J. KSIAM Vol.14, No.2, 113–124, 2010

SOLVING A MATRIX POLYNOMIAL BY NEWTON'S METHOD*

YIN-HUAN HAN AND HYUN-MIN KIM†

DEPARTMENT OF MATHEMATICS, PUSAN NATIONAL UNIVERSITY, SOUTH KOREA
E-mail address: hanyinhuan@hotmail.com, hyunmin@pusan.ac.kr

ABSTRACT. We consider matrix polynomial which has the form

$$P_1(X) = A_0X^m + A_1X^{m-1} + \cdots + A_m = 0$$

where X and A_i are $n \times n$ matrices with real elements. In this paper, we propose an iterative method for the symmetric and generalized centro-symmetric solution to the Newton step for solving the equation $P_1(X)$. Then we show that a symmetric and generalized centro-symmetric solvent of the matrix polynomial can be obtained by our Newton's method. Finally, we give some numerical experiments that confirm the theoretical results.

1. INTRODUCTION

One of the most well-known and studied nonlinear matrix equation is the matrix polynomial which can be defined by

$$P_1(X) = A_0X^m + A_1X^{m-1} + \cdots + A_m = 0, \qquad A_i, X \in R^{n\times n}. \tag{1.1}$$

If the leading coefficient A_0 is the $n \times n$ identity matrix, which has the form

$$P_2(X) = X^m + A_1X^{m-1} + \cdots + A_m = 0, \tag{1.2}$$

is called to be monic. Matrix polynomials appear in the theories of differential equations, system, network, stochastic and other areas [1, 2, 3, 9, 15, 14].

For solving the matrix polynomial (1.1) Newton's method and Berlloulli's iteration were considered by Kratz and Stickel [13], and Dennis, Jr., Traub and Weber [7, 8], respectively. Seo and Kim [19] incorporated the exact line searches into Newton's method which reduced the number of iterations for convergence and improved approaches to solve each Newton step.

However Newton's method has still a couple of weak points for solving the equation $P_1(X)$.

- It can not limit properties of solvents which converge.
- It well works when only the Fréchet derivative is nonsingular.

Received by the editors May 26 2010; Accepted June 3 2010.
2000 *Mathematics Subject Classification.* 65F30, 65H10.
Key words and phrases. matrix polynomial, solvent, Newton's method, iterative method, symmetric and generalized centro-symmetric.
* This work was supported for two years by Pusan National University Research Grant.
† Corresponding author.

Our interesting problem is how to guarantee the convergence for a particular starting matrix even if the Fréchet derivative is singular. There are a few investigations on the matrix polynomial for this topic except the problems for finding the dominant, minimal [7, 8, 10] and elementary minimal positive solvents [12] of the equation (1.1) and the quadratic matrix equation

$$Q(X) = AX^2 + BX + C = 0.$$

However, the linear matrix equation

$$AXB = C$$

has been discussed extensively with some special solutions, such as, Chu [11], Dai [4] and Peng [20] with the symmetric solution; Liang [16, 17] with the (R,S)-symmetric and generalized centro-symmetric solutions. Dehghan [6] proposed iterative algorithms to find the reflexive and anti-reflexive solutions of the matrix equation

$$A_1X_1B_1 + A_2X_2B_2 = C.$$

In this work, we apply their ideas to find the matrix polynomial (1.1) for a symmetric and generalized centro-symmetric solvent although the Fréchet derivative is singular. The definitions of a symmetric and generalized centro-symmetric matrix as follows.

Definition 1.1. [17] *For arbitrary given symmetric orthogonal matrix $P \in R^{n\times n}$, i.e., $P = P^T = P^{-1}$, we say that matrix $A \in R^{n\times n}$ is a generalized centro-symmetric(GCS) matrix with respect to the symmetric orthogonal matrix P if $A = PAP$. Especially, when A is symmetric, we call A is an symmetric and generalized centro-symmetric(SGCS) matrix.*

We review Newton's method for solving the matrix polynomial (1.1) in the next section. In the Section 3, an iterative method for an SGCS solution to the Newton step is introduced. Then the convergence results for our Newton's method will be considered. Finally, some numerical experiments are given.

2. Newton's method

Newton's method for solving the matrix polynomial (1.1) can be obtained from the expansion

$$\begin{aligned} P_1(X+H) &= P_1(X) + \sum_{i=1}^{m} \left[\left(\sum_{j=0}^{m-i} A_j X^{m-(j+i)} \right) H X^{i-1} \right] + O(H^2) \\ &= P_1(X) + D_X(H) + O(H^2), \end{aligned}$$

where $D_X(H)$ is the Fréchet derivative of P_1 at X in the direction H as the solution of the linear equation $P_1(X) + D_X(H) = 0$. Thus, the Newton iteration has the form

$$\begin{cases} \sum_{i=1}^{m} \left[\left(\sum_{j=0}^{m-i} A_j X_k^{m-(j+i)} \right) H_k X_k^{i-1} \right] = -P_1(X_k), \\ \mathrm{X}_{k+1} = X_k + H_k, \end{cases} \qquad \mathrm{w}here \quad k = 0, 1, \cdots.$$

