East Asian Mathematical Journal
Vol. 29 (2013), No. 5, pp. 511–519
http://dx.doi.org/10.7858/eamj.2013.035

THE BINOMIAL METHOD FOR A MATRIX SQUARE ROOT

Yeon-Ji Kim, Jong-Hyeon Seo, and Hyun-Min Kim†

Abstract. There are various methods for evaluating a matrix square root, which is a solvent of the quadratic matrix equation $X^2 - A = 0$. We consider new iterative methods for solving matrix square roots of M-matrices. Particulary we show that the relaxed binomial iteration is more efficient than Newton-Schulz iteration in some cases. And we construct a formula to find relaxation coefficients through statistical experiments.

1. Introduction

Any nonsingular matrix $A \in \mathbb{C}^{n\times n}$ has a square root. The number of square roots varies from two to infinite. If A is singular, the existence of a square root depends on the Jordan structure of the zero eigenvalues [1], [4], [7]. For a matrix A having no non-positive real eigenvalues, it has a unique square root for which every eigenvalue has positive real part. This square root is called the principal square root and denoted by $A^{1/2}$. In other words, $A^{1/2}$ is defined by $(A^{1/2})^2 = A$ and $\mathrm{Re}\lambda_k(A^{1/2}) > 0$ for all k, where $\lambda_k(A)$ denotes an eigenvalue of A [4], [9].

We consider some numerical methods for finding matrix square roots of M-matrices. The binomial iteration and the Newton-Schulz iteration were compared by Higham [5]. We propose a new algorithm, which is modified and is called the relaxed binomial iteration. In the Newton iteration, we must find inverses of some matrices [3]. So we apply the Schulz iteration for finding inverse matrices. Furthermore, we let the sign matrix of A as the starting matrix [8], [10]. In this way, we get the Newton-Schulz iteration. Newton-Schulz iteration for matrix square root:

$$Y_0 = A,\ Z_0 = I,$$

Received October 16, 2013; Accepted October 24, 2013.
2000 *Mathematics Subject Classification.* 58B34, 58J42, 81T75.
Key words and phrases. matrix square root, M-matrix, binomial iteration, Newton-Schulz iteration.
*This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2012R1A1A2008840).
†Corresponding author.

$$\left.\begin{aligned} Y_{k+1} &= \tfrac{1}{2}Y_k(3I - Z_kY_k) \\ Z_{k+1} &= \tfrac{1}{2}(3I - Z_kY_k)Z_k \end{aligned}\right\} \tag{1.1}$$

$$Y_k \to A^{\frac{1}{2}},\ Z_k \to A^{-\frac{1}{2}} \quad \text{as } k \to \infty.$$

Definition 1.1. [6] For any square matrices A and B we write $A \geq B(A > B)$ if $[A]_{ij} \geq [B]_{ij}([A]_{ij} > [B]_{ij})$ for all i, j. Then we call that A is nonnegative(positive) matrix if $A \geq 0(A > 0)$.

Definition 1.2. [6] The set of all $\lambda \in \mathbb{C}$ that are eigenvalues of $A \in M_n$ is called the spectrum of A and is denoted by $\sigma(A)$. The spectral radius of A is the nonnegative real number $\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$. This is just the radius of the smallest disc centered at the origin in the complex plane that includes all the eigenvalues of A.

Definition 1.3. [2] A matrix $A \in \mathbb{R}^{n\times n}$ is an M-matrix if $A = rI_n - B$ for some nonnegative matrix B and r with $r \geq \rho(B)$ where $\rho(B)$ is the spectral radius of B; it is a singular M-matrix if $r = \rho(B)$ and a (nonsingular) M-matrix if $r > \rho(B)$.

Higham [5] introduced the binomial iteration for finding a square root of M-matrices. For an M-matrix A, it can be represented by $A = rI_n - B$ for some nonnegative B and r with $r \geq \rho(B)$. Then we obtain $A^{\frac{1}{2}} = r^{\frac{1}{2}}(I_n - C)^{\frac{1}{2}}$ with $C = \frac{1}{r}B$. By the binomial expansion,

$$(I_n - C)^{\frac{1}{2}} = \sum_{j=0}^{\infty} \binom{\frac{1}{2}}{j}(-C)^j \equiv I - \sum_{j=1}^{\infty} \alpha_j C^j,\ \alpha_j > 0,$$

which is valid when $\rho(C) < 1$. If $\rho(C) \geq 1$, then we have to choose r that decrease $\rho(C) < 1$.

