NEWTON SCHULZ METHOD FOR SOLVING NONLINEAR MATRIX EQUATION $X^p + A^*XA = Q$

YOUNG-JIN KIM*, JIE MENG, AND HYUN-MIN KIM

ABSTRACT. The matrix equation $X^p + A^*XA = Q$ has been studied to find the positive definite solution in several researches. In this paper, we consider fixed-point iteration and Newton's method for finding the matrix p-th root. From these two considerations, we will use the Newton-Schulz algorithm(N.S.A). We will show the residual relation and the local convergence of the fixed-point iteration. The local convergence guarantees the convergence of N.S.A. We also show numerical experiments and easily check that the N.S. algorithm reduce the CPU-time significantly.

1. INTRODUCTION

In this paper, we will consider the matrix equation

$$(1.1) \qquad X^p + A^*XA = Q,$$

where p is a positive integer, $A, Q \in \mathbb{C}^{n \times n}$ and Q is a Hermitian positive definite matrix.

The existence of Hermitian positive definite solutions of equation (1.1) has been investigated in some cases. Put $Y = X^p$, then equation (1.1) is equivalent to $Y + A^*Y^{\frac{1}{p}}A = Q$, which is a special example of equation

$$(1.2) \qquad X + A^*\mathscr{F}(X)A = Q.$$

Ran and Reurings [4] proved that under the assumption the function $\mathscr{F}(\cdot)$ is monotone and $Q - A^*\mathscr{F}(Q)A$ is positive definite, the positive semi-definite solutions of equation (1.2) exist. EL-Sayed and Ran [1] proved that if $\mathscr{F}$ maps positive definite matrices either into positive definite matrices or into negative definite matrices, and satisfies some monotonicity property, then under some conditions an iteration method converges to a positive definite solution of equation (1.1). See also [6] for the linear matrix equation when $p = 1$.

Jia and Wei [3] studied the matrix equation $X^s + A^TX^tA = Q$ which is also equivalent to equation(1.1) if we put $Y = X^t$ and $p = s/t$, where s and t are both nonnegative integers, A and Q are $n \times n$ real matrices and Q is symmetric positive definite. They proved that the equation has a symmetric positive definite solution

if $\lambda_{\max}(A^T A) \leq \lambda_{\min}(Q)(\lambda_{\max}(Q))^{-\frac{t}{s}}$. When $Q = I$ and A is invertible, one can see the approach outlined in [2, 7, 8, 9]. They considered matrix p-th root for the way to find the solution.

In [5], Meng and Kim considered the following basic fixed-point iteration by using the matrix p-th root for finding the Hermitian positive definite solution of equation (1.1):

$$(1.3) \qquad\qquad X_{k+1} = (Q - A^* X_k A)^{\frac{1}{p}}, \quad X_0 = I.$$

The method (1.3) for finding the matrix p-th root is arbitrary, so we used Built-in function in MATLAB R2014a.

Many researchers have developed methods for finding the following p-th root of a matrix A.

$$X^p = A.$$

We call X as the principal p-th root of A when the eigenvalues of X lie in the segment $\{z \in \mathbb{C} - \{0\} : -\pi/p < arg(z) < \pi/p\}$. One of the applications of p-th root is in the computation of the matrix logarithm through the relation [20, 21]

$$\log A = p \log A^{\frac{1}{p}},$$

where p is chosen so that $A^{\frac{1}{p}}$ can be well approximated by a polynomial or rational function. Hoskins and Walton[22] consider the iteration

$$X_{k+1} = \frac{1}{p}\Big((p-1)X_k + AX_k^{1-p}\Big), \quad X_0 = A,$$

which is Newton's method for $X^p - A = 0$ simplified by using the commutativity relation $X_k A = A X_k$. They focused on symmetric positive definite matrices that X_k defined by the above iteration converges to p-th root of A. However, for more general A the iteration does not generally converge to $A^{\frac{1}{p}}$, as explained by Smith[16].

When we find p-root of a matrix, the Newton sequence is defined as following :

$$X_{k+1} = \frac{1}{p}\Big((p-1)X_k + AX_k^{1-p}\Big), \quad X_0 = I.$$

We can also find the solution of the p-th root at every step in (1.3) by using the above Newton's method.

