Available online at http://ijim.srbiau.ac.ir/
Int. J. Industrial Mathematics (ISSN 2008-5621)
Vol. 12, No. 1, 2020 Article ID IJIM-1075, 9 pages
Research Article

Science and Research Branch (IAU)

Computing the Matrix Geometric Mean of Two HPD Matrices: A Stable Iterative Method

F. Kiyoumarsi *†

Received Date: 2017-05-13 Revised Date: 2018-05-18 Accepted Date: 2019-06-18

Abstract

In this paper, a new iteration scheme for computing the sign of a matrix which has no pure imaginary eigenvalues is presented. Then, by applying a well-known identity in matrix functions theory, an algorithm for computing the geometric mean of two Hermitian positive definite matrices is constructed. Moreover, another efficient algorithm for this purpose is derived free from the computation of principal matrix square root. Finally, some tests are given to show their applicabilities.

Keywords : Iterative methods; HPD; Sign function; Stability; Convergence.

1 Introduction

Functions of matrices have attracted the attentions of many researchers in Mathematics due to several applications, refer to [9, pages 26-29] or [4, 5, 6]. Here, we focus on the computation of the geometric mean of two complex Hermitian positive definite (HPD) matrices, [10].

For two HPD matrices, the mean $GM(M,N)$ can be expressed uniquely as [17]

$$GM(M,N) = M\#N := M(M^{-1}N)^{\frac{1}{2}}. \tag{1.1}$$

This is a special form of the more general map:

$$M\#_tB := M(M^{-1}N)^t, \qquad t \in \mathbb{R}, \tag{1.2}$$

which has a geometrical interpretation as the parametrization of the geodesic joining M and N for a certain Riemannian geometry on $\mathbb{P}^n$, i.e., the set of $n \times n$ HPD matrices, [15].

It can be proved that $M\#N$ verifies all the properties required by a geometric mean [10], like

$$M\#N = N\#M, \tag{1.3}$$

and if M and N commute, then $M\#N = (AB)^{\frac{1}{2}}$. Thus, the definition is well established.

Notice that $M\#N$ solves the equation of Riccati:

$$YM^{-1}Y = N, \tag{1.4}$$

and it can be proved that it is the unique positive solution [2, page 106]. Moreover, using the properties of the principal square root one can derive

$$M\#N = M(M^{-1}N)^{\frac{1}{2}} = (NM^{-1})^{\frac{1}{2}}M =$$

$$N(NM^{-1})^{\frac{1}{2}} = (MN^{-1})^{\frac{1}{2}}N. \tag{1.5}$$

Here we refer to Y as the principal matrix square root of M and write $Y = M^{\frac{1}{2}}$ when it

[*] Corresponding author. kumarci_farshad@yahoo.com, Tel:+98(913)1819862.

[†] Department of Mathematics, Shahrekord Branch, Islamic Azad University, Shahrekord, Iran.

solves the matrix equation $F(Y) \equiv Y^2 - M = 0$. The symbol $M^{\frac{1}{2}}$ stands for the principal square root of M. Such a matrix exists and is unique if M has no nonpositive real eigenvalues, in particular if M is positive then $M^{\frac{1}{2}}$ is positive.

Bhatia in [2, page 105] proposed another formulation for computing such a mean as follows:

$$M\#N = M^{\frac{1}{2}}(M^{-1/2}NM^{-\frac{1}{2}})^{\frac{1}{2}}M^{\frac{1}{2}}, \qquad (1.6)$$

for the two complex HPD matrices, viz, M and N.

The rest of this paper is organized as follows. In Section 2, an iterative formula for computing matrix sign and its acceleration through a combination with Newton's iteration (see e.g. [14] and the references therein) are presented. We also provide some discussions and illustrate how the new scheme could be constructed and implemented. An error analysis for computing matrix sign function is brought forward in Section 3. Note that the idea of computing the geometric mean using the sign function can also be found in [9, page 131] and has recently been revived in [18]. In Section 4, we show the numerical results and highlight the benefit of the technique. Finally, several concluding comments are collected in Section 5.

