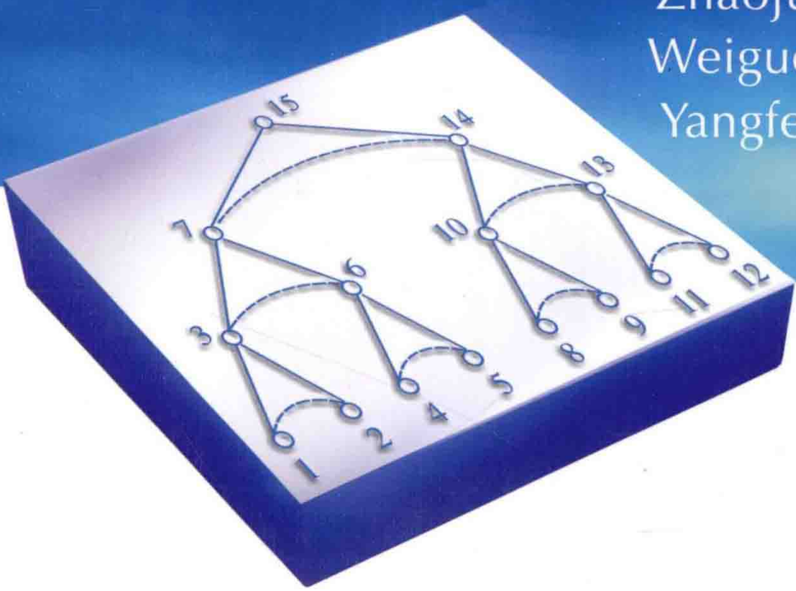
Series in Contemporary Applied Mathematics
CAM 19

Matrix Functions and Matrix Equations

矩阵函数与矩阵方程

Zhaojun Bai
Weiguo Gao
Yangfeng Su
editors



高等教育出版社

Series in Contemporary Applied Mathematics CAM 19

Matrix Functions and Matrix Equations

矩阵函数与矩阵方程

JUZHEN HANSHU YU JUZHEN FANGCHENG

editors

Zhaojun Bai
UC Davis

Weiguo Gao
Fudan University, China

Yangfeng Su
Fudan University, China

高等教育出版社·北京

Zhaojun Bai
Department of Computer Science
University of California
One Shields Avenue
Davis, CA 95616, USA
Email: zbai@ucdavis.edu

Weiguo Gao
School of Mathematical Sciences
Fudan University
Shanghai 200433, China
Email: wggao@fudan.edu.cn

Yangfeng Su
School of Mathematical Sciences
Fudan University
Shanghai 200433, China
Email: yfsu@fudan.edu.cn

图书在版编目（CIP）数据

矩阵函数与矩阵方程＝Matrix functions and matrix equations: 英文 / 柏兆俊，高卫国，苏仰锋编. -- 北京：高等教育出版社，2015.5

ISBN 978-7-04-042249-8

Ⅰ. ①矩… Ⅱ. ①柏… ②高… ③苏… Ⅲ. ①矩阵函数－英文 ②矩阵－线性方程－英文 Ⅳ. ① O151.21 ② O241.6

中国版本图书馆 CIP 数据核字（2015）第 036030 号

策划编辑 赵天夫　　责任编辑 赵天夫　　封面设计 张　楠
责任印制 韩　刚

出版发行	高等教育出版社	咨询电话	400-810-0598
社　　址	北京市西城区德外大街 4 号	网　　址	http://www.hep.edu.cn
邮政编码	100120		http://www.hep.com.cn
印　　刷	涿州市星河印刷有限公司	网上订购	http://www.landraco.com
开　　本	787 × 960　1/16		http://www.landraco.com.cn
印　　张	9.25	版　　次	2015 年 5 月第 1 版
字　　数	210 千字	印　　次	2015 年 5 月第 1 次印刷
购书热线	010-58581118	定　　价	59.00 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换

