\lambda_k) - f(\tilde{\lambda}_j)}{\lambda_k - \tilde{\lambda}_j} \right| \leq \sup_{s \in \text{co}(A, \tilde{A})} |f'(s)|, \]
Perturbations of Functions of Diagonalizable Matrices

where \( \text{co}(A, \tilde{A}) \) is the closed convex hull of the eigenvalues of the both matrices \( A \) and \( \tilde{A} \), cf. [17], we have:

**Corollary 2.4.** Let \( A \) and \( \tilde{A} \) be \( n \times n \) diagonalizable matrices and \( f \) be a function defined and differentiable on \( \text{co}(A, \tilde{A}) \). Then

\[
N_2(f(A) - f(\tilde{A})) \leq \kappa T \tilde{T} \max_{s \in \text{co}(A, \tilde{A})} |f'(s)| N_2(A - \tilde{A}).
\]

We need also the following corollary of Lemma 2.1.

**Corollary 2.5.** Under the hypotheses of Theorem 1.1, it holds that

\[
N_2(f(A) - f(\tilde{A})) \leq \|T^{-1}\| \max_{j,k} |f(\lambda_k) - f(\tilde{\lambda}_j)| \left( \sum_{j=1}^{n} \|T d_j\|^2 \sum_{k=1}^{n} \|(T^{-1})^* d_k\|^2 \right)^{1/2}.
\]

(2.7)

**3. Proof of Theorem 1.1.** Let \( Q_j \) be the Riesz projection for \( \mu_j \):

\[
Q_j = -\frac{1}{2\pi i} \int_{L_j} R_\lambda(A) d\lambda,
\]

where \( L_j := \{ z \in \mathbb{C} : |z - \mu_j| = \delta_j \} \).

**Lemma 3.1.** The inequality

\[
\|Q_j\| \leq \sum_{k=0}^{n-1} \frac{g^k(A)}{\delta_j^k \sqrt{k!}}
\]

is true.

**Proof.** By the Schur theorem, there is an orthonormal basis \( \{e_k\} \), in which \( A = D + V \), where \( D = \text{diag} \ (\lambda_j) \) is a normal matrix (the diagonal part of \( A \)) and \( V \) is an upper triangular nilpotent matrix (the nilpotent part of \( A \)). Let \( P_j \) be the eigenprojection of \( D \) corresponding to \( \mu_j \). Then by (3.1),

\[
Q_j - P_j = -\frac{1}{2\pi i} \int_{L_j} (R_\lambda(A) - R_\lambda(D)) d\lambda = \frac{1}{2\pi i} \int_{L_j} R_\lambda(A) V R_\lambda(D) d\lambda
\]

since \( V = A - D \). But

\[
R_\lambda(A) = (D + V - I\lambda)^{-1} = (I + R_\lambda(D)V)^{-1} R_\lambda(D).
\]

Consequently,

\[
R_\lambda(A) = \sum_{k=0}^{n-1} (-1)^k (R_\lambda(D)V)^k R_\lambda(D).
\]
So, we have

\[ Q_j - P_j = \sum_{k=1}^{n-1} C_k, \quad (3.2) \]

where

\[ C_k = (-1)^k \frac{1}{2\pi i} \int_{L_j} (R_\lambda(D)V)^k R_\lambda(D)d\lambda. \]

By Theorem 2.5.1 of [8], we get

\[ \| (R_\lambda(D)V)^k \| \leq \frac{N_k^k(R_\lambda(D)V)}{\delta_j^k \sqrt{k!}} \leq \frac{N_k^k(V)}{\rho(A, \lambda) \sqrt{k!}}, \]

where \( \rho(A, \lambda) = \min_k |\lambda - \lambda_k|. \) In addition, directly from the definition of \( g(A) \), when the Hilbert-Schmidt norm is calculated in the basis \( \{e_k\} \), we have \( N_2(V) = g(A) \). So,

\[ \| (R_\lambda(D)V)^k \| \leq \frac{g^k(A)}{\delta_j^k \sqrt{k!}} \]

and therefore,

\[ \| C_k \| \leq \frac{1}{2\pi} \int_{L_j} \| (R_\lambda(D)V)^k R_\lambda(D) \| |d\lambda| \leq \frac{g^k(A)}{\delta_j^k \sqrt{k!}} \]

Hence,

\[ \| Q_j - P_j \| \leq \sum_{k=1}^{n-1} \| C_k \| \leq \sum_{k=1}^{n-1} \frac{g^k(A)}{\delta_j^k \sqrt{k!}}, \quad (3.3) \]

and therefore,

\[ \| Q_j \| \leq \| P_j \| + \| Q_j - P_j \| \leq \sum_{k=0}^{n-1} \frac{g^k(A)}{\delta_j^k \sqrt{k!}}, \]

as claimed. □

Let \( \{v_k\}_{k=1}^n \) be a sequence of the eigenvectors of \( A \), and \( \{u_k\}_{k=1}^n \) be the biorthogonal sequence: \( (v_j, u_k) = 0 \) \( (j \neq k) \), \( (v_j, u_j) = 1 \) \( (j, k = 1, \ldots, n) \). So

\[ A = \sum_{k=1}^m \mu_k Q_k = \sum_{k=1}^n \lambda_k (\cdot, u_k)v_k, \quad (3.4) \]

