

# SOLVING NEARLY NON-SIMPLE MATRIX POLYNOMIAL EQUATIONS BY NEWTON'S METHOD WITH ADVANCED LINE SEARCHES

SANG-HYUP SEO

ABSTRACT.

## 1. INTRODUCTION

We consider a matrix polynomial equation(MPE) with  $n$ -degree defined by

$$(1.1) \quad P(X) = \sum_{k=0}^n A_k X^k = A_0 + A_1 X + \cdots + A_{n-1} X^{n-1} + A_n X^n = 0,$$

where the coefficient matrices  $A_k$ 's are  $m \times m$  matrices. Then, the unknown matrix  $X$  must be an  $m \times m$  matrix. A matrix  $S$  is called a solution of (1.1) if  $P(S) = 0$ .

The MPE (1.1) often occurs in the theory of differential equations, system theory, network theory, stochastic theory, quasi-birth-and-death and other areas [1–4, 10, 14, 22–24]. Specially, in quasi-birth-and-death and stochastic problems, finding the minimal nonnegative solution of a matrix equation is an important issue.

There are many researches to find the minimal nonnegative solution. Guo and Laub [13] considered a nonsymmetric algebraic Riccati equation, and they proposed iteration algorithms which converge to the minimal positive solution. In [11], Guo provided a sufficient condition for the existence of nonnegative solutions of nonsymmetric algebraic Riccati equations. Kim [21] showed that the minimal positive solutions also can be found by the Newton method with the zero initial matrices in some different types of quadratic equations. Seo and Kim [28, 31] studied the Newton iteration for matrix polynomial equations.

Newton's method is one of powerful tools to find solutions of nonlinear matrix equations. By Kantorovich theorem [17], the convergence rate of the method is quadratic if the derivative on the domain is Lipschitz continuous and at the solution is nonsingular. But, if the derivative at the solution is singular, then we cannot apply Kantorovich theorem, i.e., we cannot guarantee that the rate is quadratic. The followings are researches to analyze the problems with singular derivative at the solution and improve the method.

For general functions on Banach spaces, Reddien [26], Decker and Kelley [7, 8] gave analyses about Newton's method for singular problems. In [6], Decker, Keller, and Kelley provided an acceleration of Newton's method for singular problems and analyzed for the convergence rate of the method. Kelley and Suresh [20] suggested a new accelerated Newton's method at singular points. Decker and Kelley [9] showed an analysis for Newton's method at nearly singular roots. In [19], Kelley analyzed convergence rate of the method for functions whose high order derivatives at the

---

*Date:* February 28, 2020.

*2010 Mathematics Subject Classification.* 65H10.

*Key words and phrases.* matrix polynomial equation, positive solution, nonnegative solution,  $M$ -matrix, Newton's method, line search, acceleration of a method, nearly non-simple.

solution are singular. For specific functions on  $\mathbb{R}^{m \times n}$ , Guo and Lancaster [12] analyzed and provided a modification about Newton's method for algebraic Riccati equations at singular roots. In [29], S.-H. Seo and J.-H. Seo suggested a modified Newton method for matrix polynomial equations with  $n$ -degree for singular problems.

In [6, 20], accelerations for Newton's method at singular roots was suggested for general functions on Banach spaces. For specific functions on  $\mathbb{R}^{m \times n}$ , modifications of Newton's method at singular roots was provided in [12, 30]. But, the main hypotheses of the papers are similar and that the solution is non-simple, i.e., the derivative at the solution is singular. It means that the accelerations for Newton's method cannot be applied and the accelerated iterations cannot be guaranteed converge to the solutions if the solution is simple. Even if the iterations converge to the solutions, they cannot be guaranteed better than the pure iterations.

Otherwise, in [9], authors analyzed and suggested an acceleration for Newton's method about the case of nearly singular roots with the general bifurcation problem [18, 32]. In [28], it was suggested that a global accelerated Newton's method, which can be applied whether the solution is simple or not, for (1.1).

The aim of this paper is also to provide an improved Newton's method for (1.1) which satisfies following Assumption 1.1 at nearly singular roots. We show that the differences between the Newton sequences and the solution lie on nearly one-dimensional space. Moreover, with the exact line search idea in [27], we suggest a proper acceleration weight  $\alpha$ .

**Assumption 1.1.** For the MPE (1.1),

- 1) The coefficient matrices  $A_k$ 's are nonnegative except  $A_1$ .
- 2)  $-A_1$  is a nonsingular  $M$ -matrix.
- 3)  $A_0$ ,  $A_1$ , and  $\sum_{k=2}^n A_k$  are irreducible.

Here, we give some basic definitions and lemmas for this paper.

Let  $A, B \in \mathbb{R}^{m \times m}$  be matrices. If all elements of  $A$  is nonnegative, then we call that  $A$  is a *nonnegative* matrix and denote  $A \geq 0$ . In similar sense, we define  $A \leq 0$ ,  $A > 0$ , and  $A < 0$ . If a matrix  $A$  can be written as  $rI_m - B$  with  $B \geq 0$  and  $r \in \mathbb{R}$ , we call that  $A$  is a *Z-matrix*. Moreover, if  $A = rI_m - B$  is a *Z-matrix* and  $r \geq \rho(B)$  then  $A$  is called an *M-matrix*. The following is a basic theorem for *M-matrices*.