For solving the matrix polynomial (1.1) by Newton's method it is necessary to find the solution of the linear equation

$$\sum_{i=1}^{m} \left[\left(\sum_{j=0}^{m-i} A_j X^{m-(j+i)} \right) H X^{i-1} \right] = -P_1(X), \tag{2.1}$$

which is called the Newton step. To solve the Newton step (2.1), Kratz and Stickel [13] considered the Schur algorithm. For a given X, compute the Schur decomposition of X

$$W^*XW = U, \tag{2.2}$$

where W is a unitary and U is an upper triangular matrices. Then, substituting (2.2) into (2.1) transforms the system to

$$\sum_{i=1}^{m} \left(\sum_{j=0}^{m-i} A_j X^{m-(j+i)}\right) H'U^{i-1} = F \tag{2.3}$$

where $H' = HW$ and $F = -P_1(X)W$. Then, taking the vec function both sides of (2.3) makes a linear system such that

$$\widetilde{F}\mathrm{vec}(H') = \mathrm{vec}(F) \tag{2.4}$$

where the matrix $\widetilde{F} \in \mathbb{C}^{n^2\times n^2}$ is given by

$$\widetilde{F} = \sum_{i=1}^{m} \left(\left(U^T\right)^{i-1} \otimes \left(\sum_{j=0}^{m-i} A_j X^{m-(j+i)}\right)\right). \tag{2.5}$$

Seo and Kim [19] defined $\widetilde{F_{ij}} = \sum_{i=1}^{m} \left[U^{i-1}\right]_{ji} \left(\sum_{j=0}^{m-i} A_j X^{m-(j+i)}\right)$ to reduce the system size of the equation (2.4) to $n \times n$, then $\widetilde{F}$ in (2.5) is represented by

$$\widetilde{F} = \begin{bmatrix} \widetilde{F_{11}} & & & \\ \widetilde{F_{21}} & \widetilde{F_{22}} & & \mathbf{0} \\ \vdots & \vdots & \ddots & \\ \widetilde{F_{n1}} & \widetilde{F_{n2}} & \cdots & \widetilde{F_{nn}} \end{bmatrix}. \tag{2.6}$$

Since $\widetilde{F}$ is a block lower triangular matrix, if the matrices $\widetilde{F_{ii}}$ are nonsingular, the equation (2.4) can be changed to n linear systems with size $n \times n$ such that

$$\begin{aligned} h_1' &= \widetilde{F_{11}}^{-1} f_1 \\ h_2' &= \widetilde{F_{22}}^{-1}(f_2 - \widetilde{F_{21}}h_1') \\ &\vdots \\ h_n' &= \widetilde{F_{nn}}^{-1}(f_n - \widetilde{F_{n1}}h_1' - \cdots - \widetilde{F_{n,n-1}}h_{n-1}'), \end{aligned}$$

where h_i' and f_i are the ith columns of H' and F, respectively.

In the Fréchet derivative is nonsingular case, the Kantorovich theorem gives information on the convergence of Newton's method for solving matrix polynomial (1.1) [5].

Theorem 2.1 (Kantorovich)**.** *If there exists K such that*

$$\|D_X - D_Y\| \leq K\|X - Y\|, \qquad \textit{for all} \quad X, Y \in R^{n\times n}$$

in some closed ball $\bar{U}(X_0, r)$ and $h_0 = B_0\eta_0 K \leq \frac{1}{2}$ with $\|D_{X_0}^{-1}\| \leq B_0$ and $\|X_1 - X_0\| \leq \eta_0$, then the Newton sequence starting from X_0 will converge to a solvent S of $P_1(X)$ which exists *in* $\bar{U}(X_0, r)$*, provided that*

$$r \geq r_0 = \frac{1 - \sqrt{1 - 2h_0}}{h_0}\eta_0.$$

However, the Theorem 2.1 can not effect a settlement for the weak points of Newton's method.

3. An iterative method for solving the equation (2.1)

In this section, we introduce an iterative method to find the SGCS solution of the Newton step (2.1) which do not depend on the singularity of the Fréchet derivative. Then consider the convergence of our Newton's method for solving the matrix polynomial (1.1). Throughout this paper, $\|\cdot\|$ denotes the Euclidean norm of the matrix.

The following algorithm is the process for finding an SGCS solution of the qth Newton step (2.1).

Algorithm 3.1. *Given matrices* $A_0, A_1, \cdots, A_m \in R^{n\times n}$ *and the symmetric orthogonal matrix* $P \in R^{n\times n}$*. An SGCS matrix* $X_q \in R^{n\times n}$ *with respect to the symmetric orthogonal matrix* P *is given. Choose an SGCS matrix* H_{q_0}.

$k = 0$; $\quad R_0 = -P_1(X_q) - \left[\sum_{i=1}^{m}\left(\sum_{j=0}^{m-i} A_j(X_q)^{m-(j+i)}\right) H_{q_0}(X_q)^{i-1}\right]$

$Y_0 = \sum_{i=1}^{m}\left(\sum_{j=0}^{m-i} A_j(X_q)^{m-(j+i)}\right)^T R_0((X_q)^{i-1})^T$

$Q_0 = \frac{1}{4}[(Y_0 + Y_0^T) + P(Y_0 + Y_0^T)P]$

$\alpha_0 = \frac{\|R_0\|^2}{\|Q_0\|^2}$

while $R_k \neq 0$

$\alpha_k = \frac{\|R_k\|^2}{\|Q_k\|^2}$

$H_{q_{k+1}} = H_{q_k} + \alpha_k Q_k$

$R_{k+1} = R_k - \alpha_k\left[\sum_{i=1}^{m}\left(\sum_{j=0}^{m-i} A_j(X_q)^{m-(j+i)}\right) Q_k(X_q)^{i-1}\right]$

