2011

Norm estimates for functions of two non-commuting matrices

Michael Gil

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

Recommended Citation

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.
NORM ESTIMATES FOR FUNCTIONS OF TWO NON-COMMUTING MATRICES

MICHAEL GIL

Abstract. A class of matrix valued analytic functions of two non-commuting matrices is considered. A sharp norm estimate is established. Applications to matrix and differential equations are also discussed.

Key words. Functions of non-commuting matrices, Norm estimate, Matrix equation, Differential equation.

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

1. Introduction and statement of the main result. In the book [5], I.M. Gel’fand and G.E. Shilov have established an estimate for the norm of a regular matrix-valued function in connection with their investigations of partial differential equations. However, that estimate is not sharp; it is not attained for any matrix. The problem of obtaining a sharp estimate for the norm of a matrix-valued function has been repeatedly discussed in the literature, cf. [2]. In the paper [6] (see also [7]), the author has derived an estimate for regular matrix-valued functions, which is attained in the case of normal matrices. In [8], the results of the paper [6] were generalized to functions of two commuting matrices. In the present paper, we establish a sharp estimate for the norm of a matrix-valued function of two non-commuting matrices.

It should be noted that functions of many operators were investigated by many mathematicians, (cf. [1, 15, 16] and references therein) however the norm estimates were not considered, but as it is well-known, matrix valued functions give us representations of solutions of various differential, difference equations and matrix equations. This fact allows us to investigate stability, well-posedness and perturbations of these equations by norm estimates for matrix valued functions, cf. [2].

Let $\mathbb{C}^n$ be the Euclidean space with scalar product $(\cdot, \cdot)$, the Euclidean norm $\| \cdot \| = \sqrt{(\cdot, \cdot)}$ and the unit operator $I$. Unless otherwise stated $A$, $K$ and $\tilde{A}$ will be $n \times n$ matrices. $\|A\| = \sup_{h \in \mathbb{C}^n} \|Ah\|/\|h\|$ is the spectral (operator) norm of $A$. By $\sigma(A)$ and $R_z(A) = (A - zI)^{-1}$ ($z \notin \sigma(A)$) we denote the spectrum and resolvent of $A$, respectively.

Received by the editors on December 23, 2010. Accepted for publication on May 14, 2011. Handling Editor: Harm Bart.

Department of Mathematics, Ben Gurion University of the Negev, PO Box 653, Beer-Sheva 84105, Israel (gilmi@bezeqint.net).
Let $\Omega_A$ and $\tilde{\Omega}_A$ be open simple connected supersets of $\sigma(A)$ and $\sigma(\tilde{A})$, respectively, and $f$ be a scalar function analytic on $\Omega_A \times \tilde{\Omega}_A$. We define the matrix valued function

$$F(f, A, K, \tilde{A}) := -\frac{1}{4\pi^2} \int_{C_A} \int_{C_{\tilde{A}}} f(z, w) R_z(A) K R_w(\tilde{A}) dw \, dz,$$

where $C_A \subset \Omega_A, C_{\tilde{A}} \subset \tilde{\Omega}_A$ are closed contour surrounding $\sigma(A)$ and $\sigma(\tilde{A})$, respectively. Such functions play an essential role in the theory of matrix equations. More specifically, consider the matrix equation

$$\sum_{j=0}^{m_1} \sum_{k=0}^{m_2} c_{jk} A^j X \tilde{A}^k = K,$$

where $X$ should be found and $c_{jk}$ are complex numbers. Put

$$p(z, w) = \sum_{j=0}^{m_1} \sum_{k=0}^{m_2} c_{jk} z^j \tilde{w}^k.$$

Then by Theorem 3.1 from [2, Chapter 1] a unique solution of equation (1.2) is given by the formula

$$X = F \left( \frac{1}{p(z, w)}, A, K, \tilde{A} \right)$$

provided $\lambda_k \neq \tilde{\lambda}_j (j, k = 1, \ldots, n)$. Throughout the rest of this paper $\lambda_k$ and $\tilde{\lambda}_j$ are the eigenvalues counted with their multiplicities of $A$ and $\tilde{A}$, respectively. Equations of the type (1.2) naturally arose in various applications, cf. [2, 14, 12]. The Lyapunov equation $A^*X + XA = K$, cf. [2], and the Lyapunov type equation

$$X + A^*XA = K$$

which play an important role in the theory of difference equations, cf. [9] are the examples of equation (1.2). These equations recently attracted the attention of many mathematicians. Mainly, numerical methods for the solutions of matrix equations were developed, cf. [11, 13, 17]. In the paper [3], reflexive and anti-reflexive solutions of a linear matrix equation were explored. No estimates were established for solutions of these equations. Furthermore, suppose that

$$T(t) := -\frac{1}{4\pi^2} \int_{C_A} \int_{C_{\tilde{A}}} e^{t(z+w)} R_z(A) K R_w(\tilde{A}) dw \, dz.$$

