

## ON THE SENSITIVITY OF THE SOLUTION OF THE GENERALIZED LYAPUNOV EQUATION

HOSOO LEE

ABSTRACT. Some results on the sensitivity of the solution of the generalized Lyapunov equation

$$A^{n-1}X + A^{n-2}XA^* + \cdots + X(A^*)^{n-1} = B,$$

are shown follow easily from well-known theorems in functional analysis.

### 1. Introduction

In [3], Bhatia and Uchiyama have shown that the equation

$$A^{n-1}X + A^{n-2}XB + \cdots + AXB^{n-2} + XB^{n-1} = C,$$

has a unique solution when the eigenvalues of  $A$  and  $B$  are in the sector

$$\{w \in \mathbb{C} : w \neq 0, -\pi/n < \arg w < \pi/n\},$$

and the unique solution  $X$  can be expressed as

$$(1.1) \quad X = \frac{\sin \frac{\pi}{n}}{\pi} \int_0^\infty (t + A^n)^{-1} C (t + B^n)^{-1} t^{1/n} dt.$$

The special case when  $n = 2$ ,

$$AX + XB = C$$

---

Received September 11, 2013. Revised September 15, 2013. Accepted September 15, 2013.

2010 Mathematics Subject Classification: 15B48.

Key words and phrases: generalized Lyapunov equation.

© The Kangwon-Kyungki Mathematical Society, 2013.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

is the much studied Sylvester equation, of great interest in operator theory, numerical analysis, and engineering. There is a long tradition of finding different expressions for the solution of Sylvester equation in the form of operator integrals, some prominent examples of which occur in the works of E. Heinz, M. G. Krein, M. Rosenblum, and R. Bhatia, C. Davis and A. McIntosh.

The object of this note is the equation

$$(1.2) \quad A^{n-1}X + A^{n-2}XA^* + \cdots + X(A^*)^{n-1} = B,$$

where  $A, B$  are  $m \times m$  complex matrices, and the sensitivity of the solution of (1.2).

## 2. Main results

Let  $\mathbb{M}$  be the set of  $m \times m$  complex matrices. Denote  $\mathbb{S} = \{X \in \mathbb{M} : \sigma(X) \subset \{w \in \mathbb{C} : w \neq 0, -\pi/n < \arg w < \pi/n\}\}$ . The condition of  $A \in \mathbb{S}$  is sufficient to ensure that the equation (1.2) has a unique solution  $X$  for every  $B$ . From the identity (1.1) of Bhatia and Uchiyama, the solution of (1.2) can be expressed as

$$(2.3) \quad X = \frac{\sin \frac{\pi}{n}}{\pi} \int_0^\infty (t + A^n)^{-1} B (t + (A^*)^n)^{-1} t^{1/n} dt.$$

From this it is immediately clear that if  $B$  is positive (semidefinite) then so is  $X$ .

For  $A \in \mathbb{M}$ , let  $\|A\|$  be the norm of  $A$  as a linear operator on the Euclidean space  $\mathbb{C}^m$ . If  $\mathcal{A}$  is a linear operator on  $\mathbb{M}$ , let

$$\|\mathcal{A}\| := \sup\{\|\mathcal{A}(X)\| : X \in \mathbb{M}, \|X\| = 1\}.$$

Now, given  $A \in \mathbb{M}$ , define  $\mathcal{L}_{n,A}(X) = A^{n-1}X + A^{n-2}XA^* + \cdots + X(A^*)^{n-1}$ . This is a linear operator on  $\mathbb{M}$ , and it is invertible if the eigenvalues of  $A$  are in the sector  $\{w \in \mathbb{C} : w \neq 0, -\pi/n < \arg w < \pi/n\}$ . The equation (2.3) can be written also as

$$\mathcal{L}_{n,A}^{-1}(B) = \frac{\sin \frac{\pi}{n}}{\pi} \int_0^\infty (t + A^n)^{-1} B (t + (A^*)^n)^{-1} t^{1/n} dt.$$

The norm of the operator  $\mathcal{L}_{n,A}^{-1}$  is a measure of the sensitivity of the solution. The following theorem gives a way of evaluating it.