The Newton's method has several versions and modifications. In [14, 15], Iannazzo suggest stable versions of Newton's method. The inverse Newton's method suggested in [12, 13, 17]. The Schur Newton method was proposed by Smith[16] and it was developed by Guo and Higham[13]. Halley's method was suggested in [15], and Guo has explained the residual relation for the Newton's method and Halley's method in [10]. In that paper, Guo also showed the existence of p-th root of M-matrices and H-matrices. The residual is defined by $R(X) = I - AX^{-p}$ through $R(X) = A - X^p$ in [13].

In this paper, we will show a way of finding the sulution of 1.1. This algorithm is motivated by Newton' method for matrix p-th root. consider the following algorithm (1.4) to find the p-th root. We will set X_k and B_k as shown below and we will show the residual relation, local convergence of the fixed point iteration (1.3) and numerical experiment in later sections.

$$(1.4) \quad X_{k+1} = \frac{1}{p}\Big((p-1)X_k + B_k X_k^{1-p}\Big) \qquad \text{where} \quad B_k = Q - A^* X_k A, X_0 = I.$$

2. Residual relation for Newton Schulz Algorithm

In this section, we consider the Newton's method and show the residual relation for convergence. We also assume following conditions.

$$\begin{cases} 1. & AQ = QA \\ 2. & \rho(I - Q + A^*A) \leq 1 \end{cases}$$

We also consider the residual of the p-th root of a matrix A as following form :

$$(2.1) \qquad\qquad R(X) = I - AX^{-p}.$$

Guo showed the residual relation for the following Newton's method to find the matrix p-th root of A in [10]. So we will consider the next remark to prove the N.S. algorithm.

Remark 2.1. Assume that $\rho(I - A) \leq 1$, where $\rho(\cdot)$ denotes the spectral radius. Then the Newton sequence is well defined by (1) and

$$R(Y_{k+1}) = \sum_{i=2}^{\infty} c_i (R(Y_k))^i,$$

where $c_i > 0$ for $i \geq 2$ and $\sum_{i=2}^{\infty} c_i = 1$. Moreover, if $\|R(Y_0)\| = \|I - A\| \leq 1$ for a sub-multiplicative matrix norm $\|\cdot\|$, then for each $k \geq 0$

$$(2.2) \qquad\qquad \|R(Y_{k+1})\| \leq \|(R(Y_0))^{2^k}\| \leq \|R(Y_0)\|^{2^k}.$$

From the above remark, we can show the residual relation for (1.3) at every step. We let $Y_{k,i}$ be the Newton sequence for finding the matrix p-th root of B_k with the starting value $Y_{k,0} = X_k$.

$$(2.3) \qquad Y_{k,i+1} = \frac{1}{p}\Big((p-1)Y_{k,i} + B_k Y_{k,i}^{1-p}\Big), \quad Y_{k,0} = X_k.$$

Then X_{k+1} in (1.4) is the first element of the Newton sequence $Y_{k,i}$, i.e. $X_{k+1} = Y_{k,1}$

We define the residual for finding the matrix p-th root of B_k as following :

$$(2.4) \qquad\qquad R_k(X) = I - B_k X^{-p}.$$

At first, we show the commutativity.

Lemma 2.2. *If we set $X_0 = I$, then $X_i X_j = X_j X_i, \forall k = 0, 1, \cdots$.*

Proof. Since $X_0 = I$ and $AQ = QA$, we have

$$(Q - A^* X_k A)X_k = X_k(Q - A^* X_k A),$$

which implies that

$$X_{k+1}X_k = \frac{1}{p}\Big((p-1)X_k + (Q - A^*X_kA)X_k^{1-p}\Big)X_k = X_kX_{k+1}.$$

$\square$

Next, we will show the residual relation to prove the convergence of the sequence $\{X_k\}$.

