J. Appl. Math. & Informatics Vol. 27(2009), No. 1 - 2, pp. 105 - 121
Website: http://www.kcam.biz

ON COMPUTATION OF MATRIX LOGARITHM

NAGWA SHERIF AND EHAB MORSY*

ABSTRACT. In this paper we will be interested in characterizing and computing matrices $X \in C^{n\times n}$ that satisfy $e^X = A$, that is logarithms of A. The study in this work goes through two lines. The first is concerned with a theoretical study of the solution set, $S(A)$, of $e^X = A$. Along the second line computational approaches are considered to compute the principal logarithm of A, $LogA$.

AMS Mathematics Subject Classification : 15A24
Key words and phrases : Matrix functions, matrix logarithm, matrix theory

1. Introduction

For an n-by-n complex matrix A, $f(A)$ denotes the primary matrix function associated with the stem function $f(z)$ [11]. Computation of $f(A)$ is a frequently occurring problem in control theory [20], [25], mathematical modeling of dynamical systems [3], and other applications. Indeed, there is extensive work in the area of matrix exponential and matrix pth root computations [2], [6], [10], [13], [19], [22], [23], [26]. However as reported in [4] the computation of matrix logarithm is an important area demanding more work. This is one of our objectives in this work.

Logarithms of matrices arise in various "system identification" problems context, for example [1], [8], [24]. In this work we first study the set $S(A)$ of logarithms of an n-by-n complex matrix A, that is the solution set of the matrix equation $e^X = A$. Particular subsets of $S(A)$ are discussed and characterized. The case of real logarithm of a real matrix is of particular interest and will be investigated in a separate work. Next we propose techniques to compute $LogA$, the principal logarithm of A. These techniques have computational and theoretical advantages. Finally we study sensitivity analysis of the logarithmic function.

Received September 21, 2007. Revised March 14, 2008. Accepted March 26, 2008.
*Corresponding author.

We assemble the notation and preliminaries needed throughout the work in Section 2. The theoretical aspects of the matrix equation $e^X = A$ and its solution set $S(A)$ are investigated in Section 3. The matrix identity $LogAB = LogA + LogB$ is central to our proposed algorithms. It is discussed in Section 4. Section 5 presents scaling strategies necessary to develop and improve the algorithm proposed in Section 6.

2. Notations and preliminaries

We state briefly a number of useful concepts and facts which will be repeatedly used in this work. In this work C will denote the field of complex numbers, C^n the set of all vectors of dimension n. Let $C^{n\times n}$ is the set of all n-by-n complex matrices. For a matrix $A \in C^{n\times n}$, the spectrum of A is denoted by $\sigma(A)$, while the monic polynomial of the smallest degree that annihilates A is called the minimal polynomial of A and is denoted by $c(z)$. The spectral radius $\rho(A)$ represents the dominant eigenvalue of A, that is, $\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$.

A matrix $A \in C^{n\times n}$ is said to be normal if $A^*A = AA^*$, where A^* is the complex conjugate transpose of A. A matrix $U \in C^{n\times n}$ is called unitary if $U^*U = I$, in addition, if $U \in R^{n\times n}$, U is called orthogonal. The spectrum of the unitary and orthogonal matrices lies on the unit circle. Also an n-by-n matrix A is said to be hermitian (skew hermitian) if $A^* = A$ ($A^* = -A$). In this case $\sigma(A) \subset R$ ($\sigma(A) \subset iR$). A matrix $A \in C^{n\times n}$ is said to be positive definite if A is hermitian and $x^*Ax > 0$, for all $x \in C^n$, $x \neq 0$. The class of positive stable matrices is another important particular class of matrices. A matrix $A \in C^{n\times n}$ is said to be positive stable if $Re\lambda > 0$ for every $\lambda \in \sigma(A)$. Let $\log z$, $z \neq 0$ denote the multiple valued function, $\log z = Log\,|z| + i(\arg z + 2\pi k)$, $k \in Z$. Consider the single value defined by $Logz = Log\,|z| + i\theta$, $-\pi < \theta < \pi$. Then $Logz$ is a branch of $\log z$ in the domain $D_\pi = C\backslash\{z : z \leq 0\}$ and it is called the principal branch of $\log z$.

Let $A \in C^{n\times n}$, and let s_i be the multiplicity of $\lambda_i \in \sigma(A)$ as a root of the minimal polynomial of A. Then the scalar valued function $f(z)$ is said to be defined on $\sigma(A)$ if $f(z)$ is defined and has successive derivatives up to $s_i - 1$ at λ_i for each $\lambda_i \in \sigma(A)$.

For suitable assumptions on an n-by-n matrix A and a scalar function $f(z)$, we have different definitions of the matrix function $f(A)$, [2], [4], [12], [15], [21]. Often $f(A)$ is called primary matrix function associated with stem function $f(z)$. In our work we use the Jordan canonical form and integration definitions of such matrix function. If A is diagonalizable then the matrix function $f(A)$ is also diagonalizable. However if A is not diagonalizable, $f(A)$ may be diagonalizable. Based on the nature of the scalar function $f(z)$ and the spectrum of the underlying matrix A, the next theorem investigates the Jordan structure of $f(A)$ [12].

Theorem 1. *Let $J_{m_k}(\lambda_k)$ be an m_k-by-m_k Jordan block with eigenvalue λ_k, suppose $f(z)$ is a scalar-valued function which is $(m_k - 1)-$times differentiable at λ_k. If $f'(\lambda) \neq 0$, the Jordan canonical form of $f(J_{m_k}(\lambda_k))$ is the single block*

$J_{m_k}(f(\lambda_k))$. Let $1 \le p \le m_k$ be given and $m_k = pq + r$, with $0 \le r < p$. If $f'(\lambda_k) = f''(\lambda_k) = \ldots = f^{(p-1)}(\lambda_k) = 0$ and either $p = m_k$ or $f^{(p)}(\lambda_k) \neq 0$, the Jordan canonical form of $f(J_{m_k}(\lambda_k))$ splits into exactly p blocks, each of which has an eigenvalue $f(\lambda_k)$; there are $p - r$ blocks $J_q(f(\lambda_k))$ and r blocks $J_{q+1}(f(\lambda))$.

If $A \in C^{n\times n}$ and B is a polynomial in A, say $p(A)$, then clearly A and B are commuting. However, the converse is not necessarily true. The next theorem gives a sufficient condition for B to be a polynomial in A [11].

