Discussiones Mathematicae
Differential Inclusions, Control and Optimization 34 (2014) 191–206
doi:10.7151/dmdico.1163

NORM ESTIMATES FOR SOLUTIONS OF MATRIX EQUATIONS $AX - XB = C$ AND $X - AXB = C$

Michael I. Gil'

Department of Mathematics
Ben Gurion University of the Negev
P.O. Box 653, Beer-Sheva 84105, Israel

e-mail: gilmi@bezeqint.net

Abstract

Let A, B and C be matrices. We consider the matrix equations $Y - AYB = C$ and $AX - XB = C$. Sharp norm estimates for solutions of these equations are derived. By these estimates a bound for the distance between invariant subspaces of matrices is obtained.

Keywords: matrix equations, norm estimates, perturbations, invariant subspaces.

2010 Mathematics Subject Classification: 15A24, 93C15, 93C55, 15A41, 15A69.

1. Introduction and notation

Let A, B and C be $n \times n$ matrices. Consider the equations

$$Y - AYB = C \tag{1.1}$$

and

$$AX - XB = C. \tag{1.2}$$

Equations (1.1) and (1.2) have been widely studied in pure and applied mathematics, especially in control problems, cf. [3, 6, 7, 17, 18, 19]. In particular, the rank, image and kernel of solutions of (1.2) are important design parameters, and the role controllability and observability in the existence of solutions of (1.2) has been demonstrated in [15].

Norm estimates for solutions of equations were established when the coefficients are normal matrices [1] and diagonalizable ones [11]. Under the condition that the convex hulls of spectra A and B are disjoint, the equations have been considered in [10], but that condition is rather restrictive. Except the just pointed results, to the best of our knowledge, no estimates were established for solutions of equations (1.1) and (1.2), although such estimates are very important, in particular, for the stability analysis, as well as for linear and nonlinear perturbations, cf. [4]. In the present paper we suggest new sharp solution estimates for equations (1.1) and (1.2) with non-normal and non-diagonalizable coefficients. In Section 5 below we also consider equations whose coefficients have different dimensions and discuss applications of our results to perturbations of invariant subspaces of matrices. The importance of invariant subspaces for applications, in particular, in the control theory, is very well explained in [13].

Introduce the notations. Let $\mathbb{C}^n$ be the Euclidean space with scalar product $(\cdot,\cdot)$, the Euclidean norm $\|\cdot\| = \sqrt{(\cdot,\cdot)}$ and the unit matrix I. For an $n \times n$ matrix A, is A^* is the adjoint one; $\lambda_k(A), k = 1, \dots, n$, are the eigenvalues of A, counted with their multiplicities; $\|A\| = \sup_{h \in \mathbb{C}^n} \|Ah\|/\|h\|$ is the spectral (operator) norm of A; $r_s(A)$ is the (upper) spectral radius; $r_l(A) = \min_k |\lambda_k(A)|$ is the lower spectral radius.

The following quantity (the departure from normality of A) plays a key role hereafter:

$$g(A) = \left[|A|_F^2 - \sum_{k=1}^{n} |\lambda_k(A)|^2 \right]^{1/2},$$

where $|A|_F = (Trace\ AA^*)^{1/2}$ is the Frobenius (Hilbert-Schmidt norm) of A.

The following relations are checked in [8, Section 2.1]:

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

where $A_I = (A - A^*)/2i$. If A is a normal matrix: $AA^* = A^*A$, then $g(A) = 0$. If A_1 and A_2 are commuting matrices, then $g(A_1 + A_2) \le g(A_1) + g(A_2)$. By the inequality between geometric and arithmetic mean values we have

$$\left(\frac{1}{n} \sum_{k=1}^{n} |\lambda_k(A)|^2 \right)^n \ge \left(\prod_{k=1}^{n} |\lambda_k(A)| \right)^2.$$

So $g^2(A) \le |A|_F^2 - n(det\ A)^{2/n}$. If A is invertible, then from (1.2) we have

$$X - A^{-1}XB = A^{-1}C. \tag{1.3}$$

Taking in (1.3) A instead of A^{-1} and C instead of $A^{-1}C$, we get (1.1).

2. Statements of the main results

2.1. Equation (1.1)

Theorem 1. *Let*

$$r_s(A)r_s(B) < 1. \tag{2.1}$$

Then equation (1.1) *has a unique solution* Y *which can be represented as*

$$Y = \sum_{k=0}^{\infty} A^k C B^k. \tag{2.2}$$

Moreover,

$$\|Y\| \le \|C\| \sum_{j,k=0}^{n-1} \frac{g^k(A)g^j(B)}{(k!j!)^{3/2}} \sum_{m=0}^{\infty} \frac{(m!)^2 r_s^{m-k}(A) r_s^{m-j}(B)}{(m-k)!(m-j)!}. \tag{2.3}$$