Assume that $\rho(C) < 1$ then we obtain $(I - C)^{\frac{1}{2}} =: I - P$. Squaring the equation then we have

$$I - C = I - 2P + P^2. \tag{1.2}$$

From the equation (1.2), we obtain a iteration for computing P.

$$P_{k+1} = \frac{1}{2}(C + P_k^2),\ P_0 = 0. \tag{1.3}$$

2. Convergence of the Relaxed Binomial Iteration

The relaxed binomial iteration can be derived by choosing reasonable coefficient and has the form

$$R_\lambda(X) = (1 + \lambda)F(X) - \lambda X,\ 0 < \lambda < 1 \tag{2.1}$$

where $0 < \lambda < 1$ and $F(X) = \frac{1}{2}(X^2 + C)$.

It is well known that if $\rho(C) < 1$, then the binomial iteration converges. In this section we consider when the relaxed binomial iteration converges. First, we present some lemmas to prove the convergence of iteration (2.1).

Lemma 2.1. *Let $F(X) = \frac{1}{2}(X^2 + C)$. If $X \leq Y$, then $F(X) < F(Y)$.*

Proof.

$$\begin{aligned} F(Y) - F(X) &= \frac{1}{2}(Y^2 + C) - \frac{1}{2}(X^2 + C) \\ &= \frac{1}{2}(Y^2 - X^2) \\ &= \frac{1}{2}(Y^2 - YX + YX - X^2) \\ &= \frac{1}{2}(Y(Y - X) + (Y - X)X) \geq 0. \end{aligned}$$

□

Lemma 2.2. *Let $F(X) = \frac{1}{2}(X^2 + C)$ and $R_\lambda(X) = (1 + \lambda)F(X) - \lambda X$ where $0 < \lambda H < XH$ for $H > 0$. If $X \leq F(X)$ and $XF(X) = F(X)X$, then $F(X) \leq R_\lambda(X) \leq F(F(X))$.*

Proof.

$$\begin{aligned} R_\lambda(X) - F(X) &= (1 + \lambda)F(X) - \lambda X - F(X) \\ &= \lambda(F(X) - X) \\ &= F(X + H) \\ &\geq F(X) + XH. \end{aligned}$$

Put $H = F(X) - X$

$$\begin{aligned} F(X + H) &= F(X + F(X) - X) = F(F(X)) \\ &\geq F(X) + X(F(X) - X) \\ &\geq F(X) + \lambda(F(X) - X) \\ &= R_\lambda(X). \end{aligned}$$

□

The following theorem is related to the convergence of the binomial iteration for computing a square root of M-matrices.

Theorem 2.3. [5, Thm. 6.13] *Let $C \in \mathbb{R}^{n\times n}$ satisfy $C \geq 0$ and $\rho(C) < 1$ and write $(I - C)^{1/2} = I - P$. Then in the binomial* (1.3), *$P_k \to P$ with $0 \leq P_k \leq P_{k+1} \leq P$, $k \geq 0$; that is, P_k converges monotonically to P.*

By using Theorem 2.3, convergence of the relaxed binomial iteration can be proved.

Theorem 2.4. *Let $C \in \mathbb{R}^{n\times n}$ satisfy $C \geq 0$ and $\rho(C) < 1$ and write $(I - C)^{1/2} = I - P$. Then for the iteration (2.1) with $Y_0 = 0$ the sequence Y_k is well-defined and $Y_0 \leq Y_1 \leq \cdots \leq P$ where P is a nonnegative solvent of relaxed binomial iteration.*

Proof. For the iteration with $Y_0 = 0$, we obtain $Y_0 \leq F(Y_0) \leq Y_1$, $Y_0 \leq P_*$ and $Y_0C = CY_0$.

Therefore, the statement

$$Y_k \leq F(Y_k) \leq Y_{k+1},\ Y_k \leq P,\ Y_kC = CY_k \tag{2.2}$$

is true for $k = 0$.