Theorem 2.3. *Consider the matrix equation (1.1) and if $AQ = QA$, $\rho(I - Q + A^*A) \leq 1$ and $\rho(Q) \geq \rho(A^*A)$, then the iteration (2.3) satisfies following residual relation :*

$$\|R_k(Y_{k,1})\| \leq \|R_k(Y_{k,0})\|^2 \leq \|R_{k-1}(Y_{k-1,1})\|^2 \leq \|R_{k-1}(Y_{k-1,0})\|^{2 \cdot 2} \cdots \leq \|R_0(Y_{0,0})\|^{2^k}.$$

Which means that

$$\|R_k(X_k)\| \leq \|R_{k-1}(X_{k-1})\|^2 \leq \cdots \leq \|R_0(X_0)\|^{2^k}.$$

Proof. $Y_{0,0} = X_0 = I$ is nonsingular and $\rho(X_0) \leq 1$. Since,

$$\begin{aligned}
Y_{k+1,0} = Y_{k,1} = X_{k+1} &= \frac{1}{p}\Big((p-1)X_k + B_kX_k^{1-p}\Big) \\
&= X_k\Big(I - \frac{1}{p}I + \frac{1}{p}B_kX_k^{-p}\Big) \\
&= X_k\Big(I - \frac{1}{p}(I - B_kX_k^{-p})\Big) \\
&= X_k\Big(I - \frac{1}{p}R_k(X_k)\Big).
\end{aligned}$$

If $Y_{k,0} = X_k$ is nonsingular and $\rho(X_k) \leq 1$, then X_{k+1} is also nonsingular and $\rho(X_{k+1}) \leq \rho(X_k)$ when $\rho(R_k(X_k)) \leq 1$.

If $k = 0$, then $B_0 = Q - A^*A$ and $\rho(I - B_0) = \rho(R_0(X_0)) = \rho(I - Q + A^*A) \leq 1$ from the assumption. Thus, the Newton sequence $\{Y_{0,i}\}$ is well-defined, $Y_{0,1}(= X_1 = Y_{1,0})$ is nonsingular and $\rho(Y_{0,1})(= \rho(X_1) = \rho(Y_{1,0})) \leq 1$. Since the relation $\|R_k(Y_{k,1})\| \leq \|R_k(Y_{k,0})\|^2$ is obvious from the Remark 2.1 where $\rho(I - B_k) \leq 1$, so $\|R_0(Y_{0,1})\| \leq \|R_0(Y_{0,0})\|^2$.

We will prove that

$$\|R_{k+1}(Y_{k+1,0})\| = \|R_{k+1}(X_{k+1})\| \leq \|R_k(X_{k+1})\| = \|R_k(Y_{k,1})\|, \ \forall \ k = 0, 1, \cdots$$

where $\rho(R_k(X_k)) \leq 1$.
Since

$$\rho(A^*X_{k+1}A) = \rho(A^*X_k(I - \frac{1}{p}R_k(X_k))A) \leq \rho(A^*X_kA) \leq \rho(A^*A),$$

then

$$\rho(R_{k+1}(Y_{k+1,0})) = \rho(I - QX_{k+1}^{-p} + A^*X_{k+1}AX_{k+1}^{-p})$$
$$= \rho(I + A^*X_k(I - \frac{1}{p}R_k(X_k))AX_{k+1}^{-p} - QX_{k+1}^{-p})$$
$$\leq \rho(I + A^*X_kAX_{k+1}^{-p} - QX_{k+1}^{-p})$$
$$= \rho(R_k(Y_{k,1})).$$

Thus, $\|R_{k+1}(Y_{k+1,0})\| \leq \|R_k(Y_{k,1})\|$.

If $\rho(I - B_k) \leq 1$, then

$$\rho(I - B_{k+1}) = \rho(I - (Q - A^*X_{k+1}A))$$
$$= \rho(I - Q + A^*X_{k+1}A)$$
$$= \rho(I - Q + A^*X_k(I - \frac{1}{p}R_k(X_k))A)$$
$$\leq \rho(I - Q + A^*X_kA)$$
$$= \rho(I - B_k).$$

Thus, $\|R_k(Y_{k,1})\| \leq \|R_k(Y_{k,0})\|^2 \leq \|R_{k-1}(Y_{k-1,1})\|^2 \leq \cdots \leq \|R_0(Y_{0,0})\|^{2^k}.$ $\square$

3. Local convergence

In this section, we show the local convergence for iteration (1.3). The local convergence will guarantee the convergence for (1.4). So we consider the matrix function

$$F(X) = (Q - A^*XA)^{\frac{1}{p}},$$

where p is a positive integer, $A, Q \in \mathbb{C}^{n \times n}$ and Q is a Hermitian positive definite matrix.