This theorem is proved in the next section. Recall that $1/(m-k)! = 0$ if $m < k$. If A is normal, then $g(A) = 0$ and

$$\|Y\| \le \|C\| \sum_{j=0}^{n-1} \frac{g^j(B)}{(j!)^{3/2}} \sum_{m=0}^{\infty} \frac{m! r_s^m(A) r_s^{m-j}(B)}{(m-j)!}.$$

But

$$\begin{aligned} \sum_{m=0}^{\infty} \frac{m! r_s^m(A) r_s^{m-j}(B)}{(m-j)!} &= r_s^j(A) \sum_{m=0}^{\infty} \frac{m!(r_s(A)r_s(B))^{m-j}}{(m-j)!} \\ = r_s^j(A) \frac{d^j}{dx^j} \sum_{m=0}^{\infty} x^m &= j! r_s^j(A)(1-x)^{-j-1} \quad (x = r_s(A)r_s(B)). \end{aligned} \tag{2.4}$$

Thus, we have

$$\|Y\| \le \|C\| \sum_{j=0}^{n-1} \frac{g^j(B) r_s^j(A)}{(j!)^{1/2}(1 - r_s(A)r_s(B))^{j+1}},$$

provided A is normal. If both A and B are normal, then

$$\|Y\| \le \|C\| \sum_{m=0}^{\infty} (r_s(B)r_s(B))^m = \frac{\|C\|}{1 - r_s(A)r_s(B)}. \tag{2.5}$$

The inequality (2.5) is attained, if A, B and C are commuting normal matrices.

According to (2.4) we obtain

$$\sum_{m=0}^{\infty} \frac{(m!)^2 r_s^m(A) r_s^m(B)}{(m-k)!(m-j)!} \leq \sum_{t=0}^{\infty} \frac{t!(r_s(A)r_s(B))^{t/2}}{(t-k)!} \sum_{m=0}^{\infty} \frac{m!(r_s(A)r_s(B))^{m/2}}{(m-j)!}$$

$$= (r_s(A)r_s(B))^{(k+j)/2} \sum_{t=0}^{\infty} \frac{t!(r_s(A)r_s(B))^{(t-k)/2}}{(t-k)!} \sum_{m=0}^{\infty} \frac{m!(r_s(A)r_s(B))^{(m-j)/2}}{(m-j)!}$$

$$= \frac{j!k!(r_s(A)r_s(B))^{(k+j)/2}}{[1-(r_s(A)r_s(B))^{1/2}]^{(k+j+2)}}.$$

Now inequality (2.3) implies

Corollary 2. *Let condition* (2.1) *hold,* $r_s(A) \neq 0$ *and* $r_s(B) \neq 0$. *Then a unique solution* Y *of* (1.1) *satisfies the inequality*

$$\|Y\| \leq \|C\| \sum_{j,k=0}^{n-1} \frac{g^k(A) g^j(B)}{\sqrt{k!j!}} \frac{r_s^{(j-k)/2}(A) r_s^{(k-j)/2}(B)}{(1-\sqrt{r_s(A)r_s(B))})^{j+k+2}}.$$

Furthermore, Theorem 1 and simple calculations imply

Corollary 3. *Let* $r_s(A) < 1$. *Then the equation*

$$Y - AYA^* = C \tag{2.6}$$

has a unique solution Y_L, *which satisfies the inequalities*

$$\|Y_L\| \leq \|C\| \sum_{m=0}^{\infty} \left(\sum_{k=0}^{n-1} \frac{g^k(A) m! r_s^{m-k}(A)}{(k!)^{3/2}(m-k)!} \right)^2. \tag{2.7}$$

As it is well-known, equation (2.6) is an important tool in the theory of difference equations, cf. [9]. From (2.7) it follows

$$\|Y_L\| \leq \|C\| \left(\sum_{m=0}^{\infty} \sum_{k=0}^{n-1} \frac{g^k(A) m! r_s^{m-k}(A)}{(k!)^{3/2}(m-k)!} \right)^2.$$

Now taking into account (2.4) we get

$$\|Y_L\| \leq \|C\| \left(\sum_{k=0}^{n-1} \frac{g^k(A)}{(k!)^{1/2}(1-r_s(A))^{k+1}} \right)^2. \tag{2.8}$$

2.2. Equation (1.2)

Let us consider (1.2), assuming that

$$r_l(A) > r_s(B). \tag{2.9}$$

This inequality is equivalent the one $r_s(A^{-1})r_s(B) < 1$. Then according to (1.3) and (2.2) a solution of (1.2) is represented by

$$X = \sum_{k=1}^{\infty} A^{-1-k} C B^k. \tag{2.10}$$

This result is well-known, cf. [2].