物 料 号　42249-00

Matrix Functions and Matrix Equations

矩阵函数与矩阵方程

Series in Contemporary Applied Mathematics CAM

---

1 Jiongmin Yong, Rama Cont, eds., *Mathematical Finance: Theory and Practice* (2000)
2 F. Dubois, Huamo Wu, eds., *New Advances in Computational Fluid Dynamics: Theory, Methods and Applications* (2001)
3 Hanji Shang, Alain Tosseti, eds., *Actuarial Science: Theory and Practice* (2002)
4 Alexandre Ern, Weiping Liu, eds., *Mathematical Problems in Environmental Science and Engineering* (2002)
5 Haïm Brezis, Tatsien Li, eds., *Ginzburg-Landau Vortices* (2005)
6 Tatsien Li, Pingwen Zhang, eds., *Frontiers and Prospects of Contemporary Applied Mathematics* (2006)
7 Deguan Wang, Christian Duquennoi, Alexandre Ern, eds., *Mathematical Methods for Surface and Subsurface Hydrosystems* (2007)
8 Rolf Jeltsch, Tatsien Li, Ian H. Sloan, eds., *Some Topics in Industrial and Applied Mathematics* (2007)
9 Philippe G. Ciarlet, Tatsien Li, eds., *Differential Geometry: Theory and Applications* (2008)
10 Tatsien Li, Pingwen Zhang, eds., *Industrial and Applied Mathematics in China* (2009)
11 Zhien Ma, Yicang Zhou, Jianhong Wu, eds., *Modeling and Dynamics of Infectious Diseases* (2009)
12 Tomas Y. Hou, Chun Liu, Jianguo Liu, eds., *Multi-Scale Phenomena in Complex Fluids: Modeling, Analysis and Numerical Simulation* (2010)
13 Guiqiang Chen, Tatsien Li, Chun Liu, eds., *Nonlinear Conservation Laws, Fluid Systems and Related Topics* (2010)
14 A. Damlamian, S. Jaffard, eds., *Wavelet Methods in Mathematical Analysis and Engineering* (2010)
15 Tatsien Li, Yuejun Peng, Bopeng Rao, eds., *Some Problems on Nonlinear Hyperbolic Equations and Applications* (2010)
16 A. Damlamian, B. Miara, Tatsien Li, eds., *Multiscale Problems: Theory, Numerical Approximation and Applications* (2011)
17 Tatsien Li, Song Jiang, eds., *Hyperbolic Problems: Theory, Numerics and Applications*, Vol. 1 (2012)
18 Tatsien Li, Song Jiang, eds., *Hyperbolic Problems: Theory, Numerics and Applications*, Vol. 2 (2012)
19 Zhaojun Bai, Weiguo Gao, Yangfeng Su, *Matrix Functions and Matrix Equations* (2015)

Preface

This volume is a collection of summaries from the lectures of the Fourth Gene Golub SIAM Summer School on "Matrix Functions and Matrix Equations" held at Fudan University, Shanghai, China from July 22 to August 2, 2013. The School was in conjunction with the 3rd International Summer School on Numerical Linear Algebra and the 9th Shanghai Summer School on Analysis and Numerics in Modern Sciences. There were 45 students from 14 countries attended the School. An extra week of activities from August 5 to August 9 was organized for interested students.

Matrix functions and matrix equations are widely used in science, engineering and the social sciences, due to the succinct and insightful way in which they allow problems to be formulated and solutions to be expressed. Applications range from exponential integrators for the solution of partial differential equations to model reduction of dynamical systems. The School introduces students to underlying theory, algorithms and applications of matrix functions and matrix equations, and relevant linear solvers and eigenvalue computations. The summer school was composed of three courses:

1. Functions of matrices and exponential integrators by Nicholas Higham, The University of Manchester, United Kingdom and Marlis Hochbruck, Karlsruhe Institute of Technology, Germany.
2. Matrix equations and model reduction by Peter Benner, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany.
3. High performance linear solvers and eigenvalue computations by Ren-Cang Li, University of Texas at Arlington, United States and Xiaoye Sherry Li, Lawrence Berkeley National Laboratory, United States.

There are five chapters in this volume. Chapter 1, by Nicholas Higham and Lijing Lin, is on matrix functions: a short course. Chapter 2, by Marlis Hochbruck, is on a short course on exponential integrators. Chapter 3, by Peter Benner, Tobias Breiten and Lihong Feng, is on matrix equations and model reduction. Chapter 4, by Ren-Cang Li, is on Rayleigh quotient based optimization methods for eigenvalue problems. Chapter 5, by Xiaoye Li, is on factorization based sparse solvers and preconditioners.