Take into account that $z R_z(A) = AR_z(A) - I$. Then simple calculations show that

$$T'(t) = -\frac{1}{4\pi^2} \int_{C_A} \int_{C_{\tilde{A}}} (z + w) e^{t(z+w)} R_z(A) K R_w(\tilde{A}) dw \, dz =$$
\[
- \frac{1}{4\pi^2} \int_{C_A} \int_{C_{\tilde{A}}} e^{t(z+w)} \left[ AR_z(A)KR_w(\tilde{A}) + R_z(A)KR_w(\tilde{A})\tilde{A} \right] dw \, dz.
\]

So
\[
T'(t) = AT(t) + T(t)\tilde{A}.
\] (1.6)

Such equations arise in numerous applications, in particular in the theory of vector differential equations, cf. [10, p. 509], [2, Section VI.4, equation (4.15) and Section VI.2], [4, Section XV.10]. Additional examples are given in Section 3.

The following quantity plays a key role in this article:
\[
g(A) = \left[ N_2^2(A) - \sum_{k=1}^{n} |\lambda_k|^2 \right]^{1/2},
\]

where \( N_2(A) = (\text{Trace } AA^*)^{1/2} \) is the Frobenius (Hilbert-Schmidt norm) of \( A \). Here, \( A^* \) is adjoint to \( A \). The following relations are checked in [7, Section 2.1]:
\[
g^2(A) \leq N_2^2(A) - |\text{Trace } A^2| \quad \text{and} \quad g^2(A) \leq \frac{N_2^2(A - A^*)}{2} = 2N_2^2(A_I),
\] (1.7)

where \( A_I = (A - A^*)/2i \). If \( A \) is a normal matrix: \( AA^* = A^*A \), then \( g(A) = 0 \).

By \( \text{co}(A) \) we denote the closed convex hull of \( \sigma(A) \). Let \( f(z, w) \) be regular on a neighborhood of \( \text{co}(A) \times \text{co}(\tilde{A}) \). Put
\[
f^{(j,k)}(z, w) = \frac{\partial^{j+k} f(z, w)}{\partial z^j \partial w^k},
\]
and let the numbers \( \eta_{jk} = \eta_{jk}(f, A, \tilde{A}) \) be given by
\[
\eta_{00} = \sup_{z \in \sigma(A), w \in \sigma(\tilde{A})} |f(z, w)|; \quad \eta_{jk} = \frac{1}{(jk)!^{3/2}} \sup_{z \in \text{co}(A), w \in \text{co}(\tilde{A})} |f^{(j,k)}(z, w)|;
\]
\[
\eta_{0j} := \frac{1}{(j!)^{3/2}} \sup_{z \in \sigma(A), w \in \sigma(\tilde{A})} \left| \frac{\partial^j f(z, w)}{\partial w^j} \right|, 
\]
and
\[
\eta_{j0} := \frac{1}{(j!)^{3/2}} \sup_{z \in \text{co}(A), w \in \sigma(\tilde{A})} \left| \frac{\partial^j f(z, w)}{\partial z^j} \right| \quad (j, k \geq 1).
\]

Now we are in a position to formulate the main result of the paper.
Theorem 1.1. Let both $A$ and $\tilde{A}$ be non-normal matrices and $f(z,w)$ be regular on a neighborhood of $\text{co}(A) \times \text{co}(\tilde{A})$. Then
\[
\|F(f, A, K, \tilde{A})\| \leq N_2(K) \sum_{j,k=0}^{n-1} \eta_{jk} g^j(A)g^k(\tilde{A}).
\]
If $A$ is normal, $\tilde{A}$ is non-normal and $f(z,w)$ is regular on a neighborhood of $\sigma(A) \times \text{co}(\tilde{A})$, then
\[
\|F(f, A, K, \tilde{A})\| \leq N_2(K) \sum_{j=0}^{n-1} \eta_{0j} g^j(\tilde{A}).
\]
If $\tilde{A}$ is normal, $A$ is non-normal and $f(z,w)$ is regular on a neighborhood of $\sigma(\tilde{A}) \times \text{co}(A)$, then
\[
\|F(f, A, K, \tilde{A})\| \leq N_2(K) \sum_{j=0}^{n-1} \eta_{j0} g^j(A).
\]
If both $A$ and $\tilde{A}$ are normal and $f(z,w)$ is regular on a neighborhood of $\sigma(A) \times \sigma(\tilde{A})$, then
\[
\|F(f, A, K, \tilde{A})\| \leq N_2(K) \max_{j,k} |f(\lambda_j, \tilde{\lambda}_k)|.
\]

