LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER Linear Algebra and its Applications 330 (2001) 31–42

www.elsevier.com/locate/laa

Theoretical and numerical considerations about Padé approximants for the matrix logarithm[☆]

J.R. Cardoso [a,*], F. Silva Leite [b,1]

[a] *Instituto Superior de Engenharia de Coimbra, Quinta da Nora, 3030 Coimbra, Portugal*
[b] *Departamento de Matemática, Universidade de Coimbra, 3000 Coimbra, Portugal*

Received 22 July 1999; accepted 14 November 2000

Submitted by G. de Oliveira

Abstract

We show that for a vast class of matrix Lie groups, which includes the orthogonal and the symplectic, diagonal Padé approximants of $\log((1+x)/(1-x))$ are structure preserving. The conditioning of these approximants is analyzed. We also present a new algorithm for the Briggs–Padé method, based on a strategy for reducing the number of square roots in the inverse scaling and squaring procedure. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: P-orthogonal groups; Matrix logarithms; Padé approximants; Condition number

1. Introduction

In recent years there has been an increasing interest in computing real logarithms of real matrices. The most significant papers in this area include [3,6,8,9,13–15].

The need to find logarithms of matrices arises in many areas of engineering and control theory. The above-cited references list a number of possible applications. In addition to these, we mention a geometric scheme for generalizing Bézier curves to non-Euclidean spaces, which was the main reason for our interest in this area. Crouch et al. [4,5] presented a generalized De Casteljau algorithm to find polynomial splines on Riemannian manifolds. For the particular situation when this manifold is

☆ Work supported in part by ISR and research network contract ERB FMRXCT-970137.

* Corresponding author.

E-mail addresses: jocar@isec.pt (J.R. Cardoso), fleite@mat.uc.pt (F. Silva Leite).

1 Fax: +351-39-832568.

0024-3795/01/$ - see front matter © 2001 Elsevier Science Inc. All rights reserved.

PII: S 0 0 2 4 - 3 7 9 5 (0 1) 0 0 2 5 1 - 8

a Lie group of matrices, the implementation of this algorithm requires successive computations of matrix exponentials and logarithms. In turn, the theory of splines on Lie groups finds applications in robotics path planning and in air traffic control.

It is well known that the exponential mapping on a Lie group G is a homeomorphism from a neighborhood of $0 \in L$ (the Lie algebra of G) into a neighborhood of the identity in G. So, if we are restricted to this neighborhood, any matrix in G has a logarithm that belongs to L.

One important aspect of the problem of computing logarithms of matrices belonging to G is to consider algorithms which are structure preserving. Dieci [6] showed that the diagonal Padé approximants method for computing the logarithm of orthogonal and symplectic matrices preserves the structure. Inspired by this work, we extend some of his results to a vaster class of Lie groups that includes the orthogonal and the symplectic as particular cases.

The Briggs–Padé method, which is a combination of the Padé approximants method with an inverse scaling and squaring procedure, computes the principal logarithm of an invertible matrix with no eigenvalues on the negative real axis, by taking successive square roots until the resulting matrix is so close to the identity that the Padé approximants of the logarithmic funtions become accurate (see [14,15] and also [6,8]). The logarithm of the original matrix is then recovered using the identity $\log(T) = 2^k \log(T^{1/2^k})$. Usually, the Padé approximants method used to approximate the logarithm is related to the function $\log(1-x)$. In this paper we use instead Padé approximants associated with $\log((1+x)/(1-x))$, due to the advantages which result from simpler expression and smaller condition numbers for the latter.

It is well known that taking a great number of square roots may lead to a loss of precision in the computed result (see [8,15]). Although in theory the precision of the result increases with the number of square roots, this may not happen when working on finite precision. In this paper we also present a strategy that guarantees some desired precision but avoids taking unnecessary square roots.

