Verified computation of inverse square root and the sign functions of a matrix

Andreas Frommer[a], Behnam Hashemi[b] and Thomas Sablik[a]

[a] University of Wuppertal, Wuppertal, Germany,
[b] Shiraz University of Technology, Shiraz, Iran,

frommer@math.uni-wuppertal.de
hashemi@sutech.ac.ir
sablik@uni-wuppertal.de

Keywords: Matrix inverse square root, Matrix sign function, Interval arithmetic, Krawczyk-Rump iteration, Verified computation

1 Enclosures for the matrix inverse square root

The inverse square root of an $n \times n$ matrix A satisfies the nonlinear matrix equation

$$F(X) = XAX - I = 0. \tag{1}$$

The following result goes back to Rump's Ph.D. thesis [3] and is based on [4]. A proof can also be found in [1].

Theorem 1. *Assume that $f : D \subset \mathbb{C}^N \to \mathbb{C}^N$ is continuous in D. Let $\check{x} \in D$ and $\boldsymbol{z} \in \mathbb{IC}^N$ be such that $\check{x} + \boldsymbol{z} \subseteq D$. Moreover, assume that $\mathcal{S} \subset \mathbb{C}^{N\times N}$ is a set of matrices containing all slopes $S(\check{x}, y)$ for $y \in \check{x} + \boldsymbol{z} =: \boldsymbol{x}$. Finally, let $R \in \mathbb{C}^{N\times N}$. Denote by $\mathcal{K}_f(\check{x}, R, \boldsymbol{z}, \mathcal{S})$ the set*

$$\mathcal{K}_f(\check{x}, R, \boldsymbol{z}, \mathcal{S}) := \{-Rf(\check{x}) + (I - RS)z : S \in \mathcal{S}, z \in \boldsymbol{z}\}. \tag{2}$$

Then, if

$$\mathcal{K}_f(\check{x}, R, \boldsymbol{z}, \mathcal{S}) \subseteq \operatorname{int} \boldsymbol{z}, \tag{3}$$

the function f has a zero x^* in $\check{x} + \mathcal{K}_f(\check{x}, R, \boldsymbol{z}, \mathcal{S}) \subseteq \boldsymbol{x}$. Moreover, if $\mathcal{S}$ also contains all slope matrices $S(x, y)$ for $x, y \in \boldsymbol{x}$, then this zero is unique in $\boldsymbol{x}$.

The matrix equation (1) can be written as the vector function

$$f(x) = (X^T \otimes X)a - i,$$

where $f(x) := \operatorname{vec}(F(X))$, $a := \operatorname{vec}(A)$, $i := \operatorname{vec}(I)$ and $\otimes$ denotes the matrix Kronecker product. For any two matrices X and Y we can show that $S(x, y) = I \otimes XA + (AY)^T \otimes I$ is a slope for f. For a possible application of the Krawczyk operator we see that we can take $\boldsymbol{S} = I \otimes \boldsymbol{X}A + (A\boldsymbol{X})^T \otimes I$. Here we explicitly see that the computation of $R\boldsymbol{S}$ has complexity $\mathcal{O}(n^5)$, since $\boldsymbol{X}A$ and $A\boldsymbol{X}$ usually are dense matrices and thus $\boldsymbol{S}$ has $2n$ non-zeros per column.

We propose two methods based on interval arithmetic to compute an enclosing interval matrix for the exact inverse square root of a matrix A. Starting from an approximate inverse square root, these methods use a modification of the Krawczyk operator which reduces the computational complexity from $\mathcal{O}(n^5)$ or higher of the standard approach to $\mathcal{O}(n^3)$. Moreover, the methods use almost exclusively matrix-matrix operations and are thus particularly efficient with current implementations of machine interval arithmetic in INTLAB and C-XSC. For sake of simplicity, let us first assume that the matrix A is diagonalizable. We then can decompose A as

$$A = V\Lambda W, \text{ with } V, W, \Lambda \in \mathbb{C}^{n\times n},\ \Lambda = \mathrm{Diag}\,(\lambda_1, \ldots, \lambda_n),\ VW = I. \tag{4}$$

Of course, $W = V^{-1}$, but it will turn out useful to have this additional notation available when we have to account for the fact that inverses are usually not available exactly when computed in floating point arithmetic.