$\beta_k = \frac{\|R_{k+1}\|^2}{\|R_k\|^2}$

$Y_{k+1} = \sum_{i=1}^{m}\left(\sum_{j=0}^{m-i} A_j(X_q)^{m-(j+i)}\right)^T R_{k+1}((X_q)^{i-1})^T$

$Q_{k+1} = \frac{1}{4}[(Y_{k+1} + Y_{k+1}^T) + P(Y_{k+1} + Y_{k+1}^T)P] + \beta_k Q_k$

If $Q_{k+1} = 0$ and $R_{k+1} \neq 0$, break

end

In Algorithm 3.1, the matrices H_{q_k} and Q_k are the SGCS matrices. This iterative method shows that if $R_k = 0$, then it is terminated. And it implies that, if the matrix $Q_k = 0$ but $R_k \neq 0$ for some integer number k, then the equation (2.1) has no SGCS solution. From Algorithm 3.1, we obtain the following basic properties.

Lemma 3.2. *For the sequences* $\{Y_k\}, \{Q_k\}$ *and* $\{R_k\}$ *with* $k = 0, 1, \cdots$ *are generated by Algorithm* 3.1*, we have that*

$$\mathrm{tr}\left(Y_{k+1}^T Q_k\right) = \frac{1}{\alpha_k}\mathrm{tr}\left(R_{k+1}^T R_k\right) - \frac{1}{\alpha_k}\|R_{k+1}\|^2.$$

Proof. From Algorithm 3.1, we obtain

$$\begin{aligned}\mathrm{tr}\left(Y_{k+1}^T Q_k\right) &= \mathrm{tr}\left\{\left[\sum_{i=1}^m \left(\sum_{j=0}^{m-i} A_j (X_q)^{m-(j+i)}\right)^T R_{k+1}\left((X_q)^{i-1}\right)^T\right]^T Q_k\right\}\\ &= \mathrm{tr}\left[R_{k+1}^T\left(\sum_{i=1}^m \sum_{j=0}^{m-i} A_j (X_q)^{m-(j+i)} Q_k (X_q)^{i-1}\right)\right]\\ &= \mathrm{tr}\left[R_{k+1}^T \tfrac{1}{\alpha_k}(R_k - R_{k+1})\right]\\ &= \tfrac{1}{\alpha_k}\mathrm{tr}(R_{k+1}^T R_k) - \tfrac{1}{\alpha_k}\|R_{k+1}\|^2.\end{aligned}$$

□

Lemma 3.3. *Suppose that the qth Newton step* (2.1) *is consistent and* H_q *is an SGCS solution. Then for the starting SGCS matrix* H_{q_0}*, the sequences* $\{H_{q_k}\}, \{R_k\}$ *and* $\{Q_k\}$ *are generated by Algorithm* 3.1*, we have*

$$\mathrm{tr}\left[(H_q - H_{q_k})^T Q_k\right] = \|R_k\|^2 \qquad \text{for} \quad k = 0, 1, \cdots, n^2 - 1. \tag{3.1}$$

Proof. We prove the result (3.1) by the principle induction.
Step 1. When $k = 0$, from Algorithm 3.1

$$\begin{aligned}\mathrm{tr}\left[(H_q - H_{q_0})^T Q_0\right] &= \mathrm{tr}\left[Q_0^T (H_q - H_{q_0})\right] = \mathrm{tr}\left[Y_0^T (H_q - H_{q_0})\right]\\ &= \mathrm{tr}\left\{R_0^T\left[-P_1(X_q) - \sum_{i=1}^m \left(\sum_{j=0}^{m-i} A_j (X_q)^{m-(j+i)}\right) H_{q_0}(X_q)^{i-1}\right]\right\}\\ &= \mathrm{tr}(R_0^T R_0)\\ &= \|R_0\|^2.\end{aligned}$$

Step 2. Suppose that the statement (3.1) holds for $k = n^2 - 2$, then when $k = n^2 - 1$, we have that

$$\begin{aligned}&\mathrm{tr}\left[(H_q - H_{q_{n^2-1}})^T Q_{n^2-1}\right]\\ &= \mathrm{tr}\left[Y_{n^2-1}^T (H_q - H_{q_{n^2-1}})\right] + \beta_{n^2-2}\mathrm{tr}\left[Q_{n^2-2}^T (H_q - H_{q_{n^2-1}})\right]\\ &= \mathrm{tr}\left\{R_{n^2-1}^T\left[-P_1(X_q) - \sum_{i=1}^m \left(\sum_{l=0}^{m-i} A_j (X_q)^{m-(j+i)}\right) H_{q_{n^2-1}}(X_q)^{i-1}\right]\right\}\\ &+\beta_{n^2-2}\mathrm{tr}\left[Q_{n^2-2}^T (H_q - H_{q_{n^2-2}})\right] - \beta_{n^2-2}\alpha_{n^2-2}\mathrm{tr}\left(Q_{n^2-2}^T Q_{n^2-2}\right)\\ &= \mathrm{tr}\left(R_{n^2-1}^T R_{n^2-1}\right)\\ &= \|R_{n^2-1}\|^2.\end{aligned}$$

Hence from Step 1 and 2, $\mathrm{tr}\left[(H_q - H_{q_k})^T Q_k\right] = \|R_k\|^2$ holds for all $k = 0, 1, \cdots, n^2 - 1$. □

Lemma 3.3 implies that if there is a positive integer number l such that $Q_l \neq 0$ but $R_l = 0$, then the qth Newton step (2.1) is inconsistent.