2 A stable iterative method

As pointed out by some practitioners, see e.g. [1] and the references cited therein, one efficient way to design new iterative methods for some matrix functions is to apply the zero-finding iterative methods for solving operator equations which here is a matrix equation. Following such a strategy, we here must solve the following nonlinear matrix equation

$$F(Y) := Y^2 - I = 0, \qquad (2.7)$$

where I is an identity matrix. This could result an iterative method for calculating the matrix sign function. To this end, let us take into account the following cubically convergent scheme [13]

$$\begin{cases} y_k = l_k - s_k, \\ l_{k+1} = l_k - \left(1 + \frac{f(y_k)}{f(l_k) - \frac{5}{3}f(y_k)}\right) s_k, \end{cases} \qquad (2.8)$$

with $s_k = \frac{f(l_k)}{f'(l_k)}$.

By applying (2.8) for solving the scalar version of (2.7), we uniquely obtain the following iteration scheme (in the reciprocal form):

$$l_{k+1} = \frac{l_k(5 + 7l_k^2)}{1 + 9l_k^2 + 2l_k^4}, \quad k \geq 0. \qquad (2.9)$$

It is possible to still improve the results by performing one Newton's step at the end of a three-step iteration, after (2.8), in what follows:

$$\begin{cases} y_k = l_k - s_k, \\ z_k = l_k - \left(1 + \frac{f(y_k)}{f(l_k) - \frac{5}{3}f(y_k)}\right) s_k, \\ l_{k+1} = z_k - f'(z_k)^{-1} f(z_k). \end{cases} \qquad (2.10)$$

In this way, a sixth-order accelerated scheme is attained as comes next (again in the reciprocal form) via solving the scalar version of (2.7):

$$l_{k+1} = \frac{2l_k(5 + 7l_k^2)(1 + 9l_k^2 + 2l_k^4)}{1 + 43l_k^2 + 155l_k^4 + 85l_k^6 + 4l_k^8}, \quad k \geq 0. \qquad (2.11)$$

Now, it is pointed out that global convergence behavior of the main contributed scheme (2.11) can be investigated in a similar way to equations (17)-(21) of [19, pages 4-5].

The attraction basins of (2.11) for finding the solution of the polynomial equation $l^2 - 1 = 0$ in the square $[-2, 2] \times [-2, 2]$ of the complex plane is offered in Figure 1, which confirm the global convergence with the step size 0.01 in each dimension for discretization of the domain. The basins of attraction have been shaded according to the number of iterates for converging to the roots of $l^2 - 1 = 0$. In addition, whatever the area is darker it requires more number of iterates to converge. The two white points in the positions ±1 indicate the exact solutions. For obtaining further background about how such fractals are drawn, the reader may refer to [21].

In Figure 2, this attraction basins have been provided in the Riemann sphere. The reason of drawing the attraction basins on the Riemann sphere is to clearly show the global convergence for a very large domain of validity. Anyhow, we remark that to observe the behavior precisely, colorful printing or observing the electronic PDF version is recommended. In consequence, it is

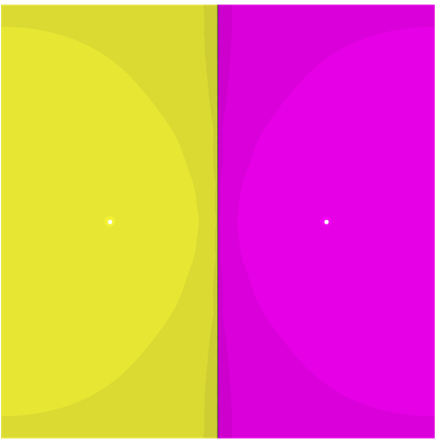

Figure 1: Attraction basins for (2.11) shaded according to the number of required iterations.

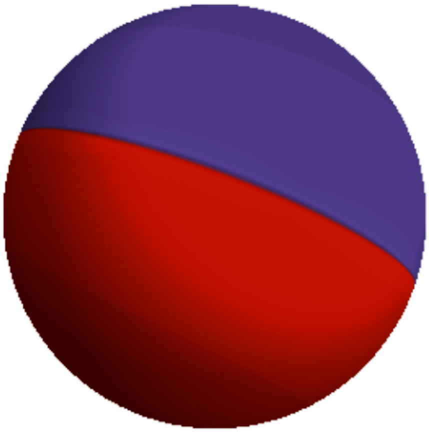

Figure 2: Attraction basins for (2.11) on the Riemann sphere.