Theorem 4. *Let condition* (2.9) *hold. Then equation* (1.2) *has a unique solution* X*, which can be represented by* (2.10)*. Moreover,*

$$\|X\| \le \|C\| \sum_{j,k=0}^{n-1} \frac{\zeta^k(A) g^j(B)}{(k!j!)^{3/2}} \sum_{m=0}^{\infty} \frac{m!(m+1)! r_s^{m-j}(B)}{r_l^{m+1-k}(A)(m-j)!(m+1-k)!}, \tag{2.11}$$

where

$$\zeta(A) := \sum_{k=0}^{n-1} \frac{g^{k+1}(A)}{r_l^{k+2}(A)(k!)^{1/2}}.$$

This theorem is also proved in the next section. If A is normal, then $g(A) = 0$ and therefore, $\zeta(A) = 0$. Thus (2.11) implies

$$\|X\| \le \|C\| \sum_{j=0}^{n-1} \frac{g^j(B)}{(j!)^{3/2}} \sum_{m=0}^{\infty} \frac{m! r_s^{m-j}(B)}{r_l^{m+1}(A)(m-j)!}.$$

But

$$\sum_{m=0}^{\infty} \frac{m! r_s^{m-j}(B)}{r_l^{m+1}(A)(m-j)!} = r_l^{-j-1}(A) \sum_{m=0}^{\infty} \frac{m! a^{m-j}}{(m-j)!} \quad (0 \le a = r_s(B)/r_l(A) < 1)$$

and

$$\sum_{m=0}^{\infty} \frac{m! a^{m-j}}{(m-j)!} = \frac{d^j}{da^j} \sum_{m=0}^{\infty} a^m = j!(1-a)^{-j-1}.$$

Consequently, if A is normal, then a solution of (1.2) under (2.9) satisfies the inequality

$$\|X\| \le \|C\| \sum_{j=0}^{n-1} \frac{g^j(B)}{(j!)^{1/2}(r_l(A) - r_s(B))^{j+1}}.$$

If both A and B are normal, then

$$\|X\| \le \|C\| \sum_{m=0}^{\infty} \frac{r_s^m(B)}{r_l^{m+1}(A)} = \frac{\|C\|}{r_l(A) - r_s(B)}. \tag{2.12}$$

Furthermore, by (1.3) and Corollary 2 with A^{-1} instead of A, and $A^{-1}C$ instead of C, applying Lemma 6 proved below we get

Corollary 5. *Let condition* (2.9) *hold. Then a unique solution X of equation* (1.2) *satisfies the inequality*

$$\|X\| \le \|A^{-1}C\| \sum_{j,k=0}^{n-1} \frac{\zeta^k(A) g^j(B)}{(k!j!)^{1/2}} \frac{(r_l(A) r_s(B))^{(k-j)/2}(B)}{[1 - (r_s(B)/r_l(A))^{1/2}]^{j+k+2}}.$$

Note that $\|A^{-1}\|$ can be estimated by inequality (3.6) presented below.

3. Proofs of Theorems 1 and 4

***Proof of Theorem* 1.** By Corollary 2.7.2 from [8] we can write

$$\|A^m\| \le \sum_{k=0}^{n-1} \frac{m! g^k(A) r_s^{m-k}(A)}{(m-k)!(k!)^{3/2}} \quad (m = 1, 2, \ldots). \tag{3.1}$$

Thus

$$\Big\| \sum_{m=0}^{\infty} A^m C B^m \Big\| \le \|C\| \sum_{m=0}^{\infty} \sum_{k=0}^{n-1} \frac{m! g^k(A) r_s^{m-k}(A)}{(m-k)!(k!)^{3/2}} \sum_{j=0}^{n-1} \frac{m! g^j(B) r_s^{m-j}(B)}{(m-j)!(j!)^{3/2}},$$

or

$$\Big\| \sum_{m=0}^{\infty} A^m C B^m \Big\| \le \|C\| \sum_{j,k=0}^{n-1} \frac{g^k(A) g^j(B)}{(k!j!)^{3/2}} \sum_{m=0}^{\infty} \frac{(m!)^2 r_s^{m-k}(A) r_s^{m-j}(B)}{(m-k)!(m-j)!}. \tag{3.2}$$

The series

$$\sum_{m=0}^{\infty} \frac{(m!)^2 r_s^{m-k}(A) r_s^{m-j}(B)}{(m-k)!(m-j)!}$$

converges. Therefore, the series in (2.2) also converges. In addition, according to (2.2)

$$\sum_{k=0}^{\infty} A^k C B^k - A \sum_{k=0}^{\infty} A^k C B^k B = C.$$

So (2.2) really solves (1.1). Now inequality (2.3) follows from (3.2). ■

To prove Theorem 4 we need to estimate the departure $g(A^{-1})$ of normality of A^{-1}. That is,

$$g(A^{-1}) = \left[|A^{-1}|_F^2 - \sum_{k=1}^{n} |\lambda_k(A^{-1})|^2 \right]^{1/2}.$$

Lemma 6. *Let $r_l(A) > 0$. Then $g(A^{-1}) \le \zeta(A)$.*

Proof. Due to the Schur theorem on the reduction of a matrix A to the triangular form, we can write

(3.3) $$A = D + V \quad (\lambda_k(A) = \lambda_k(D);\ k = 1, \dots, n)$$

with a normal (diagonal) matrix D and a nilpotent (a strictly upper-triangular) matrix V, having the same invariant subspaces. We will call equality (3.3) the triangular representation of matrix A; D and V will be called the diagonal part and the nilpotent part of A, respectively. It is not hard to check that