Theorem 2. *Let $A \in C^{n\times n}$ be a given nonderogatory matrix. A matrix $B \in C^{n\times n}$ commutes with A if and only if there is a polynomial $p(z)$ of degree at most $n - 1$ such that $B = p(A)$.*

3. Solution set of $e^X = A$

For an n-by-n matrix A, let

$$\begin{aligned} S(A) &= \{X \in C^{n\times n}: \ e^X = A\} \\ S_P(A) &= \{X \in C^{n\times n}: \ e^X = A \text{ and } X = p(A),\ p(z) \text{ is a polynomial }\} \\ S_D(A) &= \{X \in C^{n\times n}: e^X = A \text{ and } X \text{ is diagonalizable}\}. \end{aligned}$$

The solvability of the given matrix equation is completely characterized by the following theorem.

Theorem 3. *Let A be an n-by-n complex matrix. A sufficient and necessary condition for the solvability of the matrix equation $e^X = A$ is that the scalar equation $e^{x_i} = \lambda_i$ is solvable, for each $\lambda_i \in \sigma(A)$, $1 \le i \le n$. That is, the equation is solvable, if, and only if, A is nonsingular.*

Proof. Let A have the Jordan canonical form $A = SJ_AS^{-1} = Sdiag(J_{m_1}(\lambda_1), \ldots, J_{m_p}(\lambda_p))S^{-1}$. Assume that for each $1 \le i \le n$, there is an x_i such that $e^{x_i} = \lambda_i$. Set

$$J_X = diag(J_{m_1}(x_1), J_{m_2}(x_2), \ldots, J_{m_p}(x_p)).$$

We show that there exists a matrix $X \in C^{n\times n}$, that is similar to J_X and such that $e^X = A$. The Jordan structure of the matrix e^{J_X} is completely determined by the function $f(z) = e^z$; Theorem 1. Indeed the Jordan canonical form of e^{J_X} has the form $diag(J_{m_1}(e^{x_1}), J_{m_2}(e^{x_2}), \ldots, J_{m_p}(e^{x_p}))$ which by our assumption is J_A. Now, it turns out that J_A and e^{J_X} are similar, and we have $J_A = T\ e^{J_X}\ T^{-1}$, for a nonsingular matrix T. This gives

$$A = S\ J_A\ S^{-1} = ST(e^{J_X})T^{-1}S^{-1} = R\ e^{J_X} R^{-1},$$

where $R = ST$. Hence $e^X = A$ with $X = R\ J_X\ R^{-1}$. This establishes the sufficiency part, the necessary part is immediate by using the Jordan canonical form definition of the matrix exponential

$$e^X = Vdiag(J_{m_1}(e^{x_1}), J_{m_2}(e^{x_2}), \ldots, J_{m_p}(e^{x_p}))V^{-1},$$

where $X = Vdiag(J_{m_1}(x_1), J_{m_2}(x_2), \ldots, J_{m_p}(x_p))V^{-1}$. □

A solution X that is constructed in the previous theorem depends on the choices of the solutions to the scalar equation $e^{x_i} = \lambda_i$, $1 \leq i \leq n$. For any arbitrary solution set $\{x_i\}_{i=1}^{n}$, there is no guarantee that the constructed X will be a polynomial in A. In the next theorem we give the necessary and sufficient conditions for such X to be polynomial in A.

Theorem 4. *Let $A \in C^{n\times n}$, be a nonsingular matrix with Jordan canonical form*

$$A = S\ diag(J_{m_1}(\lambda_1), J_{m_2}(\lambda_2), \ldots , J_{m_p}(\lambda_p))\ S^{-1}. \tag{1}$$

Then $X \in S(A)$ is a polynomial in A if and only if the same value of the scalar logarithm is used for the same eigenvalue of A, that is, if $e^{x_k} = \lambda_k$ for every $k = 1, 2, \ldots , p$, then $\lambda_i = \lambda_j$ implies that $x_i = x_j$ for all $1 \leq i, j \leq p$.

Proof. Let X be a logarithm of A and suppose that the same branch of the scalar logarithm is used for the same eigenvalue of A. Let $x_1, x_2, \ldots , x_\mu$ be the distinct eigenvalues of X and let

$$J_c = diag(J_{n(x_1,X)}(x_1), J_{n(x_2,X)}(x_2), \ldots , J_{n(x_\mu,X)}(x_\mu)),$$

where $n(x, X)$ denotes the multiplicity of x as a root of the minimal polynomial of X. Clearly $\sigma(J_c) = \sigma(X)$ and $n(x, J_c) = n(x, X)$ for every $x \in \sigma(X)$. According to Theorem 1, no splitting of the Jordan blocks occurs due to the exponential function, so that e^{J_c} is similar to

$$diag(J_{n(x_1,X)}(e^{x_1}), J_{n(x_2,X)}(e^{x_2}), \ldots , J_{n(x_\mu,X)}(e^{x_\mu})).$$

By our choice of x_k, $1 \leq k \leq \mu$ and our assumption, $e^{x_1}, e^{x_2}, \ldots , e^{x_\mu}$ are distinct and consequently e^{J_c} is a nonderogatory matrix. On the other hand J_c and e^{J_c} are commuting. Therefore, by Theorem 2, there exists a polynomial $p(z)$ such that $J_c = p(e^{J_c})$.Consequently

$$J_{n(x,J_c)}(x) = p(e^{J_{n(x,J_c)}(x)}) \qquad \text{for all } x \in \sigma(J_c). \tag{2}$$

Hence from equation (2), and since $n(x, J_c) = n(x, X)$, we obtain $J_{n(x,X)} = p(e^{J_{n(x,X)}})$ for all $x \in \sigma(X)$, and hence $J_{n(x,X)} = p(e^{J_{n(x,X)}})$ implies that $J_{m(x,X)} = p(e^{J_{m(x,X)}})$for all $m(x, X) \leq n(x, X)$. Therefore, $X = p(e^X) = p(A)$. Conversely if $X \in S_p(A)$ then it is clear that $\lambda_1 = \lambda_2$ implies that $x_1 = x_2$. □

The next theorem describes the structure of any element $X \in S(A)$.

Theorem 5. *Let $A \in C^{n\times n}$ be a nonsingular that has the Jordan canonical form $A = SJS^{-1} = Sdiag(J_{m_1}(\lambda_1), \ldots , J_{m_p}(\lambda_p))S^{-1}$. Then all logarithms of A are given by*

$$X = SU\ diag(\log^{(j_1)}(J_{m_1}(\lambda_1)), \log^{(j_2)}(J_{m_2}(\lambda_2)), \ldots , \log^{(j_p)}(J_{m_p}(\lambda_p)))\ U^{-1}S^{-1},$$

where $j_k \in Z$, $k = 1, 2, \ldots , p$, and U is an arbitrary nonsingular matrix that commutes with J and $\log^{(j)} z$ is a branch of $\log z$.