straightforward to naïvely extend (2.11) to matrix environment and uniquely attain

$$Y_{k+1} = [10Y_k + 104Y_k^3 + 146Y_k^5 + 28Y_k^7] \\ \times [I + 43Y_k^2 + 155Y_k^4 + 85Y_k^6 + 4Y_k^8]^{-1}, \quad (2.12)$$

and further by factorizing to a more simple-to-implement version as follows:

$$Y_{k+1} = Y_k[10I + 104Y_k^2 + 146Y_k^4 + 28Y_k^6] \\ \times [I + 43Y_k^2 + 155Y_k^4 + 85Y_k^6 + 4Y_k^8]^{-1}. \quad (2.13)$$

The constructed iteration (2.13) is not a member of the Padé family of iterations introduced in [11] for computing the matrix sign function. Therefore, it is new with global convergence and worth investigating. This motivates us for further investigation of (2.13). Note that the most concise definition of the matrix sign decomposition is given by [7]:

$$M = SN = M(M^2)^{-1/2}(M^2)^{1/2}. \quad (2.14)$$

wherein $S = \text{sign}(M)$ is the matrix sign function.

A recent discussion about the link between matrix problems and nonlinear equation solvers are given in [12]. We state that all of such extensions in the scalar case and studying their orders are of symbolic computational nature [22].

It is often the case that scalar iterations involving derivatives of the scalar functions are not stable when applied to matrices in their simplified form. Due to this, we must investigate the stability of (2.13) for finding $S = \text{sign}(M)$ (see e.g. [16]). It is remarked that proving that a matrix iteration is stable boils down to showing that the Fréchet derivative in a neighborhood of the solution has bounded powers [9, Definition 4.17]. However, we here study how small perturbations can be controlled along the iterates. This is done as follows.

Lemma 2.1 *Let the complex matrix* $M = [m_{i,j}]_{n\times n}$ *have no pure imaginary eigenvalue, then the sequence* $\{Y_k\}_{k=0}^{k=\infty}$ *generated by* (2.13) *is stable using* $Y_0 = M$*.*

Proof. If Y_0 is a function of M, then the iterates from (2.13) are all functions of M and hence commute with M. Let Γ_k be a numerical perturbation introduced at the kth iterate of (2.13). Next, one has

$$\widetilde{Y}_k = Y_k + \Gamma_k. \quad (2.15)$$

Here, we perform a first-order error analysis, i.e., formally use approximations $(\Gamma_k)^i \approx 0$. The estimate is true as long as Γ_k is enough small. We also use the identity $(H + L)^{-1} \simeq H^{-1} - H^{-1}LH^{-1}$, for any nonsingular matrix H, arbitrary square matrix L. Using $\Gamma_{k+1} = \widetilde{Y}_{k+1} - Y_{k+1} = \widetilde{Y}_{k+1} - S$, one can verify (considering $Y_k \simeq \text{sign}(M) = S$, and $k \gg 1$):

$$\begin{aligned}\widetilde{Y}_{k+1} &= (10\widetilde{Y}_k + 104\widetilde{Y}_k^3 + 146\widetilde{Y}_k^5 + 28\widetilde{Y}_k^7) \\ &\quad \times \left[I + 43\widetilde{Y}_k^2 + 155\widetilde{Y}_k^4 + 85\widetilde{Y}_k^6 + 4\widetilde{Y}_k^8\right]^{-1} \\ &= (10(Y_k+\Gamma_k) + 104(Y_k+\Gamma_k)^3 \\ &\quad + 146(Y_k+\Gamma_k)^5 + 28(Y_k+\Gamma_k)^7) \\ &\quad \times [I + 43(Y_k+\Gamma_k)^2 + 155(Y_k+\Gamma_k)^4 \\ &\quad + 85(Y_k+\Gamma_k)^6 + 4(Y_k+\Gamma_k)^8]^{-1} \\ &\simeq (10(S+\Gamma_k) + 104(S+\Gamma_k)^3 \\ &\quad + 146(S+\Gamma_k)^5 + 28(S+\Gamma_k)^7) \\ &\quad \times [I + 43(S+\Gamma_k)^2 + 155(S+\Gamma_k)^4 \\ &\quad + 85(S+\Gamma_k)^6 + 4(S+\Gamma_k)^8]^{-1} \\ &\simeq \left(S + \frac{1}{2}\Gamma_k - \frac{1}{2}S\Gamma_k S\right), \end{aligned} \tag{2.16}$$

and also by applying the equalities $S^2 = I$ and $S^{-1} = S$ to

$$\Gamma_{k+1} \simeq 0.5\Gamma_k - 0.5S\Gamma_k S. \tag{2.17}$$

So because $\|\Gamma_{k+1}\| \leq 0.5\|\Gamma_0 - S\Gamma_0 S\|$, we have the stability. The proof is ended. □

