ADVANCEDMPE

SANG-HYUP SEO

Abstract. We consider the Newton iteration for a matrix polynomial equation which arises in stochastic problem. In this paper, we show that the elementwise minimal nonnegative solution of the matrix polynomial equation can be obtained using Newton's method if the equation satisfies the sufficient condition, and the convergence rate of the iteration is quadratic if the solution is simple. Moreover, we show that the convergence rate is at least linear if the solution is non-simple, but we can apply a modified Newton method whose iteration number is less than the pure Newton iteration number. Finally, we give a numerical experiment which is related with our issue.

1. Introduction

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

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

where the coefficient matrices A_k's are $m \times m$ matrices. Then, the unknown matrix X must be an $m \times m$ matrix.

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, 7, 13, 18–20].

Davis [5,6] and Higham, Kim [14,15] studied the Newton method for a quadratic matrix equation. Guo and Laub [11] considered a nonsymmetric algebraic Riccati equation, and they proposed iteration algorithms which converge to the minimal positive solution. In [8], Guo provided a sufficient condition for the existence of nonnegative solutions of nonsymmetric algebraic Riccati equations. Kim [17] 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. Hautphenne, Latouche, and Remiche [12] studied the Newton method for the Markovian binary tree.

Seo and Kim [22,24] studied the Newton iteration for a quadratic matrix equation and a matrix polynomial equation. Specially, in [22], they provided a relaxed Newton method whose convergence is faster than the pure one. Guo and Lancaster [10] analyzed and provided a modification about Newton's method for algebraic Riccati equations. They showed that the modification of Newton's method is better than the pure one if the minimal nonnegative solution is non-simple.

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.

Date: December 2, 2019.

2010 *Mathematics Subject Classification.* 65H10.

Key words and phrases. matrix polynomial equation, elementwise positive solution, elementwise nonnegative solution, M-matrix, Newton's method, convergence rate, acceleration of a method.

3) $\sum_{k=0}^{n} A_k$ are irreducible.

Our goal is to propose a singular escaping Newton method for the MPE (1.1) which satisfies Assumption 1.1. This MPE is useful for stochastic theory, quasi-birth-and-death problem, and so on. In [10], Guo and Lancaster showed that $\|Y_{i+1} - S\| < c\varepsilon$ for the modified iteration Y_{i+1}, the solution S, a constant $c > 0$, and small $\varepsilon > 0$. Similarily, Seo, Seo, and Kim [23] showed that the modified Newton iteration Y_{i+1} for the MPE is closer to the solution S than the pure Newton iteration X_{i+1}. But, in both of [10, 23], the authors showed that the modifications are better than the pure if the solution S is non-simple.

We start with some basic definitions.

Definition 1.2. Let a matrix $A \in \mathbb{R}^{m\times m}$. A is an *Z-matrix* if all its off-diagonal elements are nonpositive.

It is clear that any Z-matrix A can be written as $sI - B$ with $B \geq 0$ and $s \in \mathbb{R}$. Then M-matrix can be defined as follows.

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

The following result is well known and can be found in [9] and [21] for example.

Theorem 1.4. *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.

Definition 1.5. A positive solution S_1 of the matrix equation $P(X) = 0$ is the *elementwise minimal positive solution* and a positive solution S_2 of $P(X) = 0$ is the *elementwise maximal positive solution* if, for any positive solution S of $P(X)$,

$$S_1 \leq S \leq S_2. \tag{1.2}$$

Similarly, if nonnegative solutions S_1 and S_2 satisfy (1.2) for any nonnegative solution S, then S_1 is called the *elementwise minimal nonnegative solution* and S_2 is called the *elementwise maximal nonnegative solution.*

Definition 1.6. [16, Definition 4.2.1, Definition 4.2.9] The *Kronecker product* of $A = [a_{ij}] \in \mathbb{C}^{m\times n}$ and $B = [b_{ij}] \in \mathbb{C}^{p\times q}$ is denoted by $A \otimes B$ and is defined to be the block matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \in \mathbb{C}^{mp\times nq}.$$

The vec operator vec : $\mathbb{C}^{m\times n} \to \mathbb{C}^{mn}$ is defined by

$$\mathrm{vec}(A) = \begin{bmatrix} \mathbf{a}_1^T & \mathbf{a}_2^T & \cdots & \mathbf{a}_n^T \end{bmatrix}^T,$$

where $\mathbf{a}_i^T = \begin{bmatrix} a_{1i} & a_{2i} & \cdots & a_{ni} \end{bmatrix}$.