The summer school is hosted by School of Mathematical Sciences, Fudan University. The local organizers include Professors Tatsien Li, Jin Cheng, Weiguo Gao and Yangfeng Su of Fudan University, Professor Pingwen Zhang of Peking University and Professor Zhaojun Bai of University of California, Davis.

The School received the generous sponsorship from SIAM, US National Science Foundation, Chinese-French Institute for Applied Mathematics (ISFMA), Shanghai Center for Mathematical Sciences, National Science Foundation of China (NSFC), and NSFC 111 project. In addition, many units of Fudan University including Graduate School, DOE Key Laboratory of Nonlinear Mathematical Models and Methods, Key Laboratory of Shanghai Modern Applied Mathematics also provided financial and logistical supports. The Numerical Algorithm Group (NAG) provided the free access to numerical computing software for all participants.

We would like to express our gratitude to all authors for their contributions to this volume. Finally, the editors wish to thank Mr. Tianfu Zhao for the professional assistant on the publication of this volume.

October 2014

Zhaojun Bai
Weiguo Gao
Yangfeng Su

Contents

Matrix Functions: A Short Course

Nicholas J. Higham* Lijing Lin†

1 Introduction

A summary is given of a course on functions of matrices delivered by the first author (lecturer) and second author (teaching assistant) at the Gene Golub SIAM Summer School 2013 at Fudan University, Shanghai, China, July 22–26 2013 [34]. This article covers some essential features of the theory and computation of matrix functions. In the spirit of course notes the article is not a comprehensive survey and does not cite all the relevant literature. General references for further information are the book on matrix functions by Higham [31] and the survey by Higham and Al-Mohy [35] of computational methods for matrix functions.

2 History

Matrix functions are as old as matrix algebra itself. The term "matrix" was coined in 1850 [58] by James Joseph Sylvester, FRS (1814–1897), while the study of matrix algebra was initiated by Arthur Cayley, FRS (1821–1895) in his *A Memoir on the Theory of Matrices* (1858) [11]. In that first paper, Cayley considered matrix square roots.

Notable landmarks in the history of matrix functions include:

- Laguerre (1867) [45] and Peano (1888) [53] defined the exponential of a matrix via its power series.
- Sylvester (1883) stated the definition of $f(A)$ for general f via the interpolating polynomial [59]. Buchheim (1886) [10], [33] extended Sylvester's interpolation formula to arbitrary eigenvalues.
- Frobenius (1896) [21] showed that if f is analytic then $f(A)$ is the sum of the residues of $(zI - A)^{-1}f(z)$ at the eigenvalues of A, thereby anticipating the Cauchy integral representation, which was used by Poincaré (1899) [54].

*School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK (nick.higham@manchester.ac.uk, http://www.maths.man.ac.uk/~higham, lijing.lin@manchester.ac.uk, http://www.maths.manchester.ac.uk/~lijing).

- The Jordan form definition was used by Giorgi (1928) [22], and Cipolla (1932) [14] extended it to produce nonprimary matrix functions.
- The first book on matrix functions was published by Schwerdtfeger (1938) [56].
- Frazer, Duncan and Collar published the book *Elementary Matrices and Some Applications to Dynamics and Differential Equations* [20] in 1938, which was "the first book to treat matrices as a branch of applied mathematics" [15].
- A research monograph on functions of matrices was published by Higham (2008) [31].

3 Theory

3.1 Definitions

We are concerned with functions $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ that are defined in terms of an underlying scalar function f. Given $f(t)$, one can define $f(A)$ by substituting A for t: e.g.,

$$f(t) = \frac{1+t^2}{1-t} \quad \Rightarrow \quad f(A) = (I-A)^{-1}(I+A^2),$$

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots, \quad |x| < 1,$$

$$\Rightarrow \log(I+A) = A - \frac{A^2}{2} + \frac{A^3}{3} - \frac{A^4}{4} + \cdots, \quad \rho(A) < 1.$$

This way of defining $f(A)$ works for f a polynomial, a rational function, or a function having a convergent power series (see section 6.1). Note that f is not evaluated elementwise on the matrix A, as is the case in some programming languages.

For general f, there are various equivalent ways to formally define a matrix function. We give three definitions, based on the Jordan canonical form, polynomial interpolation, and the Cauchy integral formula.