The definition of the matrix geometric mean via (1.6) requires the computation of the matrix square root of M and its inverse. To use (2.13) for our aim, we remind an identity [9, page 108] as follows:

$$\operatorname{sign}\left(\begin{bmatrix} 0 & M \\ I & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & M^{\frac{1}{2}} \\ M^{-\frac{1}{2}} & 0 \end{bmatrix}, \tag{2.18}$$

which indicates a consequential relationship between principal matrix square root $M^{\frac{1}{2}}$ and the matrix sign function. Thus, we can compute $M^{\frac{1}{2}}$ and $M^{-\frac{1}{2}}$ at the same time using the identity (2.18) by the new scheme (2.13). In our case, M and N are HPD.

Another way for computing the matrix geometric mean of two HPD matrices without computing the principal matrix square roots, is via applying another identity [9, page 131] in what follows:

$$\operatorname{sign}\left(\begin{bmatrix} 0 & M \\ N^{-1} & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & C \\ C^{-1} & 0 \end{bmatrix}, \tag{2.19}$$

where

$$\begin{aligned} C &= M(N^{-1}M)^{-\frac{1}{2}} \\ &= M(M^{-1}N)^{\frac{1}{2}} = M\#N. \end{aligned} \tag{2.20}$$

This gives immediately the geometric mean of two HPD matrices.

In this way, one is able to propose another iteration for our purpose via matrix sign function. Although in this approach no direct computation of matrix square roots is needed, we need to compute the inverse of the HPD matrix N before starting the iteration method with a reasonable accuracy.

Now, we are able to propose a variant of the new scheme for computing the matrix geometric mean without the calculation of the matrix square root in Algorithm 1.

This algorithm must be implemented using sparse array techniques so as to save much time. To be more precise, in Algorithm 1 and even if the matrices M and N and their geometric mean are dense, but two blocks always contain zero, which this could effectively decrease the computational effort of the implementation of the proposed scheme.

It is also of requisite nature to remark that it would be favorable to still accelerate the scheme (2.13) for finding matrix sign and subsequently for matrix geometric mean. After the above discussions, it could be inferred that we could construct methods of efficient higher order or to do some scaling approach in order to accelerate the initial phase of convergence. These could be investigated for future works.

Algorithm 1. A way to calculate $M\#N$.

```
1: Considering $M$, $N$ and the initial
   approximation $Y_0 = \begin{bmatrix} 0 & M \\ N^{-1} & 0 \end{bmatrix}$
2: Apply (2.13) as long as
   $\|Y_{k+1} - Y_k\| < \epsilon$
3: Use the block $C$ based on the
   identity (2.19) using the previous step
```

3 Convergence study

This section is dedicated to the convergence properties of (2.13). Comprehensibly, when the new scheme is convergent for finding the matrix sign, then it could be used efficiently for our primary aim which is the computation of the matrix geometric mean of two complex HPD matrices.

Remark 3.1 It should be pointed out that

$\operatorname{sign}\begin{pmatrix} 0 & M+E \\ I & 0 \end{pmatrix}$ is a fixed point of the iteration (2.13) for any E such that $M+E$ is positive definite.

Theorem 3.1 *Let the complex matrix $M = [m_{i,j}]_{n\times n}$ have no pure imaginary eigenvalues. If $Y_0 = M$, then the iterative method (2.13) converges to $S = sign(M)$.*

Proof. The convergence of rational iterations can be analyzed in terms of the convergence of the eigenvalues of the matrices Y_k. The reason for this is that if Y has a Jordan decomposition $Y = ZJZ^{-1}$, then $R(Y) = ZR(J)Z^{-1}$. Let M have the following Jordan canonical form [9, page 107] $Z^{-1}MZ = \Lambda$, and the terminology $D_k = Z^{-1}Y_kZ$, and (2.13), we obtain:

$$D_{k+1} = [10D_k + 104D_k^3 + 146D_k^5 + 28D_k^7]$$
$$[I + 43D_k^2 + 155D_k^4 + 85D_k^6 + 4D_k^8]^{-1}. \quad (3.21)$$