The organization of the paper is as follows. In Section 2 we make some considerations about the real matrix logarithm and present the class of P-orthogonal matrix Lie groups that will be used in the first part of the paper. In Section 3 we prove that the Briggs–Padé method, to approximate the principal logarithm of a P-orthogonal matrix, is structure preserving if using diagonal Padé approximants. We also comment on difficulties with the numerical implementation of this method. While these two sections refer to a particular class of matrix Lie groups, the results in the rest of the paper are valid for general invertible matrices without negative eigenvalues. In Section 4, we compare the condition numbers of diagonal Padé approximants of the functions $\log((1+x)/(1-x))$ and $\log(1-x)$, to conclude that the former has advantages over the latter. We consider only diagonal Padé approximants with even degree, but the results can be easily adapted to the odd case. Finally, in Section 5, we improve the usual upper bound for the absolute error estimate of the Padé approximants for the logarithm. This is then used to reduce the number of square roots in the Briggs–Padé method. As a consequence, we propose a new algorithm for this method and make comments on the implementation based on some numerical examples.

2. Preliminaries

Let $gl(n, \mathbb{R})$ be the set of all $n \times n$ matrices with real entries and $GL(n, \mathbb{R})$ the general linear Lie group, consisting of all invertible matrices in $gl(n, \mathbb{R})$.

Given a matrix $T \in GL(n, \mathbb{R})$, all solutions (not necessarily real) of the matrix equation $e^X = T$ are called *logarithms* of T. However, if the spectrum of T, denoted by $\sigma(T)$, does not intersect $\mathbb{R}_0^-$, then T has a unique real logarithm whose spectrum lies in the strip $\{z \in \mathbb{C} : -\pi < \mathrm{Im}(z) < \pi\}$. This logarithm is called the *principal logarithm* of T and will be denoted by $\log(T)$. See, for instance, [12] for details and properties of the principal logarithm, including those listed below. Hereafter the word principal will be sometimes omitted, since this is the only logarithm to be considered in this paper.

When T satisfies $\sigma(T) \cap \mathbb{R}_0^- = \emptyset$, the principal logarithm of T is a primary matrix function and so enjoys some nice properties. See [12] for details. Among the properties of the principal logarithm, which result from the fact that it is a primary matrix function, we emphasize the following:

$$\log(T^{\mathrm{T}}) = (\log(T))^{\mathrm{T}},$$

$$\log(T^{-1}) = -\log(T) \tag{1}$$

and

$$\log(STS^{-1}) = S\log(T)S^{-1} \quad \forall S \in GL(n, \mathbb{R}).$$

(The superscript T denotes the matrix transpose.)

For any $n \times n$ orthogonal matrix P (i.e. $P^{\mathrm{T}} = P^{-1}$) define the following Lie subgroup of $GL(n, \mathbb{R})$,

$$G = \left\{T \in GL(n, \mathbb{R}) : T^{\mathrm{T}}PT = P\right\},$$

whose Lie algebra, equipped with the commutator as Lie bracket is:

$$L = \left\{X \in gl(n, \mathbb{R}) : X^{\mathrm{T}}P = -PX\right\}.$$

From now on we will refer the matrices in G as *P-orthogonal* and the matrices in L as *P-skew-symmetric*.

The particular cases when

$$P = I_n \quad \text{and} \quad P = \begin{bmatrix} 0 & I_k \\ -I_k & 0 \end{bmatrix},$$

with $2k = n$, correspond to the orthogonal/skew-symmetric and symplectic/Hamiltonian cases, respectively, studied in [6,7]. Among the infinitely many P-orthogonal

Lie groups, obtained by letting P run over the set of orthogonal matrices, we also point out the case $G = O(p, q)$, $p + q = n$, corresponding to the choice

$$P = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}.$$

The Lorentz group of all coordinate transformations in Minkowski space-time, denoted by $O(1, 3)$, is a particular case of the above and plays a very important role in physics.