Lemma 3.4. *Assume that the sequences* $\{R_i\}, \{Q_i\}$ *with* $i = 0, 1, 2, \cdots$ *are generated by Algorithm* 3.1*, i.e., there exists a positive integer number* $t \geq 1$ *such that* $R_i \neq 0$ *for all* $i = 0, 1, 2, \cdots, t$*, then*

$$\mathrm{tr}\left(R_i^T R_j\right) = 0 \quad \textit{and} \quad \mathrm{tr}\left(Q_i^T Q_j\right) = 0 \qquad \textit{for} \quad i > j = 0, 1 \cdots, t. \tag{3.2}$$

Proof. We prove the statement by principle induction.

Step 1. When $t = 1$, from Algorithm 3.1 and Lemma 3.2 we have

$$\begin{aligned}
\mathrm{tr}\left(R_1^T R_0\right) &= \mathrm{tr}\left\{\left[R_0 - \alpha_0\left(\textstyle\sum_{k=1}^m \sum_{l=0}^{m-k} A_l (X_q)^{m-(l+k)} Q_0 (X_q)^{k-1}\right)\right]^T R_0\right\} \\
&= \mathrm{tr}(R_0^T R_0) - \alpha_0 \mathrm{tr}\left[\left(\textstyle\sum_{k=1}^m \sum_{l=0}^{m-k} A_l (X_q)^{m-(l+k)} Q_0 (X_q)^{k-1}\right)^T R_0\right] \\
&= \|R_0\|^2 - \alpha_0 \mathrm{tr}\left\{Q_0^T \left[\textstyle\sum_{k=1}^m \left(\sum_{l=0}^{m-k} A_l (X_q)^{m-(l+k)}\right)^T R_0 \left((X_q)^{k-1}\right)^T\right]\right\} \\
&= \|R_0\|^2 - \alpha_0 \mathrm{tr}(Q_0^T Y_0) \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{tr}\left(Q_1^T Q_0\right) &= \mathrm{tr}\left(Y_1^T Q_0\right) + \beta_0 \|Q_0\|^2 \\
&= \tfrac{1}{\alpha_0} \mathrm{tr}\left(R_1^T R_0\right) - \tfrac{1}{\alpha_0} \|R_1\|^2 + \beta_0 \|Q_0\|^2 \\
&= 0.
\end{aligned}$$

Step 2. Suppose that (3.2) holds when $t = s$, i.e., $\mathrm{tr}\left(R_s^T R_j\right) = 0$ and $\mathrm{tr}\left(Q_s^T Q_j\right) = 0$ hold for $j = 0, 1, \cdots, s-1$. Then when $t = s+1$, we obtain

$$\begin{aligned}
\mathrm{tr}\left(R_{s+1}^T R_s\right) &= \mathrm{tr}\left\{\left[R_s - \alpha_s\left(\textstyle\sum_{k=1}^m \sum_{l=0}^{m-k} A_l (X_q)^{m-(l+k)} Q_s (X_q)^{k-1}\right)\right]^T R_s\right\} \\
&= \|R_s\|^2 - \alpha_s \mathrm{tr}\left\{Q_s^T \left[\textstyle\sum_{k=1}^m \left(\sum_{l=0}^{m-k} A_l (X_q)^{m-(l+k)}\right)^T R_s \left((X_q)^{k-1}\right)^T\right]\right\} \\
&= \|R_s\|^2 - \alpha_s \mathrm{tr}\left(Q_s^T Y_s\right) \\
&= \|R_s\|^2 - \alpha_s \|Q_s\|^2 + \alpha_s \beta_{s-1} \mathrm{tr}\left(Q_s^T Q_{s-1}\right) \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{tr}\left(Q_{s+1}^T Q_s\right) &= \mathrm{tr}\left(Y_{s+1}^T Q_s\right) + \beta_s \mathrm{tr}\left(Q_s^T Q_s\right) \\
&= -\tfrac{1}{\alpha_s} \|R_{s+1}\|^2 + \beta_s \|Q_s\|^2 \\
&= 0.
\end{aligned}$$

Furthermore, for $j = 0, 1, \cdots, s-1$

$$\begin{aligned}
\mathrm{tr}\left(R_{s+1}^T R_j\right) &= \mathrm{tr}\left\{\left[R_s + \alpha_s\left(\textstyle\sum_{k=1}^m \sum_{l=0}^{m-k} A_l (X_q)^{m-(l+k)} Q_s (X_q)^{k-1}\right)\right]^T R_j\right\} \\
&= \mathrm{tr}\left(R_s^T R_j\right) + \alpha_s \mathrm{tr}\left\{Q_s^T \left[\textstyle\sum_{k=1}^m \left(\sum_{l=0}^{m-k} A_l (X_q)^{m-(l+k)}\right)^T R_j ((X_q)^{k-1})^T\right]\right\} \\
&= \alpha_s \mathrm{tr}\left(Q_s^T Y_j\right) = \alpha_s \mathrm{tr}\left[Q_s^T (Q_j - \beta_{j-1} Q_{j-1})\right] \\
&= \alpha_s \mathrm{tr}\left(Q_s^T Q_j\right) - \alpha_s \beta_{j-1} \mathrm{tr}\left(Q_S^T Q_{j-1}\right) \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}\operatorname{tr}\left(Q_{s+1}^T Q_j\right) &= \operatorname{tr}\left(Y_{s+1}^T Q_j\right) + \beta_s \operatorname{tr}\left(Q_s^T Q_j\right) \\ &= \operatorname{tr}\left[R_{s+1}^T \left(\sum_{k=1}^{m}\sum_{l=0}^{m-k} A_l (X_q)^{m-(l+k)} Q_j (X_q)^{k-1}\right)\right] \\ &= \frac{1}{\alpha_j}\operatorname{tr}\left[R_{s+1}^T (R_j - R_{j+1})\right] = 0.\end{aligned}$$

Thus the conclusion (3.2) holds for all $i > j = 0, 1, \cdots, t$. □

If there exists a positive number k such that $R_i \neq 0$ for all $i = 0, 1, \cdots, k$ in Algorithm 3.1, then the matrices R_i and R_j are orthogonal for $i \neq j$ by Lemma 3.4.