One-step stationary iterations have the form

(3.1) $$X_{k+1} = F(X_k), \quad k = 0, 1, 2, \ldots,$$

where $F : \mathbf{D} \subset \mathbb{R}^{m \times m} \to \mathbb{R}^{m \times m}$.

Definition 3.1. A matrix function $F : \mathbf{D} \subset \mathbb{R}^{m \times m} \to \mathbb{R}^{m \times m}$ is *contractive* on a set $\mathbf{D}_0 \subset \mathbf{D}$ if there is an $\alpha < 1$ such that $\|F(X) - F(Y)\| \leq \alpha\|X - Y\|$ for all $X, Y \in \mathbf{D}_0$.

Theorem 3.2 (Contraction Mapping Theorem: version 1). *Let $F : \mathbf{D} \subset \mathbb{R}^{m \times m} \to \mathbb{R}^{m \times m}$, and suppose that F maps a closed set $\mathbf{D}_0 \subset \mathbf{D}$ into itself and is contractive. Then, F has the unique fixed point in $\mathbf{D}_0$.*

Theorem 3.3 (Local Convergence Theorem). [19] *Let $F : \mathbf{D} \subset \mathbb{R}^{m \times m} \to \mathbb{R}^{m \times m}$, and suppose that there exists a ball $\mathcal{B} := \mathcal{B}(S, \delta) \subset \mathbf{D}$ and a constant $\alpha < 1$ such that*

$$\|F(X) - S\| \leq \alpha\|X - S\|, \quad \text{for all } X \in \mathcal{B}.$$

Then, for any $X_0 \in \mathcal{B}$, the iteration from (3.1) remains in $\mathcal{B}$ and converges to S and

$$\limsup_{k \to \infty} \sqrt[k]{\|X_k - S\|} \leq \alpha.$$

In Theorem 3.3, since X_0 is arbitrary, F is invariant on $\mathcal{B}$ (i.e. F maps into itself).

From the above theorems, we have the following theorem.

Theorem 3.4. [24] *Let $S \in \mathbb{R}^{m \times m}$ be a fixed point of $F : \mathbf{D} \subset \mathbb{R}^{m \times m} \to \mathbb{R}^{m \times m}$ and suppose that there exist balls $\mathcal{B}_1 := \mathcal{B}(S, \delta) \subset \mathbf{D}$ and $\mathcal{B}_2 := \mathcal{B}(S, \delta') \subset \mathcal{B}_1$ such that*

$$
\begin{aligned}
\|F(X) - F(Y)\| &\leq& \Gamma(X,Y) \cdot \|X - Y\|, &\quad \forall\, X, Y \in \mathbf{D}, \\
\Gamma(X, S) &\leq& \bar{\mu}(\|X - S\|) < \bar{\mu}(\delta) \leq 1, &\quad \forall\, X \in \mathcal{B}_1, \\
\Gamma(X, Y) &\leq& \bar{\nu}(\delta', \delta') < 1, &\quad \forall\, Y, Z \in \mathcal{B}_2,
\end{aligned}
$$

and

$$\bar{\mu}(0) = \lim_{x \to 0} \bar{\mu}(x) = \mu < 1,$$

where $\Gamma(X, Y)$, $\bar{\mu}$ and $\bar{\nu}$ are real nonnegative valued functions on $\mathbf{D} \times \mathbf{D}$, $[0, \delta]$ and $[0, \delta'] \times [0, \delta']$ respectively, and $\bar{\mu}$ is increasing on $[0, \delta]$. Then, for any $X_0 \in \mathcal{B}_1$, the sequence $\{X_k\}$ generated by (3.1) converges to S. Moreover, S is the unique fixed point of F in $\mathcal{B}_2$ and

$$\limsup_{k \to \infty} \sqrt[k]{\|X_k - S\|} \leq \alpha.$$

Notation. For given matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{r \times q}$, the Kronecker product $A \otimes B$ is an $mr \times nq$ matrix.

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{bmatrix}.$$

The operator $\mathrm{vec}(A)$ represents a column vector from a matrix A:

$$\mathrm{vec}(A) = [a_1^T a_2^T \dots a_n^T]^T \in \mathbb{R}^{n^2 \times 1}$$

where a_i is the i-th column of A and a_i^T is the transpose of a_i.