2. Proof of Theorem 1.1. We need the following result proved in [8].

Lemma 2.1. Let $\Omega$ and $\tilde{\Omega}$ be the closed convex hulls of the complex points $x_0, x_1, \ldots, x_n$ and $y_0, y_1, \ldots, y_m$, respectively, and let a scalar-valued function $f(z,w)$ be regular on a neighborhood of $\Omega \times \tilde{\Omega}$. Additionally, let $L$ and $\tilde{L}$ be the boundaries of $\Omega$ and $\tilde{\Omega}$, respectively. Then with the notation
\[
Y(x_0, \ldots, x_n; y_0, \ldots, y_m) = -\frac{1}{4\pi^2} \int_{L} \int_{\tilde{L}} \frac{f(z,w)dz \, dw}{(z - x_0) \cdots (z - x_n)(w - y_0) \cdots (w - y_m)},
\]
we have
\[
|Y(x_0, \ldots, x_n; y_0, \ldots, y_m)| \leq \frac{1}{n!m!} \sup_{z \in \Omega, w \in \tilde{\Omega}} |f^{(n,m)}(z,w)|.
\]

Let $\{e_k\}$ and $\{\tilde{e}_k\}$ be the orthogonal normal bases of the triangular representation (Schur’s bases) to $A$ and $\tilde{A}$, respectively. So,
\[
Ae_k = \sum_{j=1}^{k} a_{jk}e_j.
\]
We can write

\[ A = D_A + V_A, \quad \tilde{A} = D_{\tilde{A}} + V_{\tilde{A}}, \]

(2.1)

where \( D_A, D_{\tilde{A}} \) are the diagonal parts, \( V_A \) and \( V_{\tilde{A}} \) are the nilpotent parts of \( A \) and \( \tilde{A} \), respectively. Namely,

\[ D_A e_k = \lambda_k e_k; \quad V_A e_k = \sum_{j=1}^{k-1} a_{jk} e_j. \]

Similarly, \( D_{\tilde{A}} \) and \( V_{\tilde{A}} \) are defined. Furthermore, let \( |V_A| \) be the operator whose entries in \( \{e_k\} \) are the absolute values of the entries of a matrix \( V_A \). That is, \( \langle |V_A| e_j, e_k \rangle = \langle (V_A e_j, e_k) \rangle \) and

\[ |V_A| = \sum_{k=1}^{n} \sum_{j=1}^{k-1} |a_{jk}| \langle \cdot, e_k \rangle e_j, \]

Similarly, \( |V_{\tilde{A}}| \) is defined with respect to \( \{\tilde{e}_k\} \). In addition, \( |K| \) is defined by

\[ |K| \tilde{e}_j = \sum_{k=1}^{n} \langle (K \tilde{e}_j, e_k) | e_k \rangle. \]

\[ \text{Lemma 2.2.} \quad \text{Under the hypothesis of Theorem 1.1, the inequality} \]

\[ \|F(f, A, K, \tilde{A})\| \leq \| |K| \| \sum_{j,k=0}^{n-1} \sqrt{k!j!\eta_{jk}} \| |V_A| \|^j \| |V_{\tilde{A}}| \|^k \]

\[ \text{is true, where} \quad V_A \quad \text{and} \quad V_{\tilde{A}} \quad \text{are the nilpotent parts of} \quad A \quad \text{and} \quad \tilde{A}, \quad \text{respectively.} \]

\[ \text{Proof.} \quad \text{It is not hard to see that the representation (2.1) implies the equality} \]

\[ (A - I\lambda)^{-1} = (D_A + V_A - \lambda I)^{-1} = (I + R_\lambda(D_A)V_A)^{-1}R_\lambda(D_A) \]

for all regular \( \lambda \). According to Lemma 1.7.1 from [7] \( R_\lambda(D_A)V_A \) is a nilpotent operator, because \( V_A \) and \( R_\lambda(D_A) \) the same invariant subspaces. Hence, \( (R_\lambda(D_A)V_A)^n = 0 \). Therefore, from (1.1) it follows

\[ F(f, A, K, \tilde{A}) = \sum_{j,k=0}^{n-1} M_{jk}, \]