**Theorem 1.2.** For a *Z-matrix*  $A$ , the following are equivalent:

- (1)  $A$  is a nonsingular *M-matrix*.
- (2)  $A^{-1}$  is nonnegative.
- (3)  $Av > 0$  for some vector  $v > 0$ .
- (4) All eigenvalues of  $A$  have positive real parts.

Let a function  $F : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$  and an equation  $F(X) = 0$  be given. If nonnegative solutions  $S_1$  and  $S_2$  of  $F(X) = 0$  are satisfy

$$(1.2) \quad S_1 \leq S \leq S_2,$$

for any nonnegative solution  $S$  of  $F(X) = 0$ , then they are called the *minimal nonnegative solution* and the *maximal nonnegative solution*, respectively. The *minimal positive solution* and the *maximal positive solution* also are can be defined, similarly. If the Fréchet derivative of  $F$  at a solution  $S$  is nonsingular, then  $S$  is called *simple*. Furthermore, we call that a not simple solution is *non-simple*. In this paper, for convenience, the notation  $\|\cdot\|$  is used instead of the Frobenius norm  $\|\cdot\|_F$  and  $\tilde{X}$  denotes  $S - X$ , the difference between  $X$  and the solution  $S$ , because they are used very frequently.

## 2. CONVERGENCE OF NEWTON'S METHOD FOR THE MPE

In this paper, we denote that

$$(2.1) \quad P'[X](H) = \sum_{k=1}^n \sum_{l=0}^{k-1} A_k X^l H X^{k-l-1}$$

is the Fréchet derivative of the MPE (1.1) at  $X$  in the direction  $H$ , and

$$(2.2) \quad P''[X](K, H) = \sum_{k=2}^n \sum_{l=0}^{k-2} \sum_{j=0}^l A_k (X^l H X^j K X^{n-l-j-2} + X^l K X^j H X^{n-l-j-2})$$

is the second Fréchet derivative at  $X$ .

For given initial guess  $X_0$ , the Newton iteration of (1.1),

$$(2.3) \quad X_{i+1} = X_i - P'[X_i]^{-1}(P(X_i)), \quad i = 0, 1, 2, \dots$$

can be separated into two parts as

$$(2.4) \quad \begin{cases} P'[X_i](H_i) = -P(X_i), \\ X_{i+1} = X_i + H_i, \end{cases} \quad i = 0, 1, 2, \dots$$

If we define that

$$(2.5) \quad \mathcal{P}'[X] = \text{vec} \circ P[X]' \circ \text{vec}^{-1} = \sum_{k=1}^n \sum_{l=0}^{k-1} (X^{k-l-1})^T \otimes A_k X^l,$$

then solving the above equation of (2.4) is equivalent to solving the following  $m^2 \times m^2$  linear system

$$(2.6) \quad \mathcal{P}'[X_i] \text{vec}(H_i) = \text{vec}(-P(X_i)),$$

where  $\text{vec}$  is the vectorization by [16, Lemma 4.3.1].

We introduce a sufficient condition of the existence of the minimal nonnegative solution of the MPE (1.1) with Assumption 1.1, and give some analysis for Newton's method.

**Theorem 2.1.** [25, Theorem 2.1] *Let the MPE (1.1) with 1) and 2) in Assumption 1.1 be given. Then, there exists the minimal nonnegative solution if*

$$(2.7) \quad -\sum_{k=0}^n A_k \text{ is a nonsingular or singular irreducible } M\text{-matrix.}$$

**Theorem 2.2.** [30, Theorem 2.2] *Suppose that the MPE (1.1) satisfies Assumption 1.1 and (2.7). Then, the Newton sequence  $\{X_i\}$  with  $X_0 = 0$  is well defined, is monotone nondecreasing, and converges to the minimal positive solution  $S$ . Furthermore,  $-\mathcal{P}'[X_i]$  is a nonsingular irreducible  $M$ -matrix for  $i \in \mathbb{N}$ , and  $-\mathcal{P}'[S]$  is an irreducible  $M$ -matrix.*

3. THE NEARLY SINGULAR DERIVATIVE  $-\mathcal{P}'[S]$ 

Generally, by Kantorovich theorem [17], the Newton iteration (2.3) converges quadratically if  $-\mathcal{P}'[S]$  is nonsingular. But, if  $-\mathcal{P}'[S]$  becomes nearly singular, the quadratic convergence radius will be very small. In this section, we will see the converging tendency of (2.3) when  $-\mathcal{P}'[S]$  is nearly singular.

Let the MPE (1.1) satisfy Assumption 1.1,  $S$  be the minimal positive solution of (1.1). Then,  $-\mathcal{P}'[S]$  is an irreducible  $M$ -matrix. Therefore, the eigenvalue  $\mu$  of  $-\mathcal{P}'[S]$  whose absolute value is the smallest in absolute values of all eigenvalues is nonnegative and simple by [15, Perron-Frobenius Theorem].