Notice that if D_0 is a diagonal matrix then based on an inductive proof, all successive D_k are diagonal too. Now the relation (3.21) can be rewritten as n uncoupled scalar iterations to solve $g(x) = x^2 - 1 = 0$ in what follows

$$u_{k+1}^i = \frac{10u_k^i + 104{u_k^i}^3 + 146{u_k^i}^5 + 28{u_k^i}^7}{1 + 43{u_k^i}^2 + 155{u_k^i}^4 + 85{u_k^i}^6 + 4{u_k^i}^8}, \quad (3.22)$$

where $u_k^i = (D_k)_{i,i}$ and $1 \le i \le n$. From (3.21) and (3.22), it is enough to study the convergence of $\{u_k^i\}$ to $\operatorname{sign}(\lambda_i)$, for all $1 \le i \le n$. From (3.22) and since the eigenvalues of M are not pure imaginary, it is clear that $\operatorname{sign}(\lambda_i) = s_i = \pm 1$. As such, we have:

$$\frac{u_{k+1}^i - \operatorname{sign}(\lambda_i)}{u_{k+1}^i + \operatorname{sign}(\lambda_i)} =$$
$$-\frac{(-\operatorname{sign}(\lambda_i) + u_k^i)^6(\operatorname{sign}(\lambda_i) - 2u_k^i)^2}{(\operatorname{sign}(\lambda_i) + u_k^i)^6(\operatorname{sign}(\lambda_i) + 2u_k^i)^2}. \quad (3.23)$$

As $|u_0^i| = |\lambda_i| > 0$, and

$$\left|\frac{u_0^i - \operatorname{sign}(\lambda_i)}{u_0^i + \operatorname{sign}(\lambda_i)}\right| < 1. \quad (3.24)$$

This yields:

$$\lim_{k\to\infty}\left|\frac{u_{k+1}^i - \operatorname{sign}(\lambda_i)}{u_{k+1}^i + \operatorname{sign}(\lambda_i)}\right| = 0, \quad (3.25)$$

and

$$\lim_{k\to\infty} |u_k^i| = 1 = |\operatorname{sign}(\lambda_i)|. \quad (3.26)$$

As long as (3.24) is broken, we can write the following:

$$\frac{\lambda_{k+1}^i - \operatorname{sign}(\lambda_i)}{\lambda_{k+1}^i + \operatorname{sign}(\lambda_i)} =$$
$$-\frac{(-\operatorname{sign}(\lambda_i) + \lambda_k^i)^6(\operatorname{sign}(\lambda_i) - 2\lambda_k^i)^2}{(\operatorname{sign}(\lambda_i) + \lambda_k^i)^6(\operatorname{sign}(\lambda_i) + 2\lambda_k^i)^2}. \quad (3.27)$$

It reveals that $\{u_k^i\}$ has convergence. So, we obtain $\lim_{k\to\infty} D_k = \operatorname{sign}(\Lambda)$. This reveals a sixth rate of speed for the discussed iteration scheme and the proof is ended.

The iteration (2.13) requires one matrix inversion pre computing step and by using (2.18) obtains both $M^{\frac{1}{2}}$ and $M^{-\frac{1}{2}}$, which are of interest in (1.6). Notice that for computing the principal matrix square root of $(M^{-\frac{1}{2}}NM^{-\frac{1}{2}})^{\frac{1}{2}}$, we use the Jordan Canonical Form [9, page 3].

4 Numerical experiments

The high-order globally-convergent scheme (2.13) denoted by PM (via computing the $(M^{-\frac{1}{2}}NM^{-\frac{1}{2}})^{\frac{1}{2}}$) and Algorithm 1 which is denoted by AL2, are tested using Mathematica 10 in machine precision [8, chapters 2-3]. The following method (DB) [3] is also employed for the sake of comparisons

$$\begin{cases} X_0 = M, \quad L_0 = I, \quad k = 0, 1, \cdots, \\ X_{k+1} = \frac{1}{2}[X_k + L_k^{-1}], \\ L_{k+1} = \frac{1}{2}[L_k + X_k^{-1}]. \end{cases} \quad (4.28)$$

This method generates the sequences $\{X_k\}$ and $\{L_k\}$ which converge to $M^{\frac{1}{2}}$ and $M^{-\frac{1}{2}}$, respectively. Then, one may use (1.5) for computing matrix geometric mean.