Proof. Suppose that X is a logarithm of A, that is a solution of the matrix equation $e^X = A$. The scalar function $f(z) = e^z$ never vanishes, this implies that X and A will have the same number and sizes of Jordan blocks, Theorem 1. So the Jordan canonical form of X is given by

$$J_X = diag(J_{m_1}(x_1), J_{m_2}(x_2), \ldots, J_{m_p}(x_p)),$$

where x_k satisfies the scalar equation $e^{x_k} = \lambda_k$, $1 \le k \le p$. Suppose that $x_k = \log^{(j_k)} \lambda_k$, where $\log^{(j_k)}(\lambda_k)$ denotes a branch of the logarithm in the neighborhood of λ_k. Set

$$Y = diag(L_1^{(j_1)}, L_2^{(j_2)}, \ldots, L_p^{(j_p)}),$$

where $L_k^{(j_k)} = \log^{(j_k)}(J_{m_k}(\lambda_k))$, $j_k \in Z$. It is clear from the structure of $L_k^{(j_k)}$ that it does not have any splitting, hence Y and X will have the same Jordan canonical forms, that is, there exists a nonsingular matrix $T \in C^{n\times n}$ such that $X = TYT^{-1}$, but $e^X = A$, hence

$$\begin{aligned} A &= e^{TYT^{-1}} = Te^YT^{-1} = Tdiag(e^{L_1^{(j_1)}}, e^{L_2^{(j_2)}}, \ldots, e^{L_1^{(j_p)}})T^{-1} \\ &= Tdiag(J_{m_1}(\lambda_1), J_{m_2}(\lambda_2), \ldots, J_{m_k}(\lambda_k))T^{-1} = TJT^{-1}. \end{aligned}$$

It is known that if $A = SJS^{-1}$ and $A = TJT^{-1}$, then $T = SU$ where U commutes with J. Consequently we have

$$X = SUdiag(\log^{(j_1)}(J_{m_1}(\lambda_1)), \ldots, \log^{(j_p)}(J_{m_p}(\lambda_p)))U^{-1}S^{-1}. \tag{3}$$

□

A direct consequence of Theorem 1 and the previous theorem is that any $X \in S(A)$ will have the same Jordan structure as A. In particular if A is nonsingular matrix then X is diagonalizable if and only if A is diagonalizable. This implies that $S(A) = S_D(A)$ if and only if A is diagonalizable.

For a matrix $A \in C^{n\times n}$ with $S(A) \neq \phi$, we discuss the cardinality of $S(A)$. Clearly the scalar equation $e^x = \lambda$ has a countable number of solutions in C. The question arises whether $S(A)$ is also countable. To discuss the countability of $S(A)$ we need the following lemma.

Lemma 1. *Let $J_{m_k}(\lambda_k)$ be a Jordan block of order m_k, $\lambda_k \neq 0$. Then $J_{m_k}(\lambda_k)$ has a countable set of logarithms, each is given by*

$$\log^{(j)}(J_{m_k}(\lambda_k)) = \begin{bmatrix} \log^{(j)}\lambda_k & 1/\lambda_k & \ldots & -\frac{(-\lambda_k)^{-(m_k-1)}}{m_k-1} \\ 0 & \log^{(j)}\lambda_k & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \log^{(j)}\lambda_k \end{bmatrix}, \tag{4}$$

where $\log^{(j)}\lambda_k$ is a branch of $\log z$ in a neighborhood of λ_k ($log^{(j)}\lambda_k = Log\,|\lambda_k| + (Arg\lambda_k + 2\pi j)i$ and $j \in Z$). Furthermore each logarithm is a polynomial in A.

Proof. Suppose that X is a logarithm of $J_{m_k}(\lambda_k)$, that is, $e^X = J_{m_k}(\lambda_k)$. Since X commutes with $J_{m_k}(\lambda_k)$, and since $J_{m_k}(\lambda_k)$ is nonderogatory matrix, so by Theorem 2 there is a polynomial $p(z)$ such that $X = p(J_{m_k}(\lambda_k))$. $J_{m_k}(\lambda_k)$ is upper triangular, which implies that X is upper triangular with equal diagonal elements. From the commutativity of X and $J_{m_k}(\lambda_k)$, we conclude that X takes the particular form $X = D + N = x_1 I + N$, where $e^{x_1} = \lambda_k$ and

$$N = \begin{bmatrix} 0 & x_2 & x_3 & \dots & x_{m_k} \\ 0 & 0 & x_2 & \dots & x_{m_k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & x_2 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

There are countably many solutions of $e^{x_1} = \lambda_k$, $x_1 = \log^{(j)} \lambda_k$, $j \in Z$ where $\log^{(j)} \lambda_k$ is a branch of the logarithm in the neighborhood of λ_k. It remains to show that X has the form (4). Clearly $N^{m_k} = 0$, then $J_{m_k}(\lambda_k)$ can be written in the form

$$J_{m_k}(\lambda_k) = e^X = \lambda_k\, e^N = \lambda_k \left[I + N + \frac{1}{2!}N^2 + \frac{1}{3!}N^3 + \ldots + \frac{1}{(m_k-1)!}N^{m_k-1} \right].$$

For n-by-n matrix B with ones in the first superdiagonal and zeros otherwise, $J_{m_k}(\lambda_k)$ can be written as

$$J_{m_k}(\lambda_k) = \lambda_k I + \lambda_k \sum_{v=1}^{m_k-1} x_{v+1} B^v + \ldots + \frac{\lambda_k}{(m_k-1)!} (\sum_{v=1}^{m_k-1} x_{v+1} B^v)^{m_k-1},$$

equating the coefficients of the powers of B in both sides, we get

$$\lambda_k I + B = \lambda_k I + (\lambda_k x_2)B + \lambda_k(x_3 + \frac{1}{2}x_2^2)B^2 + \lambda_k(x_4 + x_2 x_3 + \frac{1}{3!}x_2^3)B^3 + \ldots.$$

This in turn implies that $x_q = (-1)^q/(q-1)\ \lambda_k^{q-1}$ for all $q = 2, 3, \ldots, m_k$. Now it follows from Corollary 5 that $S(A) = S_p(A)$, and indeed every logarithm of a Jordan block is a polynomial in it. □

Theorem 6. *Let $A \in C^{n\times n}$ be a nonsingular matrix with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_\mu$ and let A have Jordan canonical form*

$$A = SJS^{-1} = Sdiag(J_{m_1}(\lambda_1), \ldots, J_{m_p}(\lambda_p))S^{-1}.$$

If $\mu < p$, then there exists an uncountable number of logarithms of A, among those there exists a countable set of logarithms that are polynomials in A. If $\mu = p$, that is, if A is nonderogatory, then the logarithms of A are countable and each of them is polynomial in A.

