Two Results About The Matrix Exponential

Hongguo Xu *

Abstract

Two results about the matrix exponential are given. One is to characterize the matrices A which satisfy $e^A e^{A^H} = e^{A^H} e^A$, another is about the upper bounds of trace $e^A e^{A^H}$. When A is stable, the bounds preserve the asymptotic stability.

1 Introduction

Let A be an n-by-n matrix, the exponential of A is defined as follows.

$$e^A = I_n + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{A^k}{k!}. \tag{1}$$

The matrix exponential plays an important role in linear control systems and ordinary differential equations, see [1, 2, 8, 9, 11, 13] and the references therein. However, the theoretical analysis as well as the numerical computation of e^A are still under investigations, see [1, 2, 5, 10]. In [1, 2], Bernstein proposed many open problems arising from linear control systems, which include some matrix exponential problems. Here we will consider two of them.

Problem

1. Is there any nonnormal matrix such that $e^A e^{A^H} = e^{A^H} e^A$ or $e^A e^{A^H} = e^{A+A^H}$?

2. Can we derive a bound of trace $e^A e^{A^H}$ in stead of trace e^{A+A^H} such that when A is stable and under asymptotic case the bound can show the stable behavior?

*Department of Mathematics, Fudan University, Shanghai, 200433, P. R. China. Current address: Fakultät für Mathematik, TU Chemnitz-Zwickau, D-09107 Chemnitz, FRG.
The author is supported by Chinese National Natural Science Foundation.

Certainly the above problems are conventional when A is normal i.e., $AA^H = A^H A$, so we only consider the nonnormal case of A. Further it is implicitly proved in [12] that $e^A e^{A^H} = e^{A+A^H}$ if and only if A is normal. We just need to consider the first part of the first problem.

We use $\lambda(A)$ to denote the eigenvalue set of square matrix A, $\|\cdot\|$ to denote $2-norm$ and $\kappa(X) = \|X\| \, \|X^{-1}\|$ for nonsingular X. we let $\sigma_{\min}(Y)$ represent the minimum singular value of the matrix Y, the superscript H and $-H$ represent the conjugate transpose and that with inverse.

We will answer the first question in Section 2 and the second one in Section 3. We will address our conclusion remarks in Section 4.

2 A Positive Answer For The First Problem

As commented in Section 1, the first problem has just the first part left, i.e., whether there is any nonnormal matrix A such that $e^A e^{A^H} = e^{A^H} e^A$. We will give a positive result.

Theorem 1 *A matrix $A \in \mathcal{C}^{n\times n}$ satisfies $e^A e^{A^H} = e^{A^H} e^A$ if and only if there is a unitary matrix Q such that*

$$A = Q \operatorname{diag}(A_1, \ldots, A_s) Q^H, \tag{2}$$

where A_m, $m = 1, \ldots, s$ are diagonalizable and

$$\forall \lambda_j^m, \lambda_l^m \in \lambda(A_m): \; \lambda_j^m - \lambda_l^m = 2ik_{j,l}^m \pi, \textit{ for integer } k_{j,l}^m, \tag{3}$$

$$\forall \lambda_m \in \lambda(A_m), \, \lambda_j \in \lambda(A_j), \, m \neq j: \; e^{\lambda_m} \neq e^{\lambda_j}. \tag{4}$$

Proof. For sufficiency it is easy to check $e^A e^{A^H} = e^{A^H} e^A$ if A can be expressed as in (2)-(4).

For necessity, from $e^A e^{A^H} = e^{A^H} e^A$ we get that e^A is normal. So there is a unitary matrix Q such that $e^A = QTQ^H$, with

$$T = \operatorname{diag}(\mu_1 I_{n_1}, \ldots, \mu_s I_{n_s}),$$

$\mu_1, \mu_2, \ldots, \mu_s$ pairwise different, $\sum_{k=1}^s n_k = n$.