Lemma 1.7. [16, Lemma 4.3.1] *Let $A \in \mathbb{C}^{m\times n}, B \in \mathbb{C}^{p\times q}$, and $C \in \mathbb{C}^{m\times q}$ be given and let $X \in \mathbb{C}^{n\times p}$ be unknown. The matrix equation*

$$AXB = C \tag{1.3}$$

is equivalent to the system of qm equations in np unknowns given by

$$(B^T \otimes A)\mathrm{vec}(X) = \mathrm{vec}(C), \tag{1.4}$$

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

Definition 1.8. Let a matrix function $F : \mathbb{C}^{m\times n} \to \mathbb{C}^{m\times n}$ be given, and let a matrix equation

$$F(X) = 0 \tag{1.5}$$

be given. Then, a solution $S \in \mathbb{C}^{m\times n}$ of (1.5) is called *simple* if the Fréchet derivative at S is nonsingular. For convenience, a solution is called *non-simple* if it is not simple.

For convenience, the notation $||\cdot||$ is used instead of the Frobenius norm $||\cdot||_F$ and $\mathbb{N}_0$ is used as $\mathbb{N}\cup\{0\}$ because the Frobenius norm and $\mathbb{N}_0$ are used very frequently in this paper.

2. Analysis for Nearly-Singular Points

Assumption 2.1. We assume that

1. $F(s) = 0$;
2. N_1 is eigenspace for the smallest eigenvalue;
3. $X = N_1 \oplus X_1$ with a closed X_1;
4. $F'[s]X_1 = X_1$, $F'[s]N_1 = N_1$;
5. P_{X_1} and P_{N_1};
6. $W_{\rho,\theta}(s) = \{x \in X : \|x - s\| \le \rho, \|P_{X_1}(x - s) \le \theta\|P_{N_1}(x - s)\|\}$.

Definition 2.2. (1) $\gamma(M, N) = \inf\{\mathrm{dist}(u, N) : u \in M, \|u\| = 1\}$.

Theorem 2.3. *Let*

(1) N_1 : 1-dim.;
(2) $F''[s]N_1N_1 \cap X_1 = \{0\}$;
(3) $\|F''[s]nx\| \ge c_1\|n\|\|x\|$ for all $n \in N_1, x \in X$ with $c_1 > 0$.

Then,

(1) $\exists\rho > 0,\ \theta > 0 : \exists F'[x]^{-1}$ for all $x(\ne s) \in W_{\rho,\theta}(s)$;
(2) $\|F'[x]^{-1}\| \le c_2\|x - s\|^{-1}$;
(3) $GW_{\rho,\theta}(s) \subset W_{\rho,\theta}(s)$ where $Gx \equiv x - F'[x]^{-1}F(x)$;
(4) $F^{-1}(0) \cap B_\rho(s) = s$;
(5) $x_i = Gx_{i-1} \Rightarrow x_i \to s$ and $\|P_{X_1}(x_i - s)\| \le c\|x_{i-1} - s\|^2$.

Proof. Let $\theta > 0$ and $\rho > 0$ be chosen. Since $F(x) = F(s) + F'[s](x - s) + \beta_2(x)$, then $F'[x]^{-1}F(x) = F'[x]^{-1}F'[s]P_{X_1}(x - s) + \beta_1(x) = P_{X_1}(x - s) + \beta_1(x)$ for $x \ne s$ in W. Thus $Gx \to s$ as $x \to s$ in W. Hence we define $Gs = s$.

$$F'[s]y = F'[x_0]y + F''[x_0](s - x_0)y + \frac{1}{2}F'''[x_0](s - x_0)^2y + \cdots$$

$$F''[x_0](x_0 - s)y = F'[x_0]y - F'[s]y + \frac{1}{2}F'''[x_0](s - x_0)^2y + \cdots$$

$$F''[x_0](x_0 - s)P_{N_1}(x_0 - s) = (F'[x_0] - F'[s])P_{N_1}(x_0 - s) + \beta_3$$

Choose $x_0 \in W$. Then $x_1 = Gx_0$, the first Newton iterate, is defined and

$$F'[x_0](x_1 - s) = F'[x_0](x_0 - s) - F(x_0).$$

Now

$$F'[x_0]^{-1}F(x_0) = -(s - x_0) - F'[x_0]^{-1}\frac{F''[x_0]}{2}(s - x_0)^2 + \beta_2(x_0),$$