Rearranging these biorthogonal sequences, we can write

\[ Q_j = \sum_{k=1}^{p_j} (\cdot, u_{jk})v_{jk}, \]
where \(\{u_{jk}\}_{k=1}^{p_j}\) and \(\{v_{jk}\}_{k=1}^{p_j}\) are biorthogonal: 
\[
(v_{jk}, u_{jm}) = 0 \quad (m \neq k), \quad (v_{jk}, u_{jk}) = 1 \quad (m, k = 1, \ldots, p_j).
\]
Observe that we can always choose these systems so that \(\|u_{jk}\| = \|v_{jk}\|\).

**Lemma 3.2.** Let \(\|u_{jk}\| = \|v_{jk}\|\). Then we have
\[
\|u_{jk}\|^2 \leq \|Q_j\|\quad (k = 1, \ldots, p_j).
\]

**Proof.** Clearly, \(Q_j u_{jk} = (Q_j u_{jk}, u_{jk}) v_{jk}\). So,
\[
(Q_j u_{jk}, v_{jk}) = (u_{jk}, u_{jk})(v_{jk}, v_{jk}) = \|u_{jk}\|^4.
\]
Hence, \(\|u_{jk}\|^4 \leq \|Q_j\| \|u_{jk}\|^2\), as claimed. \(\square\)

Lemmas 3.1 and 3.2 imply:

**Corollary 3.3.** The inequalities
\[
\|u_{js}\|^2 \leq \sum_{k=0}^{n-1} \frac{g_k(A)}{\delta_j \sqrt{k!}} \quad (s = 1, 2, \ldots, p_j)
\]
are valid.

Again, let \(\{d_k\}\) be the standard orthonormal basis.

**Lemma 3.4.** Let \(A\) be an \(n \times n\) diagonalizable matrix. Then the operator
\[
T = \sum_{k=1}^{n} (\cdot, d_k)v_k
\]
has the inverse one defined by
\[
T^{-1} = \sum_{k=1}^{n} (\cdot, u_k)d_k,
\]
and (2.1) holds.

**Proof.** Indeed, we can write out
\[
T^{-1}T = \sum_{j=1}^{n} d_j \sum_{k=1}^{n} (\cdot, d_k)(v_k, u_j) = \sum_{j=1}^{n} d_j (\cdot, d_j) = I
\]
and
\[
TT^{-1} = \sum_{k=1}^{n} (\cdot, u_k) \sum_{j=1}^{n} (d_k, d_j)v_j = \sum_{k=1}^{n} (\cdot, u_k)v_k = I.
\]
Moreover,
\[ AT = \sum_{j=1}^{n} v_j \lambda_j (\cdot, d_j) \quad \text{and} \quad S = \sum_{k=1}^{n} \lambda_k (\cdot, d_k) d_k. \]

Hence,
\[ T^{-1} AT = \sum_{k=1}^{n} d_k \sum_{j=1}^{n} (v_j, u_k) \lambda_j (\cdot, d_j) = \sum_{j=1}^{n} d_j \lambda_j (\cdot, d_j) = S. \]

So, (2.1) really holds. \( \square \)

**Lemma 3.5.** Let \( T \) be defined by (3.5). Then
\[ \| T \|^2 \leq \sum_{j=1}^{m} p_j \| Q_j \| \quad \text{and} \quad \| T^{-1} \|^2 \leq \sum_{j=1}^{m} p_j \| Q_j \|, \]
and therefore,
\[ \kappa_T \leq \sum_{j=1}^{m} p_j \| Q_j \|. \]

**Proof.** Due to the previous lemma
\[ \| T^{-1} x \|^2 = \sum_{k=1}^{n} |(x, u_k)|^2 \leq \| x \|^2 \sum_{k=1}^{n} \| u_k \|^2 \quad (x \in \mathbb{C}^n). \] \( (3.7) \)

By the Schwarz inequality,
\[ (T x, T x) = \sum_{k=1}^{n} \sum_{s=1}^{n} (x, d_k)(v_k, v_s)(x, d_s) \]
\[ \leq \sum_{k=1}^{n} |(x, d_k)|^2 \left[ \sum_{s,k=1}^{n} |(v_k, v_s)|^2 \right]^{1/2} \leq \| x \|^2 \sum_{s=1}^{n} \| v_s \|^2. \] \( (3.8) \)

But by Lemma 3.2,
\[ \sum_{s=1}^{n} \| u_s \|^2 = \sum_{s=1}^{n} \| v_s \|^2 = \sum_{j=1}^{m} p_j \sum_{k=1}^{n} \| v_{jk} \|^2 \leq \sum_{j=1}^{m} p_j \| Q_j \|. \]

Now (3.7) and (3.8) yield the required result. \( \square \)

Lemmas 3.1 and 3.5 imply:
Corollary 3.6. The inequality $\kappa_T \leq \beta(A)$ is true.