Remark 3.5. The matrix equation

$$AXB = C$$

is equivalent to the system of $q \times m$ equations in $n \times p$ unknowns given by

$$(B^T \otimes A)\mathrm{vec}(X) = \mathrm{vec}(C)$$

that is, $\mathrm{vec}(AXB) = (B^T \otimes A)\mathrm{vec}(X)$.

Suppose A is nonsingular, let $M = Q - A^*XA$ and $N = Q - A^*YA$, then

$$F(X) - F(Y) = M^{\frac{1}{p}} - N^{\frac{1}{p}},$$

and

$$X - Y = (A^*)^{-1}(N - M)A^{-1}$$
$$= \sum_{i=0}^{p-1}(A^*)^{-1}N^{\frac{p-1-i}{p}}(N^{\frac{1}{p}} - M^{\frac{1}{p}})M^{\frac{i}{p}}A^{-1}.$$

Thus

$$\mathrm{vec}(X - Y) = (A^T \otimes A^*)^{-1}\Big(\sum_{i=0}^{p-1}(M^{\frac{i}{p}})^T \otimes N^{\frac{p-1-i}{p}}\Big)\mathrm{vec}(F(Y) - F(X)).$$

which implies that

$$\mathrm{vec}\Big(F(Y) - F(X)\Big) = \Big(\sum_{i=0}^{p-1}(M^{\frac{i}{p}})^T \otimes N^{\frac{p-1-i}{p}}\Big)^{-1}(A^T \otimes A^*)\mathrm{vec}(X - Y).$$

So we can get

(3.2) $$\|F(Y) - F(X)\|_F \leq \Big\|(\sum_{i=0}^{p-1}(M^{\frac{i}{p}})^T \otimes N^{\frac{p-1-i}{p}})^{-1}\Big\|_2 \|A\|_2^2 \|X - Y\|_F.$$

Now we define $\Gamma(X, Y)$ by

$$\Gamma(X, Y) = \Big\|\Big(\sum_{i=0}^{p-1}(M^{\frac{i}{p}})^T \otimes N^{\frac{p-1-i}{p}}\Big)^{-1}\Big\|_2 \|A\|_2^2.$$

Theorem 3.6. *Let S be a solvent of (1.1), and let $a = \|A\|_2$, $s = \sigma_{\min}(S)$. If $s^p - \left(\frac{a^2}{p}\right)^{\frac{p}{p-1}} > 0$, then*

$$\Gamma(X, Y) \leq \bar{\nu}(\delta', \delta') < 1, \quad \forall X, Y \in \mathcal{B}_2 := \mathcal{B}\left(S, \delta'\right),$$

where

$$\delta' < \frac{s^p - \left(\frac{a^2}{p}\right)^{\frac{p}{p-1}}}{a^2}$$

and

$$\bar{\nu}(\|X - S\|_2, \|Y - S\|_2) = \frac{a^2}{\sum_{i=0}^{p-1}\left(s^p - \|X - S\|_2 a^2\right)^{\frac{i}{p}}\left(s^p - \|Y - S\|_2 a^2\right)^{\frac{p-1-i}{p}}}.$$

Proof. Let $X, Y \in \mathcal{B}_2$, then we have

$$\Gamma(X, Y) = \Big\|\Big(\sum_{i=0}^{p-1}\left((Q - A^*XA)^{\frac{i}{p}}\right)^T \otimes (Q - A^*YA)^{\frac{p-1-i}{p}}\Big)^{-1}\Big\|_2 \|A\|_2^2$$

$$= \Big\|\Big(\sum_{i=0}^{p-1}\left((Q - A^*(X - S)A - A^*SA)^{\frac{i}{p}}\right)^T \otimes \Big(Q - A^*(Y - S)A - A^*SA\Big)^{\frac{p-1-i}{p}}\Big)^{-1}\Big\|_2 \|A\|_2^2$$

$$= \Big\|\Big(\sum_{i=0}^{p-1}\left((S^p - A^*(X - S)A)^{\frac{i}{p}}\right)^T \otimes \Big(S^p - A^*(Y - S)A\Big)^{\frac{p-1-i}{p}}\Big)^{-1}\Big\|_2 \|A\|_2^2$$

$$\leq \frac{a^2}{\sum_{i=0}^{p-1}\left((s^p - \sigma_{\max}(A^*(X - S)A))^{\frac{i}{p}}(s^p - \sigma_{\max}(A^*(Y - S)A))^{\frac{p-1-i}{p}}\right)}$$

$$\leq \frac{a^2}{\sum_{i=0}^{p-1}\left((s^p - \|X - S\|_2 a^2)^{\frac{i}{p}}(s^p - \|Y - S\|_2 a^2)^{\frac{p-1-i}{p}}\right)}.$$