Theorem 3.5. *When the qth Newton step* (2.1) *is consistent, then for any SGCS starting matrix H_{q_0}, an SGCS solution can be obtained within n^2 iterative steps.*

Proof. Assume that there exists a positive integer number $n^2 - 1$ such that $R_i \neq 0$ for all $i = 0, 1, \cdots, n^2 - 1$. Then from Lemma 3.4, $\operatorname{tr}\left(R_i^T R_j\right) = 0$ for $i > j = 0, 1, \cdots, n^2 - 1$. Since the qth Newton step (2.1) has an SGCS solution, $Q_i \neq 0$ for all $i = 0, 1, \cdots, n^2 - 1$ by Lemma 3.2. Therefore, Algorithm 3.1 generates $H_{q_{n^2}}$ and R_{n^2} which satisfy that $\operatorname{tr}\left(R_{n^2}^T R_j\right) = 0$, where $j = 0, 1, \cdots, n^2 - 1$. Thus R_{n^2} is comprised in an orthogonal basis $\{R_0, R_1, \cdots, R_{n^2-1}\}$ of matrix space $R^{n\times n}$. It implies that $R_{n^2} = 0$, so $H_{q_{n^2}}$ is a solution of the equation (2.1). From Algorithm 3.1, we know that $H_{q_{n^2}}$ is SGCS for the SGCS starting matrix H_{q_0}. □

The following theorem is the existence theory of solvent. It gives the idea to solve the first weak point of Newton's method for solving the equation (1.1).

Theorem 3.6. *Suppose the matrix polynomial* (1.2) *satisfies assumptions of Theorem* 2.1 *for a GCS matrix X_0. If the coefficients of the monic matrix polynomial* (1.2) *are all GCS matrices, then the Newton sequence starting from X_0 will converge to a GCS solvent S.*

Proof. From Kantorovich theorem 2.1, a solvent S of the monic matrix polynomial (1.2) can be obtained. Now we will show that S is a GCS matrix.
Since D_{X_0} is nonsingular for the GCS starting matrix X_0, there exists $H \in R^{n\times n}$ satisfying

$$\begin{aligned}\sum_{i=1}^{m}(X_0)^{m-i}H(X_0)^{i-1} + \sum_{i=1}^{m}\left(\sum_{j=1}^{m-i} A_j (X_0)^{m-(j+i)}\right) H (X_0)^{i-1} \\ = -P_2(X_0) = -(X_0)^m - \sum_{j=1}^{m} A_j (X_0)^{m-j}.\end{aligned} \tag{3.3}$$

Furthermore X_0 and A_j are GCS matrices for all $j = 1, 2, \cdots, m$, with respect to the symmetric orthogonal matrix

$$P = U\begin{bmatrix} I_r & 0 \\ 0 & -I_{n-r}\end{bmatrix}U^T, \tag{3.4}$$

where $U = (U_1, U_2)$ is an orthogonal matrix and $U_1 \in R^{n\times r}, U_2 \in R^{n\times(n-r)}$. Therefore we can express X_0 and A_j by

$$X_0 = U\begin{bmatrix} X_{01} & 0 \\ 0 & X_{04}\end{bmatrix}U^T, \quad A_j = U\begin{bmatrix} A_{j1} & 0 \\ 0 & A_{j4}\end{bmatrix}U^T,$$

where $X_{01}, A_{j1} \in R^{r\times r}, X_{04}, A_{j4} \in R^{(n-r)\times(n-r)}$ and U, U^T are as in (3.4), see [6]. Thus the equation (3.3) can rewrite as

$$\begin{aligned}
&\sum_{i=1}^{m}\sum_{j=1}^{m-i} U \begin{bmatrix} A_{j1} & 0 \\ 0 & A_{j4} \end{bmatrix} U^T U \begin{bmatrix} X_{01} & 0 \\ 0 & X_{04} \end{bmatrix}^{m-(j+i)} U^T H U \begin{bmatrix} X_{01} & 0 \\ 0 & X_{04} \end{bmatrix}^{i-1} U^T \\
&+\sum_{i=1}^{m} U \begin{bmatrix} X_{01} & 0 \\ 0 & X_{04} \end{bmatrix}^{m-i} U^T H U \begin{bmatrix} X_{01} & 0 \\ 0 & X_{04} \end{bmatrix}^{i-1} U^T \\
&= -U \begin{bmatrix} X_{01} & 0 \\ 0 & X_{04} \end{bmatrix}^{m} U^T - \sum_{j=1}^{m} U \begin{bmatrix} A_{j1} & 0 \\ 0 & A_{j4} \end{bmatrix} U^T U \begin{bmatrix} X_{01} & 0 \\ 0 & X_{04} \end{bmatrix}^{m-j} U^T.
\end{aligned} \tag{3.5}$$

