Reading : high97

Zhengbo Zhou*

February 5, 2023

Abstract

Newton's method has been used for computing the matrix square root. However, this method is unstable. The paper rederive the stable DB iteration and derive the coupled Newton–Schulz iteration for matrix square root. Scaling method is briefly discussed in context of the new derivation. A new way for computing the square root of the positive definite matrix is given.

```
- Higham, N. J. (1997).
  Stable iterations for the matrix square root.
  Numerical Algorithms, 15(2), 227-242.
  http://dx.doi.org/10.1023/a:1019150005407
```

The Newton iteration for $A^{1/2}$ is

$$X_{k+1} = \frac{1}{2}(X_k + X_k^{-1}A), \quad X_0 = A. \qquad \lim_{k\to\infty} X_k \to A^{1/2}.$$

Although this method is converged quadratically, poor numerical stability makes it useless for practical use. The instability is explained by [2]. DB iteration [1] is derived by using $Y_k = X_k$ and $Z_k = A^{-1}X_k$, and $Y_k \to A^{1/2}$ and $Z_k \to A^{-1/2}$.

This paper rederived DB iteration and combined this with the Schulz iteration [5] to get a new iteration by using a simple identity:

Main result, [3, Lemma 1]

$$\operatorname{sign}\left(\begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & A^{1/2} \\ A^{-1/2} & 0 \end{bmatrix}$$

Then based on the standard Newton iteration for computing $\operatorname{sign}(B)$:

$$X_{k+1} = \frac{1}{2}(X_k + X_k^{-1}), \qquad X_0 = B, \qquad B = \begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix}. \tag{1}$$

This rederived the DB iteration: Define and notice

$$X_k = \begin{bmatrix} 0 & Y_k \\ Z_k & 0 \end{bmatrix} \implies X_k^{-1} = \begin{bmatrix} 0 & Z_k^{-1} \\ Y_k^{-1} & 0 \end{bmatrix}.$$

*Department of Mathematics, University of Manchester, Manchester, M13 9PL, England (zhengbo.zhou@postgrad.manchester.ac.uk).

Plug this into the (1), we have

$$X_{k+1} = \begin{bmatrix} 0 & Y_{k+1} \\ Z_{k+1} & 0 \end{bmatrix} = \frac{1}{2}\left(\begin{bmatrix} 0 & Y_k + Z_k^{-1} \\ Z_k + Y_k^{-1} & 0 \end{bmatrix}\right) \xrightarrow{k\to\infty} \begin{bmatrix} 0 & A^{1/2} \\ A^{-1/2} & 0 \end{bmatrix}$$

which is precisely the DB iteration.

Using the Schulz iteration

$$X_{n+1} = X_n(2I - AX_n)$$

which computes the A^{-1} where X_0 is any approximation of A^{-1}. By the Newton–Schulz iteration for the matrix sign function,

$$X_{k+1} = \frac{1}{2}X_k(3I - X_k^2), \quad X_0 = B,$$

and Corollary 1.34 from [4], the paper derived the coupled Newton–Schulz iteration for matrix square root

$$Y_{k+1} = \frac{1}{2}Y_k(3I - Z_kY_k), \qquad Z_{k+1} = \frac{1}{2}(3I - Z_kY_k)Z_k.$$

> Although the method is at least quadratic rates of convergence. Rapid convergence is not guaranteed in practice because the error can take may iterations to become small enough for quadratic convergence to be observed.

This can be accelerated by Newton iteration with scaling. In this paper, it provides the determinantal scaling $\gamma_k = |\det(X_k)^{-1/n}|$. This can be easily formed during the formation of X_k^{-1} by using LU decomposition. A scaled version DB iteration is also provided by using (1).

An algorithm for computing the matrix square root is provided [3, Algorithm 2] for a Hermitian positive definite matrix A: First computing the Cholesky factorisation $A = RR^*$, then compute the polar decomposition of $R = UH$ [4, Chapter 8]. The Hermitian positive definite factor of this decomposition is the principal square root of A.

Finally, the paper also provide a "condition number" of the matrix square root, $\alpha(X)$ which arises in the situation that the matrix A is not symmetric.

Consider the relative residual:

$$\mathrm{res}(X_k) = \frac{\|A - X_k^2\|_2}{\|A\|_2}.$$

We shouldn't expect the res to reach the roundoff error since if we perturb X by E where $\|E\|_2 \le \epsilon\|X\|_2$, then

$$\mathrm{res}(X + E) = \frac{\|-XE - EX - E^2\|_2}{\|A\|_2} \le \frac{(2\epsilon + \epsilon^2)\|X\|_2^2}{\|A\|_2}.$$

For symmetric matrix, res $\le 2\epsilon + O(\epsilon^2)$. For nonsymmetric matrices $\alpha(X) = \|X\|_2^2/\|A\|_2$ can be arbitrarily large.

References

[1] Eugene D. Denman and Alex N. Beavers. The matrix sign function and computations in systems. *Applied Mathematics and Computation*, 2(1):63–94, 1976. (Cited on p. 1.)

[2] Nicholas J. Higham. Newton's method for the matrix square root. *Mathematics of Computation*, 46(174):537–549, 1986. (Cited on p. 1.)

[3] Nicholas J. Higham. Stable iterations for the matrix square root. *Numerical Algorithms*, 15(2):227–242, 1997. (Cited on pp. 1 and 2.)

[4] Nicholas J. Higham. *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008. xx+425 pp. ISBN 978-0-898716-46-7. (Cited on p. 2.)

[5] Günther Schulz. Iterative Berechung der reziproken Matrix. *ZAMM - Zeitschrift für Angewandte Mathematik und Mechanik*, 13(1):57–59, 1933. (Cited on p. 1.)