1.1 Reducing the wrapping effect

In several lines of our algorithm we perform two-sided interval matrix multiplications. The result is an interval matrix which can be quite substantially larger than the interval hull of all point matrices due, in particular, to the so-called wrapping effect of interval arithmetic. If we can avoid multiplications with $\boldsymbol{I}_W$ and $\boldsymbol{I}_V$, we might succeed more often and with narrower intervals. The key is to consider the linearly transformed function

$$\hat{f}(\hat{x}) = (V^T \otimes W)\ \ f\left((V^{-T} \otimes W^{-1})\hat{x}\right),$$

or, equivalently, in matrix form

$$\hat{F}(\hat{X}) = W \cdot F\left(W^{-1}\hat{X}V^{-1}\right) \cdot V.$$

The slope $\hat{S}(\hat{x}, \hat{y})$ of this linear transformation of f can therefore be computed as follows.

$$\hat{F}(\hat{X}) - \hat{F}(\hat{Y}) = W \cdot \left(F(W^{-1}\hat{X}V^{-1}) - F(W^{-1}\hat{Y}V^{-1})\right) \cdot V$$

If $X := W^{-1}\hat{X}V^{-1}$ and $Y := W^{-1}\hat{Y}V^{-1}$, we have

$$\begin{aligned}\hat{F}(\hat{X}) - \hat{F}(\hat{Y}) =& WXAXV - WYAYV \pm WXAYV\\ =& WXA(X-Y)V + W(X-Y)AYV.\end{aligned}$$

So,

$$\begin{aligned}\hat{f}(\hat{x}) - \hat{f}(\hat{y}) &= \left(V^T \otimes WXA + (AYV)^T \otimes W\right)(x-y)\\ &= \left(V^T \otimes WXA + (AYV)^T \otimes W\right)(V^{-T} \otimes W^{-1}) \cdot (\hat{x} - \hat{y})\\ &= \left(I \otimes (WXAW^{-1}) + (V^{-1}AYV)^T \otimes I\right) \cdot (\hat{x} - \hat{y})\end{aligned}$$

which means that

$$\hat{S}(\hat{x},\hat{y}) = I \otimes (WXAW^{-1}) + (V^{-1}AYV)^T \otimes I.$$

Given $\check{x} = \mathrm{vec}(\check{X} \approx A^{-1/2})$ and $(V^T \otimes W)\check{x}$ the corresponding approximate zero of $\hat{f}$, the interval matrix

$$\boldsymbol{S} = I \otimes (W\boldsymbol{X}A\boldsymbol{I}_W) + (\boldsymbol{I}_V A\boldsymbol{X}V)^T \otimes I,$$

where $\boldsymbol{X} = \check{X} + \boldsymbol{I}_W\hat{\boldsymbol{Z}}\boldsymbol{I}_V$ contains all slopes from $\hat{\mathcal{S}} = \{\hat{S}(\hat{x},\hat{y}),\, \hat{x},\hat{y} \in \hat{\boldsymbol{x}} = (V^T \otimes W)\check{x} + \hat{\boldsymbol{z}}\}$.

We can therefore compute a superset for $\mathcal{K}_{\hat{f}}((V^T \otimes W)\check{x}, \Delta^{-1}, \hat{\boldsymbol{z}}, \hat{\mathcal{S}})$ as

$$\Delta^{-1}\left(-(V^T \otimes W)f(\check{x}) + (\Delta - \boldsymbol{S})\hat{\boldsymbol{z}}\right).$$

1.2 Modification when A cannot be stably diagonalized

In the case that the matrix A cannot be stably diagonalized, we have the option of using a block diagonalization of A, and our suggested methods can, in principle, still be applied. Our algorithms can be modified in the case that V is ill-conditioned (and also when A is not diagonizable at all) in the following manner: Instead of the spectral decomposition we use the *block diagonalization* of Bavely and Stewart to control the condition number of V at the expense of having D with (hopefully small) blocks along the diagonal. The block diagonal factorization can be written as

$$A = V^{-1}\Gamma V, \tag{5}$$

where Γ is *block* diagonal with each diagonal block being triangular.