THEOREM 1. For every matrix  $A$  in  $\mathbb{S}$ ,

$$(2.4) \quad \|\mathcal{L}_{n,A}^{-1}\| = \|\mathcal{L}_{n,A}^{-1}(I)\|.$$

*Proof.* Let  $H = \mathcal{L}_{n,A}^{-1}(I)$ . As noted above,  $H$  is a positive matrix. Define a map  $\Gamma_{n,A}$  on  $\mathbb{M}$  as

$$(2.5) \quad \Gamma_{n,A}(Y) := H^{-1/2}\mathcal{L}_{n,A}^{-1}(Y)H^{-1/2}.$$

Then  $\Gamma_{n,A}$  is a *positive map* (it takes positive matrices to positive matrices) and is *unital* (it takes  $I$  to  $I$ ). By the theorem of Russo and Dye [7] such a map on any  $C^*$ -algebra has norm 1. So  $\|\Gamma_{n,A}\| = 1$  and hence  $\|\mathcal{L}_{n,A}^{-1}\| \leq \|H\|$ . But, by the definition of  $H$ , we must then have  $\|\mathcal{L}_{n,A}^{-1}\| = \|H\|$ . This proves (2.4).  $\square$

Let  $\langle X, Y \rangle = \text{tr}(X^*Y)$  be the Frobenius inner product on  $\mathbb{M}$ . If  $\mathcal{A}$  is a linear map on  $\mathbb{M}$ , its adjoint  $\mathcal{A}^*$ , defined with respect to this inner product, is the map satisfying

$$\langle \mathcal{A}(X), Y \rangle = \langle X, \mathcal{A}^*(Y) \rangle \quad \text{for all } X, Y.$$

LEMMA 2. Let  $A \in \mathbb{S}$ . Then

$$(2.6) \quad \mathcal{L}_{n,A}^* = \mathcal{L}_{n,A^*}$$

$$(2.7) \quad \Gamma_{n,A}^*(X) = \mathcal{L}_{n,A^*}^{-1}(H^{-1/2}XH^{-1/2}).$$

*Proof.* By a simple calculation with traces, we have

$$\begin{aligned} \langle \mathcal{L}_{n,A}^*(X), Y \rangle &= \langle X, \mathcal{L}_{n,A}(Y) \rangle \\ &= \langle X, A^{n-1}Y + A^{n-2}YA^* + \dots + Y(A^*)^{n-1} \rangle \\ &= \langle X, A^{n-1}Y \rangle + \langle X, A^{n-2}YA^* \rangle + \dots + \langle X, Y(A^*)^{n-1} \rangle \\ &= \langle (A^{n-1})^*X, Y \rangle + \langle (A^{n-2})^*XA, Y \rangle + \dots + \langle XA^{n-1}, Y \rangle \\ &= \langle (A^{n-1})^*X + (A^{n-2})^*XA + \dots + XA^{n-1}, Y \rangle \\ &= \langle \mathcal{L}_{n,A^*}(X), Y \rangle \end{aligned}$$

and

$$\begin{aligned} \langle \Gamma_{n,A}^*(X), Y \rangle &= \langle X, \Gamma_{n,A}(Y) \rangle \\ &= \langle X, H^{-1/2}\mathcal{L}_{n,A}^{-1}(Y)H^{-1/2} \rangle \\ &= \langle H^{-1/2}XH^{-1/2}, \mathcal{L}_{n,A}^{-1}(Y) \rangle \\ &= \langle (\mathcal{L}_{n,A}^{-1})^*(H^{-1/2}XH^{-1/2}), Y \rangle \\ &\stackrel{(2.6)}{=} \langle (\mathcal{L}_{n,A^*}^{-1})(H^{-1/2}XH^{-1/2}), Y \rangle. \end{aligned}$$

□

A linear map  $\mathcal{A}$  on  $\mathbb{M}$  is called *doubly stochastic* if it is (i) positive, (ii) unital, and (iii) trace preserving, i.e.,  $\text{tr } \mathcal{A}(X) = \text{tr } X$  for all  $X$ . The third condition is equivalent to the condition that  $\mathcal{A}^*$  is unital (see [1]). It is natural to ask whether the map  $\Gamma_{n,A}$  defined in (2.5) is doubly stochastic.