From this, we get

$$\Gamma(X,Y) \le \bar{\nu}(\|X - S\|_2, \|Y - S\|_2),$$

and

$$\bar{\nu}(\|X - S\|_2, \|Y - S\|_2) \le \bar{\nu}(\delta', \delta') = \frac{a^2}{p\big(s^p - a^2\delta'\big)^{\frac{p-1}{p}}},$$

where $\delta' = \min\{\|X - S\|_2, \|Y - S\|_2\}$.

Since $s^p - \big(\frac{a^2}{p}\big)^{\frac{p}{p-1}} > 0$ and $\delta' < \frac{s^p - \big(\frac{a^2}{p}\big)^{\frac{p}{p-1}}}{a^2}$, then

$$\left(\frac{a^2}{p}\right)^{\frac{p}{p-1}} < s^p - a^2\delta'.$$

which implies

$$\frac{a^2}{p} < \big(s^p - a^2\delta'\big)^{\frac{p-1}{p}}.$$

Thus

$$\bar{\nu}(\delta', \delta') = \frac{a^2}{p\big(s^p - a^2\delta'\big)^{\frac{p-1}{p}}} < 1.$$

$\square$

4. Numerical Experiments

In this section, we give two numerical experiments in [5] to show the efficiency of the iterations for finding the positive definite solution of equation (1.1). We also give two experiments to find the convergence radius of local convergence. Our experiments were done in Matlab R2014a and each time for iteration is the average of 10 times.

Example 4.1. [5]Let matrix $A = rand(10) \times 10^{-2}$, $Q = eye(10)$ and $p = 2, 3, \cdots, 10$.

	B.I.-iter	B.I.-time	N.S.-iter	N.S.-time
$p = 2$	5	0.0014	5	0.0004
$p = 3$	5	0.0013	5	0.0003
$p = 4$	5	0.0010	5	0.0003
$p = 5$	5	0.0010	5	0.0003
$p = 6$	5	0.0010	5	0.0003
$p = 7$	5	0.0010	5	0.0003
$p = 8$	5	0.0010	5	0.0003
$p = 9$	5	0.0010	5	0.0003
$p = 10$	4	0.0008	5	0.0003

Example 4.2. [5]Let matrix $A = rand(50) \times 10^{-2}$, $Q = eye(50)$ and $p = 2, 3, \cdots, 10$.
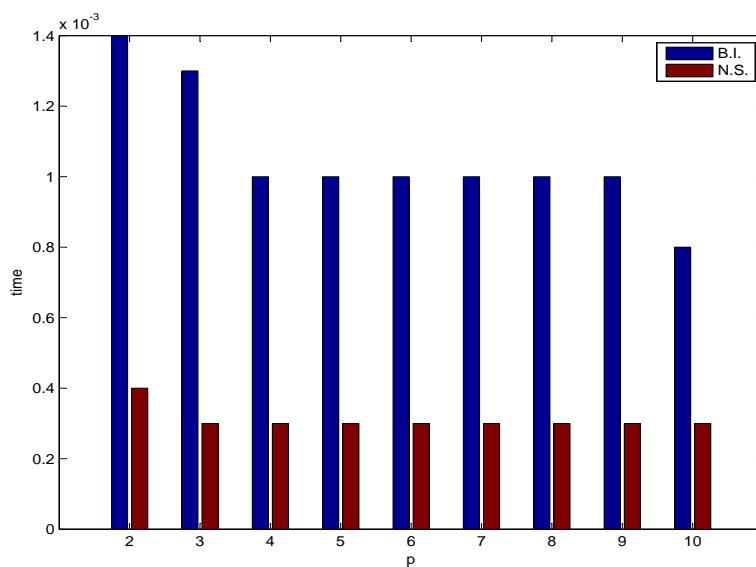
FIGURE 4.1. Time in example 4.1