3.1.1 Definition via Jordan canonical form

Any matrix $A \in \mathbb{C}^{n\times n}$ can be expressed in the Jordan canonical form

$$Z^{-1}AZ = J = \mathrm{diag}(J_1, J_2, \ldots, J_p), \tag{3.1a}$$

$$J_k = J_k(\lambda_k) = \begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix} \in \mathbb{C}^{m_k\times m_k}, \tag{3.1b}$$

where Z is nonsingular and $m_1 + m_2 + \cdots + m_p = n$. Denote by

- $\lambda_1, \ldots, \lambda_s$ the distinct eigenvalues of A,
- n_i the order of the largest Jordan block in which λ_i appears, which is called the *index* of λ_i.

We say the function f is *defined on the spectrum* of A if the values

$$f^{(j)}(\lambda_i), \qquad j = 0\colon n_i - 1, \quad i = 1\colon s \tag{3.2}$$

exist.

Definition 3.1 (matrix function via Jordan canonical form). *Let f be defined on the spectrum of $A \in \mathbb{C}^{n\times n}$ and let A have the Jordan canonical form* (3.1). *Then*

$$f(A) := Zf(J)Z^{-1} = Z\mathrm{diag}(f(J_k))Z^{-1}, \tag{3.3}$$

where

$$f(J_k) := \begin{bmatrix} f(\lambda_k) & f'(\lambda_k) & \cdots & \dfrac{f^{(m_k-1)}(\lambda_k)}{(m_k-1)!} \\ & f(\lambda_k) & \ddots & \vdots \\ & & \ddots & f'(\lambda_k) \\ & & & f(\lambda_k) \end{bmatrix}. \tag{3.4}$$

The definition yields a matrix $f(A)$ that can be shown to be independent of the particular Jordan canonical form.

In the case of multivalued functions such as $\sqrt{t}$ and $\log t$ it is implicit that a single branch has been chosen in (3.4) and across the Jordan blocks with the same eigenvalue; the resulting function is called a *primary* matrix function. If an eigenvalue occurs in more than one Jordan block and a different choice of branch is made in two different blocks then a *nonprimary* matrix function is obtained (see section 3.1.5).

3.1.2 Definition via interpolating polynomial

Before giving the second definition, we recall some background on polynomials at a matrix argument.

- The *minimal polynomial* of $A \in \mathbb{C}^{n\times n}$ is defined to be the unique monic polynomial ϕ of lowest degree such that $\phi(A) = 0$. The existence of the minimal polynomial is proved in most textbooks on linear algebra.
- By considering the Jordan canonical form it is not hard to see that $\phi(t) = \prod_{i=1}^{s}(t-\lambda_i)^{n_i}$, where $\lambda_1, \ldots, \lambda_s$ are the distinct eigenvalues of A and n_i is the index of λ_i. It follows immediately that ϕ is zero on the spectrum of A (that is, the values (3.2) are all zero for $f(t) = \phi(t)$).

- Given any polynomial p and any matrix $A \in \mathbb{C}^{n\times n}$, p is clearly defined on the spectrum of A and $p(A)$ can be defined by substitution.
- For polynomials p and q, $p(A) = q(A)$ if and only if p and q take the same values on the spectrum [31, Thm. 1.3]. Thus the matrix $p(A)$ is completely determined by the values of p on the spectrum of A.

The following definition gives a way to generalize the property of polynomials in the last bullet point to arbitrary functions and define $f(A)$ in terms of the values of f on the spectrum of A.

Definition 3.2 (matrix function via Hermite interpolation). *Let f be defined on the spectrum of $A \in \mathbb{C}^{n\times n}$. Then $f(A) := p(A)$, where p is the unique polynomial of degree less than $\sum_{i=1}^{s} n_i$ (which is the degree of the minimal polynomial) that satisfies the interpolation conditions*

$$p^{(j)}(\lambda_i) = f^{(j)}(\lambda_i), \qquad j = 0\colon n_i - 1, \quad i = 1\colon s.$$

The polynomial p specified in the definition is known as the Hermite interpolating polynomial.