Proof. Theorem 5 implies that any logarithm of A has the form

$$X = SUdiag(\log^{(j_1)}(J_{m_1}(\lambda_1)), \ldots, \log^{(j_p)}(J_{m_p}(\lambda_p)))U^{-1}S^{-1}, \tag{5}$$

where U is any nonsingular matrix that commutes with J. When $\mu < p$, it may happen that U does not commute with $\log J$ and in this case there are uncountable logarithms of A since there are uncountable matrices that commute with J . On the other hand if the same branch of the scalar logarithm is used for Jordan blocks with the same eigenvalues, that is, $j_{t_1} = j_{t_2}$ when $\lambda_{t_1} = \lambda_{t_2}$, then $\log J$ is a polynomial in J, Theorem 4. In this case U commutes with $\log J$, so (5) implies that $X = S\log^{(j)} JS^{-1}$. Now this in turn implies that $S_p(A)$ is countable.

If $\mu = p$ then J is nonderogatory matrix and any logarithm $\log J$ of J is a polynomial in J. It follows that U commutes with $\log J$, and hence all the logarithms of A are countable and polynomials in A. □

If two or more Jordan blocks of an n-by-n nonsingular matrix A correspond to the same eigenvalue, we get a parameterized family of solutions as indicated by the following corollary.

Corollary 1. *Let $A \in C^{n\times n}$ be a nonsingular matrix, with Jordan canonical form $A = SJS^{-1} = S\ diag(J_{m_1}(\lambda_1), \dots, J_{m_p}(\lambda_p))\ S^{-1}$, where $\lambda_1, \lambda_2, \dots, \lambda_p$ are not necessarily distinct. For a fixed $j^{(\circ)} = (j_1^{(\circ)}, j_2^{(\circ)}, \dots, j_p^{(\circ)})$, $j_k^{(\circ)} \in Z$, $0 \le k \le p$, there exists an uncountable set of logarithms of A, each is given by*

$$X_{j^{(\circ)}}(U) = SU\ diag(\log^{j_1^{(\circ)}}(J_{m_1}(\lambda_1)), \dots, \log^{j_p^{(\circ)}}(J_{m_p}(\lambda_p)))U^{-1}\ S^{-1}$$

where U is any nonsingular matrix that commutes with J. The members of this parametrized family have a common spectrum.

We end this section by characterizing those matrices A for which the matrix equation $e^X = A$ has a solution belonging to a particular class of matrices. Results in this direction for general $f(A)$ can be found in [11].

Lemma 2. *Let $A \in C^{n\times n}$ be a nonsingular matrix, then A has a normal logarithm if and only if A is normal.*

Proof. Suppose that A has a normal logarithm X . Then there exists a unitary matrix $U \in C^{n\times n}$, such that $X = U\ diag(x_1, x_2, \dots, x_n)\ U^*$, where $x_1, x_2, \dots, x_n \in \sigma(X)$, and $e^{x_i} = \lambda_i$ for all $i = 1, 2, \dots, n$. Hence, we have

$$A = e^X = U\ diag(e^{x_1}, e^{x_2}, \dots, e^{x_n})\ U^*$$

therefore A is normal with eigenvalues $e^{x_1}, e^{x_2}, \dots, e^{x_n}$.

Conversely, let A be normal and for every $\lambda \in \sigma(A)$, there exists $x_i \in C$ such that x_i is a scalar logarithm of λ_i for all $i = 1, 2, \dots, n$. Then there exists a unitary matrix $U \in C^{n\times n}$ such that $A = U\ diag(\lambda_1, \lambda_2, \dots, \lambda_n)\ U^*$, where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of A. Let $X = U\ diag(\ x_1, x_2, \dots, x_n)\ U^*$; it follows that X is normal and $e^X = U\ diag(\ e^{x_1}, e^{x_2}, \dots, e^{x_n})\ U^* = A$. □

As a result of the previous lemma we can conclude that a normal matrix A has a normal logarithm X with prescribed spectrum $\sigma(X) \subset K$, $K \subset C$ if and only if the equation $e^x = \lambda$ has solution in K for each $\lambda \in \sigma(A)$. In what follows we characterize A so that $LogA$ belongs to certain classes of matrices.

Theorem 7. *Let $A \in C^{n\times n}$ be a nonsingular matrix, then*

(a) *There exists a hermitian logarithm of A if and only if A is positive definite. This hermitian logarithm is unique.*
(b) *There exists a skew-hermitian logarithm of A if and only if A is unitary.*
(c) *There exists a unique positive definite (semi-definite) logarithm of A if and only if $A - I$ is positive definite (positive semidefinite).*
(d) *There exists a unitary logarithm of A if and only if A is normal and for every $\lambda \in \sigma(A)$ there exists $x \in C$ such that x is the scalar logarithm of λ and $|x| = 1$.*

Proof. Let $A \in C^{n\times n}$ be nonsingular. It is well known that A is hermitian, skew hermitian, positive definite or unitary if $\sigma(A) \subset R$, iR, R^+, or unit circle respectively. In all the previous cases, A is normal. First we prove (a). By using Lemma 2 and the obvious fact that the scalar equation $e^x = \lambda$ has a unique real solution if and only if λ is positive real number, that is, $\lambda \in R^+$. Hence the matrix equation $e^X = A$ has a hermitian solution if and only if A is positive definite. The uniqueness follows from Theorem 5. Next we prove (b). There exists $x \in iR$ such that $e^x = \lambda$ if and only if λ lies on the unit circle, that is, if and only if A is unitary. The theorem is completely proved by taking $K = R_+^*(R_+)$ and $K = \{z \in C \setminus |z| = 1\}$ in (c) and (d) respectively in the previous remark, where $R_+ = \{z : z \geq 0\}$ and $R_+^* = \{z : z > o\}$. □

In developing our algorithms for computing $LogA$, a need arises for the study of the matrix identity $LogAB = LogA + LogB$. We study this matrix identity in the next section. The discussion reveals different sets of sufficient conditions for its validity.

4. On the identity LogAB = LogA + LogB

Let $A, B \in C^{n\times n}$ be nonsingular matrices. We seek conditions on the matrices A and B such that the identity

$$LogAB = LogA + LogB \tag{6}$$

is valid. The corresponding identity for the exponential function, $e^{A+B} = e^A e^B$ has been thoroughly investigated by many authors, for example [12], [27]. However there is no known set of sufficient and necessary conditions to ensure that (6) is satisfied. Clearly a necessary condition is that A and B are commuting. The following example shows that commutativity of A and B is not sufficient. Let $A = diag(i, 1)$ and $B = diag(-1 + i, 1 + i)$. The commuting A, B do not satisfy (6). It is well known that this identity, (6), is not always true even in the scalar case. Indeed for $z_1, z_2 \in D_\pi$ one can show that

$$Logz_1z_2 = Logz_1 + Logz_2 \text{ if and only if } -\pi < Argz_1 + Argz_2 < \pi. \tag{7}$$

The next theorem gives a natural extension of (7) to the matrix case.