Here, we think over an other MPE,

$$(3.1) \quad P_\varepsilon(X) = (A_0 - \varepsilon S) + (A_1 + \varepsilon I_m)X + \sum_{k=2}^n A_k X^k = 0,$$

where  $\varepsilon \leq \mu$ . Clearly,  $S$  is a solution of (3.1). Since the representation of the Fréchet derivative of  $P_\varepsilon$  at  $S$  is that

$$\begin{aligned} \mathcal{P}'_\varepsilon[S] &= I_m \otimes (A_1 + \varepsilon I_m) + \sum_{k=2}^n \sum_{l=0}^{k-1} (S^{k-l-1})^T \otimes A_k S^l \\ &= \mathcal{P}'[S] + \varepsilon I_{m^2}, \end{aligned}$$

$-\mathcal{P}'_\varepsilon[S]$  is an  $M$ -matrix by Theorem 1.2, and the nonnegative number  $\mu - \varepsilon$  is the smallest eigenvalue of  $-\mathcal{P}'_\varepsilon[S]$ . In addition,  $-A_1 - \varepsilon I_m$  is also an  $M$ -matrix.

**Lemma 3.1.** *For  $\varepsilon \leq \mu$ , (3.1) has  $S$  as the minimal positive solution.*

*Proof.* Clearly,  $S$  is a solution of (3.1). □

#### 4. WHAT

This paragraph is with reference to [9]. Let  $F_\varepsilon(X) = 0$  be the general bifurcation problem where  $\varepsilon$  is a chosen continuation parameter and  $\varepsilon = 0$  denotes the bifurcation point and let  $S_\varepsilon$  denote a solution arc through the branch point. We assume that the followings hold,

$$(4.1) \quad \begin{aligned} \mathcal{N}_0 &= \text{Ker}(-F'_0[S_0]) = \text{span}\{E_0\}, \quad \|E_0\| = 1, \\ -F'_0[S_0](\mathcal{M}_0) &= \mathcal{M}_0, \quad \mathbb{R}^{m \times m} = \mathcal{M}_0 \oplus \mathcal{N}_0. \end{aligned}$$

Then, we can define  $\mathbf{P}_{\mathcal{N}_0}$  to be the projection onto  $\mathcal{N}_0$  parallel to  $\mathcal{M}_0$  and  $\mathbf{P}_{\mathcal{M}_0} = I - \mathbf{P}_{\mathcal{N}_0}$ . Clearly,  $F'_0[S_0]$  is a Fredholm operator of index zero and hence will remain so for  $\varepsilon$  near zero. In this case, it can be shown [5] that the structure of (4.1) is also preserved for small  $\varepsilon$ , i.e., we can solve

$$(4.2) \quad -F'_\varepsilon[S_\varepsilon](E_\varepsilon) = \mu_\varepsilon E_\varepsilon$$

for a continuous eigenvalue-eigenvector pair  $(\mu_\varepsilon, E_\varepsilon)$  where  $\mu_0 = 0$ .

We assume that

$$(4.3) \quad \begin{aligned} -P'[S](E) &= \varepsilon E, \quad \|E\| = 1, \quad \mathcal{N}_\varepsilon = \text{span}\{E\}, \\ \mathbb{R}^{m \times m} &= \mathcal{M}_\varepsilon \oplus \mathcal{N}_\varepsilon, \quad P'[S](\mathcal{M}_\varepsilon) = \mathcal{M}_\varepsilon, \end{aligned}$$

where  $\varepsilon$  is a simple eigenvalue of  $-P'[S]$  whose absolute value is the smallest. Then, we can define  $\mathbf{P}_{\mathcal{N}_\varepsilon}$  to be the projection onto  $\mathcal{N}$  parallel to  $\mathcal{M}$  and  $\mathbf{P}_{\mathcal{M}_\varepsilon} = I - \mathbf{P}_{\mathcal{N}_\varepsilon}$ . If  $\varepsilon = 0$ , then the following theorem shows the converging tendency of (2.3) for the singular  $-P'_0[S]$ .

**Theorem 4.1.** (cf. [12, Theorem 1.5], [19, Theorem 1.1]) *Let  $\mathbf{P}_{\mathcal{N}_0} P''[S](E, E)$  be nonzero and let*

$$(4.4) \quad W(\rho, \theta) = \left\{ X \mid 0 < \|\tilde{X}\| < \rho, \|\mathbf{P}_{\mathcal{M}_0}(\tilde{X})\| \leq \theta \|\mathbf{P}_{\mathcal{N}_0}(\tilde{X})\| \right\}.$$

*If  $X_0 \in W(\rho_0, \theta_0)$  for  $\rho_0, \theta_0$  sufficiently small, then the Newton sequence  $\{X_i\}$  is well defined and  $\|P'[X_i]^{-1}\| \leq c \|\tilde{X}_i\|^{-1}$  for all  $i \geq 1$  and some constant  $c > 0$ . Moreover,*

$$(4.5) \quad \lim_{i \rightarrow \infty} \frac{\|\tilde{X}_{i+1}\|}{\|\tilde{X}_i\|} = \frac{1}{2}, \quad \lim_{i \rightarrow \infty} \frac{\|\mathbf{P}_{\mathcal{M}_0}(\tilde{X}_i)\|}{\|\mathbf{P}_{\mathcal{N}_0}(\tilde{X}_i)\|^2} = 0.$$