For an example, let $f(t) = t^{1/2}$ (the principal branch of the square root function, so that $\operatorname{Re} t^{1/2} \ge 0$), $A = \left[\begin{smallmatrix} 2 & 2 \\ 1 & 3 \end{smallmatrix}\right]$, $\lambda(A) = \{1, 4\}$. Seeking $p(t)$ with $p(1) = f(1)$ and $p(4) = f(4)$, we obtain

$$p(t) = f(1)\frac{t-4}{1-4} + f(4)\frac{t-1}{4-1} = \frac{1}{3}(t+2).$$

$$\Rightarrow \quad A^{1/2} = p(A) = \frac{1}{3}(A + 2I) = \frac{1}{3}\begin{bmatrix} 4 & 2 \\ 1 & 5 \end{bmatrix}.$$

Several properties follow immediately from this definition:

- $f(A) = p(A)$ is a polynomial in A, where the polynomial p depends on A.
- $f(A)$ commutes with A.
- $f(A^T) = f(A)^T$.

Because the minimal polynomial divides the characteristic polynomial, $q(t) = \det(tI - A)$, it follows that $q(A) = 0$, which is the Cayley-Hamilton theorem. Hence A^n can be expressed as a linear combination of lower powers of A: $A^n = \sum_{k=0}^{n-1} c_k A^k$. Using this relation recursively we find that any power series collapses to a polynomial. For example, $e^A = \sum_{k=0}^{\infty} A^k/k! = \sum_{k=0}^{n-1} d_k A^k$ (but the d_k depend on A).

3.1.3 Definition via Cauchy integral theorem

Definition 3.3 (matrix function via Cauchy integral). *For* $A \in \mathbb{C}^{n\times n}$,

$$f(A) = \frac{1}{2\pi i}\int_{\Gamma} f(z)(zI - A)^{-1}dz,$$

where f *is analytic on and inside a closed contour* Γ *that encloses* $\lambda(A)$.

3.1.4 Multiplicity and equivalence of definitions

Definitions 3.1, 3.2, and 3.3 are equivalent, modulo the analyticity assumption for the Cauchy integral definition [31, Thm. 1.12]. Indeed this equivalence extends to other definitions, as noted by Rinehart [55]:

> "There have been proposed in the literature since 1880 eight distinct definitions of a matric function, by Weyr, Sylvester and Buchheim, Giorgi, Cartan, Fantappiè, Cipolla, Schwerdtfeger and Richter ... All of the definitions, except those of Weyr and Cipolla are essentially equivalent."

The definitions have different strengths. For example, the interpolation definition readily yields some key basic properties (as we have already seen), the Jordan canonical form definition is useful for solving matrix equations (e.g., $X^2 = A$, $e^X = A$) and for evaluation when A is normal, and the Cauchy integral definition can be useful both in theory and in computation (see section 7.2).

3.1.5 Nonprimary matrix functions

Nonprimary matrix functions are ones that are not obtainable from our three definitions, or that violate the single branch requirement in Definition 3.1. Thus a nonprimary matrix function of A is obtained from Definition 3.1 if A is derogatory and a different branch of f is taken in two different Jordan blocks for λ. For example, the 2×2 identity matrix has two primary square roots and an infinity of nonprimary square roots:

$$\begin{aligned} I_2 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}^2 && \text{(primary)} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^2 = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}^2 && \text{(nonprimary).} \end{aligned}$$

In general, primary matrix functions are expressible as polynomials in A, while nonprimary ones are not. The 2×2 zero matrix 0_2 is its own primary square root. Any nilpotent matrix of degree 2 is also a nonprimary square root of 0_2, for example $[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}]$, but the latter matrix is not a polynomial in 0_2.

The theory of matrix functions is almost exclusively concerned with primary matrix functions, but nonprimary functions are needed in some applications, such as the embeddability problem in Markov chains [31, Sec. 2.3].

3.1.6 Principal logarithm, root, and power

Let $A \in \mathbb{C}^{n\times n}$ have no eigenvalues on $\mathbb{R}^-$ (the closed real axis). We need the following definitions.

Principal log: $X = \log A$ denotes the unique X such that $e^X = A$ and $-\pi < \operatorname{Im} \lambda_i < \pi$ for every eigenvalue λ_i of X.

