Linear Algebra and its Applications 456 (2014) 199–213

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Computing enclosures for the inverse square root and the sign function of a matrix

Andreas Frommer [a,*], Behnam Hashemi [b,c,1], Thomas Sablik [d]

[a] Department of Mathematics, Bergische Universität Wuppertal, 42097 Wuppertal, Germany
[b] School of Mathematics, Institute for Research in Fundamental Sciences (IPM), P.O. Box: 19395-5746, Tehran, Iran
[c] Department of Mathematics, Faculty of Basic Sciences, Shiraz University of Technology, Modarres Boulevared, Shiraz 71555-313, Iran
[d] Faculty of Electrical, Information and Media Engineering, Bergische Universität Wuppertal, 42097 Wuppertal, Germany

ARTICLE INFO

Article history:
Received 1 September 2013
Accepted 28 November 2013
Available online 17 December 2013
Submitted by C.H. Guo

MSC:
65F05
65G20

Keywords:
Matrix inverse square root
Matrix sign function
Interval arithmetic
Krawczyk's method
Verified computation

ABSTRACT

We study computational methods for obtaining rigorous a posteriori error bounds for the inverse square root and the sign function of an $n \times n$ matrix A. Given a computed approximation for the inverse square root of A, our methods work by using interval arithmetic to obtain a narrow interval matrix which, with mathematical certainty, is known to contain the exact inverse square root. Particular emphasis is put on the computational efficiency of the method which has complexity $\mathcal{O}(n^3)$ and which uses almost exclusively matrix–matrix operation, a key to the efficient use of available software for interval computations. The standard formulation of the method assumes that A can be diagonalized and that the eigenvector matrix of A is well-conditioned. A modification relying on a stable similarity transformation to block diagonal form is also developed.

1. Introduction

Given a function $f : \Omega \subseteq \mathbb{C} \to \mathbb{C}$, the matrix function $f(A)$ for $A \in \mathbb{C}^{n\times n}$ is defined as soon as $\operatorname{spec}(A) \subseteq \Omega$ and f is $s(\lambda) - 1$ times differentiable at any eigenvalue λ of A, where $s(\lambda)$ denotes the

* Corresponding author.
E-mail addresses: frommer@math.uni-wuppertal.de (A. Frommer), hashemi@sutech.ac.ir (B. Hashemi), sablik@uni-wuppertal.de (T. Sablik).

[1] This research was in part supported by a grant from IPM (No. 91650055).

0024-3795/$ – see front matter
http://dx.doi.org/10.1016/j.laa.2013.11.047

index of λ, i.e. the size of the largest Jordan block belonging to λ, and λ lies in the interior of Ω whenever $s(\lambda) > 1$, see [1, Definitions 1.1 and 1.2] or [2, Definition 6.2.4]. In the present work we consider the matrix inverse square root, $f(z) = z^{-1/2}$. This matrix function has applications, e.g., in the optimal symmetric orthogonalization of a set of vectors [3] and the generalized eigenvalue problem [4]. It also appears in the matrix sign function $\mathrm{sign}(A)$ which can be defined as $A(A^2)^{-1/2}$ and which arises in the solution of algebraic Riccati equations [5] and also in applications from Theoretical Physics [6].

The complex function, $f(z) = z^{-1/2}$ has two different branches. Referring to [1] for details,[2] it suffices here to indicate that there exist several (primary) inverse square roots of A depending on which branch of f is used for the different Jordan blocks. If A has no eigenvalues on $(-\infty, 0]$, the so-called *principal* inverse square root (see also [3]) is uniquely defined by requiring that for all $\lambda \in \mathrm{spec}(A)$ the branch of f is chosen such that $f(\lambda)$ lies in the open right half plane. For the matrix square root and A non-singular, it has been shown [1, Thm. 1.26] that all primary square roots are isolated (and non-singular) solutions of the matrix equation $X^2 - A = 0$. Observing that for A non-singular any solution X of one of the equations

$$X^2 - A^{-1} = 0, \tag{1a}$$

$$X^{-2} - A = 0, \tag{1b}$$

$$XAX - I = 0, \tag{1c}$$

$$X^2 A - I = 0, \tag{1d}$$

$$AX^2 - I = 0 \tag{1e}$$

is non-singular and satisfies $(X^{-1})^2 - A = 0$, we conclude that a primary inverse square roots is always an *isolated* solution of any of the above equations.

Several different iterative schemes are available in the literature for computing the matrix inverse square root, see e.g. [7–10,3]. All these classical, floating point numerical methods will always yield a result which is not an exact inverse square root of A but rather an approximation to it. The purpose of this paper is to propose and analyze an efficient interval-arithmetic based method which, given a relatively accurate floating point approximation to a primary (usually the principal) inverse square root, obtains a narrow enclosure for the *exact* inverse square root. It does so by proving that one of the equations in (1) has exactly one solution X within a whole set of matrices, represented as an interval matrix, i.e. a matrix with interval entries. Our approach is based on techniques developed earlier for the square root in [11].

The paper is organized as follows: In Section 2 we introduce some notation and standard results which are at the basis of our method. In particular, we discuss the necessary background from interval analysis. In Section 3 we develop our method for computing enclosures for the inverse square root, investigate its algorithmic complexity and shortly discuss some variants. We then explain in some detail how the method can be used to also obtain enclosures for the matrix sign function in Section 4. Section 5 contains results of numerical experiments, and our conclusions are summarized in Section 6.

2. Preliminaries

Throughout this paper lower case letters are used for scalars and vectors and uppercase letters for matrices. For $A \in \mathbb{C}^{m\times n}$, the vector $a = \mathrm{vec}(A) \in \mathbb{C}^{mn}$ is obtained by stacking the columns of A. We will systematically—and often implicitly—use the convention that a lower case letter denotes the vector obtained via vec from a matrix denoted by the respective uppercase letter.

The point-wise division $A./B$ of two matrices $A, B \in \mathbb{C}^{m\times n}$ is given by

$$A./B = C \in \mathbb{C}^{m\times n}, \quad \text{where } C = (c_{ij}) \text{ with } c_{ij} = a_{ij}/b_{ij}.$$

[2] [1] treats the matrix square root, but all considerations there immediately carry over to the inverse square root.