Defining

$$(4.6) \quad \begin{aligned} \lambda E &= \mathbf{P}_{\mathcal{N}_\varepsilon} P''[S](E, E), \\ \mu E &= \mathbf{P}_{\mathcal{N}_\varepsilon} \tilde{X}, \\ s &= \text{sign}(\varepsilon \lambda), \end{aligned}$$

we consider the conical region

$$(4.7) \quad W_s(\rho, \theta) = \left\{ X \mid 0 < \|\tilde{X}\| < \rho, \|\mathbf{P}_{\mathcal{M}_\varepsilon}(\tilde{X})\| \leq \theta \|\mathbf{P}_{\mathcal{N}_\varepsilon}(\tilde{X})\|, \text{sign}(\mu) = s \right\}.$$

**Theorem 4.2.** (cf. [9, Theorem 3.22]) *Assume  $\lambda \neq 0$ , and  $0 < \varepsilon < \bar{\varepsilon}$ . Then, there are continuous functions  $\rho = \rho(\varepsilon)$ ,  $\theta = \theta(\varepsilon)$ , monotonically increasing as  $\varepsilon \downarrow 0$  such that if  $X_0 \in W_s(\rho, \theta)$  then  $F_{X_0}'^{-1}$  exists and all subsequent Newton iterates remain in this set and converges to  $S = S(\varepsilon)$ . Further,*

$$(4.8) \quad \|\mathbf{P}_{\mathcal{N}} \tilde{X}_{i+1}\| < \frac{3}{4} \|\mathbf{P}_{\mathcal{N}} \tilde{X}_i\|, \quad \|\mathbf{P}_{\mathcal{M}}(\tilde{X}_{i+1})\| \leq K \|\tilde{X}_i\|^2,$$

for  $i \in \mathbb{N} \cup \{0\}$ , some  $K > 0$ , and  $\rho(\varepsilon)$ ,  $\theta(\varepsilon)$  may be chosen such that  $\rho(0) = \rho_0$ ,  $\theta(0) = \theta_0$  where  $\rho_0$ ,  $\theta_0$  are values for which the conclusions of Theorem 4.1 hold.

**Lemma 4.3.** *Suppose the matrix polynomial equation (1.1) satisfies Assumption 1.1. If the matrix  $-\mathcal{P}'_S$  is a singular M-matrix, then 0 is a simple eigenvalue of  $-\mathcal{P}'_S$ ,  $\mathcal{N} \oplus \mathcal{M} = \mathbb{R}^{m \times m}$ ,  $\mathcal{N}$  is one-dimensional and the map  $\mathcal{B}_{N_0}$  is invertible for some nonzero  $N_0 \in \mathcal{N}$ .*

*Proof.* Since  $S$  is positive and  $A_k$ 's are irreducible,  $-\mathcal{P}'_S$  is irreducible. Then, by Perron-Frobenius Theorem [15, Theorem 8.4.4], 0 is a simple eigenvalue of  $\mathcal{P}'_S$  with a positive eigenvector. Thus, we can find  $n^2$  linearly independent vectors  $\chi_1, \chi_2, \dots, \chi_{n^2}$  such that  $\chi_1 > 0$  and

$$(4.9) \quad \mathcal{X}^{-1} \mathcal{P}'_S \mathcal{X} = \begin{bmatrix} 0 & 0 \\ 0 & \mathcal{D} \end{bmatrix}, \text{ where } \mathcal{X} = \left[ \begin{array}{c|c|c|c} \chi_1 & \chi_2 & \cdots & \chi_{n^2} \end{array} \right]$$

and  $\mathcal{D}$  is an  $(n^2 - 1) \times (n^2 - 1)$  nonsingular matrix. By the same way, we also have a positive vector  $\psi$  such that  $\psi^T \mathcal{P}'_S = 0$ . Now,  $\mathcal{P}'_S(N) = 0$  if and only if  $\mathcal{P}'_S \text{vec}(N) = 0$ . From (4.9),  $\mathcal{P}'_S \text{vec}(N) = 0$  if and only if  $\text{vec}(N) \in \text{span}(\chi_1)$ , in which case we write  $N = a \text{vec}^{-1}(\chi_1)$  for some nonzero  $a \in \mathbb{R}$ . Thus,  $\mathcal{N} = \text{span}(\text{vec}^{-1}(\chi_1))$ . Similarly,  $\mathcal{M} = \text{span}(\text{vec}^{-1}(\chi_2), \dots, \text{vec}^{-1}(\chi_{n^2}))$ . Therefore,  $\mathcal{N}$  is one-dimensional and  $\mathbb{R}^{m \times m} = \mathcal{N} \oplus \mathcal{M}$ .