Since $Q^H AQ$ and T commute, we can verify that

$$Q^H AQ := \operatorname{diag}(A_1, \ldots, A_s),$$

where A_k is $n_k \times n_k$, and $e^{A_k} = T_k$, $k = 1, \ldots, s$. With the property of the exponential function, the eigenvalues of A_k must satisfy (3). Moreover A_k

must be diagonalizable, since any Jordan block in A_k will cause the same order Jordan block in e^{A_k}, e. g., [3, Corollary 3.8,Theorem 3.3] or [7, Section 6.4]. The property $\mu_k \neq \mu_l$ with $k \neq l$ implies the condition (4). $\square$

Unlike the case of $e^A e^{A^H} = e^{A+A^H}$, this theorem shows that there do exist nonnormal matrices which satisfy $e^A e^{A^H} = e^{A^H} e^A$, furthermore the structure of such matrices is quite simple. Let us look at two examples.

Example 1 [12] *Let*

$$A = \begin{bmatrix} i\pi & 1 \\ 0 & -i\pi \end{bmatrix}.$$

Clearly the eigenvalues of A satisfy the conditions of Theorem 1 and A is diagonalizable, so we have

$$e^A e^{A^H} = e^{A^H} e^A = I_2.$$

However $e^{A+A^H} = \begin{bmatrix} 0 & e \\ e & 0 \end{bmatrix}$.

Example 2 *Let*

$$A = \begin{bmatrix} 1 & \frac{\pi}{2} & 0 & \pi \\ -\frac{\pi}{2} & 1 & -\pi & 0 \\ 0 & 0 & 1 & \frac{5\pi}{2} \\ 0 & 0 & -\frac{5\pi}{2} & 1 \end{bmatrix}.$$

This is a real matrix. With simple calculations,

$$A = Q\,\mathrm{diag}(\begin{bmatrix} 1+\frac{\pi}{2}i & \pi i \\ 0 & 1+\frac{5\pi}{2}i \end{bmatrix}, \begin{bmatrix} 1-\frac{\pi}{2}i & -\pi i \\ 0 & 1-\frac{5\pi}{2}i \end{bmatrix})Q^H$$

$$:= Q\,\mathrm{diag}(T_1, T_2)Q^H := QTQ^H,$$

where

$$Q = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 & 0 & i & 0 \\ i & 0 & 1 & 0 \\ 0 & 1 & 0 & i \\ 0 & i & 0 & 1 \end{bmatrix}$$

is unitary and T_1, T_2 are diagonalizable,

$$T_k = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix} \hat{T}_k \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix}, \quad k = 1, 2.$$

with

$$\hat{T}_1 = \mathrm{diag}(1 + \frac{\pi}{2}i, 1 + \frac{5\pi}{2}i),$$

$$\hat{T}_2 = \mathrm{diag}(1 - \frac{\pi}{2}i, 1 - \frac{5\pi}{2}i).$$

So A satisfies the condition of Theorem 1, and actually we have

$$e^A e^{A^H} = e^{A^H} e^A = e^2 I_4,$$

but

$$e^{A+A^H} = \begin{bmatrix} \alpha & 0 & 0 & \beta \\ 0 & \alpha & -\beta & 0 \\ 0 & -\beta & \alpha & 0 \\ \beta & 0 & 0 & \alpha \end{bmatrix},$$

where $\alpha = \frac{1}{2}(e^{2+\pi} + e^{2-\pi}),\ \beta = \frac{1}{2}(e^{2+\pi} - e^{2-\pi})$.

The first question is completely resolved by Theorem 1 together with the results in [12].

3 Generalized Bounds For $\mathrm{trace}\, e^A e^{A^H}$

The inequality

$$\mathrm{trace}\, e^A e^{A^H} \leq \mathrm{trace}\, e^{A+A^H} \tag{5}$$

is well known, see [2] and the references therein. In control theory one usually estimates $\mathrm{trace}\, e^{At} e^{A^H t}$ when A is stable, i.e., $\forall \lambda \in \lambda(A) :\ \mathrm{Re}\, \lambda < 0$. Hence $\lim_{t\to+\infty} \mathrm{trace}\, e^{At} e^{A^H t} = 0$. However if (5) is applied, $\mathrm{trace}\, e^{At} e^{A^H t} \leq \mathrm{trace}\, e^{(A+A^H)t}$, but $A + A^H$ may not be stable when A is nonnormal, which means the asymptotic stability can be destroyed. Consequently in such a case the upper bound in (5) is not good.