Principal p th root: For integer $p > 0$, $X = A^{1/p}$ is the unique X such that $X^p = A$ and $-\pi/p < \arg \lambda_i < \pi/p$ for every eigenvalue λ_i of X.

Principal power: For $s \in \mathbb{R}$, the principal power is defined as $A^s = e^{s \log A}$, where $\log A$ is the principal logarithm. An integral representation is also available:

$$A^s = \frac{\sin(s\pi)}{s\pi} A \int_0^\infty (t^{1/s} I + A)^{-1} dt, \qquad s \in (0,1).$$

3.2 Properties and formulas

Three basic properties of $f(A)$ were stated in section 3.1.2. Some other important properties are collected in the following theorem.

Theorem 3.4 ([31, Thm. 1.13]). *Let $A \in \mathbb{C}^{n\times n}$ and let f be defined on the spectrum of A. Then*

(a) $f(XAX^{-1}) = Xf(A)X^{-1}$;

(b) *the eigenvalues of $f(A)$ are $f(\lambda_i)$, where the λ_i are the eigenvalues of A;*

(c) *if X commutes with A then X commutes with $f(A)$;*

(d) *if $A = (A_{ij})$ is block triangular then $F = f(A)$ is block triangular with the same block structure as A, and $F_{ii} = f(A_{ii})$;*

(e) *if $A = \operatorname{diag}(A_{11}, A_{22}, \ldots, A_{mm})$ is block diagonal then $f(A) = \operatorname{diag}(f(A_{11}), f(A_{22}), \ldots, f(A_{mm}))$.*

Some more advanced properties are as follows.

- $f(A) = 0$ if and only if (from Definition 3.1 or 3.2) $f^{(j)}(\lambda_i) = 0$, $j = 0 : n_i - 1$, $i = 1 : s$.

- The sum, product, composition of functions work "as expected":
  - $(\sin + \cos)(A) = \sin A + \cos A$,
  - $f(t) = \cos(\sin t) \quad \Rightarrow \quad f(A) = \cos(\sin A)$.
- Polynomial functional relations generalize from the scalar case. For example: if $G(f_1, \dots, f_m) = 0$, where G is a polynomial, then $G(f_1(A), \dots, f_m(A)) = 0$. E.g.,
  - $\sin^2 A + \cos^2 A = I$,
  - $(A^{1/p})^p = A$ for any integer $p > 0$,
  - $e^{iA} = \cos A + i \sin A$.
- However, other plausible relations can fail:
  - $f(A^*) \neq f(A)^*$ in general,
  - $e^{\log A} = A$ but $\log e^A \neq A$ in general,
  - $e^A \neq (e^{A/\alpha})^\alpha$ in general,
  - $(AB)^{1/2} \neq A^{1/2}B^{1/2}$ in general,
  - $e^{(A+B)t} = e^{At}e^{Bt}$ for all t if and only if $AB = BA$.

  Correction terms involving the matrix unwinding function can be introduced to restore equality in the second to fourth cases [7].

3.3 Fréchet derivative and condition number

3.3.1 Relative condition number

An important issue in the computation of matrix functions is the conditioning. The data may be uncertain and rounding errors from finite precision computations can often be interpreted via backward error analysis as being equivalent to perturbations in the data. So it is important to understand the sensitivity of $f(A)$ to perturbations in A. Sensitivity is measured by the condition number defined as follows.

Definition 3.5. *Let $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ be a matrix function. The relative condition number of f is*

$$\mathrm{cond}(f, A) := \lim_{\epsilon\to 0} \sup_{\|E\|\le\epsilon\|A\|} \frac{\|f(A+E) - f(A)\|}{\epsilon\|f(A)\|},$$

where the norm is any matrix norm.