(2.2)

where

\[ M_{jk} = (-1)^{k+j} \frac{1}{4\pi^2} \int_{C_A} \int_{C_{\tilde{A}}} f(z, w)(R_z(D_A)V_A)^j R_z(D_A)K(R_{w}(D_{\tilde{A}})V_{\tilde{A}})^k R_w(D_{\tilde{A}}) \, dz \, dw. \]
Since $D_A$ is a diagonal matrix with respect to the Schur basis $\{e_k\}$ and its diagonal entries are the eigenvalues of $A$, we obtain

$$R_z(D_A) = \sum_{j=1}^{n} \frac{Q_j}{\lambda_j(A) - z},$$

where $Q_k = (\cdot, e_k)e_k$. Similarly,

$$R_z(D_A) = \sum_{j=1}^{n} \frac{\tilde{Q}_j}{\lambda_j(A) - z},$$

where $\tilde{Q}_k = (\cdot, \tilde{e}_k)\tilde{e}_k$. Taking into account that $Q_s V_A Q_m = 0$, $\tilde{Q}_s V_A \tilde{Q}_m = 0$ ($s \geq m$), we get

$$M_{jk} = \sum_{1 \leq s_1 < s_2 < \cdots < s_{j+1} \leq n} Q_{s_1} V_A Q_{s_2} V_A \cdots V_A Q_{s_{j+1}} K \times$$

$$\times \sum_{1 \leq m_1 < m_2 < \cdots < m_{k+1} \leq n} \tilde{Q}_{m_1} V_A \tilde{Q}_{m_2} V_A \cdots V_A \tilde{Q}_{m_{k+1}} I(s_1, \ldots, s_{j+1}, m_1, \ldots, m_{k+1}),$$

where $0 \leq j, k \leq n - 1$ and

$$I(s_1, \ldots, s_{j+1}, m_1, \ldots, m_{k+1}) =$$

$$\frac{(-1)^{k+j}}{4\pi^2} \int_{C_A} \int_{C_A} \frac{f(z, w)dz \; dw}{(\lambda_{s_1}(A) - z) \cdots (\lambda_{s_{j+1}}(A) - z)(\lambda_{m_1}(\tilde{A}) - w) \cdots (\lambda_{m_{k+1}}(\tilde{A}) - w)}.$$

Hence, with $M_{jk} = M$, we have

$$| \langle M \tilde{e}_m, e_s \rangle | = | \sum_{s < s_2 < \cdots < s_{j+1} \leq n} \sum_{1 \leq m_1 < m_2 < \cdots < m} I(s_1, \ldots, s_{j+1}, m_1, \ldots, m) \times$$

$$(Q_s V_A Q_{s_2} V_A \cdots V_A Q_{s_{j+1}} K \tilde{Q}_{m_1} V_A \tilde{Q}_{m_2} V_A \cdots V_A \tilde{Q}_{m_{k+1}} \tilde{e}_m, e_s) | \leq J_{jk} \sum_{s < s_2 < \cdots < s_{j+1} \leq n}$$

$$\times \sum_{1 \leq m_1 < m_2 < \cdots < m} (Q_s |V_A| Q_{s_2} |V_A| \cdots Q_{s_{j+1}} |K| \tilde{Q}_{m_1} |V_A| \tilde{Q}_{m_2} |V_A| \cdots \tilde{Q}_{m_{k+1}} \tilde{e}_m, e_s),$$

where

$$J_{jk} := \max_{1 \leq s_1 < \cdots < s_{j+1} \leq n, 1 \leq m_1 < \cdots < m_{k+1} \leq n} |I(s_1, \ldots, s_{j+1}, m_1, \ldots, m_{k+1})|. $$
Thus $|(M \tilde{e}_m, e_s)| \leq (T \tilde{e}_m, e_s)$, where
\[
T = J_{jk} \sum_{s_1 < s_2 < \cdots < s_{j+1} \leq n} \sum_{1 \leq m_1 < m_2 < \cdots < m_{k+1} \leq n} Q_{s_1} |V_A| Q_{s_2} |V_A| \cdots |V_A| Q_{s_{j+1}} |K| \times \\
\times \tilde{Q}_{m_1} |V_A| \tilde{Q}_{m_2} |V_A| \cdots |V_A| \tilde{Q}_{m_{k+1}}.
\] (2.3)

Take into account that
\[
M x = \sum_{k=1}^{n} (x, \tilde{e}_k) M \tilde{e}_k = \sum_{j=1}^{n} \sum_{k=1}^{n} (x, \tilde{e}_k)(M \tilde{e}_k, e_j) e_j \quad (x \in \mathbb{C}^n).
\]