We try to use (5) to set up more generalized bounds and then applying Lyapunov theory to give better estimators for $\mathrm{trace}\, e^A e^{A^H}$ which are suitable for asymptotic stability analysis. We begin the discussions with a lemma.

Lemma 2 *Let $A \in \mathcal{C}^{n\times n}$, $X = BB^H$, with $B \in \mathcal{C}^{n\times n}$ an arbitrary nonsingular matrix. Then*

$$\mathrm{trace}\, e^A e^{A^H} \leq \kappa(X)\, \mathrm{trace}\, e^{B^{-1}(AX + XA^H)B^{-H}}. \tag{6}$$

Proof. Since $X = BB^H$,

$$B^{-1}(AX + XA^H)B^{-H} = B^{-1}AB + (B^{-1}AB)^H.$$

Using (5) we have

$$\begin{aligned}
& \operatorname{trace} e^{B^{-1}(AX+XA^H)B^{-H}} \\
= \; & \operatorname{trace} e^{B^{-1}AB+(B^{-1}AB)^H} \geq \operatorname{trace} e^{B^{-1}AB} e^{(B^{-1}AB)^H} \\
= \; & \operatorname{trace} B^{-1}e^A BB^H e^{A^H} B^{-H} = \operatorname{trace}(BB^H)^{-1} e^A BB^H e^{A^H} \\
\geq \; & \sigma_{\min}((BB^H)^{-1}) \operatorname{trace} e^A BB^H e^{A^H} = \frac{1}{\|X\|} \operatorname{trace} BB^H e^{A^H} e^A \\
\geq \; & \frac{\sigma_{\min}(BB^H)}{\|X\|} \operatorname{trace} e^{A^H} e^A = \frac{1}{\|X\| \, \|X^{-1}\|} \operatorname{trace} e^A e^{A^H}.
\end{aligned}$$

This proves (6). □

If $B = I_n$ in (6), then the inequality is just that in (5). So the result in Lemma 2 is more general and has the freedom of choosing the matrix B.

Now we consider the case when A is stable. At first we recall Lyapunov's Theorem.

Theorem 3 (Lyapunov) *Let $A \in \mathcal{C}^{n\times n}$ be stable. Then for an arbitrary negative definite Hermitian matrix C, there exists a unique positive definite Hermitian solution X of the Lyapunov equation*

$$AX + XA^H = C. \tag{7}$$

Proof. See [7, pages 96-99]. □

When A is stable, combining Theorem 3 and Lemma 2 we get the following results.

Theorem 4 *Let $A \in \mathcal{C}^{n\times n}$ be stable, and let $C \in \mathcal{C}^{n\times n}$ be a negative definite Hermitian matrix. Let X be the positive definite solution of the Lyapunov equation (7), and express X as $X = BB^H$. Then*

$$\operatorname{trace} e^A e^{A^H} \leq \kappa(X) \operatorname{trace} e^{B^{-1}(AX+XA^H)B^{-H}} \leq \kappa(X) \operatorname{trace} e^{\frac{1}{\|X\|}C}. \tag{8}$$

Proof. Using the inequality in Lemma 2 we get

$$\operatorname{trace} e^A e^{A^H} \leq \kappa(X) \operatorname{trace} e^{B^{-1}(AX+XA^H)B^{-H}} = \kappa(X) \operatorname{trace} e^{B^{-1}CB^{-H}}. \tag{9}$$