While the exponential of a P-skew-symmetric matrix is P-orthogonal, a real logarithm of a P-orthogonal matrix may not be P-skew-symmetric. However, providing that $\sigma(T) \cap \mathbb{R}_0^- = \emptyset$, the principal logarithm is always P-skew-symmetric. In fact, since T is P-orthogonal, then $T^{-1} = P^{-1}T^{\mathrm{T}}P$ and using (1), it follows immediately that

$$-\log(T) = P^{-1}(\log(T))^{\mathrm{T}}P,$$

which implies that $\log(T)$ is P-skew-symmetric.

3. Structure preserving Padé approximants for the logarithm of *P*-orthogonal matrices

We refer to [1,2] for more details concerning the general theory of Padé approximants.

It is well known that diagonal Padé approximants $R_{mm}(A)$ of the matrix function $f(A) = \log(I - A)$ may be used to approximate the principal logarithm of any matrix $T := I - A$, with $\|I - T\| < 1$ (see, for instance, [13]).

According to Theorem 1.5.2 in [2],

$$R_{mm}(A) = S_{mm}(B), \tag{2}$$

where $S_{mm}(B)$ denotes the Padé approximant of $g(B) = \log\,[(I + B)(I - B)^{-1}]$, with $B = A(A - 2I)^{-1}$.

It turns out that some important simplifications take place if one uses $S_{mm}(B)$ instead of $R_{mm}(A)$, to approximate the principal logarithm of $T = I - A$. For this reason we proceed our work with the former. One particular aspect of this simplification, which becomes clearer from expressing $\log((1 + x)/(1 - x))$ as the continued fraction expansion:

$$g(x) = \log\left(\frac{1+x}{1-x}\right) = \cfrac{2x}{1 - \cfrac{x^2}{3 - \cfrac{4x^2}{5 - \cdots \cfrac{k^2x^2}{2k+1-\cdots}}}}, \tag{3}$$

is that the numerator of S_{mm} has only odd powers in x, while the denominator has only even powers.

Some analysis (see [2, Chapter 4]) shows that the sequence $(S_{mm}(B))_{m\in\mathbb{N}}$ converges to $g(B)$, for all matrices B whose spectrum does not intersect the set $\{\lambda \in \mathbb{R} : |\lambda| \geqslant 1\}$, that is, when the matrix T satisfies $\sigma(T) \cap \mathbb{R}_0^- = \emptyset$.

In the next theorem, we show that diagonal Padé approximants of any order, to approximate the logarithm of a matrix in the Lie group of P-orthogonal matrices are structure preserving, in the sense that they always produce a matrix belonging to the corresponding Lie algebra.

Theorem 3.1. *If T is P-orthogonal and $\sigma(T) \cap \mathbb{R}_0^- = \emptyset$, then $S_{mm}(B)$, where $B = (T - I)(T + I)^{-1}$, is P-skew-symmetric.*

Proof. We first show that if T is P-orthogonal, that is, $T^{\mathrm{T}} = PT^{-1}P^{-1}$, then B is P-skew-symmetric. Indeed,

$$\begin{aligned} B^{\mathrm{T}} &= (T^{\mathrm{T}} - I)(T^{\mathrm{T}} + I)^{-1} = (PT^{-1}P^{-1} - I)(PT^{-1}P^{-1} + I)^{-1} \\ &= P(T^{-1} - I)(T^{-1} + I)^{-1}P^{-1} \\ &= P(T^{-1} - I)TT^{-1}(T^{-1} + I)^{-1}P^{-1} \\ &= -P(T - I)(T + I)^{-1}P^{-1} = -PBP^{-1}. \end{aligned}$$

To conclude the proof it is enough to observe that $S_{mm}(B)$ is a primary matrix function and $S_{mm}(B) = -S_{mm}(-B)$. $\square$

Due to relationship (2), this result generalizes Theorem 2.2 in [6], which is valid for the orthogonal and the symplectic groups only.