(3.4) $$g(A) = |V|_F,$$

cf. [8, Lemma 2.3.2]. Furthermore, making use (3.3), we have $A^{-1} = D^{-1} + W$, where

$$W = A^{-1} - D^{-1} = -A^{-1}(A - D)D^{-1} = -A^{-1}VD^{-1}$$

is a nilpotent matrix, since A^{-1}, V, D^{-1} and W have the same invariant subspaces. But D^{-1} is normal, and therefore, W is the nilpotent part of A^{-1}, and according to (3.4)

(3.5)
$$g(A^{-1}) = |W|_F = |A^{-1}VD^{-1}|_F \le \|A^{-1}\| |V|_F \|D^{-1}\| = \|A^{-1}\| g(A)\, r_l^{-1}(A).$$

Due to Corollary 2.1.2 from [8],

(3.6) $$\|A^{-1}\| \le \sum_{k=0}^{n-1} \frac{g^k(A)}{r_l^{k+1}(A)(k!)^{1/2}}.$$

This and (3.5) prove the result. ■

Inequality (3.1) and the previous lemma imply

Corollary 7. *Let $r_l(A) > 0$. Then*

$$\|A^{-m}\| \le \sum_{k=0}^{n-1} \frac{m! g^k(A^{-1})}{r_l^{m-k}(A)(m-k)!(k!)^{3/2}}$$
$$\le \sum_{k=0}^{n-1} \frac{m! \zeta^k(A)}{r_l^{m-k}(A)(m-k)!(k!)^{3/2}} \quad (m = 1, 2, \ldots).$$

***Proof of Theorem* 4.** Due to (2.8), (3.1) and the previous corollary we can write

$$\|X\| = \|\sum_{m=0}^{\infty} A^{-m-1} C B^m\| \le$$

$$\|C\| \sum_{m=0}^{\infty} \sum_{k=0}^{n-1} \frac{(m+1)! \zeta^k(A)}{r_l^{m+1-k}(A)(m+1-k)!(k!)^{3/2}} \sum_{j=0}^{n-1} \frac{m! g^j(B) r_s^{m-j}(B)}{(m-j)!(j!)^{3/2}},$$

or

$$\|X\| \le \|C\| \sum_{j,k=0}^{n-1} \frac{\zeta^k(A) g^j(B)}{(j!k!)^{3/2}} \sum_{m=0}^{\infty} \frac{m!(m+1)! r_s^{m-j}(B)}{r_l^{m+1-k}(A)(m+1-k)!(m-j)!}.$$

The series

$$\sum_{m=0}^{\infty} \frac{m!(m+1)! r_s^{m-j}(B)}{r_l^{m+1-k}(A)(m+1-k)!(m-j)!}$$

converges. Therefore, the series in (2.9) also converges. This proves the theorem. ∎

4. Additional estimates. Lyapunov's equation

Condition (2.7) does not allow us to consider equation (1.2) with $B = -A^*$, since $r_s(A^*) = r_s(A)$. Because of this we are going to derive estimates under other conditions. To this end put $\alpha(A) = \max \; Re \; \lambda_k(A)$, $\beta(A) = \min \; Re \; \lambda_k(A)$ and

$$\gamma(a, A, B) := \sum_{j,k=0}^{n-1} \frac{(k+j)! g^k(A) g^j(B)}{a^{k+j+1}(k!j!)^{3/2}}$$

for a constant $a > 0$.

First, we need the following

Lemma 8. *Let the condition*

$$\beta(A) > \alpha(B) \tag{4.1}$$

hold. Then

$$\int_0^\infty \|e^{-At}\| \|e^{Bt}\| dt \le \gamma(\eta, A, B),$$

where $\eta := \beta(A) - \alpha(B)$.

Proof. Due to Example 1.10.3 from [8] we have

$$\|e^{As}\| \le e^{\alpha(A)s} \sum_{k=0}^{n-1} \frac{s^k g^k(A)}{(k!)^{3/2}} \quad (s \ge 0). \tag{4.2}$$

For the brevity put $b = \alpha(B)$ and $c = \eta/2$. Then $\beta(A) = b + 2c$ and $\alpha(B - I(b + c)) = -c$. Moreover, $\alpha(-(A - I(b + c)) = \alpha(-A) + b + c = -\beta(A) + b + c = -c$. In addition, $e^{-At}e^{Bt} = e^{-(A-(b+c)I)t}e^{(B+I(b+c))t}$. Take into account that, $g(A) = g(A - I(b + c))$ and $g(B) = g(B - I(b + c))$. From (4.2) we have,

$$\|e^{-(A-I(b+c))s}\| \le e^{-cs} \sum_{k=0}^{n-1} \frac{s^k g^k(A)}{(k!)^{3/2}} \text{ and } \|e^{(B-I(b+c))s}\| \le e^{-cs} \sum_{k=0}^{n-1} \frac{s^k g^k(B)}{(k!)^{3/2}} \ (s \ge 0).$$