To prove the map  $\mathcal{B}_{N_0}$  is invertible for a nonzero matrix  $N_0 \in \mathcal{N}$ , we only need to show that

$$\mathbf{P}_{\mathcal{N}}(P''_S(N_0, N)) \neq 0,$$

for all nonzero  $N \in \mathcal{N}$  because  $\mathcal{B}_{N_0}$  is linear and  $\mathcal{N}$  is one-dimensional. Since  $\text{vec}^{-1}(\chi_1) > 0$  and  $S > 0$ , we have

$$\begin{aligned} P''_S(N_0, N) &= \sum_{k=2}^n \sum_{l=0}^{k-2} \sum_{j=0}^l A_k (S^l N_0 S^j N S^{n-l-j-2} + S^l N S^j N_0 S^{n-l-j-2}) \\ &= 2ab \sum_{k=2}^n \sum_{l=0}^{k-2} \sum_{j=0}^l A_k (S^l \text{vec}^{-1}(\chi_1) S^j \text{vec}^{-1}(\chi_1) S^{n-l-j-2}) \neq 0 \end{aligned}$$

where  $N = a \text{vec}^{-1}(\chi_1)$  and  $N_0 = b \text{vec}^{-1}(\chi_1)$ . Moreover,  $P''_S(N_0, N)$  is either positive or negative.

On the other hand,

$$\text{vec}(P''_S(N_0, N)) = k_1 \chi_1 + k_2 \chi_2 + \cdots + k_{n^2} \chi_{n^2}$$

for some real numbers  $k_1, k_2, \dots, k_{n^2}$ . By Fundamental theorem of linear algebra in [33] and Lemma 6.3.10 in [15], we have

$$(4.10) \quad \psi^T \text{vec}(P_S''(N_0, N)) = k_1 \psi^T \chi_1.$$

Since  $P_S''(N_0, N)$  is either positive or negative and  $\psi$  is positive, the left side of (4.10) is also either positive or negative. So,  $k_1$  cannot be zero. Therefore,

$$\mathbf{P}_N(P_S''(N_0, N)) = k_1 \text{vec}^{-1}(\chi_1) \neq 0.$$

□

**Lemma 4.4.** *Let  $S$  be the minimal positive solution of (1.1) with Assumption 1.1, and let  $\{X_i\}_{i=0}^\infty$  be a Newton sequence in (2.3). Then,*

$$\|P(X_i)\| \leq a\|\tilde{X}_i\|^2 + b\|\tilde{X}_i\|\|\tilde{X}_{i-1}\| + c\|\tilde{X}_{i-1}\|^2$$

for some positive real numbers  $a, b, c$ .

*Proof.* From Taylor's Theorem and putting  $S = X_{i-1} - \tilde{X}_{i-1}$ , we have

$$(4.11) \quad P(X_i) = P(S) + P'_S(X_i - S) + O(\|X_i - S\|^2),$$

and

$$0 = P(S) = P(X_{i-1} - \tilde{X}_{i-1}) = P(X_{i-1}) - P'_{X_{i-1}}(\tilde{X}_{i-1}) + O(\|\tilde{X}_{i-1}\|^2),$$

which is equivalent to

$$(4.12) \quad -P(X_{i-1}) + P'_{X_{i-1}}(\tilde{X}_{i-1}) = O(\|\tilde{X}_{i-1}\|^2).$$

From (2.4), we have

$$(4.13) \quad 0 = P'_{X_{i-1}}(X_i - X_{i-1}) + P(X_{i-1}),$$

and clearly

$$(4.14) \quad X_i - X_{i-1} = \tilde{X}_i - \tilde{X}_{i-1}.$$

If we subtract (4.13) from (4.11) and substitute (4.12), we obtain

$$\begin{aligned} P(X_i) &= P(S) + P'_S(\tilde{X}_i) - P(X_{i-1}) - P'_{X_{i-1}}(\tilde{X}_i - \tilde{X}_{i-1}) + O(\|\tilde{X}_i\|^2) \\ &= P'_S(\tilde{X}_i) - P'_{X_{i-1}}(\tilde{X}_i) + O(\|\tilde{X}_i\|^2) + O(\|\tilde{X}_{i-1}\|^2). \end{aligned}$$

Putting  $S = X_{i-1} - \tilde{X}_{i-1}$  in the previous equality,

$$\begin{aligned} P(X_i) &= \sum_{k=1}^n \sum_{l=0}^{k-1} A_k (X_{i-1} - \tilde{X}_{i-1})^l \tilde{X}_i (X_{i-1} - \tilde{X}_{i-1})^{k-l-1} \\ &\quad - P'_{X_{i-1}}(\tilde{X}_i) + O(\|\tilde{X}_i\|^2) + O(\|\tilde{X}_{i-1}\|^2) \\ &= \sum_{k=1}^n \sum_{l=0}^{k-1} A_k X_{i-1}^l \tilde{X}_i X_{i-1}^{k-l-1} + O(\|\tilde{X}_i\|\|\tilde{X}_{i-1}\|) \\ &\quad - P'_{X_{i-1}}(\tilde{X}_i) + O(\|\tilde{X}_i\|^2) + O(\|\tilde{X}_{i-1}\|^2) \\ &= P'_{X_{i-1}}(\tilde{X}_i) + O(\|\tilde{X}_i\|\|\tilde{X}_{i-1}\|) - P'_{X_{i-1}}(\tilde{X}_i) + O(\|\tilde{X}_i\|^2) + O(\|\tilde{X}_{i-1}\|^2) \\ &= O(\|\tilde{X}_{i-1}\|^2) + O(\|\tilde{X}_i\|\|\tilde{X}_{i-1}\|) + O(\|\tilde{X}_i\|^2). \end{aligned}$$