Suppose the statement is true for $k = i$

$$\begin{aligned} Y_{i+1}C &= R_\lambda(Y_i)C \\ &= \{(1+\lambda)F(Y_i) - \lambda Y_i\}C \\ &= (1+\lambda)F(Y_i)C - \lambda Y_iC \\ &= (1+\lambda)CF(Y_i) - \lambda CY_i \\ &= C\{(1+\lambda)F(Y_i) - \lambda Y_i\} \\ &= CY_{i+1}. \end{aligned}$$

By Lemma 2.1 and Lemma 2.2, we have $F(F(Y_i)) \leq F(Y_{i+1})$ and $F(Y_i) \leq R_\lambda(Y_i) \leq F(F(Y_i))$. Since $R_\lambda(Y_i) = Y_{i+1}$, then $Y_{i+1} \leq F(Y_{i+1})$. It follows form Lemma 2.2 that

$$F(Y_{i+1}) \leq R(Y_{i+1}) \leq F(F(Y_{i+1})).$$

Therefore, $Y_{i+1} \leq Y_{i+2}$.

Since $Y_i \leq Y_*$, $F(P) = P$ and Lemma 2.1, $F(Y_i) \leq F(F(Y_i)) \leq P$. By Lemma2.2, we have

$$F(Y_i) \leq Y_{i+1} \leq F(F(Y_i)) \leq P.$$

Therefore, $Y_{i+1} \leq P$.

Since the sequence $\{Y_k\}$ is monotone nondeceasing and bounded above, then the sequence has a limit, P_*. This limit satisfies $P_* = \frac{1}{2}(C + P_*^2)$. By squaring this equation, $(I - P_*)^2 = I - C$. Since $\rho(P_k) \leq \rho(P) < 1$ for all k, so $\rho(P_*) < 1$. Therefore, $I - P_*$ is the principal square root. In other words, $P_* = P$. □

3. Finding Relaxation coefficients

Higham [5] demonstrated the usefulness of the Newton-Schulz iteration for solving a square root of M-matrices . In Section 2, we confirm that the relaxed binomial iteration converges with some relaxation coefficients. Our purpose in this section is to show experimentally the benefits of the relaxed binomial iteration and the method of finding an appropriate relaxation coefficient. Our experiments were done in MATLAB. Through some experiments we see that there is correlation between relaxation coefficients and $\rho(C)$. So we construct a relation through statistical experiments. In the tables, we use "IT" to denote the number of iteration steps and $\mu = \frac{|\lambda_{max}|+|\lambda_{min}|}{2}$. The stopping criterion for each iteration is $\|X_k - X_{k-1}\| < 10^{-5}$, where X_k is the k-th iteration value.

Figure 3.1 and Figure 3.2 give the result that the relaxed binomial method can be more efficient according to the value of relaxation coefficient, in the sense of flops.

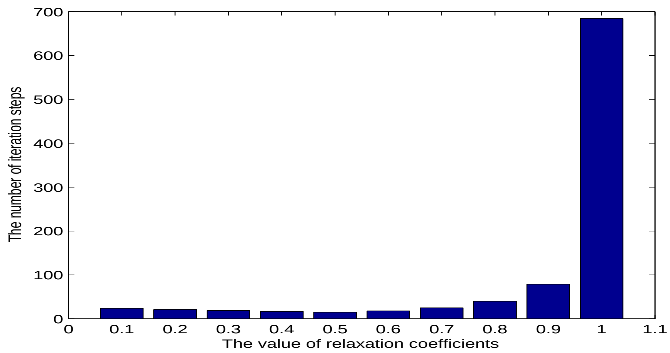

FIGURE 3.1. γ and the number of iteration steps

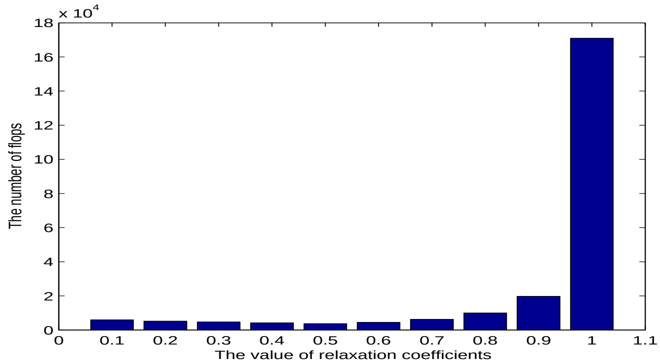

FIGURE 3.2. γ and the number of flops

So we see that it is important to find an appropriate relaxation coefficient. Like Figure 3.1 and Figure 3.2, we can obtain a proper relaxation coefficient by repeating experiments to substitute γ from 0.0001 to 0.9999 in the order.