So

$$\|e^{-As}\| \, \|e^{Bs}\| \le e^{-2sc} \sum_{j,k=0}^{n-1} \frac{s^{k+j} g^k(A) g^j(B)}{(k!j!)^{3/2}}. \tag{4.3}$$

But

$$\int_0^\infty s^{k+j} e^{-2cs} ds = \frac{(k+j)!}{(2c)^{k+j+1}}.$$

Thus (4.3) proves the lemma. ∎

Theorem 9. *Let condition* (4.1) *hold. Then equation* (1.2) *has a unique solution, which can be represented as*

$$X = \int_0^\infty e^{-At} C e^{Bt} dt. \tag{4.4}$$

Moreover,

$$\|X\| \le \|C\| \gamma(\eta, A, B). \tag{4.5}$$

Proof. Again put $c=\eta/2$. Equation (1.2) is equivalent to the following one:

$$(A-(\alpha(B)+c)I)X - X(B-(\alpha(B)+c)I) = C$$

As it was shown in the previous lemma $\alpha(B-I(b+c)) = -c$ and $\alpha(-(A-I(b+c)) = -c$ with $b=\beta(B)$. Due to Theorem 9.2 from [2] a solution of (1.2) is defined by the equality

$$X = \int_0^\infty e^{-(A-I(\alpha(B)+c))t} C e^{(B_F-(\alpha(B)+c)I)t} dt.$$

Hence (4.4) follows. Now the previous lemma proves the theorem. ■

If A is normal, then $g(A)=0$ and therefore,

$$\gamma(\eta, A, B) := \sum_{j=0}^{n-1} \frac{g^j(B)}{\eta^{j+1}(j!)^{1/2}}.$$

If both A and B are normal, then $\gamma(\eta, A, B) = \frac{1}{\eta}$. It is not hard to check that in (4.5) we have the equality, provided A, B and C are commuting normal matrices.

Consider the Lyapunov equation for differential equations

$$AX + XA^* = -C. \tag{4.6}$$

Taking in Theorem 4.1 $-A$ instead of A and A^* instead of B, since $\alpha(A^*)=\alpha(A)$ and $\beta(-A) = -\alpha(A)$, we get

Corollary 10. *Let $\alpha(A)<0$. Then equation (4.6) has a unique solution, which satisfies the inequality*

$$\|X\| \le \|C\| \sum_{j,k=0}^{n-1} \frac{(k+j)!g^{k+j}(A)}{(2|\alpha(A)|)^{k+j+1}(k!j!)^{3/2}}.$$

5. Estimates via Kronecker's sums of matrices

In this section we consider the Sylvester equation

$$A_1X + XA_2 = C, \tag{5.1}$$

whose coefficients have different dimensions. Namely A_l are $n_l \times n_l$ $(l=1,2)$ matrices, C is an $n_1 \times n_2$ matrix. In this case the results are more complicated than Theorems 4 and 9. Put

$$M = A_1 \otimes I_2 + I_1 \otimes A_2,$$

where $\otimes$ means the tensor product [16], I_l is the unit $n_l \times n_l$ matrix ($l = 1, 2$). The eigenvalues of M are $\lambda_{jk}(M) = \lambda_j(A_1) + \lambda_k(A_2)$ ($j = 1, \dots, n_1;\ k = 1, \dots, n_2$) and

$$g^2(M) = |M|_F^2 - \sum_{j,k=1}^{n} |\lambda_{jk}(M)|^2 = |M|_F^2 - \sum_{j=1}^{n_1} \sum_{k=1}^{n_2} |\lambda_k(A_2) + \lambda_j(A_1)|^2$$

is the departure from normality of M. Let $\sigma(A)$ denote the spectrum of a matrix A. Due to the classical Schur theorem, A_l admits the triangular representation

$$A_l = D_l + V_l \quad (\sigma(A_l) = \sigma(D_l), \quad l = 1, 2), \tag{5.2}$$

where D_l is the diagonal (normal) part, and V_l is the nilpotent (strictly upper triangular) part of A_l. So

$$M = I_1 \otimes (D_2 + V_2) + (D_1 + V_1) \otimes I_2 = D_M + V_M.$$

Here $D_M = I_1 \otimes D_2 + D_1 \otimes I_2$ is the diagonal part of M, and $V_M = I_1 \otimes V_2 + V_1 \otimes I_2$ is the nilpotent part of M. With $m = n_1 + n_2 - 1$ we obtain

$$V_M^m = \sum_{k=0}^{m} \frac{m!}{(m-k)!k!} (I_1 \otimes V_2)^k (V_1 \otimes I_2)^{m-k} = 0, \tag{5.3}$$

since $(I_1 \otimes V_2)^{n_2} = (V_1 \otimes I_2)^{n_1} = 0$, and either $k \ge n_2$ or $m - k \ge n_1$. We thus have proved

Lemma 11. *One has* $V_M^{n_1+n_2-1} = 0$.

Put $\rho(M, \lambda) := \min_{j,k} |\lambda_{jk}(M) - \lambda|$.