We also could similarly use the Newton's method (NB) given by

$$Y_{k+1} = 0.5\left[Y_k + Y_k^{-1}\right], \quad (4.29)$$

for matrix sign with an application of (2.18). This scheme is implemented along with steps 1-3 of Algorithm 1 for finding matrix geometric mean (use (4.29) instead of (2.13)).

Since the new methods are based on the computation of matrix sign function, we compare our results with NB and DB only. Accordingly methods based on polar decomposition [10, formula (7)] are not considered for comparisons. They can be considered for numerical comparisons if we extend our high-order globally-convergent method (2.13) for polar decomposition. Furthermore, we state that there is no breakdown in case of inverting matrices in (2.13) or (4.29) since the input matrices are HPD and therefore there is no eigenvalue on the imaginary axis to make the process breakdown.

Now, we compare the behavior of different methods and report the numerical results using l_∞ for all norms involved with the stopping criterion

$$\|Y_{k+1} - Y_k\|_\infty \leq \epsilon = 10^{-6}. \tag{4.30}$$

The computer specifications are Windows 7 Ultimate, Intel(R) Core(TM) i5-4440 CPU 3.10GHz, 8.00 GB of RAM with 64-bit Operating System.

Example 4.1 *[20] We consider two HPD matrices as follows:*

$$M = \begin{pmatrix} 2 & 0 & 1 & & \\ 0 & 2 & 0 & 1 & \\ 1 & 0 & 2 & 0 & \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ & & 1 & 0 & 2 \end{pmatrix}_{n\times n},$$

$$N = \begin{pmatrix} 1.5 & \frac{2}{3} & & \\ \frac{2}{3} & \ddots & \ddots & \\ & \ddots & \ddots & \frac{2}{3} \\ & & \frac{2}{3} & 1.5 \end{pmatrix}_{n\times n},$$

when $n = 100$. The results of comparisons for different methods are given in Table 1. In the meantime, we re-examine the above matrices for higher dimensional cases. Results for these cases are gathered up in Tables 2-4.

From the numerical results presented in Tables 1-4, we observe that the accuracy of approximations to the solution increases as the iteration process (2.13) proceed, showing the stable character of the proposed method. The acquired numerical results agree with the theoretical discussions given in the Sections 2-3. The accuracy was measured via (4.30) for this test. Note also that in general the geometric mean of sparse matrices could be dense.

Table 4 includes a comparison of computational times. Accordingly, we can state that AL2 reduces the number of iterations and the computational time in finding the geometric mean favorably. The execution time of the new method AL2 in Table 4 seems to grow faster in contrast to NB, thus we expect the latter to be asymptotically better for very large scale matrices. But our proposed algorithm AL2 is a good choice for the moderate sizes. Another point is that although the two methods PM and AL2 are basically the same method applied to different matrices, they achieve much different accuracies in the computation of the geometric mean. This phenomenon occurs since the implementations (structures) of these schemes are totally different to each other as it was seen in Section 2.

	$\|Y_2 - Y_1\|$	$\|Y_3 - Y_2\|$	$\|Y_4 - Y_3\|$
PM	9.20731	0.0000554761	2.7929×10^{-14}
AL2	8.95253	0.000105972	5.81497×10^{-14}
DB	0.566718	0.126108	0.0608576
NB	80.5	39.2501	17.6667

Table 1: Numerical simulations for Experiment 4.1.

	$\|Y_3 - Y_2\|$	$\|Y_4 - Y_3\|$	$\|Y_5 - Y_4\|$
PM	0.0000554761	2.5209×10^{-14}	
AL2	0.000105972	3.85714×10^{-14}	
DB	0.126108	0.0608576	0.0218202
NB	39.2501	17.6667	5.84346

Table 2: Results of comparisons for Experiment 4.1 when $n = 200$.

	$\|Y_3 - Y_2\|$	$\|Y_4 - Y_3\|$	$\|Y_5 - Y_4\|$
PM	0.0000554761	2.5209×10^{-14}	
AL2	0.000105972	5.712×10^{-14}	
DB	0.126108	0.0608576	0.0218202
NB	39.2501	17.6667	5.84346

Table 3: Computational evidences in Experiment 4.1 for $n = 300$.