We now can proceed in exactly the same manner as outlined in section 1.1 with $\Lambda^{1/2}$ replaced by $\Gamma^{1/2}$ everywhere, where for each diagonal block of Γ we obtain its square root, a triangular matrix, just approximately via some floating point algorithm. Occurrences of Δ^{-1} must be replaced by a forward substitution with the large, sparse triangular matrix $I \otimes \Gamma^{1/2} + \Gamma^{1/2} \otimes I$. This forward substitution cannot be cast into a matrix-matrix operation, making the modified algorithm substantially slower when implemented in INTLAB or C-XSC. Also, the diagonal blocks should all be small in size, because otherwise the dependence property of interval arithmetic will yield very large intervals as a result of the substitution process. We refer to [1] for further details and a discussion of why a standard Schur decomposition of A, i.e. an orthogonal reduction to (full) triangular form, is not a viable approach for an enclosure method based on machine interval arithmetic. Future research will address the question whether in this case we can modify our approach in a manner that it again uses only matrix-matrix operations and to which extent the dependence problem when solving linear recurrences can be avoided.

2 Enclosures for the matrix sign function

Let the non-singular matrix A have no purely imaginary eigenvalues. The matrix inverse square root can be expressed as

$$\operatorname{sign}(A) = A\ (A^2)^{-1/2}. \tag{6}$$

Here, $(A^2)^{1/2}$ denotes the principal square root of A^2. Note that A having no purely imaginary eigenvalues is equivalent to A^2 having no eigenvalues on $\mathbb{R}^-$.

As opposed to the (inverse) square root, there seems to be no nonlinear function for which $\operatorname{sign}(A)$ would appear directly as a non-isolated zero and on which we could use interval arithmetic to compute an enclosure for its zero. We can, however, use the relation (6) and proceed as follows:

> Step 1. Use interval arithmetic to compute an interval matrix $\boldsymbol{A}_2$ that contains the exact result of the matrix-matrix multiplication $A^2 = A \cdot A$.
>
> Step 2. Use a modification of the above-mentioned algorithms to compute an interval matrix $\boldsymbol{X}$ that satisfies
>
> $$\boldsymbol{X} \supseteq \{B^{-1/2} \mid B \in \boldsymbol{A}_2\}.$$
>
> Step 3. Use interval arithmetic to compute an interval matrix $\boldsymbol{S} = A\boldsymbol{X}$ which, by (6) and the enclosure property of interval arithmetic, contains $\operatorname{sign}(A)$.

Due to the enclosure property of interval arithmetic, a modified Krawczyk-Rump type iterative algorithm computes an interval matrix $\boldsymbol{K}$ such that

$$\operatorname{vec}(\boldsymbol{K}) \supseteq \mathcal{K}_{f_{A_2}}(\check{x}, R, \boldsymbol{z}, \boldsymbol{S}) \text{ for all } A_2 \in \boldsymbol{A}_2,$$

where $F_{A_2}(X) = XA_2X - I$ and $\boldsymbol{S}$ is an interval matrix which contains all slopes $S_{f_{A_2}}(x, y)$ for all $A_2 \in \boldsymbol{A}_2$ and all $x, y \in \boldsymbol{X} = \check{X} + \boldsymbol{Z}$. Therefore, if $\boldsymbol{K} \subset \operatorname{int} \boldsymbol{Z}$ we can apply Brouwer's fixed point theorem to each of the functions f_{A_2} individually. The enclosure property of interval arithmetic implies that $\operatorname{sign}(A) = A(A^2)^{-1/2} \in A\boldsymbol{X}$. See [2] for more details.

References:

[1] A. Frommer and B. Hashemi, Verified Computation of Square Roots of a Matrix, *SIAM J. Matrix Anal. Appl.*, 31, pp. 1279–1302, 2009.

[2] A. Frommer, B. Hashemi and T. Sablik, Computing enclosures for the inverse square root and the sign function of a matrix, *Linear Algebra Appl.*, Accepted for publication in the special issue on matrix functions, 2014.

[3] Siegfried M. Rump, *Kleine Fehlerschranken bei Matrixproblemen*, Fakultät für Mathematik, Universität Karlsruhe, 1980.

[4] R. Krawczyk, Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken, *Computing*, 4, pp. 187–201, 1969.