and so

$$\begin{aligned}
x_1 - s &= x_0 - s - F'[x_0]^{-1}F(x_0) \\
&= \frac{1}{2}F'[x_0]^{-1}F''[x_0](x_0-s)^2 + \beta_2(x_0) \\
&= \frac{1}{2}F'[x_0]^{-1}F''[x_0](x_0-s)P_{X_1}(x_0-s) \\
&\quad + \frac{1}{2}F'[x_0]^{-1}F''[x_0](x_0-s)P_{N_1}(x_0-s) + \beta_2(x_0) \\
&= \frac{1}{2}F'[x_0]^{-1}F''[x_0](x_0-s)P_{X_1}(x_0-s) \\
&\quad + \frac{1}{2}F'[x_0]^{-1}(F'[x_0]-F'[s])P_{N_1}(x_0-s) + \beta_2(x_0) \\
&= \frac{1}{2}F'[x_0]^{-1}F''[x_0](x_0-s)P_{X_1}(x_0-s) \\
&\quad + \frac{1}{2}P_{N_1}(x_0-s) - \frac{1}{2}F'[x_0]^{-1}F'[s]P_{N_1}(x_0-s) + \beta_2(x_0) \\
&= \frac{1}{2}F'[x_0]^{-1}F''[x_0](x_0-s)P_{X_1}(x_0-s) \\
&\quad + \frac{1}{2}P_{N_1}(x_0-s) - \frac{1}{2}\lambda F'[x_0]^{-1}P_{N_1}(x_0-s) + \beta_2(x_0)
\end{aligned}$$

$$\begin{aligned}
P_{X_1}(x_1-s) &= \frac{1}{2}(P_{X_1}F'[x_0]^{-1})F''[x_0](x_0-s)P_{X_1}(x_0-s) \\
&\quad - \frac{1}{2}\lambda P_{X_1}F'[x_0]^{-1}P_{N_1}(x_0-s) + \beta_2(x_0)
\end{aligned}$$

$$\begin{aligned}
\|P_{X_1}(x_1-s)\| &\le c^*\|x_0-s\|^2 + |\lambda|c^{**}\|x_0-s\| \\
\frac{\|P_{X_1}(x_1-s)\|}{\|x_0-s\|} &\le c^*\|x_0-s\| + |\lambda|c^{**}
\end{aligned}$$

□

Lemma 2.4. *(1) There exists a constant $c_3 > 0$ such that $\|F'[x]y\| \ge c_3\|y\|$ holds for all $y \in X$ and $x \in B_\rho(s)$.*
(2) There exists a constant $c_4 > 0$ so that $\|(F'[x]-\lambda I)m\| \ge c_4\|x-s\|\|m\|$ holds for all $m \in N_1$ and all $x \in B_\rho(s)$.

Proof. (2)

$$\begin{aligned}
F'[x]m &= F'[s]m + F''[s]m(x-s) + \frac{1}{2}F'''[s]m(x-s)(x-s) + O(\|m\|\|x-s\|^3) \\
&= \lambda m + F''[s]m(x-s) + O(\|m\|\|x-s\|^2) \\
&= \lambda m + O(\|m\|\|x-s\|)
\end{aligned}$$

$$\begin{aligned}
F'[x]m - \lambda m &= O(\|m\|\|x-s\|) \\
(F'[x]-\lambda I)m &= O(\|m\|\|x-s\|)
\end{aligned}$$

□

Lemma 2.5. *There exist positive constants c_5, ρ, and θ such that $\gamma(F'[x]N_1, F'[x]X_1) \ge c_5$ for $x \ne s$ in $W_{\rho,\theta}(s)$.*

Proof. Fix $\rho > 0$. Suppose the conclusion is false. Then we have sequences $\{x_i\} \in B_\rho(s)$, $\{m_i\} \in N_1$, and $\{y_i\} \in X_1$ such that $\|F'[x_i]m_i\| = 1$, $\|F'[x_i]m_i - F'[x_i]y_i\| \equiv \varepsilon_i \to 0$, and $\|P_{X_i}(x_i - s)\| \leq \theta_i \|P_{N_1}(x_i - s)\|$ with $\theta_i \to 0$. Note for ρ small $F'[x]N_1 \neq \{0\}$, $x \neq s$ in $B_\rho(s)$. Anyway, $\gamma(F'[x]N_1, F'[x]X_1)$ is defined. Note also that there exist $a_1, a_2 > 0$ such that $a_1 \leq \|y_i\| \leq a_2$. Since $F''[s]N_1N_1$ is 1-dim. of X such that $F''[s]N_1N_1 \cap X_1 = \{0\}$, $\|w - P_{X_1}w\| \geq \alpha\|w\|$ for some $\alpha > 0$ and all w in $F''[s]N_1N_1$. We finally note that since $F'[x_i]m_i = F'[s]m_i + F''[s](x_i - s)m_i + \beta_2(x_i)m_i$, then