Since C is negative definite, so is $B^{-1}CB^{-H}$. We denote the eigenvalues of C and $B^{-1}CB^{-H}$, respectively, by

$$0 > \lambda_1 \geq \cdots \geq \lambda_n; \;\; 0 > \mu_1 \geq \cdots \geq \mu_n,$$

and denote by $\mathcal{V}_i$ an arbitrary $i-$dimensional subspace. Using Minimax Theorem (see [4]) we have

$$\begin{aligned}
\mu_i &= \max_{\mathcal{V}_i} \min_{0\neq x\in\mathcal{V}_i} \frac{x^H B^{-1}CB^{-H}x}{x^H x} = \max_{\mathcal{V}_i} \min_{0\neq y\in\mathcal{V}_i} \frac{y^H C y}{y^H y}\frac{y^H y}{y^H BB^H y} \\
&= \max_{\mathcal{V}_i} \min_{0\neq y\in\mathcal{V}_i} \frac{y^H C y}{y^H y}\frac{y^H y}{y^H X y} \leq \frac{1}{\|X\|} \max_{\mathcal{V}_i} \min_{0\neq y\in\mathcal{V}_i} \frac{y^H C y}{y^H y} \\
&= \frac{1}{\|X\|}\lambda_i, \;\; i = 1, \ldots, n.
\end{aligned}$$

Hence

$$\begin{aligned}
\operatorname{trace} e^{B^{-1}CB^{-H}} &= \sum_{i=1}^{n} e^{\mu_i} \\
&\leq \sum_{i=1}^{n} e^{\frac{\lambda_i}{\|X\|}} = \operatorname{trace} e^{\frac{1}{\|X\|}C}.
\end{aligned}$$

Substituting this into (9) we get (8). □

Inequality (8) reflects the asymptotic stability. Actually we can get

$$\operatorname{trace} e^{At} e^{A^H t} \leq \kappa(X) \operatorname{trace} e^{\frac{t}{\|X\|}C}. \tag{10}$$

When C is selected to be negative definite, then the right side of (10) tends to zero as $t \to +\infty$.

When $A + A^H$ is negative definite, set $C = A + A^H$, in this case we have $X = I$ and (8) becomes (5).

The negative definite matrix C can be arbitrary selected in Theorem 4, so we have the following results.

Corollary 5 *Let $A \in \mathcal{C}^{n\times n}$ and $\mathcal{D} := \{C | C \in \mathcal{C}^{n\times n}, \text{negative definite}\}$. Then*

$$\operatorname{trace} e^{A} e^{A^H} \leq \inf_{C\in\mathcal{D}} (\kappa(X) \operatorname{trace} e^{\frac{1}{\|X\|}C}), \tag{11}$$

where X satisfies (7) corresponding to C.

Unfortunately we do not know how to obtain the infimum of (11). But with different choices of the negative definite matrix C we can get different upper bounds. Here are two of them.

Corollary 6 *Let A be stable and have the Jordan canonical form*

$$A = PJP^{-1} = P\,\mathrm{diag}(J_1, \ldots, J_s)P^{-1},$$

with

$$J_i = \mathrm{diag}(\underbrace{J_{i,1}, \ldots, J_{i,1}}_{n_{i,1}}, \ldots, \underbrace{J_{i,t_i}, \ldots, J_{i,t_i}}_{n_{i,t_i}}),\ i = 1, \ldots, s,$$

where

$$J_{i,j} = \begin{bmatrix} \lambda_i & -\mathrm{Re}\,\lambda_i & & \\ & \ddots & \ddots & \mathbf{O} \\ & \mathbf{O} & \ddots & -\mathrm{Re}\,\lambda_i \\ & & & \lambda_i \end{bmatrix}$$

is a modified j-by-j Jordan block for λ_i. Then

$$\mathrm{trace}\, e^A e^{A^H} \le \kappa^2(P) \sum_{i=1}^{s} e^{2\,\mathrm{Re}\,\lambda_i} (\sum_{j=1}^{t_i} n_{i,j} (\sum_{k=1}^{j} e^{2\,\mathrm{Re}\,\lambda_i \cos\frac{k\pi}{j+1}})). \tag{12}$$

Proof. Each $J_{i,j} + J_{i,j}^H$ has the eigenvalues $2\,\mathrm{Re}\,\lambda_i(1 + \cos\frac{k\pi}{j+1})$, $k = 1, \ldots, j$. Because A is stable, $\mathrm{Re}\,\lambda_i < 0$, $1 + \cos\frac{k\pi}{j+1} > 0$ for all j, so $J_{i,j} + J_{i,j}^H$ is negative definite and then $C = P(J + J^H)P^H$ is also negative definite. Therefore (7) has the unique positive definite solution $X = PP^H$. Applying the first part of (8) to these X and C we arrive at (12). $\square$

We can get various bounds in such a way, i.e., first use similarity transformations with diagonal matrices to J to reduce the magnitudes of the subdiagonal elements, such that $J + J^H$ is negative definite, then determine C and X, and finally get an upper bound by applying Theorem 4 to C and X. But this is not a practical way to estimate $\mathrm{trace}\, e^A e^{A^H}$, because if we know the Jordan canonical form of A, we can explicitly compute it.