Theorem 8. *Let $A, B \in C^{n\times n}$ be nonsingular simultaneously diagonalizable matrices. Let $S \in C^{n\times n}$ be nonsingular such that $A = S\ diag(\lambda_1, \lambda_2, \dots, \lambda_n)$*

S^{-1} *and* $B = S\ diag(\lambda_1', \lambda_2', \ldots, \lambda_n')\ S^{-1}$. *Then* $LogAB = LogA + LogB$ *if and only if* $-\pi < Arg\lambda_i + Arg\lambda_i' < \pi$, $1 \le i \le n$.

Proof. It is clear from equation (7) and the conditions in the theorem that

$$LogAB = Sdiag(Log\lambda_1 + Log\lambda_1', \ldots, Log\lambda_n + Log\lambda_n')S^{-1} = LogA + LogB.$$

□

This theorem and [11, Theorem 1.3.12] imply the following result.

Corollary 2. *Let* $A, B \in C^{n\times n}$, *be nonsingular diagonalizable commuting matrices such that for any* $\lambda_i \in \sigma(A)$, *and* $\lambda_i' \in \sigma(B)$ *we have* $-\pi < Arg\lambda_i + Arg\lambda_i' < \pi$. *Then* $LogAB = LogA + LogB$

The condition of simultaneously diagonalizable in Theorem 8 is sufficient but not necessary. It straight forward to check that $LogJ_2(\lambda)J_2(\mu) \neq LogJ_2(\lambda) + LogJ_2(\mu)$ for arbitrary Jordan blocks. Indeed a sufficient and necessary condition for $LogJ_2(\lambda)J_2(\mu) \neq LogJ_2(\lambda) + LogJ_2(\mu)$ is that $-\pi < Arg\lambda + Arg\mu < \pi$.

An instance where we can drop the condition that A is diagonalizable is established in the following theorem.

Theorem 9. *Let* $A \in C^{n\times n}$, $\sigma(A) \subset D_\pi$ *and* $B = g(A)$ *where* $g(z)$ *is an analytic function on* D_π *with range* $g(z) \subset D_\pi$. *Assume that the inequality*

$$-\pi < Arg\lambda + Arg(g(\lambda)) < \pi, \tag{8}$$

is valid on $\sigma(A)$. *Then* $LogAB = LogA + LogB$.

Proof. It is known [15, Theorem 5.7.1] that if $G(u_1, u_2, \ldots, u_s)$ is a polynomial in $u_1, u_2, \ldots, u_s$ and if $f_1(z), f_2(z), \ldots f_s(z)$ are functions defined on the spectrum of the matrix $A \in C^{n\times n}$ such that the function $g(z) = G(f_1(z), f_2(z), \ldots f_s(z))$ is zero on the spectrum of A, then $G(f_1(A), f_2(A), \ldots f_s(A)) = 0$. For $G(u_1, u_2, u_3) = u_1 - u_2 - u_3$ where $f_1(z) = Log(zg(z))$, $f_2(z) = Logz$, and $f_3(z) = Log(g(z))$, and using (8) we get

$$G(f_1(z), f_2(z), f_3(z)) = Log(zg(z)) - Logz - Log(g(z)) = 0$$

on $\sigma(A)$. Hence $G(f_1(A), f_2(A), f_3(A)) = 0$, that is, $Log(Ag(A)) = LogA + Log(g(A))$. □

In the next section we develop particular techniques to transform a given matrix A, so that the spectrum of the transformed matrix satisfies certain inclusion properties. We call these techniques scaling strategies.

5. Scaling strategies

Let A be an n-by-n complex matrix with $\sigma(A) \subset D_\pi$. We discuss different scaling strategies to transform A to A_{sc}, so that $\sigma(A_{sc})$ is clustered around 1 and $diam(\sigma(A_{sc}))$ is "small". In particular we require that $\sigma(A_{sc}) \subset \{z : |z-1| < 1\}$. Identity (6), when applicable, will be essential so that $LogA_{sc}$ can be related to

LogA. An efficient scaling strategy will be shown to rely on the structure of the matrix A, in particular on the distribution of eigenvalues of A.

Next, we develop simple and successive scaling strategies. We discuss the range of applicability and computational merits of each one. Also we comment on the actual implementation of such strategies.

5.1. Simple scaling

Let $A \in C^{n\times n}$ with $\sigma(A) \subset R_+^*$, suppose that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0$, which is the case for particular classes of matrices (positive definite and M-matrices). An estimate for a scaling factor, m_{sc}, so that $A_{sc} = \frac{1}{m_{sc}} A$ has the desired spectral properties, is given in the next lemma.

Lemma 3. *Let $A \in C^{n\times n}$ with $\sigma(A) \subset R_+^*$ and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0$ be the eigenvalues of A. Choose $m_{sc} = \frac{\lambda_1+\lambda_n}{2}$, then the transformed matrix $A_{sc} = \frac{1}{m_{sc}} A$ has spectrum $\sigma(A_{sc})$ with $diam(\sigma(A_{sc})) < 2$ and $\sigma(A_{sc}) \subset \left[\frac{2}{a+1}, \frac{2a}{a+1}\right]$, with $a = \lambda_1/\lambda_n$.*

Proof. The proof is immediate since it is easy to check that scaling A by the positive real number $m_{sc} = (\lambda_1 + \lambda_n)/2$, forces the eigenvalues of the scaled matrix to lie in the interval $\left[\frac{2}{a+1}, \frac{2a}{a+1}\right]$. We note that this interval is symmetric around 1. □

If $\sigma(A)$ is not a subset of R_+^*, but A is positive stable, we can devise a technique for estimating m_{sc} as follows.

If α is a positive real number such that $0 < \alpha \leq Re\lambda \leq |\lambda| \leq \rho(A)$ for all $\lambda \in \sigma(A)$, a suitable choice of m_{sc} satisfies $m_{sc} > \rho(A)^2/(2\alpha)$. For this choice of m_{sc}, we have for any $\lambda \in \sigma(A)$

$$\frac{|\lambda|^2}{m_{sc}^2} \leq \frac{\rho(A)^2}{m_{sc}^2} < \frac{2\alpha}{m_{sc}} \leq \frac{2(Re\lambda)}{m_{sc}}$$

and hence $|\lambda/m_{sc} - 1| < 1$. Thus, we have proved the following result.