Example 4.2 *In this experiment another example of different nature has been tested. Let us consider M and N to be two Hermitian positive definite matrices obtained by taking covariance matrix of two random complex matrices of the size 50×50 in what follows (in the style of the programming package Mathematica):*

Methods	PM	AL2	DB	NB
$n = 100$	0.66	0.10	0.56	0.55
$n = 150$	1.50	0.25	1.20	1.33
$n = 200$	2.84	0.56	2.19	2.30
$n = 250$	4.01	1.02	3.60	3.44
$n = 300$	6.71	1.96	5.07	5.28
$n = 350$	8.57	3.15	6.70	7.14
$n = 400$	11.72	6.73	8.82	9.58
$n = 450$	15.58	9.79	13.06	12.23
$n = 500$	20.89	13.36	17.25	15.51
$n = 550$	24.63	17.54	20.25	18.33
$n = 600$	31.31	21.95	27.69	22.24

Table 4: The elapsed times in Experiment 4.1.

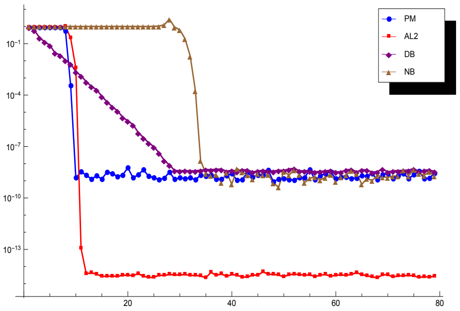

Figure 3: Error decay in Experiment 4.2.

```
n = 50; SeedRandom[1];
M = Covariance@RandomComplex[1
        + I, {n, n}];
N = Covariance@RandomComplex[2
        + I, {n, n}];
```

Here, we compare the results of different schemes in Figure 3 in terms of the residual $\frac{\|Y_{k+1}-Y_k\|_\infty}{\|Y_{k+1}\|_\infty}$ *after 80 iterations without considering a tolerance to check the stability behavior as well. The quantity which is plotted in Figure 3 is* $\frac{\|Y_{k+1}-Y_k\|_\infty}{\|Y_{k+1}\|_\infty}$.

Results are in agreement with the above discussions. The interesting point is that only AL2 can reach accuracies higher than 10^{-9}, which is an advantage of the proposed algorithm in contrast to the existing ones.

The higher accuracy of AL2 is due to its high rate of convergence while its computational time is reasonable since the process converges quickly to the matrix geometric mean and this yields in fewer computational effort in the whole of the implementation.

Example 4.3 *Using the same conditions and criterions as in Experiment 4.2, here we compare the convergence history of different methods for the following two random complex matrices of the size* 200×200*:*

```
n = 200; SeedRandom[123456];
M = Covariance@RandomComplex[300
        - 4 I, {n, n}];
N = Covariance@RandomComplex[
        20, {n, n}];
```

The results are furnished in Figure 4 in terms of the residual $\frac{\|Y_{k+1}-Y_k\|_\infty}{\|Y_{k+1}\|_\infty}$.

The results once again show that the presented schemes in this work are good choices for finding the matrix geometric mean of two HPD matrices and they support the theoretical results exposed previously.

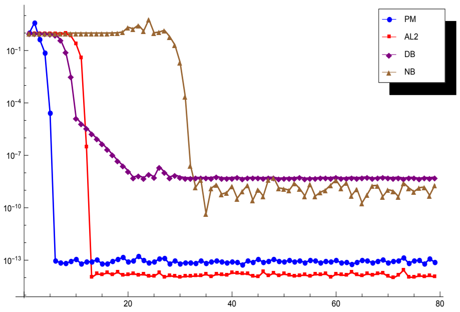

Figure 4: Error decay in Experiment 4.3.

5 Conclusion

We have extended a third-order method for solving (2.7) and obtained an iteration scheme for finding the sign function. Combining this iteration with a Newton's step, a sixth-order method (2.13) had been derived. Next, it was discussed that how the new method owns global convergence for finding matrix sign function. Using a well-known identity in matrix functions theory, we designed an algorithm for computing the geometric mean of two HPD matrices. As a matter of fact, the application of matrix sign function in computing principal matrix square root had been applied. A new algorithm has also been proposed. The asymptotical stability and error analysis of the proposed method have been studied as well. To illustrate the new technique some numerical experiments were furnished. Computational results have justified the effective convergence behavior of AL2. At last, it is stated that the extension of the new scheme for computing geometric mean of more than two matrices can be taken into account for future works.