Although the previous theorem solves, in theory, the problem of approximating the matrix logarithm of P-orthogonal matrices, in practice it does not. In fact, if a matrix is too far from the identity, it is known that one needs to use high-order Padé approximants to compute the matrix logarithm, which is not an efficient procedure. In such a case, one possible alternative is to use the Briggs–Padé method, which consists in combining an inverse scaling and squaring process with the Padé approximants method. Basically one takes sucessive square roots of the given matrix T, until the resulting matrix is in a small neighborhood of the identity, and then recovers the logarithm of the original matrix through the identity

$$\log(T) = 2^k \log(T^{1/2^k}).$$

It is also true, and easy to prove, that if T is P-orthogonal and $\sigma(T) \cap \mathbb{R}_0^- = \emptyset$, then $T^{1/2}$ is also P-orthogonal, where $T^{1/2}$ is the unique square root whose eigenvalues satisfy

$$-\pi/2 < \arg(\lambda) < \pi/2,$$

So, according to these observations, it seems that the Briggs–Padé method is the key to approximate the principal logarithm of a P-orthogonal matrix. Unfortunately, this is only true from a theoretical point of view, since no numerical methods to compute matrix square roots are known to be structure preserving. In order to minimize this difficulty, we propose, in Section 5, a procedure to reduce the number of square roots in the Briggs–Padé method.

In Section 4, we compare the conditioning of S_{mm} with that of R_{mm}, in order to emphasize another advantage in using the former Padé approximant instead of the classical one.

4. Conditioning of S_{mm}

It is well known that the conditioning of a Padé approximant depends on the conditioning of its denominator (see [8,13]).

In this section, we give an upper bound for the condition number of the denominator of $S_{mm}(B)$, which is much better than the one presented in [13] for the classical Padé approximants $R_{mm}(A)$, at least for the values of m pertinent to the present work. For the sake of simplicity, we consider Padé approximants of even order only. We first recall that the condition number for any invertible matrix M is defined by

$$\text{cond}\,(M) = \|M\|\,\|M^{-1}\|.$$

Theorem 4.1. *Let l be a positive integer. If Q_{2l} is the denominator of $S_{2l,2l}$ and B is a matrix satisfying $\|B^2\| < 1$, then*

$$\text{cond}(Q_{2l}(B)) \leqslant \frac{q_l(-\|B^2\|)}{q_l(\|B^2\|)},$$

where q_l is a polynomial of degree l which satisfies $q_l(x^2) = Q_{2l}(x)$.

Proof. To understand the arguments we have to recall that the function $\log((1+x)/(1-x))$ may be written in terms of a hypergeometric series in the following way:

$$\log\left(\frac{1+x}{1-x}\right) = 2x\ {}_2F_1\left(\frac{1}{2}, 1, \frac{3}{2}; x^2\right). \tag{4}$$

Now, the denominator of the Padé approximants of a power series can be written in terms of orthogonal polynomials [1, Chapter 7], which in turn allows us to locate the poles of these approximants. For the particular case above, it follows from arguments in [1, p. 87], that $Q_{2l}(x)$ has real simple roots in the interval $]-\infty, -1[\cup]1, +\infty[$. So, $q_l(y)$ has real simple roots lying in the interval $]1, +\infty[$. But since $q_l(y)$ is also the denominator of the $(l-1, l)$ Padé approximant of ${}_2F_1\left(\frac{1}{2}, 1, \frac{3}{2}; y\right)$, we can apply Lemma 3 in [13] to write the following inequality:

$$\text{cond}\ (Q_{2l}(B)) = \text{cond}\ \left(q_l(B^2)\right) \leqslant \frac{q_l\left(-\|B^2\|\right)}{q_l\left(\|B^2\|\right)},$$

which completes the proof. □

Comparing the upper bound given in the previous theorem with that given in [13] for the conditioning of the denominator of the Padé approximant $R_{2l,2l}(A)$, our tests for the cases $l = 1, \ldots, 5$ show that the conditioning of the Padé approximant for $\log((1+x)/(1-x))$ is much better than that of $\log(1-x)$. For the case $l = 4$, we have