Lemma 4. *Let $A \in C^{n\times n}$ be a positive stable matrix. If α is a positive real number such that $0 < \alpha \leq Re\lambda \leq |\lambda| \leq \rho(A)$ for every $\lambda \in \sigma(A)$, then a scaling factor m_{sc} for which $\left|\frac{\lambda}{m_{sc}} - 1\right| < 1$, $\lambda \in \sigma(A)$ is given by $m_{sc} > \rho(A)^2/(2\alpha)$.*

If we assume for the moment that $LogA_{sc}$ can be computed by some means, then upon invoking Theorem 9 we can recover $LogA$ by the matrix identity

$$LogA = LogA_{sc} + (Logm_{sc})I. \tag{9}$$

In some appropriate classes of matrices the value of α in the above lemma is easily obtained. For example if A is a nonsingular normal matrix, then for $H = (A^* + A)$ we have $\sigma(H) = \{2Re\lambda : \lambda \in \sigma(A)\}$ [5]. Consequently $2\alpha = \min\{\mu : \mu \in \sigma(H)\}$.

5.2. Successive scaling

A powerful technique of scaling $A \in C^{n\times n}$ with spectrum $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n| > 0$ is to start by scaling A with m_{sc} to give $A_1 = \frac{1}{m_{sc}}A$, then scaling A_1 using inverse scaling and squaring technique; Kenney and Loab [14], namely

$$A_{sc} = A_1^{1/2^j} = (A/m_{sc})^{1/2^j} \qquad for\ some\ j \in N.$$

In this case the scaling factors m_{sc} and j can be chosen so that the scaled matrix A_{sc} has eigenvalues close enough to 1. Finally we compute $LogA$ from the relation

$$LogA = 2^j LogA_{sc} + (Logm_{sc})I. \tag{10}$$

Matrix identities in (9) and (10) are special cases from the identity $LogAB = LogA + LogB$, Theorem 9.

Let $A \in C^{n\times n}$ with spectrum $\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. One method of computing α in Lemma 4 is to let $B = A^{1/2}$, the principal square root of A, so the corresponding $\sigma(B) = \left\{\lambda_1', \lambda_2', \ldots, \lambda_n'\right\}$, where $\lambda_i' = \lambda_i^{1/2}$ for all $i = 1, 2, \ldots, n$. Since the spectrum of A lies in the left half plane. A is positive stable necessitates that

$$\lambda_i' = |\lambda_i'|\ e^{i\phi_i} \qquad -\pi/4 < \phi_i < \pi/4.$$

Now $Re\lambda_i' = |\lambda_i'|\cos\phi_i > |\lambda_i'|/\sqrt{2}$, and α can be taken to be $\min|\lambda_i'|/\sqrt{2}$. Therefore B can be scaled by the scaling factor of the Lemma 4 and then $LogB$ is computed via relation (9). Finally $LogA$ is given by

$$LogA = 2LogA^{1/2} = 2LogB.$$

Successive scaling will be repeatedly used in our developed algorithms to compute $LogA$. This strategy is summarized in the following algorithm, scaling algorithm.

Algorithm.1 (Scaling algorithm) If $A \in C^{n\times n}$ and m_{sc} and j are given, the following algorithm computes the scaled matrix $A_{sc} = (A/m_{sc})^{1/2^j}$.

A_{sc} =scale(A, m_{sc}, j)
Input A, m_{sc}, j
$A_1 = \frac{1}{m_{sc}}A$
$A_{sc} = A_1^{1/2^j}$
end.

In implementing the previous algorithm, the scaling factor m_{sc} is suggested to be $(|\lambda_1| + |\lambda_n|)/2$, where λ_1 is the dominant eigenvalue of A computed using power method [16, p. 364] and the minimum eigenvalue λ_n is the reciprocal of the dominant eigenvalue of A^{-1}. The square roots are computed by using Schur decomposition [2].

6. Computation of *LogA*

In this section we are interested in computing a particular element of $S(A)$, the principal logarithm $LogA$. We develop different approaches to compute $LogA$. These approaches have theoretical as well as computational advantages. In fact, one of these approaches establish new representations of $LogA$, Section 6.2. Our proposed algorithms are iterative in nature, hence we provide an error estimate for the approximation of $LogA$.

6.1. Series and rational approximation

One effective method for approximating general matrix function $f(A)$ is through the truncation of its Taylor series. This general technique can be adapted to compute $LogA$. The Taylor series approximation of $LogA$ for a nonsingular A is given by

$$LogA = -\sum_{k=1}^{\infty} \frac{(I-A)^k}{k} \tag{11}$$

provided this matrix series is convergent, that is, if and only if

$$\rho(A - I) < 1. \tag{12}$$

Condition (12) necessitates that $\sigma(A) \subset \{z : |z-1| < 1\}$. For an arbitrary nonsingular matrix A, this condition can be realized for the scaled matrix A_{sc}. Here A_{sc} is obtained by the scaling algorithm, Section 5.

For a nonsingular matrix A Golub and Loan [4] give an error bound on approximating $f(A)$ by its truncated Taylor series. Mathias [18] improves this error bound to be independent of the size of the underlying matrix. The latter result, when adapted to our case, yields the following error estimate.

Theorem 10. *Let $A \in C^{n\times n}$ be nonsingular with $\rho(A-I) < 1$, then*

$$\left\| LogA - \sum_{k=1}^{q} \frac{-1}{k}(I-A)^k \right\| \le \frac{1}{(q+1)} \max_{0\le s\le 1} \left\| (I-A)^{q+1} \left[I - s(I-A)\right]^{-(q+1)} \right\|. \tag{13}$$

Another general technique for computing $f(A)$ is by rational approximation. The Padé approximation of $LogA$ can be developed as follows. Let $f(z) = Log(1-z) = -\sum_{k=1}^{\infty} \frac{z^k}{k}$. Set the rational function

$$R_{km}(z) = \frac{p_k(z)}{q_m(z)},$$

where p_k and q_m are polynomials in z of degree at most k and m, respectively. Then $R_{km}(z)$ is called (k,m) Padé approximant of f if R_{km} agrees with f up through order $k+m$. Now, we can get the Padé approximant of $LogA$ as

$$R_{km}(A) = p_k(I-B)(q_m(I-B))^{-1}$$

where $B = I - A$. Several methods are analyzed for evaluating the Padé approximant of $LogA$, for examples, Horner's method, Paterson-Slockmeyer method, Van Loan's method, and methods based on continued fraction and partial fraction expansions [8], [9].

It is well known that Padé approximation works well when the norm $\|I - A\|$ is closer to zero [14], [19]. This can be ensured by using scaling strategies to bring the matrix close to the identity, computes a Padé approximant of the scaled matrix, and then scales back.