LEMMA 3. *If  $A$  is a normal matrix in  $\mathbb{S}$ , then the operator  $\Gamma_{n,A}$  is doubly stochastic.*

*Proof.* By the definition of  $H$  and equation (2.7), we have

$$\begin{aligned} \Gamma_{n,A}^*(I) &= \mathcal{L}_{n,A^*}^{-1}(H^{-1}) \\ &= \frac{\sin \frac{\pi}{n}}{\pi} \int_0^\infty (t + (A^*)^n)^{-1} H^{-1} (t + A^n)^{-1} t^{1/n} dt \\ &= \int_0^\infty (t + (A^*)^n)^{-1} \left( \int_0^\infty (s + A^n)^{-1} (s + (A^*)^n)^{-1} s^{1/n} ds \right)^{-1} \\ &\quad (t + A^n)^{-1} t^{1/n} dt. \end{aligned}$$

This shows that if  $A$  is normal then  $\Gamma_{n,A}^*$  is unital and hence  $\Gamma_{n,A}$  is doubly stochastic. □

Doubly stochastic maps have several special properties. The one relevant to our discussion is that they are contractive with respect to every unitarily invariant norm on  $\mathbb{M}$  (see [1]).

THEOREM 4. *Let  $A$  be a normal matrix in  $\mathbb{S}$ . Then*

$$(2.8) \quad \|\|\mathcal{L}_{n,A}^{-1}(X)\|\| \leq \|\|\mathcal{L}_{n,A}^{-1}(I)\|\| \|\|X\|\|$$

for every unitarily invariant norm  $\|\|\cdot\|\|$ .

*Proof.* Let  $A$  be a normal matrix. Then by Lemma 3,  $\Gamma_{n,A}$  is doubly stochastic and hence

$$(2.9) \quad \|\|\Gamma_{n,A}(X)\|\| \leq \|\|X\|\|$$

for every unitarily invariant norm  $\|\|\cdot\|\|$ . All unitarily invariant norms satisfy the inequality  $\|\|XYZ\|\| \leq \|\|X\|\| \|\|Y\|\| \|\|Z\|\|$  for any three matrices  $X, Y, Z$ . So from (2.9), we get

$$(2.10) \quad \|\|\mathcal{L}_{n,A}^{-1}(X)\|\| = \|\|H^{1/2}\Gamma_{n,A}(X)H^{1/2}\|\| \leq \|\|H\|\| \|\|X\|\| = \|\|\mathcal{L}_{n,A}^{-1}(I)\|\| \|\|X\|\|.$$

□

It is a basic fact from [5] that if  $T$  is an operator between normed spaces and  $T^*$  is its adjoint, then  $\|T\| = \|T^*\|$ . Let

$$H_* = (\mathcal{L}_{n,A}^{-1})^*(I).$$

Then from (2.6) we have

$$H_* = \mathcal{L}_{n,A^*}^{-1}(I) = \frac{\sin \frac{\pi}{n}}{\pi} \int_0^\infty (t + (A^*)^n)^{-1}(t + A^n)^{-1}t^{1/n}dt.$$

**THEOREM 5.** *Let  $A \in \mathbb{S}$ , and let  $B \geq O$  (that is,  $B$  is a positive semidefinite matrix). Let  $X$  be the solution of (1.2). Then we have*

$$(2.11) \quad \lambda_{\min}(H_*) \operatorname{tr}(B) \leq \operatorname{tr}(X) \leq \lambda_{\max}(H_*) \operatorname{tr}(B).$$

*Proof.* Let  $X$  be the solution of (1.2) when  $A \in \mathbb{S}$  and  $B \geq O$ . Then

$$(2.12) \quad \begin{aligned} \operatorname{tr}(X) &= \langle X, I \rangle = \langle \mathcal{L}_{n,A}^{-1}(B), I \rangle = \langle B, (\mathcal{L}_{n,A}^{-1})^*(I) \rangle \\ &= \langle B, H_* \rangle = \operatorname{tr}(BH_*) \end{aligned}$$