$$q_4(y) = 1 - \frac{28}{15}y + \frac{14}{13}y^2 - \frac{28}{143}y^3 + \frac{7}{1287}y^4,$$

which satisfies $q_4(x^2) = Q_8(x)$. Assuming $\|A\| = \|I - T\| = 0.9$ and since, according to [8], $\|(T + I)^{-1}\| \leqslant (2 - \|A\|)^{-1}$, it follows that

$$\|B\| = \|(T - I)(T + I)^{-1}\| \leqslant \frac{\|A\|}{2 - \|A\|} \quad \text{and} \quad \|B^2\| \leqslant \|B\|\|B\| \leqslant 0.669.$$

Now, a few calculations show that

$$\operatorname{cond}(Q_8(B)) \leqslant 15.8867.$$

On the other hand, if F_8 denotes the denominator of the $(8, 8)$ Padé approximant of $\log(1-x)$, it has been proved in [13] that

$$\operatorname{cond}(F_8(A)) \leqslant \frac{F_8(-\|A\|)}{F_8(\|A\|)}.$$

So, under the above assumptions, we have

$$\operatorname{cond}(F_8(A)) \leqslant 11045.8.$$

5. A new algorithm for the Briggs–Padé method

To approximate $\log(I - A)$, the usual practice is to determine a number of square roots of $T = I - A$ until the following condition is satisfied:

$$\|I - T^{1/2^k}\| \leqslant \epsilon < 1. \tag{5}$$

This, in turn, is related with the inequality

$$\|R_{mm}(A) - \log(I - A)\| \leqslant R_{mm}(\|A\|) - \log(1 - \|A\|), \tag{6}$$

that measures the error of the Padé approximants of $\log(I - A)$, when $\|A\| < 1$ [13, Corollary 4]. However, condition (5) is not necessary to guarantee high precision of the Padé approximants method. For instance, as pointed out in [9], the requirement that all diagonal blocks in the real Schur form of $T^{1/2^k}$ are close to the identity, performs better than condition (5).

In this section we take a somewhat similar route and show that full accuracy may be obtained even when $\|I - T^{1/2^k}\| \gg 1$. This will be a consequence of an improvement in the error bounds given by (6).

Theorem 5.1. *Suppose that*

$$H(y) = {}_2F_1\left(\frac{1}{2}, 1, \frac{3}{2}; y\right) \quad \textit{and} \quad t_{l-1,l}(y)$$

is the associated $(l-1, l)$ *Padé approximant. If B is a matrix such that* $\|B^2\| < 1$, *then*

$$\begin{aligned} &\|\log[(I+B)(I-B)^{-1}] - S_{2l,2l}(B)\| \\ &\quad \leqslant 2\|B\|[H(\|B^2\|) - t_{l-1,l}(\|B^2\|)]. \end{aligned} \tag{7}$$

Proof. Replacing X by B^2 in Corollary 4 of [13], it follows that

$$\|H(B^2) - t_{l-1,l}(B^2)\| \leqslant H(\|B^2\|) - t_{l-1,l}(\|B^2\|).$$

This inequality together with

$$H(B^2) = (2B)^{-1}\log[(I+B)(I-B)^{-1}]$$

and

$$S_{2l,2l}(B) = 2Bt_{l-1,l}(B^2),$$

yields

$$\begin{aligned} &\|\log[(I+B)(I-B)^{-1}] - S_{2l,2l}(B)\| \\ &\quad = \|2BH(B^2) - 2Bt_{l-1,l}(B^2)\| \\ &\quad \leqslant 2\|B\|[H(\|B^2\|) - t_{l-1,l}(\|B^2\|)]. \qquad \square \end{aligned}$$

Now, we are going to compare the upper bound given in (7) with the one in (6).