6.2. Algebraic approach

In this subsection we present a technique to compute $LogA$ that is different in nature from that developed in Section 6.1. This technique is based on the analytic properties of $Logz$ and is valid for any A with $\sigma(A) \subset D_\pi$. In fact we show that $LogA$ can be realized as the limit of a sequence of matrices $\{A_r\}$, where each A_r is an algebraic function of A. Namely

$$LogA = \lim_{r\to\infty} A_r \tag{14}$$

The above relation (14) suggests an iterative technique to compute $LogA$. We start our discussion of this technique by considering the class of positive definite matrices where the theory becomes mathematically rich. The following result is essential to our development.

Theorem 11. *Let $a \in R_+^*$ be given and define the sequence $\{a_r\}$, $a_r = r(a^{1/r} - 1)$. Then*

(i) $\{a_r\}$ is monotonically decreasing and $\lim_{r\to\infty} a_r = Loga$.

(ii) $\frac{|a_r - Loga|}{|Loga|} < \frac{|a-1|}{r}$.

Proof. A direct application of the generalized mean value theorem yields the relation,

$$Loga = r(a^{\frac{1}{r}} - 1)\frac{1}{\sqrt[r]{\zeta}}. \tag{15}$$

where ξ lies between a and 1. Now, it follows that

$$Loga = \lim_{r\to\infty} a_r = \lim_{r\to\infty} r(a^{\frac{1}{r}} - 1).$$

To show that $\{a_r\}$ is monotonically decreasing, we need to show that for $f(x) = x(a^{1/x} - 1)$, we have $f'(x) < 0$. We consider two cases $a < 1$, and $a > 1$. As for the case $a < 1$, we have $a^{1/x} < 1$ and $f'(x) = a^{1/x} - 1 - a^{1/x}\frac{Loga}{x} = u[1 - Logu] - 1$, where $u = a^{1/x}$. However, $Logu = -\sum_{k=1}^{\infty}(1-u)^k/k$, and hence $1 - Logu < \sum_{k=0}^{\infty}(1-u)^k = 1/u$, so finally $f'(x) = u[1 - Logu] - 1 < 0$. As for $a > 1$, we have

$$xf'(x) = \sum_{k=1}^{\infty}\left[\frac{1}{k!} - \frac{1}{(k-1)!}\right]\frac{(Loga)^k}{x^k k!} < 0,$$

and consequently $f'(x) < 0$. To prove the second assertion, we have

$$a_r - Loga = (\xi^{1/r} - 1)Loga.$$

If $a > 1$, we have $\xi^{1/r} < a^{1/r}$ and $Loga > 0$, then

$$\frac{a_r - Loga}{Loga} < \frac{a-1}{1 + a^{1/r} + a^{2/r} + \ldots + a^{(r-1)/r}}, \quad (16)$$

and so $(a_r - Loga)/Loga < (a-1)/r$. Now $\{a_r\}$ being monotone decreasing and we have $|a_r - Loga|/|Loga| < (a-1)/r$.

On the other hand if $a < 1$, we have $Loga < 0$ and $\xi^{1/r} > a^{1/r}$, so

$$\begin{aligned} \frac{a_r - Loga}{Loga} &= \xi^{1/r} - 1 > a^{1/r} - 1 \\ &= \frac{a-1}{1 + a^{1/r} + a^{2/r} + \ldots + a^{(r-1)/r}}, \end{aligned} \quad (17)$$

and indeed $(a_r - Loga)/Loga > (a-1)/r$. Thus $|a_r - Loga|/|Loga| < (1 - a)/r$. □

The sequence $\{a_r\}$ can be shown to be linearly convergent, this can be seen by estimating the limit $\lim_{r\to\infty} \frac{|e_{r+1}|}{|e_r|}$where $e_r = a_r - Loga$..

The next theorem generalizes the previous result to the corresponding matrix case. The theorem is valid for a wider class than those of positive definite matrices.

Theorem 12. *Let $A \in C^{n\times n}$ be a nonsingular diagonalizable matrix with $\sigma(A) \subset R_+^*$. Then, for $A = Sdiag(\lambda_1, \lambda_2, \ldots, \lambda_n)S^{-1}$, the principal logarithm $LogA$ is given by*

$$LogA = \lim_{r\to\infty} r(A^{1/r} - I).$$

Moreover, $\|A_r - LogA\| \leq \kappa(S) \max_{1\leq i\leq n} \frac{|\lambda_i - 1|}{r} . \max_{1\leq i\leq n} |Log\lambda_i|$.

Proof. A is diagonalizable, then there exists a nonsingular matrix S such that $A = Sdiag(\lambda_1, \ldots, \lambda_n)S^{-1}$, $\lambda_i \in R^+$ $1 \leq i \leq n$. Let $A^{1/r}$ denote the pricipal rth root of A then $r(A^{1/r} - I) = Sdiag(r(\lambda_1 - 1), \ldots, r(\lambda_n - 1))S^{-1}$. Applying Theorem 11 will complete the proof. □

For the case of a positive definite matrix A, Theorems 11 and 12 imply the following result.

Theorem 13. *Let $A \in C^{n\times n}$ be a positive definite matrix A. $LogA$ is realized as the limit of monotone mutually commuting decreasing sequence $\{A_r\}$, where $A_r = r(A^{1/r} - I)$. Furthermore for any unitarily invariant norm*

$$\frac{\|A_r - LogA\|}{\|LogA\|} \leq \frac{\rho(A)}{r}.$$

The previous results can be extended to a general matrix $A \in C^{n\times n}$. Firstly it is easy to check that for $\lambda \in C$, $\lambda \in D_\pi$, then $r(\lambda^{1/r} - 1)$ converges to $Log\lambda$.

Lemma 5. *Let $\lambda \in D_\pi$, then $Log\lambda = \lim_{r\to\infty} r(\lambda^{1/r} - 1)$.*

The next theorem extends our results to any nonsingular matrix $A \in C^{n\times n}$ with spectrum $\sigma(A) \in D_\pi$.