Corollary 7 *Let A be stable and let $\hat{X}$ be the solution of $AX + XA^H = -I_n$. Then*

$$\mathrm{trace}\, e^A e^{A^H} \le n\kappa(\hat{X}) e^{-\frac{1}{\|\hat{X}\|}}. \tag{13}$$

E	B_{old}	$B_{new,1}$	$B_{new,2}$
$3e^{-2\epsilon}$	$(e+e^{-1})e^{-2\epsilon}$	$\max\{\epsilon^2, \epsilon^{-2}\}(e^{-\epsilon}+e^{-3\epsilon})$	$\frac{2(\sqrt{(2\epsilon)^2+1}+1)^2}{(2\epsilon)^2} e^{-(1-\frac{1}{\sqrt{(2\epsilon)^2+1}})2\epsilon}$

Table 1: Exact value and bounds of trace $e^A e^{A^H}$

$E(t)$	$B_{old}(t)$	$B_{new,1}(t)$	$B_{new,2}(t)$
$(2+t^2)e^{-2\epsilon t}$	$(e^t+e^{-t})e^{-2\epsilon t}$	$\max\{\epsilon^2, \epsilon^{-2}\}(e^{-\epsilon t}+e^{-3\epsilon t})$	$\frac{2(\sqrt{(2\epsilon)^2+1}+1)^2}{(2\epsilon)^2} e^{-(1-\frac{1}{\sqrt{(2\epsilon)^2+1}})2\epsilon t}$

Table 2: Exact value and bounds of trace $e^{At} e^{A^H t}$

Proof. Use (8) with $C = -I_n$ and $X = \hat{X}$. $\square$

From a result in [6], in such a case $1/\left\|\hat{X}\right\| = sep(A, A^H)$, where

$$sep(A, A^H) = \min_{\substack{X \neq 0 \\ X\,Hermitian}} \frac{\left\|AX + XA^H\right\|}{\|X\|}.$$

For an arbitrary negative definite C and the solution X, $\|C\| / \|X\| \geq 1/\left\|\hat{X}\right\|$, so the bound in (13) maybe in general is also the worst one in asymptotic analysis. However, it is a cheaper estimator in practice. To get this bound only the cost of solving a Lyapunov equation is needed.

Finally we give a simple example to compare the new and old bounds.

Example 3 *Let*

$$A = \begin{bmatrix} -\epsilon & 1 \\ 0 & -\epsilon \end{bmatrix}, \ \epsilon > 0.$$

We list the exact value of trace $e^A e^{A^H}$ and its bound in (5), (12) and (13) in Table 1, which are denoted by E, B_{old}, $B_{new,1}$ and $B_{new,2}$, respectively. Also we list them with time t in Table 2 denoted by $E(t)$, $B_{old}(t)$, $B_{new,1}(t)$ and $B_{now,2}(t)$, respectively. When ϵ is sufficiently small B_{old} is a better estimator for E than $B_{new,1}$ and $B_{new,2}$. But in the asymptotic case ($\epsilon < 1/2$) $\lim_{t\to+\infty} B_{old}(t) = +\infty$, while $\lim_{t\to+\infty} B_{new,1} = 0$ with order of $e^{-\epsilon t}$ and $\lim_{t\to+\infty} B_{new,2} = 0$ with order of $e^{-4\epsilon^3 t}$. So $B_{new,1}$ and $B_{new,2}$ are successful to show the asymptotic stability of trace $e^{At} e^{A^H t}$, while B_{old} fails. Furthermore, $B_{new,1}$ is better than $B_{new,2}$.