	B.I.-iter	B.I.-time	N.S.-iter	N.S.-time
$p = 2$	9	0.0358	9	0.0043
$p = 3$	8	0.0267	7	0.0038
$p = 4$	8	0.0275	8	0.0042
$p = 5$	8	0.0282	8	0.0048
$p = 6$	7	0.0239	7	0.0046
$p = 7$	7	0.0239	7	0.0044
$p = 8$	7	0.0250	7	0.0053
$p = 9$	7	0.0245	7	0.0052
$p = 10$	7	0.0246	7	0.0044

In section 3, we showed the local convergence of the method (1.3). Now, we will calculate the convergence radius of the following examples. We used matrix 2-norm for calculation.

Example 4.3. Let $A = \begin{pmatrix} 0.5 & -0.45 \\ 0.45 & 0 \end{pmatrix}$, $Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $p = 2, 3, 4, 5, 6$.

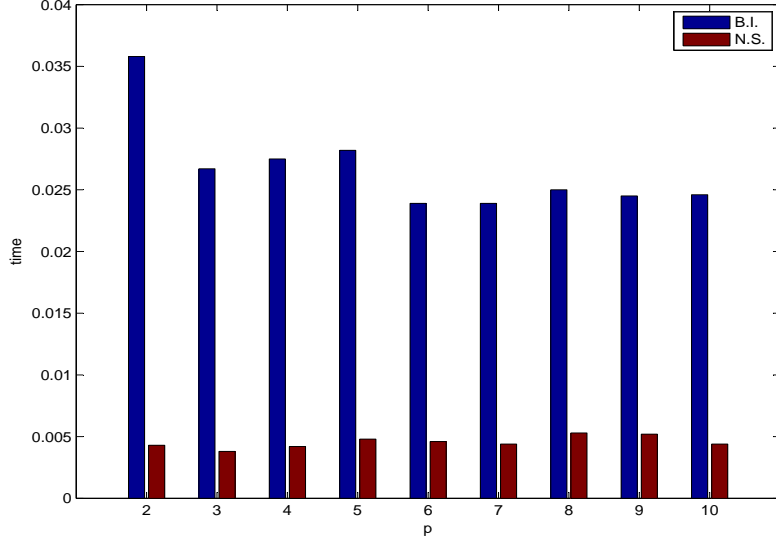|	s	a	δ
$p = 2$	0.6902	0.7648	0.6682
$p = 3$	0.7713	0.7648	0.6372
$p = 4$	0.8186	0.7648	0.6361
$p = 5$	0.8497	0.7648	0.6403
$p = 6$	0.8717	1.0215	0.6453

FIGURE 4.2. Time in example 4.2

Example 4.4. Let $A = \begin{pmatrix} 0.2 & 0.4 \\ 0.05 & 0.25 \end{pmatrix}$, $Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $p = 2, 3, 4, 5, 6$.

	s	a	δ
$p = 2$	0.8716	0.5114	2.8393
$p = 3$	0.9099	0.5114	2.7818
$p = 4$	0.9306	0.5114	2.7663
$p = 5$	0.9435	0.5114	2.7632
$p = 6$	0.9524	0.5114	2.7641

The iteration number of N.S. algorithm is equal to or greater than Built-in function in MATLAB when we compare only the iteration numbers. We don't check the iteration number of Built-in function when we use it to find the p-th root. We reduce CPU-time for finding the p-th root at each steps. By the N.S. algorithm, we save the CPU time to find the solution. We can consider other methods which evaluate the matrix p-th root. Some other methods also reduce the CPU-time. For example, we also checked the Halley method in [10] and it also reduces CPU-time, but it is slower than N.S. algorithm. It is possible to find suitable algorithms for solving the equation (1.1) in various conditions of coefficient matrices.

References

REFERENCES

[1] S. M. EL-Sayed and A. C. M. Ran, On an iteration method for solving a class of nonlinear matrix equautions, SIAM J. Matrx Anal. Appl. 23 (2001) 632-645.