If denote $U^T H U \equiv \widetilde{H}$ in (3.5), then we simplify it as

$$\begin{aligned}
&\sum_{i=1}^{m}\sum_{j=1}^{m-i} U \begin{bmatrix} A_{j1}(X_{01})^{m-(j+i)} & 0 \\ 0 & A_{j4}(X_{04})^{m-(j+i)} \end{bmatrix} \widetilde{H} \begin{bmatrix} X_{01} & 0 \\ 0 & X_{04} \end{bmatrix}^{i-1} U^T \\
&+\sum_{i=1}^{m} U \begin{bmatrix} X_{01} & 0 \\ 0 & X_{04} \end{bmatrix}^{m-i} \widetilde{H} \begin{bmatrix} X_{01} & 0 \\ 0 & X_{04} \end{bmatrix}^{i-1} U^T \\
&= -U \begin{bmatrix} X_{01} & 0 \\ 0 & X_{04} \end{bmatrix}^{m} U^T - \sum_{j=1}^{m} U \begin{bmatrix} A_{j1}(X_{01})^{m-j} & 0 \\ 0 & A_{j4}(X_{04})^{m-j} \end{bmatrix} U^T.
\end{aligned} \tag{3.6}$$

Premultiplying U^T and postmultiplying U in (3.6), it transforms

$$\begin{aligned}
&\sum_{i=1}^{m} \begin{bmatrix} \sum_{j=1}^{m-i} A_{j1}(X_{01})^{m-(j+i)} & 0 \\ 0 & \sum_{j=1}^{m-i} A_{j4}(X_{04})^{m-(j+i)} \end{bmatrix} \widetilde{H} \begin{bmatrix} X_{01} & 0 \\ 0 & X_{04} \end{bmatrix}^{i-1} \\
&+\sum_{i=1}^{m} \begin{bmatrix} X_{01} & 0 \\ 0 & X_{04} \end{bmatrix}^{m-i} \widetilde{H} \begin{bmatrix} X_{01} & 0 \\ 0 & X_{04} \end{bmatrix}^{i-1} \\
&= -\begin{bmatrix} (X_{01})^m + \sum_{j=1}^{m} A_{j1}(X_{01})^{m-j} & 0 \\ 0 & (X_{04})^m + \sum_{j=1}^{m} A_{j4}(X_{04})^{m-j} \end{bmatrix}.
\end{aligned} \tag{3.7}$$

From the linear equation (3.7), we can obtain the following system:

$$\begin{bmatrix} X_{11} & 0 & 0 & 0 \\ 0 & X_{22} & 0 & 0 \\ 0 & 0 & X_{33} & 0 \\ 0 & 0 & 0 & X_{44} \end{bmatrix} \begin{bmatrix} \widetilde{H_1} \\ \widetilde{H_2} \\ \widetilde{H_3} \\ \widetilde{H_4} \end{bmatrix} = \begin{bmatrix} B_{11} \\ 0 \\ 0 \\ B_{44} \end{bmatrix},$$

where

$$\begin{aligned}
X_{11} &= \sum_{i=1}^{m} \left(X_{01}^T\right)^{i-1} (X_{01})^{m-i} + \sum_{i=1}^{m} \left(X_{01}^T\right)^{i-1} \sum_{j=1}^{m-i} A_{j1}(X_{01})^{m-(j+i)}, \\
X_{22} &= \sum_{i=1}^{m} \left(X_{01}^T\right)^{i-1} (X_{04})^{m-i} + \sum_{i=1}^{m} \left(X_{01}^T\right)^{i-1} \sum_{j=1}^{m-i} A_{j1}(X_{04})^{m-(j+i)}, \\
X_{33} &= \sum_{i=1}^{m} \left(X_{04}^T\right)^{i-1} (X_{01})^{m-i} + \sum_{i=1}^{m} \left(X_{04}^T\right)^{i-1} \sum_{j=1}^{m-i} A_{j1}(X_{01})^{m-(j+i)}, \\
X_{44} &= \sum_{i=1}^{m} \left(X_{04}^T\right)^{i-1} (X_{04})^{m-i} + \sum_{i=1}^{m} \left(X_{04}^T\right)^{i-1} \sum_{j=1}^{m-i} A_{j1}(X_{04})^{m-(j+i)}, \\
B_{11} &= -(X_{01})^m - \sum_{j=1}^{m} A_{j1}(X_{01})^{m-(j+i)}, \\
B_{44} &= -(X_{04})^m - \sum_{j=1}^{m} A_{j4}(X_{04})^{m-(j+i)}.
\end{aligned}$$

and

$$\widetilde{H} = \begin{bmatrix} \widetilde{H_1} & \widetilde{H_3} \\ \widetilde{H_2} & \widetilde{H_4} \end{bmatrix}.$$

Hence $H = U \begin{bmatrix} X_{11}^{-1}B_{11} & 0 \\ 0 & X_{44}^{-1}B_{44} \end{bmatrix} U^T$, which is a GCS matrix with respect to the symmetric orthogonal matrix P in (3.4), the next approximation $X_1 = X_0 + H$ is also a GCS.

Furthermore, since the Fréchet derivative $D_{X_k}(H_k)$ is nonsingular for all $k = 0, 1, \cdots$ by Theorem 2.1, we can assume there exists a matrix E such that

$$\sum_{i=1}^{m}(X_q)^{m-i}E(X_q)^{i-1} + \sum_{i=1}^{m}\left(\sum_{j=1}^{m-i} A_j(X_q)^{m-(j+i)}\right) E(X_q)^{i-1} = -P_2(X_q), \tag{3.8}$$

for the GCS matrix X_q. Similar with the first step, we easy to check that the matrix E is GCS. So the sequence $\{X_k\}$ is the set of the GCS matrices. Hence, this matrix polynomial has a GCS solvent. □

The main result is that we can have the SGCS solvent by our Newton's method even if the Fréchet derivative is singular. This theorem is the answer of the first and second weak points of Newton's method.