Lemma 12. *We have*

$$\|(M - \lambda I)^{-1}\| \le \sum_{k=0}^{n_1+n_2-2} \frac{g^k(M)}{\sqrt{k!}\rho^{k+1}(M, \lambda)} \quad (\lambda \notin \sigma(M)). \tag{5.4}$$

Proof. Let $R_\lambda(M) = (M - \lambda I)^{-1}$ ($\lambda \notin \sigma(M)$) be the resolvent of M. Clearly,

$$R_\lambda(M) = (D_M + V_M - \lambda I)^{-1} = (I + R_\lambda(D_M) V_M)^{-1} R_\lambda(D_M) \tag{5.5}$$

for all λ regular for M. Denote by $|A|_{Sb}$ the matrix whose entries in its Schur basis (the orthonormal basis of the triangular representation) are the absolute values of A. Then $|R_\lambda(D_M) V_M|_{Sb} \le |R_\lambda(D_M)|_{Sb} |V_M|_{Sb}$. The inequalities are understood in the entry-wise sense. Since a normal operator is diagonal in its

Schur basis, we can write $|R_\lambda(D_M)|_{Sb} \leq \rho^{-1}(D_M,\lambda)I$ and thus $|R_\lambda(D_M)V_M|_{Sb} \leq \rho^{-1}(D_M,\lambda)|V_M|_{Sb}$. This implies

$$|(R_\lambda(D_M)V_M)^k|_{Sb} \leq (\rho^{-1}(D_M,\lambda))^k|V_M|^k_{Sb}.$$

According to Lemma 11 $|V_M|^m_{Sb} = 0$ with $m = n_1+n_2-1$. Hence, $(R_\lambda(D_M)V_M)^m = 0$. Now (5.5) implies

$$R_\lambda(M) = \sum_{k=0}^{m-1}(R_\lambda(D_M)V_M)^k(-1)^kR_\lambda(D_M). \tag{5.6}$$

Due to Theorem 2.5.1 from [8], for any nilpotent $n\times n$ matrix V we have $\|V^k\| \leq \frac{|V|^k_F}{\sqrt{k!}}$ $(k = 1,\dots,n-1)$. Therefore

$$\|R_\lambda(M)\| \leq \sum_{k=0}^{m-1}\frac{|R_\lambda(D_M)V_M|^k_F}{\sqrt{k!}}\|R_z(D_M)\|.$$

But $|R_\lambda(D_M)V_M|_F \leq \|R_\lambda(D_M)\||V_M|_F$, $\|R_\lambda(D_M)\| = \rho^{-1}(D_M,\lambda)$ and thanks to Lemma 2.3.2 from [8], $|V_M|_F = g(M)$. Thus

$$\|R_\lambda(M)\| \leq \sum_{k=0}^{m-1}\frac{g^k(M)}{\rho^{k+1}(D_M,\lambda)\sqrt{k!}}.$$

Taking into account that $\sigma(M) = \sigma(D_M)$, we get the required result. ∎

Lemma 13. *One has* $|M|^2_F = |A_1|^2_F n_2 + |A_2|^2_F n_1 + 2Re\ (Trace\ A_1)\ (Trace\ A_2^*)$.

Proof. Obviously,

$$\begin{aligned} M^*M &= (A_1^*\otimes I_2 + I_1\otimes A_2^*)(A_1\otimes I_2 + I_1\otimes A_2) \\ &= A_1^*A_1\otimes I_2 + A_1^*\otimes A_2 + I_1\otimes A_2^*A_2 + A_2\otimes A_1^*. \end{aligned}$$

Take into account that $Trace\ (A_1\otimes A_2) = (Trace\ A_1)(Trace\ A_2)$. Then

$$\begin{aligned} |M|^2_F = Trace\ (M^*M) = n_2 Trace\ A_1^*A_1 + n_1 Trace\ A_2^*A_2 \\ + (Trace\ A_1^*)(Trace\ A_2) + (Trace\ A_1)\ (Trace\ A_2^*). \end{aligned}$$

This proves the lemma. ∎

Since V_1 and V_2 are nilpotent, the previous lemma yields $|V_M|^2_F = n_2|V_1|^2_F + n_1|V_2|^2_F$. But $|V_l|_F = g(A_l)$, $|V_M|_F = g(M)$.Thus we get

Corollary 14. *The equality* $g(M) = g(A_1, A_2)$, *is valid, where* $g(A_1, A_2) := (n_2 g^2(A_1) + n_1 g^2(A_2))^{1/2}$.