Theorem 14. *Let $A \in C^{n\times n}$, $\sigma(A) \subset D_\pi$, then $LogA = \lim_{r\to\infty} r(A^{1/r} - 1)$.*

Proof. Let $A = SJS^{-1} = S\ diag(J_{m_1}(\lambda_1), \dots, J_{m_p}(\lambda_p))\ S^{-1}$ be the Jordan canonical form of A, then

$$r(A^{1/r} - I) = S\ diag(r(J_{m_1}^{1/r}(\lambda_1) - I_{m_1}), \dots, r(J_{m_p}^{1/r}(\lambda_p) - I_{m_p}))\ S^{-1} \tag{18}$$

where I_{m_k} is an m_k-by-m_k identity matrix. Now for the Jordan block $J_{m_k}(\lambda_k)$, the matrix $r(J_{m_k}^{1/r}(\lambda_k) - I_{m_k})$ becomes

$$\begin{bmatrix} r(\lambda_k^{1/r}-1) & \lambda_k^{\frac{1}{r}-1} & \frac{(\frac{1}{r}-1)\lambda_k^{\frac{1}{r}-2}}{2!} & \dots & \frac{(\frac{1}{r}-1)\dots(\frac{1}{r}-(m_k-2))\lambda_k^{\frac{1}{r}-(m_k-1)}}{(m_k-1)!} \\ 0 & r(\lambda_k^{1/r}-1) & \lambda_k^{\frac{1}{r}-1} & \dots & \frac{(\frac{1}{r}-1)\dots(\frac{1}{r}-(m_k-3))\lambda_k^{\frac{1}{r}-(m_k-2)}}{(m_k-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_k^{\frac{1}{r}-1} \\ 0 & 0 & 0 & \dots & r(\lambda_k^{1/r}-1) \end{bmatrix},$$

by taking the limit as $r \to \infty$, we get $\lim_{r\to\infty} r(J_{m_k}^{1/r}(\lambda_k) - I_{m_k}) = Log(J_{m_k}(\lambda_k))$. Then from (18) we conclude that

$$\begin{aligned} \lim_{r\to\infty} r(A^{1/r} - I) &= S\ diag(Log(J_{m_1}(\lambda_1)), \dots, Log(J_{m_p}(\lambda_p)))\ S^{-1} \\ &= S\ LogJ\ S^{-1} = LogA. \quad \square \end{aligned}$$

One shortcome of this iterative technique (from the computational point of view) is that the convergence is slow. There are well known techniques for the remedy of this shortcome [?]. We will not pursue this point here any further. Another shortcoming of this algorithm is the fact that the formula $A_r = r(A^{1/r} - I)$ suffers cancellation for large r. One way for overcoming this problem is by using the fact that

$$2^r(A^{1/2^r} - I) = 2^r(A^{1/2^{r+1}} - I)(A^{1/2^{r+1}} + I),$$

that is, $A_{r+1} = 2A_r(A^{1/2^{r+1}} + I)^{-1}$.

References

1. R. C. Allen and S. A. Pruess, *An analysis of an inverse problem in ordinary differential equations*, SIAM J. Sci. Statist. Comput., **2**(1981), 176-185.
2. A. Bjorck and S. Hammarling, *A Schur method for the square root of a matrix*, Linear Algebra Appl., **52**(1983), 127-140.

3. W. J. Culver, *it An analytic theory of modeling for a class of minimal energy control systems*, SIAM journal control, **2**(1964), 267-294.
4. G. H. Golub and C. V. Loan, *Matrix Computations*, The Johns Hopkins University Press., Baltimore, 1989.
5. R. Grone, C. R. Johnson and E. M. Sa, *Normal matrices*, Linear Algebra Appl., **87**(1987), 213-225.
6. N. J. Higham, *Computing real square roots of a real matrix*, Linear Algebra Appl., **88**(1987), 405-430.
7. N. J. Higham, *A new Sqrtm for MATLAP*, Numerical Analysis report 336, Manchester centre for computational mathematics, Manchester, January, England, 1999.
8. Sheung, Higham, Kenney, and Laub, *Approximating the logarithm of a matrix to specified accuracy*, SIAM J. Matrix Anal. Appl., **22**(2001), No.4, 1112-1125.
9. N. J. Higham, *Evaluating Pade approximants of the matrix logarithm*, SIAM J. Matrix Anal. Appl., **22**(2001), No.4, 1126-1135.
10. M. Hochbruch and C. Lubich, *On Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., **34**(1997), No.5, 1911-1925.
11. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press., Cambridge, 1991.
12. R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press., Cambridge, 1991.
13. W. D. Hoskins and D. J. Walton, *A faster, more stable method for computing the Pth roots of positive definte matrices*, Linear Algebra Appl., **26**(1979), 139-163.
14. C. Kenney and A. Laub, *Condition estimates for matrix functions*, SIAM J. Matrix Anal. Appl., **10**(1989), No.2, 191-209.
15. P. Lancaster, *Theory of Matrices*, Academic Press., New York, 1969.
16. G. R. Lindfield and J. E. Penny, *Microcomputers in Numerical Analysis*, John Wiley and Sons, New York, 1989.
17. R. Mathias, *Evaluating the Frechet derivative of the matrix exponential*, Numerische Mathematik, **63**(1992), 213-226.
18. R. Mathias, *Approximation of matrix valued functions*, SIAM J. Matrix Anal. Appl., **14**(1993), No.4, 1061-1063.
19. C. Moler and C. V. Loan, *Nineteen dubious ways to compute the exponential of a matrix*, SIAM Rev., **20**(1978), No.4, 801-834.
20. S. Puthenpura and N. Sinha, *Transformation of continuous-time model of a linear multivariable system from its discrete-time model*, Electronic Letters, **20**(1984), 737-738.
21. R. F. Rinehart, *The equivalence of definitions of a matric function*, Amer. Math. Monthly, *62*(1955), 395-414.
22. N. Sherif, *On the computation of a matrix inverse square root*, Computing, **46**(1989), 295-305.
23. N. Sherif, *On the computation of roots of a unitary matrix*, The Arabian J. Sc. Enging, **16**(1991), No.3, 427-434.
24. B. Singer and S. Spilerman, *The representation of social processes by Markov models*, Amer. J. Sociology, **82**(1976), 1-54.
25. N. Sinha and G. Lastman, *Transformation algorithm for identification of continuous-time multivariable systems from discrete data*, Electronic Letters, **17**(1981), 779-780.
26. R. C. Ward, *Numerical computation of the matrix exponential with accuracy estimate*, SIAM J. Numer. Anal. Appl., **14**(1977), No.4, 600-610.
27. E. M. Wermuth, *Two remarks on matrix exponentials*, Linear Algebra Appl., **117**(1989), 127-132.

Nagwa Sherif: received her B.SC from Cairo University (Egypt), M.SC and PhD at McMaster University (Canada). She occupied the position of assistant professor in King Saud University and Kuwait University. In 1991 she joined Qatar University as associate professor. Since 1995 she has been at the Suez Canal University. Her research interests focus on invariant subspace theory and related differential geometry approach. Also her research interests include the field of computational linear algebra.

Department of Mathematics, Suez Canal University, Ismailia 22541, Egypt
e-mail: nasherif@yahoo.com

Ehab Morsy: was born in El-Beheira, Egypt, On January 8, 1974. He received the B.A. and M.E. degrees from Suez Canal University, Egypt, in 1996, and in 2002, respectively. He is a PhD student in the Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University. His research interests include graph algorithms and network routing problems.

Department of Mathematics, Faculty of Science, Suez Canal University, Ismailia 22541, Egypt
e-mail: ehabmorsy@gmail.com