References

[1] S. Amat, J. A. Ezquerro, M. A. Hernández-Verón, On a new family of high-order iterative methods for the matrix pth root, *Numer. Linear Algebra Appl.* 22 (2015) 585-595.

[2] R. Bhatia, Positive definite matrices, Princeton Series in Applied Mathematics, *Princeton University Press, Princeton,* (2007).

[3] E. Denman, A. Beavers, The matrix sign function and computations in systems, *Appl. Math. Comput.* 2 (1976), 63-94.

[4] A. Frommer, B. Hashemi, T. Sablik, Computing enclosures for the inverse square root and the sign function of a matrix, *Linear Algebra Appl.* 456 (2014) 199-213.

[5] O. Gomilko, D. B. Karp, M. Lin, K. Ziętak, Regions of convergence of a Padé family of iterations for the matrix sector function and the matrix p-th root, *J. Comput. Appl. Math.* 236 (2012) 4410-4420.

[6] F. Greco, B. Iannazzo, F. Poloni, The Padé iterations for the matrix sign function and their reciprocals are optimal, *Linear Algebra Appl.* 436 (2012) 472-477.

[7] N. J. Higham, The matrix sign decomposition and its relation to the polar decomposition, *Lin. Algebra Appl.* 212 (1994) 3-20.

[8] N. J. Higham, Accuracy and Stability of Numerical Algorithms, *Second Edition, SIAM, UK,* 2002.

[9] N. J. Higham, Functions of Matrices: Theory and Computation, *Society for Industrial and Applied Mathematics, Philadelphia, PA, USA,* 2008.

[10] B. Iannazzo, The geometric mean of two matrices from a computational viewpoint,*Numer. Linear Algebra Appl.* 23 (2016) 208-229.

[11] C. S. Kenney, A. J. Laub, Rational iterative methods for the matrix sign function, *SIAM J. Matrix Anal. Appl.* 12 (1991) 273-291.

[12] F. Khaksar Haghani, F. Soleymani, An improved Schulz-type iterative method for matrix inversion with application, *Tran. Inst. Measur. Cont.* 36 (2014) 983-991.

[13] D. Li, P. Liu, J. Kou, An improvement of Chebyshev-Halley methods free from second derivative, *Appl. Math. Comput.* 235 (2014) 221-225.

[14] T. Lotfi, F. Soleymani, M. Ghorbanzadeh, P. Assari, On the construction of some triparametric iterative methods with memory, *Numer. Algorithms* 70 (2015) 835-845.

[15] J. D. Lawson, Y. Lim, The geometric mean, matrices, metrics and more, *Amer. Math. Month.* 108 (2001) 797-812.

[16] M. Sh. Misrikhanov, V. N. Ryabchenko, Matrix sign function in the problems of analysis and design of the linear systems, *Auto. Remote Cont.* 69 (2008) 198-222.

[17] G. Pusz, S. L. Woronowicz, Functional calculus for sesquilinear forms and the purification map, *Rep. Math. Phys.* 8 (1975) 159-170.

[18] F. Soleymani, E. Tohidi, S. Shateyi, F. Khaksar Haghani, Some matrix iterations for computing matrix sign function, *J. Appl. Math.* (2014) 9 pages.

[19] F. Soleymani, P. S. Stanimirović, I. Stojanović, A novel iterative method for polar decomposition and matrix sign function, *Disc. Dyn. Nat. Soc.* (2015).

[20] F. Soleymani, M. Sharifi, S. Shateyi, F. Khaksar Haghani, An algorithm for computing geometric mean of two Hermitian positive definite matrices via matrix sign, *Abst. Appl. Anal.* Volume 2014, Article ID 978629, 6 pages.

[21] J. L. Varona, Graphic and numerical comparison between iterative methods, *Math. Intell.* 24 (2002) 37-46.

[22] Wolfram Research, Inc., Mathematica, Version 10.1, *Champaign, IL* (2015).

Farshad Kiyoumarsi is Assistant Professor at Shahrekord Branch, Islamic Azad University, Shahrekord, Iran. He has got his PhD degree from Shahid Bahonar University of Kerman in 2009. His research interests include computational Mathematics, numerical methods for solving integral equations, Fuzzy Data Envelopment Analysis, Ranking and fuzzy systems. Optimization, Automatic Text Summarization, and Artificial Intelligent, etc