In the expression of $E(t)$ there is a quadratic polynomial coefficient $2+t^2$ varying for t. With the elementary mathematical result that for arbitrary polynomial $p(t)$ and $\alpha > 0$, $\lim_{t\to+\infty} p(t)e^{-\alpha t} = 0$, we can replace the variable coefficient by a scalar. For example since $\max_{t\in[0,+\infty)} t^2 e^{-\epsilon t} = \frac{4}{e^2}\epsilon^{-2}$, $E(t) \le \frac{4}{e^2}\epsilon^{-2}e^{-\epsilon t} + 2e^{-2\epsilon t}$, which is just equivalent to $B_{new,1}$.

In general, by using the Schur form of A we must have the following form

$$\operatorname{trace} e^{At}e^{A^H t} = \sum_{i=1}^{s}\sum_{j=1}^{s} p_{ij}(t)e^{(\operatorname{Re}\lambda_i + \operatorname{Re}\lambda_j)t} \le p(t)e^{(2\max_i \operatorname{Re}\lambda_i)t},$$

where $p(t)$, $p_{ij}(t)$, $i,j = 1,\ldots,s$ are polynomials and λ_i, $i = 1,\ldots,s$ are the pairwise different eigenvalues of A. Roughly speaking, our bounds are just to try to replace the polynomial coefficient by a scalar. The cost is that we have to take some order $e^{-\delta t}$, with $\delta < -2\max_i \operatorname{Re}\lambda_i$, from the exponential part, to keep the form $p(t)e^{-\delta t}$ bounded and keep the bounds asymptotic stable.

4 Conclusion

In this paper we have achieved two things. The first one is that we have given sufficient and necessary conditions for a matrix A which satisfies $e^A e^{A^H} = e^{A^H}e^A$. The second is that we give a general upper bound for $\operatorname{trace} e^A e^{A^H}$ and two particular bounds for that when A is stable. These two bounds are sometimes weaker than $\operatorname{trace} e^{A+A^H}$, but the importance is that they reveal the asymptotic stability of $\operatorname{trace} e^{At}e^{A^H t}$. Furthermore, the general bound has a freedom to choose the negative definite matrix C, which maybe is useful in practical applications.

Acknowledgements *The author thanks Prof. V. Mehrmann for his valuable comments and helps. He also thanks the referee for his(her) useful suggestions to make the proof of Theorem 1 simple.*

References

[1] D. Bernstein. Orthogonal matrices and the matrix exponential. Problem 90-18, *SIAM Review,* 32: 673, 1990.

[2] D. Bernstein. Some open problems in matrix theory arising in linear system and control. *Linear Algebra and Its Applications,* 162-164: 409-432, 1992.

[3] Jean-Claude Evard and F. Uhlig. On the matrix equation $f(X) = A$. *Linear Algebra and Its Applications,* 162-164: 447-519, 1992.

[4] E. Fischer. Über quadratsche Formen mit reelen Koffizienten. *Monatshefte für Mathematik und Physiks,* 16:234-249, 1905.

[5] G. Golub and C. Van Loan. *Matrix Computations.* Second Edition. The Johns Hopkins University Press, Baltimore and London, 1989.

[6] G. Hewer and C. Kenney. The sensitivity of the stable Lyapunov operator. *SIAM J. Contr. and Optim.,* 26: 321-344,1988.

[7] R. Horn and C. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, 1991.

[8] H. Kwakernaak and R. Sivan. *Linear Optimal Control Systems.* Wiley Press, 1972.

[9] P. Lancaster and M. Tismenetsky. *The Theory of Matrices.* Academic Press, New York, NY, 1985.

[10] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix. *SIAM Review,* 20: 801-836, 1978.

[11] J. Snyders and M. Zakai. On nonnegative solutions of the equation $AD + DA' = -C$. *SIAM J. Applied Mathematics,* 18: 704-714, 1970.

[12] W. So. Equality cases in matrix exponential inequalities. *SIAM J. Matrix Anal. Appl.,* 13:1154-1158, 1992.

[13] W. M. Wonham. *Linear Multivariable Control: A Geometric Approach.* Springer-Verlag, Berlin and New York, 1979.