3.3.2 Fréchet derivative

To obtain explicit expressions for $\mathrm{cond}(f, A)$, we need an appropriate notion of derivative for matrix functions. The *Fréchet derivative* of a matrix function $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ at a point $A \in \mathbb{C}^{n\times n}$ is a linear mapping $L_f(A, \cdot) : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ such that for all $E \in \mathbb{C}^{n\times n}$

$$f(A+E) = f(A) + L_f(A, E) + o(\|E\|).$$

It is easy to show that the condition number $\mathrm{cond}(f, A)$ can be characterized as

$$\mathrm{cond}(f, A) = \frac{\|L_f(A)\| \, \|A\|}{\|f(A)\|},$$

where

$$\|L_f(A)\| := \max_{Z \neq 0} \frac{\|L_f(A, Z)\|}{\|Z\|}.$$

3.3.3 Condition number estimation

Since L_f is a linear operator,

$$\mathrm{vec}(L_f(A, E)) = K_f(A)\mathrm{vec}(E)$$

where $K_f(A) \in \mathbb{C}^{n^2 \times n^2}$ is a matrix independent of E known as the *Kronecker form* of the Fréchet derivative. It can be shown that $\|L_f(A)\|_F = \|K_f(A)\|_2$ and that $\|L_f(A)\|_1$ and $\|K_f(A)\|_1$ differ by at most a factor n. Hence estimating $\mathrm{cond}(f, A)$ reduces to estimating $\|K_f(A)\|$ and this can be done using a matrix norm estimator, such as the block 1-norm estimator of Higham and Tisseur [39].

4 Applications

Functions of matrices play an important role in many applications. Here we describe some examples.

4.1 Toolbox of matrix functions

In software we want to be able to evaluate interesting f at matrix arguments as well as scalar arguments. For example, trigonometric matrix functions, as well as matrix roots, arise in the solution of second order differential equations: the initial value problem

$$\frac{d^2y}{dt^2} + Ay = 0, \quad y(0) = y_0, \quad y'(0) = y_0'$$

has solution

$$y(t) = \cos(\sqrt{A}t)y_0 + (\sqrt{A})^{-1}\sin(\sqrt{A}t)y_0',$$

where $\sqrt{A}$ denotes any square root of A. On the other hand, the differential equation can be converted to a first order system and then solved using the exponential:

$$\begin{bmatrix} y' \\ y \end{bmatrix} = \exp\left(\begin{bmatrix} 0 & -tA \\ tI_n & 0 \end{bmatrix}\right)\begin{bmatrix} y_0' \\ y_0 \end{bmatrix}.$$

4.2 Nuclear magnetic resonance

In nuclear magnetic resonance (NMR) spectroscopy, the Solomon equations

$$\frac{dM}{dt} = -RM, \quad M(0) = I$$

relate a matrix of intensities $M(t)$ to a symmetric, diagonally dominant matrix R (known as the relaxation matrix). Hence $M(t) = e^{-Rt}$. NMR workers need to solve both forward and inverse problems:

- in simulations and testing, compute $M(t)$ given R;
- determine R from observed intensities: estimation methods are used since not all the m_{ij} are observed.

4.3 Phi functions and exponential integrators

The φ functions are defined by the recurrence $\varphi_{k+1}(z) = \frac{\varphi_k(z)-1/k!}{z}$ with $\varphi_0(z) = e^z$, and are given explicitly by

$$\varphi_k(z) = \sum_{j=0}^{\infty} \frac{z^j}{(j+k)!}.$$

They appear in explicit solutions to certain linear differential equations:

$$\begin{aligned}
\frac{dy}{dt} &= Ay, \quad y(0) = y_0 &&\Rightarrow \quad y(t) = e^{At}y_0,\\
\frac{dy}{dt} &= Ay + b, \quad y(0) = 0 &&\Rightarrow \quad y(t) = t\varphi_1(tA)b,\\
\frac{dy}{dt} &= Ay + ct, \quad y(0) = 0 &&\Rightarrow \quad y(t) = t^2\varphi_2(tA)c,
\end{aligned}$$

and more generally provide an explicit solution for a differential equation with right-hand side $Ay + p(t)$ with p a polynomial.

Consider an initial value problem written in the form

$$u'(t) = Au(t) + g(t, u(t)), \quad u(t_0) = u_0, \quad t \geq t_0, \tag{4.1}$$

where $u(t) \in \mathbb{C}^n$, $A \in \mathbb{C}^{n\times n}$, and g is a nonlinear function. Spatial semidiscretization of partial differential equations leads to systems in this form with A representing a discretized linear operator. Thus A may be large and sparse. The solution can be written as [47, Lem. 5.1]

$$u(t) = e^{(t-t_0)A}u_0 + \sum_{k=1}^{\infty} \varphi_k\big((t-t_0)A\big)(t-t_0)^k u_k, \tag{4.2}$$