Choose an orthonormal basis in which  $H_*$  is diagonal, and calculate the traces. Since  $H_*$  is positive, this gives (2.11).  $\square$

Recall that the norm  $\|A\|_p$  for  $1 \leq p \leq \infty$  is defined as

$$\|A\|_p = \left( \sum_{j=1}^m (s_j(A))^p \right)^{1/p},$$

where  $s_j(A), 1 \leq j \leq m$ , are the singular values of  $A$  arranged in decreasing order. Note that the operator norm  $\|A\| := \|A\|_\infty = s_1(A)$ . If  $\mathcal{A}$  is a linear operator on  $\mathbb{M}$ , let

$$\|\mathcal{A}\|_{p \rightarrow p} = \sup\{\|\mathcal{A}\|_p : \|X\|_p = 1\}.$$

This is the norm of the operator  $\mathcal{A}$  when the underlying space  $\mathbb{M}$  is equipped with the norm  $\|\cdot\|_p$ .

A well-known theorem, called the Calderon-Lions interpolation theorem [5], implies that for  $1 \leq p \leq \infty$  and a linear operator  $\mathcal{A}$  we have

$$(2.13) \quad \|\mathcal{A}\|_{p \rightarrow p} \leq \|\mathcal{A}\|_{1 \rightarrow 1}^{1/p} \|\mathcal{A}\|_{\infty \rightarrow \infty}^{1-1/p}.$$

Using this we can prove the following.

**THEOREM 6.** *Let  $A \in \mathbb{S}$  and*

$$H = \frac{\sin \frac{\pi}{n}}{\pi} \int_0^\infty (t + A^n)^{-1}(t + (A^*)^n)^{-1}t^{1/n}dt,$$

$$H_* = \frac{\sin \frac{\pi}{n}}{\pi} \int_0^\infty (t + (A^*)^n)^{-1} (t + A^n)^{-1} t^{1/n} dt.$$

Then, for  $1 \leq p \leq \infty$ ,

$$(2.14) \quad \|\mathcal{L}_{n,A}^{-1}\|_{p \rightarrow p} \leq \|H_*\|^{1/p} \|H\|^{1-1/p}.$$

*Proof.* In this notation, Theorem 1 can be restated as

$$\|\mathcal{L}_{n,A}^{-1}\|_{\infty \rightarrow \infty} = \|H\|.$$

Since the norm  $\|\cdot\|_1$  is dual to  $\|\cdot\|_\infty$ , we have

$$\|\mathcal{L}_{n,A}^{-1}\|_{1 \rightarrow 1} = \|(\mathcal{L}_{n,A}^{-1})^*\|_{\infty \rightarrow \infty} = \|\mathcal{L}_{n,A^*}^{-1}\|_{\infty \rightarrow \infty} = \|\mathcal{L}_{n,A^*}^{-1}(I)\| = \|H_*\|.$$

This shows that for  $p = 1, \infty$ , we have equality in (2.14). For other values of  $p$ , we get this inequality using the interpolation result (2.13) cited above.  $\square$

## References

- [1] T. Ando, Majorization, *Doubly stochastic matrices and comparison of eigenvalues*, Linear Algebra Appl. **118** (1989), 163–248.
- [2] R. Bhatia, *A note on the Lyapunov equation*, Linear Algebra Appl. **259** (1997), 71–76.
- [3] R. Bhatia and M. Uchiyama, *The operator equation  $\sum_{i=0}^n A^{n-i} X B^i = Y$* , Expo. Math. **27** no.3 (2009), 251–255.
- [4] G. Hewer and C. Kenney, *The sensitivity of the stable Lyapunov equation*, SIAM J. Control Optim. **26** (1988), 321–344.
- [5] M. Reed and B. Simon, *Methods of Morder Mathematical Physics II*, Academic, New York, 1975
- [6] W. Rudin, *Functional Analysis*, McGraw-Hill, New York, 1973
- [7] B. Rosso and H. A. Dye, *A note on unitary operator in  $C^*$ -algebras*, Duke Math. J. **33** (1966), 413–416.

Department of Mathematics  
Louisiana State University  
Baton Rouge  
LA 70803 U.S.A.  
*E-mail:* hosoo9181@gmail.com