First we observe that since $A = I - T$ and $B = A(A-2I)^{-1} = (T-I)(T+I)^{-1}$, then

$$\|\log[(I+B)(I-B)^{-1}] - S_{2l,2l}(B)\| = \|R_{2l,2l}(A) - \log(I-A)\|.$$

We now assume that $\|A\| \leqslant 1$ and set $u = 2\|B\|[H(\|B^2\|) - t_{l-1,l}(\|B^2\|)]$ and $v = R_{2l,2l}(\|A\|) - \log(1-\|A\|)$. From the positivity of the error expansion coefficients, as observed in [13], from Lemma 1 in [13], and (2), and since

$$\|B\| \leqslant \frac{\|A\|}{2-\|A\|},$$

it follows that

$$\begin{aligned} u &\leqslant 2\|B\|\,[H(\|B\|^2) - t_{l-1,l}(\|B\|^2)] \\ &\leqslant \log\left(\frac{1+\|B\|}{1-\|B\|}\right) - S_{2l,2l}(\|B\|) \\ &\leqslant v. \end{aligned}$$

Using u instead of v is clearly advantageous when $\|B^2\| < 1$ but $\|A\| \gg 1$, due to a reduction in the number of square roots.

To summarize, we propose the following algorithm for the Briggs–Padé method.

Algorithm. T is real such that $\sigma(T) \cap \mathbb{R}_0^- = \emptyset$, ε is a given tolerance and l is a positive integer. Set

$$B_j := (T^{1/2^j} - I)(T^{1/2^j} + I)^{-1}$$
$$u_j := 2\|B_j\|\,[H(\|B_j^2\|) - t_{l-1,l}(\|B_j^2\|)], \quad j \in \mathbb{N}.$$

1. Compute k successive square roots of T until $\|B_k^2\| < 1$ and $u_k < \varepsilon$;
2. Compute $S_{2l,2l}(B_k)$;
3. Approximate $\log(T)$ using the relations

$$\log(T) = 2^k \log(T^{1/2^k}) \approx 2^k S_{2l,2l}(B_k).$$

Next we compare numerically three different algorithms to compute the matrix logarithm.

- Algorithm 1 is the one we have just presented.
- Algorithm 2 is the classical one, where v is used instead of u to estimate the error and $R_{2l,2l}$ replaces $S_{2l,2l}$.
- Algorithm 3 is Algorithm 1 with $S_{2l,2l}$ replaced by $R_{2l,2l}$. This avoids difficulties that may arise in the computation of $(T^{1/2^k} + I)^{-1}$ in step 2.

We next present four tests with ill-conditioned full matrices.

- Test 1 reports on the matrix

  $$T = \frac{e^\alpha}{2}\begin{bmatrix} 2+\beta & \beta \\ -\beta & 2-\beta \end{bmatrix},$$

  for the values $\alpha = 0.1$, $\beta = 10^6$. In this case

  $$\log(T) = \frac{1}{2}\begin{bmatrix} 2\alpha+\beta & \beta \\ -\beta & 2\alpha-\beta \end{bmatrix}, \quad T^{1/2^k} = \frac{e^{\alpha/2^k}}{2}\begin{bmatrix} 2+\frac{\beta}{2^k} & \frac{\beta}{2^k} \\ -\frac{\beta}{2^k} & 2-\frac{\beta}{2^k} \end{bmatrix},$$

  and $\mathrm{cond}(T) \approx 10.0 \times 10^{11}$. (The matrix T results from the one in Example 2 of [15] by an orthogonal similarity transformation.)
- Test 2 refers to a matrix obtained from the example in [8, p. 591] by orthogonal similarity transformation. For the exact logarithm we consider the one computed using the formula of Theorem 11.1.3 in [10] valid for upper triangular matrices, and then recover the original logarithm through the corresponding orthogonal similarity. The square roots in the inverse scaling and squaring procedure were computed using a similar idea.The condition number is $\mathrm{cond}(T) \approx 1.1 \times 10^{15}$.
- Test 3 is for the matrix $A = T\tilde{A}T^{-1}$, where T and $\tilde{A}$ are as in Example 6 of [15], with $\alpha = 10$ and $\theta = 3.14159265 \approx \pi$. The exact logarithm is known and the square roots were computed using