Since  $\|\cdot\|$  is a multiplicative matrix norm on  $\mathbb{R}^{m \times m}$ , we have required result. □

**Lemma 4.5.** *For any fixed  $\theta > 0$ , let*

$$\mathcal{Q} = \{i \in \mathbb{N} \cup \{0\} \mid \|\mathbf{P}_M(\tilde{X}_i)\| > \theta \|\mathbf{P}_N(\tilde{X}_i)\|\}$$

where  $\{X_i\}$  is the Newton sequence in Theorem 2.2. Then, there exist an integer  $i_0$  and a constant  $c > 0$  such that  $\|\tilde{X}_i\| \leq c\|\tilde{X}_{i-1}\|^2$  for all  $i \geq i_0$  in  $\mathcal{Q}$ .

*Proof.* Using Taylor's Theorem and the fact that  $P'_S(\mathbf{P}_\mathcal{N}(\tilde{X}_i)) = 0$ ,

$$(4.15) \quad P(X_i) = P(S) + P'_S(\tilde{X}_i) + O(\|\tilde{X}_i\|^2) = P'_S(\mathbf{P}_\mathcal{M}(\tilde{X}_i)) + O(\|\tilde{X}_i\|^2).$$

Since  $P'_S|_{\mathcal{M}} : \mathcal{M} \rightarrow \mathcal{M}$  is invertible,  $\|P'_S(\mathbf{P}_\mathcal{M}(\tilde{X}_i))\| \geq c_1 \|\mathbf{P}_\mathcal{M}(\tilde{X}_i)\|$  for some constant  $c_1 > 0$ . For  $i \in \mathcal{Q}$ , we have

$$\|\tilde{X}_i\| \leq \|\mathbf{P}_\mathcal{M}(\tilde{X}_i)\| + \|\mathbf{P}_\mathcal{N}(\tilde{X}_i)\| \leq (\theta^{-1} + 1) \|\mathbf{P}_\mathcal{M}(\tilde{X}_i)\|.$$

Thus by (4.15),

$$(4.16) \quad \|P(X_i)\| \geq c_1 \|\mathbf{P}_\mathcal{M}(\tilde{X}_i)\| - c_2 \|\tilde{X}_i\|^2 \geq c_1(\theta^{-1} + 1)^{-1} \|\tilde{X}_i\| - c_2 \|\tilde{X}_i\|^2.$$

On the other hand, from Lemma 4.4, we have

$$\|P(X_i)\| \leq c_3 \|\tilde{X}_i\|^2 + c_4 \|\tilde{X}_{i-1}\| \|\tilde{X}_i\| + c_5 \|\tilde{X}_{i-1}\|^2.$$

From (4.16) and the fact that  $X_i \neq S$  for any  $i$ , we have

$$c_1(\theta^{-1} + 1)^{-1} - c_2 \|\tilde{X}_i\| \leq c_3 \|\tilde{X}_i\| + c_4 \|\tilde{X}_{i-1}\| + c_5 \frac{\|\tilde{X}_{i-1}\|^2}{\|\tilde{X}_i\|}.$$

Since  $\tilde{X}_i$  converges to 0 by Theorem 2.2, we can find an  $i_0$  such that  $\|\tilde{X}_i\| \leq c \|\tilde{X}_{i-1}\|^2$  for all  $i \geq i_0$ .  $\square$

**Corollary 4.6.** *Assume that, for given  $\theta > 0$ ,  $\|\mathbf{P}_\mathcal{M}(\tilde{X}_i)\| > \theta \|\mathbf{P}_\mathcal{N}(\tilde{X}_i)\|$  for all  $i$  large enough. Then  $X_i \rightarrow S$  quadratically.*

When  $P'_S$  is singular practically the Newton sequence converges linearly, according to the corollary we conclude that *the error will generally be dominated by its  $\mathcal{N}$  component* [12]. From Lemmas 4.3 and 4.5 we have the following main theorem.

**Theorem 4.7.** *If  $-P'_S$  is a singular  $M$ -matrix and the convergence rate of the Newton sequence  $\{X_i\}$  in Theorem 2.2 is not quadratic, then  $\|P'_{X_i}\| \leq c \|\tilde{X}_i\|^{-1}$  for all  $i \geq 1$  and some constant  $c > 0$ . Moreover,*

$$\lim_{i \rightarrow \infty} \frac{\|\tilde{X}_{i+1}\|}{\|\tilde{X}_i\|} = \frac{1}{2}, \quad \lim_{i \rightarrow \infty} \frac{\|\mathbf{P}_\mathcal{M}(\tilde{X}_i)\|}{\|\mathbf{P}_\mathcal{N}(\tilde{X}_i)\|^2} = 0.$$

