CONSTRUCTIVE THEORY OF FUNCTIONS'87

SOFIA, 1988

Numerical Computation of the Matrix Exponential

Guido Walz

In the following we present a new algorithm for the computation of the matrix exponential function. Applications of this algorithm to other matrix functions as well as a lot of numerical comparisons with other methods can be found in [5].

For $A \in \mathbb{C}^{n \times n}$ the *matrix exponential (function)* $\exp(A)$ is defined as

$$\exp(A) := I + \sum_{\nu=1}^{\infty} \frac{1}{\nu!} A^{\nu} \quad . \tag{1}$$

The function $\exp(A)$ is by (1) well-defined for all quadratic matrices A, because for each submultiplicative matrix norm the inequality

$$\|I + \sum_{\nu=1}^{\infty} \frac{1}{\nu!} A^{\nu}\| \le 1 + \sum_{\nu=1}^{\infty} \frac{1}{\nu!} \|A\|^{\nu} = \exp(\|A\|) < \infty$$

is valid.

So, what we *have* is a matrix function, which is well-defined for all quadratic matrices A; what we *need* is a numerical procedure which can compute the values of this function also for all quadratic matrices A.

In fact there exist a lot of methods for the approximation of $\exp(A)$; most of them can be found in the paper of C. MOLER and C. v.LOAN [1], another one is proposed by E. STICKEL [2]. However, all of these methods have some disadvantage: e.g. numerical instabilities, long computation time or applicability to only a small class of matrices. Therefore MOLER and v.LOAN had good reasons to give their paper [1] the title "Nineteen *Dubious* Ways to Compute the Exponential of a Matrix".

What we want to do in the following is to present a method which is not "dubious", what means, which is numerically very stable and there are no restrictions concerning the matrix A. To this end we define for all $n \in \mathbb{N}$ the matrix functions

$$\sigma_n(A) := (I + \frac{1}{n} A)^n \quad . \tag{2}$$

It is well-known that

$$\lim_{n \to \infty} \sigma_n(A) = \exp(A) \ ,$$

and in fact (2) is sometimes used for the approximation of exp(A); but because of the relation

$$\sigma_n(A) = \exp(A) + O(n^{-1}) \qquad \text{for } n \to \infty \tag{3}$$

the convergence of this sequence is very slow and using (2) would be no good approximation method. But in the following we will show how to enlarge the order of convergence of the σ_n.

First we need the following Lemma 1, which gives deeper insight into the structure of this convergence:

Lemma 1: For all $n \in \mathbb{N}$ the σ_n possess the expansion

$$\sigma_n(A) = \exp(A) + \sum_{\nu=1}^{\infty} \frac{1}{n^\nu} c_\nu(A) \tag{4}$$

with matrix functions $c_\nu(A)$, independent of n.

The proof of this lemma can be executed in exact analogy to the scalar case treated in [4].

Now we need the following Lemma 2, which represents a special case of a much greater theory concerning the approximation of scalar and matrix functions:

Lemma 2: Let $S_n(A)$ be a sequence of matrix functions such that for all $n \in \mathbb{N}$ the following relation holds:

$$S_n(A) = f(A) + \sum_{\nu=1}^{\infty} \frac{1}{n^{\rho_\nu}} c_\nu(A) \tag{5}$$

with matrix functions c_ν, $\nu \in \mathbb{N}$, and f, all independent of n, and a sequence $\{\rho_\nu\}$ of real numbers with

$$0 < \rho_1 < \rho_2 < \cdots \quad .$$

Then we can apply the following *linear elimination procedure:*

$$Y_i^{(0)} = S_{2^i}(A) \quad , \qquad i = 0(1)\infty$$

$$Y_i^{(k)} = Y_{i+1}^{(k-1)} + \frac{1}{2^{\rho_k} - 1} \cdot (Y_{i+1}^{(k-1)} - Y_i^{(k-1)}) \quad \begin{cases} k = 1(1)\infty \\ i = 0(1)\infty \end{cases}$$

and get for all i and k the asymptotic expansions

$$Y_i^{(k)} = f(A) + \sum_{\nu=k+1}^{\infty} c_\nu^{(k)}(A) \cdot 2^{-i \cdot \rho_\nu} \qquad \text{for } i \to \infty \ .$$

Combining these two lemmata leads us to the following

Algorithm for the numerical computation of the matrix exponential:

1. Choose a maximal index k_{max}
2. For $i = 0(1)k_{max}$ compute
   $$Y_i^{(0)} := \sigma_{2^i}(A)$$
3. For $k = 1(1)k_{max}$ and $i = 0(1)k_{max} - k$ compute
   $$Y_i^{(k)} := Y_{i+1}^{(k-1)} + \frac{1}{2^{\rho_k} - 1} \cdot (Y_{i+1}^{(k-1)} - Y_i^{(k-1)}) \ .$$
4. Use $Y_0^{(k_{max})}$ as approximation to $\exp(A)$.

This method can be applied to all quadratic matrices A; furthermore, it is highly linear (the only point, where matrix multiplications are needed, is the computation of the starting values $Y_i^{(0)}$), what results in a high numerical stability, see the examples given in [5]. Hence one can say that our algorithm is – in contrast to all others known until now – not a dubious one at all.

Finally we note that our method can also be applied to other matrix functions. All one needs is a sequence $\{S_n(A)\}$ which has the convergence structure given by (5). There exist methods for the *construction* of such sequences for given function $f(A)$ (cp. [5] and [3, 4] for the scalar case), but this will not be discussed here.

Refences

[1] Moler, C. and v.Loan, C.:
Nineteen Dubious Ways to Compute the Exponential of a Matrix
SIAM Review 20 (1978), S.801 – 836

[2] Stickel, E.:
Über die Berechnung elementarer Matrixfunktionen
Dissertation, Universität Ulm 1985

[3] Walz, G.:
Effiziente Berechnung spezieller Funktionen mittels asymptotischer Entwicklungen und Eliminationsprozeduren
Journal of Approximation Theory, to appear

[4] Walz, G.:
Asymptotische Entwicklungen
Research Report A, DFG project Me 66/5
Universität Mannheim 1987 (to appear)

[5] Walz, G.:
Computing the Matrix Exponential and Other Matrix Functions
Submitted to the Journal of Comp. and Appl. Maths.

Guido Walz
Lehrstuhl für Mathematik IV
Universität Mannheim
Seminargebäude A5
D–6800 MANNHEIM 1

West Germany