$$A^{1/2^k} = T \begin{bmatrix} e^{\alpha/2^k} & 0 & 0 & 0 \\ 0 & e^{-\alpha/2^k} & 0 & 0 \\ 0 & 0 & \cos(\frac{\theta}{2^k}) & -\sin(\frac{\theta}{2^k}) \\ 0 & 0 & \sin(\frac{\theta}{2^k}) & \cos(\frac{\theta}{2^k}) \end{bmatrix} T^{-1};$$

cond$(A) \approx 6.0 \times 10^{11}$.

- Test 4 is for the matrix in *gallery* (3) of Matlab, for which the exact value of the logarithm is unknown and the square roots were computed using the algorithm proposed in [11]; cond$(T) \approx 2.8 \times 10^5$.

These tests were performed in Matlab (with relative machine epsilon $\epsilon \approx 2.2 \times 10^{-16}$) on a Pentium II. We used the Frobenius norm and $(8, 8)$ Padé approximants. We considered in all cases the tolerance $\varepsilon = \|T\| \times 10^{-16}$, which seemed to be a good choice in order to obtain maximum accuracy in the final result.

For Algorithm 1,

$$S_{88}(x) = \frac{-2x(15\,159x^6 - 147\,455x^4 + 345\,345x^2 - 225\,225)}{35(35x^8 - 1260x^6 + 6930x^4 - 12\,012x^2 + 6435)}$$

and $u = 2\|B\|\,[H(\|B^2\|) - t_{34}(\|B^2\|)]$, where

$$H(x) = \frac{1}{2\sqrt{x}} \log\left(\frac{1+\sqrt{x}}{1-\sqrt{x}}\right) \quad \text{and} \quad t_{34}(x) = \frac{1}{2\sqrt{x}} S_{88}(\sqrt{x}).$$

For Algorithm 2, $v = R_{88}(\|A\|) - \log(1 - \|A\|)$ and the expression of R_{88} may be found in [13] or [14].

In Table 1, k denotes the minimum number of necessary square roots in step 1, $n = \|I - T^{1/2^k}\|$ and e measures the relative error. In tests 1–3,

$$e = \frac{\|\bar{X} - \log T\|}{\|\log T\|},$$

where $\bar{X}$ is the approximation obtained for the logarithm and $\log(T)$ is the exact value given as above. For test 4, since the exact value for the logarithm is unknown, the error was estimated using

$$e = \frac{\|e^{\bar{X}} - T\|}{\|T\|}.$$

Based on the information presented in Table 1, we observe that the result for test 1 in Algorithm 1 is somewhat surprising when compared with Algorithm 2. After 10 square roots we have got $u_{10} \approx 10^{-13}$. And u_{11} was already the machine zero, which means that, at least theoretically, no more square roots are needed in the inverse scaling and squaring procedure. However, in practice, we observed that taking more square roots increases the accuracy. In our computer full accuracy was obtained when $k = 20$, and started decreasing for $k > 20$. It is unclear why this situation has occurred. We think that this problem is due to the computation of the inverse of

Table 1

	Algorithm 1			Algorithm 2			Algorithm 3
	k	n	e	k	n	e	e
Test 1	10	976.7	2.1×10^{-12}	21	0.48	3.0×10^{-16}	4.6×10^{-7}
Test 2	20	4.8×10^{3}	3.8×10^{-11}	34	0.29	5.3×10^{-11}	4.0×10^{-6}
Test 3	6	12.2	5.3×10^{-15}	11	0.38	3.8×10^{-14}	5.4×10^{-14}
Test 4	6	7.9	4.2×10^{-12}	11	0.25	7.3×10^{-11}	7.3×10^{-13}

$T^{1/2^k} + I$, which has a condition number of about 2.4×10^5. When $\|I - T^{1/2^k}\| \gg 1$, like in test 1, we have no guarantee that $T^{1/2^k} + I$ is well conditioned. Taking more square roots to bring T near the identity may be advantageous since $T^{1/2^k} + I$ becomes a well-conditioned matrix. We note that in all the other tests 2–4, this situation has not occurred, even though the matrix $T^{1/2^k} + I$ has a large condition number.