But this method is too expansive when matrix size is significantly large. So, we want to estimate $\hat{\gamma}$. To find it, we construct a least square problem. We want to find the output γ to be a linear function of the input μ. That is, We find a straight line $\gamma = \alpha + \beta\mu$ from solving the least square problem [11].

$$\begin{bmatrix} 1 & \mu_1 \\ 1 & \mu_2 \\ \vdots & \vdots \\ 1 & \mu_k \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_k \end{bmatrix}$$

In the equation, $\mu_i = \frac{|\lambda_{i_{max}}|+|\lambda_{i_{min}}|}{2}$ and γ_i is the optimal relaxation coefficient that is most reduce the number of iteration steps. The optimal γ can be obtained by repetition method.

Through 100 times experiments, we can get $\hat{\gamma} = -0.5621 + 1.9848\mu$. Then, we need to confirm that $\hat{\gamma}$ obtained from linear function is a good estimate from an experiment.

Example 3.1.

$$A = \begin{bmatrix} 2.052 & -0.24106 & -0.021699 & -0.9913 & -0.28753 \\ -0.13479 & 1.4345 & -0.15953 & -0.71203 & -0.060941 \\ -0.22333 & -0.3911 & 1.5175 & -0.87136 & -0.26247 \\ -0.39655 & -0.51126 & -0.87915 & 1.8824 & -0.18626 \\ -0.13514 & -0.092896 & -0.18699 & -0.496 & 1.4449 \end{bmatrix}$$

From definition of M-matrix, A can be represented as $A = 2.2140(I - C)$, then $C = I - \frac{1}{2.2140}A$.

$$C = \begin{bmatrix} 0.073194 & 0.10888 & 0.0098006 & 0.44774 & 0.12987 \\ 0.060879 & 0.35209 & 0.072056 & 0.3216 & 0.027525 \\ 0.10087 & 0.17665 & 0.3146 & 0.39356 & 0.11855 \\ 0.17911 & 0.23092 & 0.39708 & 0.1498 & 0.084127 \\ 0.061039 & 0.041958 & 0.084456 & 0.22403 & 0.34738 \end{bmatrix}$$

Applying the relaxed binomial iteration to this matrix with the optimal γ and $\hat{\gamma}$ respectively. From Table 1, we see the optimal γ and $\hat{\gamma}$ are not same but they have an effect on method similarly. Thus, we know $\hat{\gamma}$ is a good estimate.

	$\mu = 0.5044$	
Method	IT	Flops
RBI with $\gamma_{opt} = 0.4327$	21	5250
RBI with $\hat{\gamma} = 0.4390$	22	5500

TABLE 1. The result of Example 3.1

4. Experiments

In this section, we give some examples and compare the Newton-Schulz iteration, the binomial iteration and the relaxed binomial iteration.

Example 4.1. Consider the matrix A of size $n = 5$, we write $C = I - \frac{1}{3.0655}A$ to make $\rho(C) < 1$.

$$A = \begin{bmatrix} 2.3351 & -0.7984 & -0.5456 & -0.8364 & -0.1340 \\ -0.9249 & 2.1492 & -0.2843 & -0.1453 & -0.8848 \\ -0.6295 & -0.9811 & 2.2134 & -0.1715 & -0.5147 \\ -0.8783 & -0.0960 & -0.0647 & 2.5162 & -0.9636 \\ -0.6417 & -0.5275 & -0.5448 & -0.8240 & 2.4637 \end{bmatrix}$$

$$C = \begin{bmatrix} 0.2383 & 0.2604 & 0.1780 & 0.2728 & 0.0437 \\ 0.3017 & 0.2989 & 0.0928 & 0.0474 & 0.2886 \\ 0.2054 & 0.3201 & 0.2780 & 0.0560 & 0.1679 \\ 0.2865 & 0.0313 & 0.0211 & 0.1792 & 0.3144 \\ 0.2093 & 0.1721 & 0.1777 & 0.2688 & 0.1963 \end{bmatrix}$$

Then three methods apply for the matrix C, we can get the result Table 2.