Moreover, the previous corollary and Lemma 12 imply

Corollary 15. *One has*

$$\|(M - \lambda I)^{-1}\| \le \sum_{k=0}^{n_1+n_2-2} \frac{g^k(A_1, A_2)}{\sqrt{k!}\rho^{k+1}(M, \lambda)} \quad (\lambda \notin \sigma(M)).$$

This corollary considerably improves inequality (6.2) from [12] (which is proved in the case $n_1 = n_2$). It is assumed that

$$\rho_0(A_1, A_2) := \min_{j,k=1,\dots,n} |\lambda_k(A_1) + \lambda_j(A_2)| > 0. \tag{5.7}$$

Put $M' = A_1 \otimes I_2 - I_1 \otimes A_2^T$, where A_2^T is the matrix transposed to A_2. As it is well known, [16, p. 255], equation (3.1) is equivalent to the following one: $M'\ vec\ X = vec\ C$, where $vec\ X$ for an $n_1 \times n_2$ matrix $X = (x_{jk})$ is defined as

$$vec\ X = column\ (x_{11}, x_{12}, \dots, x_{1n_2}, x_{21}, \dots, x_{2n_2}, \dots, x_{n_1 1}, \dots, x_{n_1 n_2}).$$

Hence, $vec\ X = (M')^{-1}\ vec\ C$. Obviously, $\|vec\ X\| = |X|_F$ and $\|vec\ C\| = |C|_F$. So we get $|X|_F \le \|(M')^{-1}\||C|_F$. But $\rho_0(A_1, A_2^T) = \rho_0(A_1, A_2)$ and $g(A_1, A_2^T) = g(A_1, A_2)$. Now Corollary 15 implies

Theorem 16. *Let condition* (5.7) *hold. Then a solution* X *of* (5.6) *satisfies the inequality*

$$|X|_F \le |C|_F \sum_{k=0}^{n_1+n_2-2} \frac{g^k(A_1, A_2)}{\sqrt{k!}\rho_0^{k+1}(A_1, A_2)}.$$

Let us apply Theorem 16 to perturbations of invariant subspaces. Recall that a subspace S of a Euclidean space is invariant for an operator A if $AS \subset S$.

For two $n \times n$ matrices A and B suppose that

$$\sigma(A) = \sigma_1(A) \cup \sigma_2(A) \text{ and } \sigma(B) = \sigma_1(B) \cup \sigma_2(B)$$

with $\sigma_1(A) \cap \sigma_2(A) = \emptyset$, $\sigma_1(B) \cap \sigma_2(B) = \emptyset$, and

$$\delta := \text{dist}(\sigma_2(A), \sigma_1(B)) > 0. \tag{5.8}$$

Denote by Q_A the orthogonal projection onto the invariant subspace of A, corresponding to $\sigma_1(A)$, and by Q_B the orthogonal projection onto the invariant subspace of B, corresponding to $\sigma_1(B)$. That is, $AQ_A = Q_A A Q_A$, $BQ_B = Q_B A Q_B$, $\sigma(AQ_A) = \sigma_1(A)$ and $\sigma(BQ_B) = \sigma_1(B)$, and put $\hat{Q}_A = I - Q_A$.

The gap or distance between these subspaces is defined as $\|Q_A - Q_B\|$, cf. [1, p. 202]. In the spectral norm, one has [1, Exercise VII.1.11] $\|Q_A - Q_B\| = \|\hat{Q}_A Q_B\|$. Therefore, to bound $\|Q_A - Q_B\|$ it is sufficient to bound $\|\hat{Q}_A Q_B\|$.

The literature devoted to the quantity $\|\hat{Q}_A Q_B\|$ is rather rich. Mainly, the normal and Hermitian matrices are considered, cf. [1, 5, 14] and references therein.

Put $n_1 = \text{rank } Q_B$ and $n_2 = \text{rank } \hat{Q}_A$.

Theorem 17. *Let condition* (5.8) *hold. Then*

$$|\hat{Q}_A Q_B|_F \le |A - B|_F \sum_{k=0}^{n_1+n_2-1} \frac{(n_1 g^2(A) + n_2 g^2(B))^{k/2}}{\sqrt{k!}\,\delta^{k+1}}. \tag{5.9}$$

If A is normal, then $g(A) = 0$ and

$$|\hat{Q}_A Q_B|_F \le |A - B|_F \sum_{k=0}^{n_1+n_2-1} \frac{n_2^k g^k(B)}{\sqrt{k!}\,\delta^{k+1}}. \tag{5.10}$$

If both A and B are normal, then (5.9) yields $|\hat{Q}_A Q_B|_F \le \frac{|A-B|_F}{\delta}$. This is the well-known result of Davis and Kahan [5]. To prove Theorem 17 we need

Lemma 18. *Let Q be an (orthogonal) projection onto an invariant subspace of of A. Then $g(AQ) \le g(A)$ and $g((I - Q)A) \le g(A)$.*

Proof. Recall that $A = D_A + V_A$, where D_A is the diagonal part and V_A is the nilpotent part of A. We have $AQ = D_A Q + V_A Q$. It is not hard to check $D_A Q$ is normal and $V_A Q$ is nilpotent, since A, D_A and V_A have the same invariant subspaces. Consequently, $g(AQ) = |V_A Q|_F \le |V_A|_F = g(A)$. Similarly the second inequality can be proved. ■