Comparing Algorithm 1 with Algorithm 3, which involves the same number of square roots, our tests showed that Algorithm 1 is slightly more accurate. The computation of $(T^{1/2^k} + I)^{-1}$ is not necessary but, in contrast, when $\|I - T^{1/2^k}\| \gg 1$, the condition number of the denominator of $R_{2l,2l}$, which needs to be inverted, is usually very large.

We have also tested several triangular matrices, including ill-conditioned cases such as the matrix in [15, Example 2] with $\alpha = 0.1$ and $\beta = 10^6$, and the matrix in [8, p. 591]. For these cases our numerical experiments showed that Algorithm 1 is more accurate than Algorithm 2, especially in ill-conditioned problems. This is due to a reduction in the number of square roots in our algorithm. For well-conditioned matrices, both algorithms revealed a similar accuracy. However, we note that for triangular matrices the strategy presented in [9] performs better than the algorithms tested here.

References

[1] G.A. Baker, Essentials of Padé Approximants, Academic Press, New York, 1975.
[2] G.A. Baker, P. Graves-Morris, Padé Approximants, Part I, Encyclopedia Math. 13, Addison-Wesley, Reading, MA, 1981.
[3] W.J. Culver, On the existence and uniqueness of the real logarithm of a matrix, Proc. Amer. Math. Soc. 17 (1966) 1146–1151.
[4] P. Crouch, G. Kun, F. Silva Leite, The De Casteljau algorithm on Lie groups and spheres, J. Dyn. Control Syst. 5 (3) (1999) 397–429.
[5] P. Crouch, G. Kun, F. Silva Leite, De Casteljau algorithm for cubic polynomials on the rotation group, in: Proceedings of 2nd Portuguese Conference on Automatic Control, vol. II, Coimbra, 9–11 September 1996, pp. 547–552.
[6] L. Dieci, Considerations on computing real logarithms of matrices, Hamiltonian logarithms, and skew-symmetric logarithms, Linear Algebra Appl. 244 (1996) 35–54.

[7] L. Dieci, Real Hamiltonian logarithm of a symplectic matrix, Linear Algebra Appl. 281 (1998) 227–246.
[8] L. Dieci, B. Morini, A. Papini, Computational techniques for real logarithms of matrices, SIAM J. Matrix Anal. Appl. 17 (3) (1996) 570–593.
[9] L. Dieci, A. Papini, Conditioning and Padé approximation of the logarithm of a matrix, SIAM J. Matrix Anal. Appl. 31 (3) (2000) 913–930.
[10] G. Golub, C. Van Loan, Matrix Computations, third ed., Johns Hopkins University Press, Baltimore, MD, 1996.
[11] N.J. Higham, A new sqrtm for Matlab, Numerical Analysis Report, vol. 336, University of Manchester, 1999.
[12] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.
[13] C. Kenney, A.J. Laub, Padé error estimates for the logarithm of a matrix, Int. J. Control 50 (3) (1989) 707–730.
[14] C. Kenney, A.J. Laub, Condition estimates for matrix functions, SIAM J. Matrix Anal. Appl. 10 (1989) 191–209.
[15] C. Kenney, A.J. Laub, A Schur–Frechet algorithm for computing the logarithm and exponential of a matrix, SIAM J. Matrix Anal. Appl. 19 (3) (1998) 640–663.