Theorem 3.7. *Assume that the matrix polynomial* (1.1) *has an SGCS solvent and each Newton step* (2.1) *is consistent for the SGCS starting matrix X_0. The sequence $\{X_k\}$ is generated by Newton's method with X_0 such that*

$$\lim_{k\to\infty} X_k = S,$$

and the matrix S satisfies $P_1(S) = 0$, then S is an SGCS solvent.

Proof. Suppose that we already got a SGCS solution H_0 of

$$\sum_{i=1}^{m}(X_0)^{m-i}H_0(X_0)^{i-1} + \sum_{i=1}^{m}\left(\sum_{j=1}^{m-i} A_j(X_0)^{m-(j+i)}\right) H_0(X_0)^{i-1} = -P_1(X_0)$$

with SGCS starting matrix X_0 by Algorithm 3.1. Then from our Newton's method and Theorem 3.5, we know that the SGCS matrix

$$\begin{aligned} X_{i+1} &= X_i + H_i \\ &= X_0 + H_0 + \cdots + H_i \end{aligned}$$

always generated for all $i = 1, 2, \cdots$, with SGCS starting matrix X_0. So, by the convergence of Newton's method we can obtain an SGCS solvent of the matrix polynomial (1.1). □

4. Numerical Experiments and Conclusion

In this section, we illustrate that our Newton's method converges to the SGCS solvent. And we compare some experimental results using our Newton's method and Kratz's method. All experiments are done in MATLAB 7.1 and all iterations are terminated when the relative residual

$\rho_P(X_k)$ satisfies

$$\rho_P(X_k) = \frac{\|fl(P(X_k))\|}{\|A_0\|\|X_k\|^m + \|A_1\|\|X_k\|^{m-1} + \cdots + \|A_m\|} \leq n\varepsilon$$

where $\varepsilon = 1.0\text{e} - 015$.

First we consider the cubic matrix equation which the Fréchet derivative is singular for given starting matrix. From it, we assured the result of Theorem 3.7.

Example 4.1. *Let*

$$P_1(X) \equiv \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} X^3 + \begin{bmatrix} -4 & 0 & -4 \\ -4 & 0 & -4 \\ -4 & 0 & -4 \end{bmatrix} X + \begin{bmatrix} 32 & 0 & -32 \\ 32 & 0 & -32 \\ 32 & 0 & -32 \end{bmatrix} = 0. \tag{4.1}$$

The symmetric orthogonal matrix $P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ *is given. We choose SGCS starting matrix* $X_0 = \begin{bmatrix} 3 & 0 & -2 \\ 0 & -6400 & 0 \\ -2 & 0 & 3 \end{bmatrix}$ *with respect to the matrix* P*, then the Fréchet derivative of the problem* (4.1) *is*

$$\begin{aligned} D_{X_0} &= I \otimes \left[A_0(X_0)^2 + A_2\right] + X_0^T \otimes A_0 X_0 + \left(X_0^T\right)^2 \otimes A_0 \\ &= \begin{bmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{bmatrix}, \end{aligned}$$

where $D_{11} = D_{33} = \begin{bmatrix} -22 & 0 & 31 \\ -22 & 0 & 31 \\ -22 & 0 & 31 \end{bmatrix}$, $D_{22} = \begin{bmatrix} 12784 & 0 & 40940809 \\ 12784 & 0 & 40940809 \\ 12784 & 0 & 40940809 \end{bmatrix}$, $D_{12} = D_{21} = D_{23} = D_{32} = 0$*, and* $D_{13} = D_{31} = \begin{bmatrix} 4 & 0 & -18 \\ 4 & 0 & -18 \\ 4 & 0 & -18 \end{bmatrix}$*, is singular. In this case, Kratz's method can not be applied to solve the equation* (4.1)*. But our Newton's method with Algorithm* 3.1 *gives a solvent of the problem* (4.1) *as follows*

$$X_5 = \begin{bmatrix} 3 & 0 & -1 \\ 0 & -38400 & 0 \\ -1 & 0 & 3 \end{bmatrix}.$$

We can easy to check that the solvent X_5 *is an SGCS matrix because* $PX_5P = X_5$ *holds. Figure* 1 *shows that the convergence result of the problem* (4.1)*.*

Example 4.2. *The cubic matrix equation is*

$$P_2(X) \equiv X^3 + \begin{bmatrix} 1 & -4 \\ -4 & 1 \end{bmatrix} X^2 + \begin{bmatrix} 0 & -4 \\ -4 & 0 \end{bmatrix} X = 0. \tag{4.2}$$

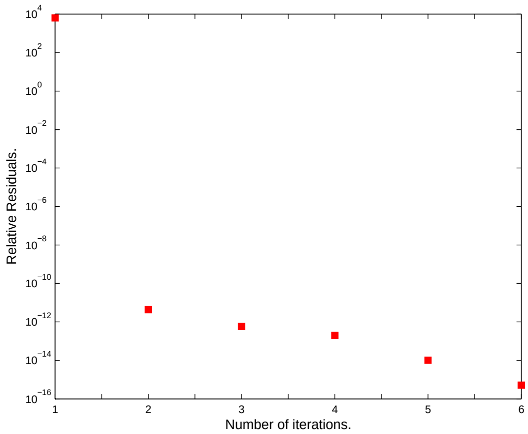

FIGURE 1. Convergence for the problem (4.1).