$$\begin{aligned}
-A_1 &= rI_m - N \\
-A_1^{-1} &= \frac{1}{r} \sum_{n=0}^{\infty} \left( \frac{N}{r} \right)^n \\
-A_1^{-1} A_0 &= \frac{1}{r} \sum_{n=0}^{\infty} \left( \frac{N}{r} \right)^n A_0 = X_1 > 0 \\
-A_1 - \varepsilon I_m &= (r - \varepsilon)I_m - N \\
(-A_1 - \varepsilon I_m)^{-1} &= \frac{1}{r - \varepsilon} \sum_{n=0}^{\infty} \left( \frac{N}{r - \varepsilon} \right)^n \\
&= \sum_{n=0}^{\infty} (-\varepsilon A_1^{-1})^n (-A_1^{-1}) \\
&= \sum_{n=0}^{\infty} \varepsilon^n (-A_1^{-1})^{n+1} \\
&= -A_1^{-1} + \sum_{n=1}^{\infty} \varepsilon^n (-A_1^{-1})^{n+1} \\
(-A_1 - \varepsilon I_m)^{-1} (A_0 - \varepsilon S) &= \sum_{n=0}^{\infty} \varepsilon^n (-A_1^{-1})^{n+1} (A_0 - \varepsilon S) \\
&= -A_1^{-1} (A_0 - \varepsilon S) + \sum_{n=1}^{\infty} \varepsilon^n (-A_1^{-1})^{n+1} (A_0 - \varepsilon S) \\
(-A_1 - \varepsilon I_m)^{-1} (A_0 - \varepsilon S) &= \frac{1}{r - \varepsilon} \sum_{n=0}^{\infty} \left( \frac{N}{r - \varepsilon} \right)^n A_0 - \frac{\varepsilon}{r - \varepsilon} \sum_{n=0}^{\infty} \left( \frac{N}{r - \varepsilon} \right)^n S
\end{aligned}$$

$$(5.1) \quad P_\varepsilon(X) = (A_0 - \varepsilon S) + (A_1 + \varepsilon I_m)X + \sum_{k=2}^n A_k X^k = 0, \quad \varepsilon \leq \mu$$

$$\begin{aligned}
\mathcal{P}'_\varepsilon[S] &= I_m \otimes (A_1 + \varepsilon I_m) + \sum_{k=2}^n \sum_{l=0}^{k-1} (S^{k-l-1})^T \otimes A_k S^l \\
&= \mathcal{P}'[S] + \varepsilon I_{m^2},
\end{aligned}$$

$$\begin{aligned}
X_1 &= X_0 - P'_\varepsilon[X_0]^{-1} (P(X_0)) \\
P'_\varepsilon[X_0](H_0) &= -P(X_0)
\end{aligned}$$

$$\begin{aligned}
&\sum_{k=2}^n \sum_{l=0}^{k-1} A_k X_0^l H_0 X_0^{k-l-1} + (A_1 + \varepsilon I_m) H_0 \\
&= -A_0 + \varepsilon S - A_1 X_0 - \varepsilon X_0 - \sum_{k=2}^n A_k X_0^k
\end{aligned}$$



If  $X_0 = 0$ ,

$$\begin{aligned}(A_1 + \varepsilon I_m)H_0 &= -A_0 + \varepsilon S \\ X_1 - 0 &= H_0 = -(A_1 + \varepsilon I_m)^{-1}(A_0 - \varepsilon S) \\ &= -(A_1 + \varepsilon I_m)^{-1}A_0 - (A_1 + \varepsilon I_m)^{-1}(-\varepsilon S).\end{aligned}$$

If  $X_0 = \varepsilon S$ ,

$$\begin{aligned}&\sum_{k=2}^n \sum_{l=0}^{k-1} A_k(\varepsilon S)^l H_0(\varepsilon S)^{k-l-1} + (A_1 + \varepsilon I_m)H_0 \\ &= -A_0 + \varepsilon S - A_1(\varepsilon S) - \varepsilon I_m(\varepsilon S) - \sum_{k=2}^n A_k(\varepsilon S)^k \\ &\sum_{k=2}^n \sum_{l=0}^{k-1} \varepsilon^{k-1} A_k S^l H_0 S^{k-l-1} + (A_1 + \varepsilon I_m)H_0 \\ &= -A_0 + \varepsilon S - \varepsilon A_1 S - \varepsilon^2 S - \sum_{k=2}^n \varepsilon^k A_k S^k\end{aligned}$$