[2] J. C. Engwerda, A. C. M. Ran and A. L. Rijkeboer, Necessary and sufficient conditions for the existence of a positive solution of the matrix equation $X + A^* X^{-1} A = Q$, Linear Algebra Appl. 186 (1993) 255-275.

[3] Z. Jia and M. Wei, Solvability and sensitivity theory of polynomial matrix equation $X^s + A^T X^t A = Q$, Appl. Math. Comput. 209 (2009) 230-237.

[4] A. C. M. Ran and M. C. B. Reurings, On the nonlinear matrix equation $X + A^* \mathscr{F}(X) A = Q$: solution and perturbation theory, Linear Algebra Appl. 346 (2002) 15-26.

[5] J. Meng and H.-M. Kim, The positive definite solution to a nonlinear matrix equation, Linear and Multilinear Algebra, Vol. 64, No. 4, (2016) 653-666.

[6] A. C. M. Ran and M. C. B. Reurings, The symmetric linear matrix equation, Electron. J. Linear Algebra 9 (2002) 93-107.

[7] A. C. M. Ran and M. C. B. Reurings, A nonlinear matrix equation connected to interpolation theory, Linear Algebra Appl. 379 (2004) 289-302.

[8] M. C. B. Reurings, Contractive maps on normed linear spaces and their applications to nonlinear matrix equations, Linear Algebra Appl. 418 (2006) 292-311.

[9] M. C. B. Reurings, Symmetric Matrix Equations, Ph.D. thesis, Vrije Universiteit Amsterdam, 2003. <http://www.math.vu.nl/proefschriften/ Reurings>.

[10] C.-H. Guo, On Newton's method and Halley's method for the principal p-th root of a matrix, Linear Algebra Appl. 432 (2010) 1905-1922.

[11] N. J. Higham, Functions of Matrices : Theory and Computation, SIAM, Philadelphia,2008.

[12] D. A. Bini, N. J. Higham and B. Meini, Algorithms for the matrix pth root, Numer. Algorithms 39 (2005) 349-378.

[13] C.-H.Guo and N. J. Higham, A Schur-Newton method for the matrix pth root and its inverse, SIAM J. Matrix Anal. Appl. 29 (2007) 396-412.

[14] B. Iannazzo, On the Newton method for the matrix pth root, SIAM J. Matrix Anal. Appl. 28 (2006) 503-523.

[15] B. Iannazzo, A family of rational iterations and its applications to the computation of the matrix pth root, SIAM J. Matrix Anal. Appl. 30 (2008) 1445-1462.

[16] M. I. Smith, A Schur Algorithm for computing matrix pth roots, SIAM J. Matrix Anal. Appl. 24 (2003) 971-989.

[17] R. A. Smith, Infinite product expansions for matrix n-th roots, J. Austral. Math. Soc. 8 (1968) 242-249.

[18] A. Sadeghi, A. I. Md. Ismail and A. Ahmad, Computing the pth Roots of a matrix with Repeated Eigenvalues, Applied Mathematical Sciences, Vol. 5, No. 53, (2011) 2645-2661.

[19] J. M. Ortega and W. C. Rheinboldt, Iterative solution of nonlinear equations in several variables, Academic Press, New York, (2000)

[20] S. H. Cheng, N. J. Higham, C. S. Kenney and A. J. Laub, Approximating the logarithm of a matrix to specified accuracy, SIAM J. Matrix Anal. Appl. 22(4) (2001) 1112-1125.

[21] C. S. Kenney and A. J. Laub, Condition estimates for matrix functions, SIAM J. Matrix Anal. Appl. 10(2) (1989) 191-209.

[22] W. D. Hoskins and D. J. Walton, A faster, more stable method for computing the pth roots of positive definite matrices, Linear Algebra Appl. 26 (1979) 139-163.

[23] L.-S. Shieh, Y. T. Tsay and R. E. Yates, Computation of the principal nth roots of complex matrices, IEEE Trans. Automat. Control 30(6) (1985) 606-608.

[24] Y.-J. Kim, J.-H. Seo and H.-M. Kim, Local convergence of functional iterations for solving a quadratic matrix equation, accepted to B. Korean Math. Soc.

Department of Mathematics, Pusan National University, Busan, 609-735, Republic of Korea

E-mail address: kyj2806@hanmail.net