Evaluating the Frechet derivative of the matrix exponential*

Roy Mathias
Department of Mathematics, College of William and Mary, Williamsburg, VA 23187, USA

Received February 1, 1992

Summary. Let $M_n$ denote the space of $n \times n$ matrices. Given $X, Z \in M_n$ define

$$L(X, Z) = \int_0^1 e^{sX} Z e^{(1-s)X} ds,$$

and given a norm $\| \cdot \|$ on $M_n$ define

$$\| L(X, \cdot) \| = \max \{ \| L(X, Y) \| : \| Y \| \leq 1, Y \in M_n \}.$$

The quantity $\| L(X, \cdot) \|$ is important in the study of the sensitivity of the matrix exponential. Given an integration rule $R$ and a positive integer $m$ let $L_{R,m}(X, Z)$ denote the approximation to $L(X, Z)$ given by the composite rule $R$ applied with $m$ subintervals. We give bounds on $\| L(X, \cdot) - L_{R,m}(X, \cdot) \|$ that are valid for any unitarily invariant norm $\| \cdot \|$ and any integration rule $R$.

We show that in a particular situation, that arises in practice, the composite Simpson's rule has a better error bound than the composite trapezoidal rule, and that it can be evaluated with the same number of matrix multiplications.

We also note a symmetry property of the function $L(X, \cdot)$ and compare the performance of the power method and two variants of the Lanczos algorithm for estimating $\| L(X, \cdot) \|$.

Mathematics Subject Classification (1991): 15A12, 65F35, 65F30

1 Introduction

Let $M_n$ denote the space of $n \times n$ complex matrices. Given $X \in M_n$ we define

$$e^X \equiv \sum_{k=0}^{\infty} \frac{X^k}{k!}.$$ 

*Research supported by a Summer Research Grant from the College of William and Mary
A norm \( \| \cdot \| \) on \( M_n \) is \textit{unitarily invariant} if \( \| UAV \| = \| A \| \) for all \( A \in M_n \) and all unitary \( U, V \in M_n \). We will let \( \| \cdot \| \) denote a general unitarily invariant norm on \( M_n \) and use \( \| \cdot \|_2 \) to denote the spectral norm (i.e., \( \| X \| _2^2 = \rho(X^*X) \)). We will frequently use the fact that for any unitarily invariant norm \( \| \cdot \| \)

\[
\| AB \| \leq \| A \|_2 \| B \|.
\]

We will use \( \| \cdot \|_F \) to denote the Frobenius norm (i.e., \( \| X \| _F^2 = \text{tr}X^*X \)). It can be shown that the Fréchet derivative of the function \( e^X \) at \( X_0 \) in the direction \( Z \) is

\[
L(X_0, Z) = \frac{1}{0} e^{tx_0} Ze^{(1-t)x_0} dt.
\]

Either use the argument in [1, Sect. 10.14] or replace \( e^{tx} \) and \( e^{(1-t)x} \) in (1.2) by their series representations, integrate term by term, and note that the result is the same as the series representation of the derivative given in [5, (2.2)].

Given a linear operator \( T: M_n \rightarrow M_n \) and a norm \( \| \cdot \| \) on \( M_n \) we define the induced norm of \( T \) to be

\[
\| T \| = \max \{ \| T(X) \| : X \in M_n, \| X \| \leq 1 \}.
\]

We use the same symbol as no confusion will arise.

One is interested in \( \| L(X, \cdot) \| \) since it can be shown to be \( \kappa_{\exp}(X) \), the condition number (with respect to the norm \( \| \cdot \| \)) of the matrix exponential at \( X \), defined by

\[
\kappa_{\exp}(X) \equiv \lim_{\delta \rightarrow 0} \max_{E=\delta} \frac{\| e^X - e^{X+E} \|}{\delta}.
\]

In the case of the Frobenius norm Kenny and Laub [5] have shown how to estimate \( \| L(X, \cdot) \|_F \) using the power method. The only problem is that it is not known how to compute \( L(X, Z) \) except when \( X \) or \( Z \) is of some special form. One approach is to approximate the integral in (1.2) by the use of composite numerical integration rules.

In Sect. 2 we show how to compute the accuracy of these estimates and how to evaluate them efficiently. The composite trapezoidal rule was analyzed in [5]. We will show that the composite Simpson's rule can be evaluated using no more matrix multiplications than the trapezoidal rule and that it generally gives more accurate results.

The power method suggested by Kenny and Laub is satisfactory for condition estimation since in this context it is sufficient to estimate \( \| L(X, \cdot) \| \) to within an order of magnitude and one iteration of the power method is usually sufficient to ensure this (see the results in [5]). For a more exact estimate it would be more efficient to use the Lanczos algorithm to compute the largest singular value of the linear operator \( L(X, \cdot) \). In Sect. 3 we discuss the power method and correct a minor omission in [5]. In Sect. 4 we consider two Lanczos algorithms and in Sect. 5 we give numerical results and compare the power method and these two Lanczos algorithms.

An extensive bibliography on the computation and sensitivity of the matrix exponential is given in [5].

The rest of this section is devoted to definitions. Let \( R \) be a given integration rule on \([0, 1]\) and let \( m \) be a positive integer, then let \( L_{R,m}(X, Z) \) denote the