## REFERENCES

- [1] Attahiru Sule Alfa. Combined elapsed time and matrix-analytic method for the discrete time  $GI/G/1$  and  $GI^X/G/1$  systems. *Queueing Syst.*, 45:5–25, 2003.
- [2] Nigel G. Bean, Leslie W. Bright, Guy Latouche, Charles E. M. Pearce, Philip K. Pollett, and Peter G. Taylor. The quasi-stationary behavior of quasi-birth-and-death processes. *Ann. Appl. Probab.*, 7(1):134–155, Feb 1997.
- [3] Dario A. Bini, Guy Latouche, and Beatrice Meini. *Numerical Methods for Structured Markov Chains*. Oxford University Press Oxford, 2005.
- [4] Geoffrey. J. Butler, Charles R. Johnson, and Henry Wolkowicz. Nonnegative solutions of a quadratic matrix equation arising from comparison theorems in ordinary differential equations. *SIAM J. Algebraic Discrete Methods*, 6(1):47–53, January 1985.
- [5] Michael G. Crandall and Paul H. Rabinowitz. *Bifurcation, Perturbation of Simple Eigenvalues and Linearized Stability*. University of Wisconsin-Madison, Mathematics Research Center, 1973.
- [6] D. W. Decker, H. B. Keller, and C. T. Kelley. Convergence rates for Newton's method at singular points. *SIAM J. Numer. Anal.*, 20(2):296–314, 1983.
- [7] D. W. Decker and C. T. Kelley. Newton's method at singular points. I. *SIAM J. Numer. Anal.*, 17(1):66–70, 1980.
- [8] D. W. Decker and C. T. Kelley. Newton's method at singular points. II. *SIAM J. Numer. Anal.*, 17(3):465–471, 1980.
- [9] D. W. Decker and C. T. Kelley. Expanded convergence domains for Newton's method at nearly singular roots. *SIAM J. Sci. Statist. Comput.*, 6(4):951–966, 1985.
- [10] Israel Gohberg, Peter Lancaster, and Leiba Rodman. *Matrix Polynomials*. Academic Press, 1982.
- [11] Chun-Hua Guo. Nonsymmetric algebraic Riccati equations and Wiener-Hopf factorization for M-matrices. *SIAM J. Matrix Anal. Appl.*, 23(1):225–242, 2001.
- [12] Chun-hua Guo and Peter Lancaster. Analysis and modification of Newton's method for algebraic Riccati equations. *Math. Comp.*, 67(223):1089–1105, 1998.
- [13] Chun-Hua Guo and Alan J Laub. On the iterative solution of a class of nonsymmetric algebraic Riccati equations. *SIAM J. Matrix Anal. Appl.*, 22(2):376–391, 2000.
- [14] Qi-Ming He and Marcel F. Neuts. On the convergence and limits of certain matrix sequences arising in quasi-birth-and-death Markov chains. *J. Appl. Probab.*, 38(2):519–541, 2001.
- [15] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
- [16] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1994.
- [17] L. V. Kantorovich and G. P. Akilov. *Functional Analysis in Normed Linear Spaces*. English translation, Pergamon Press, New York, 1964.

- [18] H.B. Keller. Numerical solution of bifurcation and nonlinear eigenvalue problems. In *Applications of bifurcation theory*, pages 359–384, New York, 1977. Proc. Advanced Sem., Univ. Wisconsin, Madison, Wis., 1976, Publ. Math. Res. Center, No. 38, Academic Press.
- [19] C. T. Kelley. A shamanskii-like acceleration scheme for nonlinear equations at singular roots. *Math. Comp.*, 47(176):609–623, 1986.
- [20] C. T. Kelley and R. Suresh. A new acceleration method for Newton’s method at singular points. *SIAM J. Numer. Anal.*, 20(5):1001–1009, 1983.
- [21] Hyun-Min Kim. Convergence of Newton’s method for solving a class of quadratic matrix equations. *Honam Math. J.*, 30(2):399–409, 2008.
- [22] Peter Lancaster. *Lambda-matrices and Vibrating Systems*. Pergamon Press, 1966.
- [23] Peter Lancaster and Miron Tismenetsky. *The Theory of Matrices with Applications*. Academic Press, 2 edition, 1985.
- [24] Guy Latouche and Vaidyanathan Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modeling*. ASA-SIAM, 1999.
- [25] Jie Meng, Sang-hyup Seo, and Hyun-min Kim. Condition numbers and backward error of a matrix polynomial equation arising in stochastic models. *J. Sci. Comput.*, 76(2):759–776, January 2018. First Online : <https://doi.org/10.1007/s10915-018-0641-x>.
- [26] G. W. Reddien. On Newton’s method for singular problems. *SIAM J. Numer. Anal.*, 15(5):993–996, 1978.
- [27] Jong Hyeon Seo and Hyun-Min Kim. Solving matrix polynomials by Newton’s method with exact line searches. *J. Korean Soc. Ind. Appl. Math.*, 12(2):55–68, 2008.
- [28] Jong-Hyeon Seo and Hyun-Min Kim. Convergence of pure and relaxed Newton methods for solving a matrix polynomial equation arising in stochastic models. *Linear Algebra Appl.*, 440:34–49, 2014.
- [29] Sang-hyup Seo and Jong-Hyeon Seo. Convergence of relaxed Newton method for order-convex matrix equations. *Comput. Appl. Math.*, 39(1):1–17, 2020.
- [30] Sang-hyup Seo, Jong-Hyeon Seo, and Hyun-Min Kim. Convergence of a modified Newton method for a matrix polynomial equation arising in stochastic problem. *Electron. J. Linear Algebra*, 34:500–513, October 2018.
- [31] Sang-Hyup Seo, Jong-Hyun Seo, and Hyun-Min Kim. Newton’s method for solving a quadratic matrix equation with special coefficient matrices. *Honam Math. J.*, 35(3):417–433, 2013.
- [32] Ivar Stakgold. Branching of solutions of nonlinear equations. *SIAM Rev.*, 13(3):289–332, 1971.
- [33] Gilbert Strang. *Linear Algebra and Its Applications*. Thomson, Brooks/Cole, 4th edition, 2006.

SANG-HYUP SEO, WHERE  
 Email address: [saibie1677@gmail.com](mailto:saibie1677@gmail.com)