Lemma 3.7. Under the hypotheses of Theorem 1.1, we have

$$N_2(f(A) - f(\tilde{A})) \leq \|T^{-1}\| \|T\| \max_{j,k} |f(\lambda_k) - f(\tilde{\lambda}_j)| \left( \sum_{j=1}^{\tilde{m}} \tilde{p}_j \|\tilde{Q}_j\| \sum_{k=1}^{m} p_k \|Q_k\| \right)^{1/2},$$

where $Q_k, \tilde{Q}_j$ are the eigenprojections of $A$ and $\tilde{A}$, respectively.

Proof. By (3.6), we have

$$(T^{-1})^* = \sum_{k=1}^{n} (\cdot, d_k) u_k.$$ 

Moreover, due to Lemma 3.2,

$$\sum_{j=1}^{n} \| (T^{-1})^* d_j \|^2 = \sum_{k=1}^{n} \| u_k \|^2 \leq \sum_{k=1}^{m} p_k \|Q_k\|.$$ 

Similarly,

$$\sum_{k=1}^{n} \| \tilde{T} d_k \|^2 \leq \sum_{k=1}^{m} p_k \|\tilde{Q}_k\|.$$ 

Now (2.7) implies the required result. \(\square\)

Thanks to Lemmas 3.5 and 3.7 we get the inequality

$$N_2(f(A) - f(\tilde{A})) \leq \max_{j,k} |f(\lambda_k) - f(\tilde{\lambda}_j)| \sum_{j=1}^{\tilde{m}} \tilde{p}_j \|\tilde{Q}_j\| \sum_{k=1}^{m} p_k \|Q_k\|. \quad (3.9)$$

Proof of Theorem 1.1: Inequality (1.2) is due to (2.6) and Corollary 3.6. Inequality (1.3) is due to (3.9) and Lemma 3.1. \(\square\)

4. Applications. Let us consider the two differential equations

$$\frac{du}{dt} = Au \quad (t > 0) \quad \text{and} \quad \frac{d\tilde{u}}{dt} = \tilde{A}\tilde{u} \quad (t > 0),$$

whose solutions are $u(t)$ and $\tilde{u}(t)$, respectively. Let us take the initial conditions $u(0) = \tilde{u}(0) \in \mathbb{C}^n$. Then we have

$$u(t) - \tilde{u}(t) = (\exp [At] - \exp [\tilde{A}t])u(0).$$

By (1.2),

$$\|u(t) - \tilde{u}(t)\| \leq \|u(0)\| N_2(\exp [At] - \exp [\tilde{A}t]).$$
\[ \leq te^{t\alpha} N_2(A - \tilde{A}) \beta(A) \beta(\tilde{A}) \| u(0) \| \quad (t \geq 0), \]

where

\[ \alpha = \max \left\{ \max_k \Re \lambda_k, \max_j \Re \tilde{\lambda}_j \right\}. \]

Furthermore, let \( 0 \not\in \sigma(A) \cup \sigma(\tilde{A}) \). Then the function

\[ A^{-1/2} \sin \left( A^{1/2} t \right) \quad (t > 0) \]

is the Green function to the Cauchy problem for the equation

\[ \frac{d^2w}{dt^2} + Aw = 0 \quad (t > 0). \]

Simultaneously, consider the equation

\[ \frac{d^2\tilde{w}}{dt^2} + \tilde{A}\tilde{w} = 0 \quad (t > 0), \]

and take the initial conditions

\[ w(0) = \tilde{w}(0) \in \mathbb{C}^n \quad \text{and} \quad w'(0) = \tilde{w}'(0) = 0. \]

Then we have

\[ w(t) - \tilde{w}(t) = \left[ A^{-1/2} \sin \left( A^{1/2} t \right) - \tilde{A}^{-1/2} \sin \left( \tilde{A}^{1/2} t \right) \right] w(0). \]

By (1.2),

\[ \| w(t) - \tilde{w}(t) \| \leq \beta(A) \beta(\tilde{A}) N_2(A - \tilde{A}) \max_j \left| \frac{\sin \left( t \sqrt{\lambda_j} \right)}{\sqrt{\lambda_j}} - \frac{\sin \left( t \sqrt{\tilde{\lambda}_j} \right)}{\sqrt{\tilde{\lambda}_j}} \right| \quad (t > 0). \]

Here one can take an arbitrary branch of the roots.

**Acknowledgment.** I am very grateful to the referee of this paper for his really helpful remarks.

**REFERENCES**


Perturbations of Functions of Diagonalizable Matrices