So
\[
\|M x\|^2 = \sum_{j=1}^{n} \left| \sum_{k=1}^{n} (x, \tilde{e}_k)(M \tilde{e}_k, e_j) \right|^2 \leq \\
\sum_{j=1}^{n} \left( \sum_{k=1}^{n} (x, \tilde{e}_k)(T \tilde{e}_k, e_j) \right)^2.
\]

Since $\|x\| = \|y\|$ for
\[
y = \sum_{k=1}^{n} (x, \tilde{e}_k) \tilde{e}_k,
\]
we obtain $\|M\| \leq \|T\|$. But
\[
\sum_{1 \leq s_1 < s_2 < \cdots < s_{j+1} \leq n} Q_{s_1} |V_A| Q_{s_2} |V_A| \cdots |V_A| Q_{s_{j+1}} = |V_A|^j
\]
and
\[
\sum_{1 \leq m_1 < m_2 < \cdots < m_{k+1} \leq n} \tilde{Q}_{m_1} |V_A| \tilde{Q}_{m_2} |V_A| \cdots |V_A| \tilde{Q}_{m_{k+1}} = |V_A|^k.
\]

So by (2.3)
\[
\|M_{jk}\| \leq \|T\| \leq J_{jk} \|V_A|^j |K| |V_A|^k \| \quad (j, k \geq 0). \quad (2.4)
\]

Due to Lemma 2.1
\[
J_{jk} \leq \sup_{z \in \text{co}(A), w \in \text{co}(\tilde{A})} \frac{|f^{(j,k)}(z, w)|}{j!k!} = \sqrt{j!k!} \eta_{jk} \quad (j, k \geq 1).
\]
Thus,
\[ \| M_{jk} \| \leq \sqrt{j!k!} \eta_{jk} \| V_A \|^j \| K \| |V_A|^k \| \quad (j, k \geq 0). \] (2.5)

This inequality and (2.2) imply the required result. \[ \square \]

**Proof of Theorem 1.1.** Theorem 2.5.1 from [7] implies
\[ \| W_k \| \leq \frac{1}{\sqrt{k!}} N_2^k(W) \] (2.6)
for any \( n \times n \) nilpotent matrix \( W \). Take into account that \( N_2(\| V_A \|) = N_2(V_A) \).
Moreover, by Lemma 2.3.2 from [7], \( N_2(V_A) = g(A) \). Thus,
\[ \| V_A \|^k \| \leq \frac{1}{\sqrt{k!}} g^k(A) \quad (k = 1, \ldots, n - 1). \]

The similar inequality holds for \( V_\tilde{A} \). In addition,
\[ N_2^2(\| K \|) = \sum_{j=1}^{n} \| K |\vec{e}_j \|^2 = \sum_{j=1}^{n} \sum_{k=1}^{n} \| (K |\vec{e}_j, e_k) \|^2 = \sum_{j=1}^{n} \sum_{k=1}^{n} \| K |\vec{e}_j \|^2 = N_2^2(K). \]

Now the previous lemma yields the required result. \[ \square \]

**3. Examples.** Consider the equation
\[ AX - X \tilde{A} = K \] (3.1)
assuming that
\[ \delta := \text{dist} \ (\text{co}(A), \text{co}(\tilde{A})) > 0. \]

Take \( f(z, w) = \frac{1}{z-w} \). Then
\[ \eta_{jk} \leq \frac{(k+j)!}{\delta^{j+k+1} (k! j!)^{3/2}} \quad (j, k = 0, 1, \ldots, n - 1). \]

Hence, by Theorem 1.1 and (1.3) a solution of (3.1) satisfies the inequality
\[ \| X \| \leq N_2(K) \sum_{j,k=0}^{n-1} \frac{(k+j)!}{\delta^{j+k+1} (k! j!)^{3/2}} g^j(A) g^k(\tilde{A}). \]

Finally, consider the function
\[ S(x) := -\frac{1}{4\pi^2} \int_{C_A} \int_{C_\tilde{A}} \sin (x(z+w)) R_z(A)R_w(\tilde{A}) \, dw \, dz \quad (x \in \mathbb{R}). \]
We have
\[ S''(x) = \frac{1}{4\pi^2} \int_{C_A} \int_{C_A} (z + w)^2 \sin (x(z + w)) R_z(A) K R_w(\tilde{A}) dw \, dz. \]
But \( z R_z(A) = AR_z(A) - I \) and therefore,
\[ z^2 R_z(A) = zAR_z(A) - zI = A(AR_z(A) - I) - zI = A^2 R_z(A) - I - zI. \]
So, \( S(x) \) is a solution of the equation
\[ S'' = A^2 S + AS\tilde{A} + S\tilde{A}^2. \]

REFERENCES