***Proof of Theorem* 17.** Denote $A_2 = \hat{Q}_A A, B_1 = BQ_B, C_{21} = \hat{Q}_A(A - B)Q_B$ and $X_{21} = \hat{Q}_A Q_B$. Then

$$C_{21} = \hat{Q}_A A \hat{Q}_A Q_B - \hat{Q}_A Q_B B Q_B = \hat{Q}_A A X_{21} - X_{21} B Q_B,$$

and therefore,

$$A_2 X_{21} - X_{21} B_1 = C_{21}. \tag{5.11}$$

Clearly,

$$|C_{21}|_F = |\hat{Q}_A(A-B)Q_B|_F \le |A-B|_F \|\hat{Q}_A\| \|Q_B\| = |A-B|_F. \tag{5.12}$$

Apply to (5.11) Theorem 16 with $A_1 = -B_1$. Taking into account that

$$g^2(B_1, A_2) = n_1 g^2(A_2) + n_2 g^2(B_1) \text{ and } \rho_0(A_2, -B_1) = \delta,$$

we have

$$|X_{21}|_F \le |C_{21}|_F \sum_{k=0}^{n_1+n_2-1} \frac{(n_1 g^2(A_2) + n_2 g^2(B_1))^{k/2}}{\sqrt{k!}\delta^{k+1}}.$$

Due to the previous lemma, $g(A_2) \le g(A), g(B_1) \le g(B)$. This and equalities $X_{21} = \hat{Q}_A Q_B$ and (5.12) prove the theorem. ∎

References

[1] R. Bhatia, Matrix Analysis (Springer, NY, 1997). doi:10.1007/978-1-4612-0653-8

[2] R. Bhatia and P. Rosenthal, *How and why to solve the operator equation $AX - XB = Y$*, Bull. London Math. Soc. **29** (1997) 1–21. doi:10.1112/S0024609396001828

[3] R.R. Bitmead and H. Weiss, *On the solution of the discrete time Lyapunov matrix equation in controllable canonical form*, IEEE Trans. Automat. Control AC-24 (1979) 481–482. doi:10.1109/TAC.1979.1102064

[4] Yu. L. Daleckii and M.G. Krein, *Stability of Solutions of Differential Equations in Banach Space*, Amer. Math. Soc., Providence, R.I, 1971.

[5] C. Davis and W. Kahan, *The rotation of eigenvectors by a perturbation* III, SIAM J. Numer. Anal. **7** (1970) 146–162. doi:10.1137/0707001

[6] M. Dehghan and M. Hajarian, *The reflexive and anti-reflexive solutions of a linear matrix equation and systems of matrix equations*, Rocky Mountain J. Math. **40** (3) (2010) 825–848. doi:10.1216/RMJ-2010-40-3-825

[7] B.W. Dickinson, *Analysis of the Lyapunov equation using generalized positive real matrices*, IEEE Trans. Autumut. Control AC-25 (1980) 560–563. doi:10.1109/TAC.1980.1102391

[8] M.I. Gil', Operator Functions and Localization of Spectra, Lecture Notes In Mathematics vol. 1830 (Springer-Verlag, Berlin, 2003).

[9] M.I. Gil', Difference Equations in Normed Spaces. Stability and Oscillations, North-Holland, Mathematics Studies **206** (Elsevier, Amsterdam, 2007).

[10] M.I. Gil', *Norm estimates for functions of two non-commuting matrices*, Electronic Journal of Linear Algebra **22** (2011) 504–512.

[11] M.I. Gil', *Matrix equations with diagonalizable coefficients*, Gulf J. Math. **1** (2013) 98–104.

[12] M.I. Gil, *Bounds for spectra of operators on tensor of Euclidean spaces*, PanAmerican Math. J. **24** (3) (2014) 35–47.

[13] I. Gohberg, R. Lancaster and L. Rodman, Invariant Subspaces of Matrices with Applications (Wiley, New York, 1986).

[14] L. Grubisic, N. Truhar and K. Veseli, *The rotation of eigenspaces of perturbed matrix pairs*, Linear Algebra and Appl. **436** (2012) 4161–4178. doi:10.1016/j.laa.2012.01.026

[15] J.Z. Hearon, *Nonsingular solutions of* $TA - BT = C$, Linear Algebra and Appl. **16** (1977) 5783. doi:10.1016/0024-3795(77)90019-2

[16] R. Horn and C. Johnson, Topics in Matrix Analysis (Cambridge University Press, Cambridge, 1991). doi:10.1017/CBO9780511840371

[17] M. Konstantinov, Da-Wei Gu, V. Mehrmann and P. Petkov, Perturbation Theory for Matrix Equations, Studies in Computational Mathematics, 9 (North Holland, Amsterdam, 2003).

[18] A.G. Mazko, Matrix Equations, Spectral Problems and Stability of Dynamic Systems. Stability, Oscillations and Optimization of Systems (Scientific Publishers, Cambridge, 2008).

[19] V. Ptak and N.J. Young, *A generalization of the zero location theorem of Schur and Cohn*, IEEE Trans. Automat. Control AC-25 (1980) 978–980. doi:10.1109/TAC.1980.1102476

Received 23 April 2014