	$\mu = 0.5227$	
Method	IT	Flops
RBI with $\hat{\gamma} = 0.4754$	37	9250
BI	57	14250
Newton-Schulz	13	9750

TABLE 2. The results of Example 4.1

In Table 2, the Newton-Schulz iteration has 13 times iteration steps small compared with the relaxed binomial iteration. But the Newton-Schulz iteration has more flops than the relaxed binomial iteration because the Newton-Schulz iteration requires three matrix multiplications per iteration versus one for the relaxed binomial iteration. Thus we can see that the relaxed binomial iteration is more efficient than Newton-Schulz iteration for calculating a square root of M-matrices.

Example 4.2. We compare three methods by applying for an M-matrix A of size $n = 20$. Table 3 presents that we get the same result when matrix size is large.

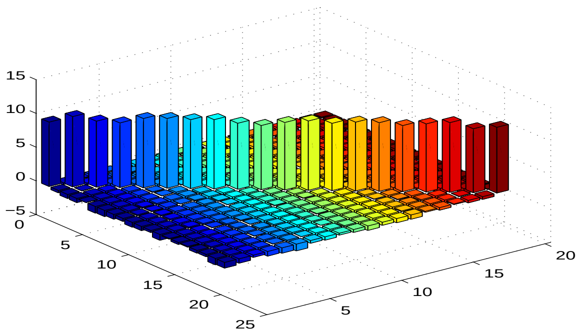

FIGURE 4.1. M-matrix A of size $n = 20$

	$\mu = 0.5059$	
Method	IT	Flops
RBI with $\hat{\gamma} = 0.4420$	17	4250
BI	28	6250
Newton-Schulz	8	6000

TABLE 3. The results of Example 4.2

References

[1] G.W. Cross and P.Lancaster. Square roots of complex matrices. *Linear and Multilinear Algebra*, 1:289–293, 1974.

[2] C.-H. Guo and N. J. Higham. Iterative solution of a nonsymmetric algebraic Riccati equation. *SIAM J. Matrix Anal. Appl.*, 29:396–412, 2007.

[3] N. J. Higham. Newton's method for the matrix square root. *Math. Comput.*, 46(174):537–549, 1986.

[4] N. J. Higham. Stable iterations for the matrix square root. *Numer. Algorithms*, 15:227–242, 1997.

[5] N. J. Higham. *Functions of matrices: theory and computation*. SIAM, 2008.

[6] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, 1985.

[7] R. A. Horn and C. R. Johnson. *Topics in matrix analysis*. Cambridge University Press, 1994.

[8] Charles Kenney and Alan J. Laub. Rational iterative methods for the matrix sign function. *SIAM Journal on Matrix Analysis and Applications*, 12:273–291, 1991.

[9] Hyun-Min Kim. Numerical methods for solving a quadratic matrix equation. Master's thesis, Department of Mathematics, University of Manchester, 2000.

[10] J. D. ROBERTS. Linear model reduction and solution of the algebraic riccati equation by use of the sign function. *International Journal of Control*, 32:677–687, 1980.

[11] Gilbert Strang. *Linear Algebra and Its Applications*. Thomson Brooks/Cole Cengage learning, 2006.

Yeon-Ji Kim
Department of Mathematics, Pusan National University, Busan, 609-735, Republic of Korea
E-mail address: kyj14jeff@pusan.ac.kr

Jong-Hyeon Seo
Department of Mathematics, Pusan National University, Busan, 609-735, Republic of Korea
E-mail address: hyeonni94@pusan.ac.kr

Hyun-Min Kim
Department of Mathematics, Pusan National University, Busan, 609-735, Republic of Korea
E-mail address: hyunmin@pnu.edu