By using MATLAB's Symbolic Math Toolbox [18], *we have eight nontrivial SGCS solvents with respect to the symmetric orthogonal matrix* $P = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ *which are*

$$S_1 = \begin{bmatrix} -0.5 & -0.5 \\ -0.5 & -0.5 \end{bmatrix}, \; S_2 = \begin{bmatrix} -2.5 & 1.5 \\ 1.5 & -2.5 \end{bmatrix}, \; S_3 = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}, \; S_4 = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix},$$
$$S_5 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \; S_6 = \begin{bmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix}, \; S_7 = \begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix}, \; S_8 = \begin{bmatrix} 1.5 & 2.5 \\ 2.5 & 1.5 \end{bmatrix}.$$

By our Newton's method with the SGCS starting matrices

$$X_0 = \begin{bmatrix} x & y \\ y & x \end{bmatrix}, \; -10 \le x \le 10, \; -10 \le y \le 10, \; x, y \in Z, \tag{4.3}$$

the SGCS solvents $S_1, \cdots, S_8$ *can be obtained.*

5. CONCLUSION

We presented an iterative method for computing the Newton step (2.1) over the SGCS matrix. Then, we solved the matrix polynomial (1.1) using Newton's method. One of the advantage of our Newton's method is that it guarantee the convergence of solvents at which the Fréchet derivative is singular. Another is that if the matrix polynomial (1.1) satisfies the assumptions of Theorem 3.7, the SGCS solvent can be obtained for any given SGCS starting matrices. But, we still need to develop the existence theories for solvents and find practical examples.

REFERENCES

[1] Attahiru Sule Alfa, *Combinded elapsed time and matrix-analytic method for the discrete time GI/G/I and GIX/G/I systems*, Qucucing Systems, **45** (2003), 5–25.

[2] N. G. Bean, L. Bright, G. Latouche, C. E. M. Pearce, P. K. Pollett, and P. G. Taylor, *The quasi-stationary behavior of quasi-birth-death processes*, The Annals of Applied Probability, **7(1)** (1997), 134–155.

[3] G. J. Butler, C. R. Johnson, and H. Wolkowicz, *Nonnegative solutions of a quadratic matrix equation arising from comparison theorems in ordinary*, SIAM J. Alg. Disc. Meth., **6(1)** (1985), 47–53.

[4] H. Dai, *On the symmetric solutions of linear matrix equations*, Linear Algebra and Appl., **131** (1990), 1–7.

[5] George J. Davis, *Numerical solution of a quadratic matrix equation*, SIAM J. Sci. Stat. Comput., **2(2)** (1981), 164–175.

[6] Mehdi Dehghan and Masoud Hajarian, *Finite iterative algorithms for the reflexive and anti-reflexive solutions of the matrix equation* $A_1X_1B_1 + A_2X_2B_2 = C$, Mathematical and Computer Modeling, **49** (2009), 1937–1959.

[7] J. E. Dennis, Jr., J. F. Traub, and R. P. Weber, *The algebraic theory of matrix polynomials*, SIAM J. Numer. Anal., **13(6)** (1976), 831–845.

[8] J. E. Dennis, Jr., J. F. Traub, and R. P. Weber, *Algorithms for solvents of matrix polynomials*, SIAM J. Numer. Anal., **15(3)** (1978), 523–533.

[9] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, Academic Press, New York, 1982.

[10] Nicholas J. Higham and Hyun-Min Kim, *Numerical analysis of a quadratic matrix equation*, IMA J. Numer. Anal., **20(4)** (2000), 499–519.

[11] Chu K.E, *Symmetric solutions of the linear matrix equations by matrix decompositions*, Linear Algebra and Appl., **119** (1989), 35–50.

[12] Hyun-Min Kim, *Convergence of newton's method for solving a class of quadratic matrix equations*, Honam Mathematical Journal, **30(2)** (2008), 399–409.

[13] W. Kratz and E. Stickel, *Numerical solution of matrix polynomial equations by Newton's method*, IMA J. Numer. Anal., **7** (1987), 355–369.

[14] P. Lancaster and M. Tismenetsky, *The Theory of Matrices with Applications, 2nd ed*, Academic Press, New York, 1982.

[15] Peter Lancaster, *Lambda-Matrices and Vibrating Systems*, Pergamon Press, Oxford, 1966.

[16] Maolin Liang, Lifang Dai, and Sanfu Wang, *An iterative method for (R,S)-symmetric solution of matrix equation* $AXB = C$, Sci. Magna., **4** (2008) , 60–70.

[17] Maolin Liang, Chuanhua You, and Lifang Dai, *An efficient algorithm for the generalized centro-symmetric solution of matrix equation* $AXB = C$, Numer. Algor., **44** (2007), 173–184.

[18] Cleve Moler and Peter J. Costa, *Symbolic Math Toolbox Version 2: User's Guide*, The MathWorks, Inc, Natick, MA, USA, 1998.

[19] Jong Hyeon Seo and Hyun-Min Kim, *Solving matrix polynomials by Newton's method with exact line searches*, J. KSIAM, **12** (2008), 55–68.

[20] X. Y. Peng, X. Y. Hu, and L. Zhang, *An iterative method for symmetric slutions and the optimal approximation solution of the matrix equation* $AXB = C$, Appl. Math. Comput., **160** (2005), 763–777.