$$
\begin{aligned}
F'[x_i]m_i &= F'[s]m_i + F''[s](x_i - s)m_i + \beta_2(x_i)m_i \\
F'[x_i]m_i - \beta_2(x_i)m_i - \lambda m_i &= F''[s](x_i - s)m_i \\
\|F'[x_i]m_i - \beta_2(x_i)m_i - \lambda m_i\| &= \|F''[s](x_i - s)m_i\| \\
\|F'[x_i]m_i\| + c_6\|x_i - s\|^2\|m_i\| + |\lambda|\|m_i\| &\geq \|F'[x_i]m_i\| + \|\beta_2(x_i)m_i\| + |\lambda|\|m_i\| \\
&\geq \|F''[s](x_i - s)m_i\| \\
&\geq c_1\|x_i - s\|\|m_i\|
\end{aligned}
$$

$$1 \geq c_1\|x_i - s\|\|m_i\| - c_6\|x_i - s\|^2\|m_i\| - |\lambda|\|m_i\| \tag{2.1}$$

and

$$1 \leq |\lambda|\|m_i\| + \|F''[s]\|\|x_i - s\|\|m_i\| + c_6\|x_i - s\|^2\|m_i\|, \tag{2.2}$$

if $\|F''[s]\|$ is bounded. Thus for some constants a_3 and a_4,

$$0 < a_3 \leq \|x_i - s\| \cdot \|m_i\| \leq a_4 \tag{2.3}$$

holds for ρ sufficiently small.

Define $w_i = F''[s]P_{N_1}(x_i - s)m_i$. Then,

$$
\begin{aligned}
\alpha c_1\|P_{N_1}(x_i - s)\| \cdot \|m_i\| &\leq \|w_i - P_{X_1}w_i\| \leq c\|w_i - F'[s]y_i\| \\
&\leq c\|F''[s](x_i - s)m_i - F'[s]y_i\| + c\|F''[s]P_{X_1}(x_i - s)m_i\| \\
&\leq c\|F''[s](x_i - s)m_i - F'[x_i]y_i\| + c\|F'[x_i]y_i - F'[s]y_i\| \\
&\qquad + c\|F''[s]P_{X_1}(x_i - s)m_i\| \\
&\leq c\varepsilon_i + cc_6\|x_i - s\|^2\|m_i\| + \|\beta_1(x_i)\|a_2 c \\
&\qquad + c\|F''[s]\|\theta_i\|P_{N_1}(x_i - s)\| \cdot \|m_i\|,
\end{aligned} \tag{2.4}
$$

where $\|P_{N_1}z\| \leq c\|z\|$. Note that $\|P_{X_1}(x_i - s)\| \leq c\theta_i\|x_i - s\|$ and $\|P_{N_1}(x_i - s)\| \geq (1 - c\theta_i)\|x_i - s\|$. Substituting into (2.4) and dividing by $\|x_i - s\| \cdot \|n_i\|$, we have

$$\alpha c_1(1 - c\theta_i) \leq c\varepsilon_i a_3^{-1} + cc_6\|x_i - s\| + c_7 a_3^{-1}\|x_i - s\| + c_5\theta_i. \tag{2.5}$$

Thus for i sufficiently large, we have $\alpha c_1 \leq 2(cc_6 + c_7 a_3^{-1})\|x - s\|$. Thus for ρ sufficiently small, we have a contradiction, completing the proof. □

3. Note

$$Q(X) = A_2X^2 + A_1X + A_0$$

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] George J. Davis. Numerical solution of a quadratic matrix equation. *SIAM J. Sci. Comput.*, 2:164–175, 1981.

[6] George J. Davis. Algorithm 598: an algorithm to compute solvent of the matrix equation $AX^2 + BX + C = 0$. *ACM Trans. Math. Software*, 9(2):246–254, June 1983.

[7] Israel Gohberg, Peter Lancaster, and Leiba Rodman. *Matrix Polynomials*. Academic Press, 1982.

[8] Chun-Hua Guo. Nonsymmetric algebraic Riccati equations and Wiener-Hopf factorization for M-matrices. *SIAM J. Matrix Anal. Appl.*, 23(1):225–242, 2001.

[9] Chun-Hua Guo and Nicholas J. Higham. Iterative solution of a nonsymmetric algebraic Riccati equation. *SIAM J. Matrix Anal. Appl.*, 29:396–412, 2007.

[10] Chun-hua Guo and Peter Lancaster. Analysis and modification of Newton's method for algebraic Riccati equations. *Math. Comp.*, 67(223):1089–1105, 1998.

[11] 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.

[12] Sophie Hautphenne, Guy Latouche, and Marie-Ange Remiche. Newton's iteration for the extinction probability of a Markovian binary tree. *Linear Algebra Appl.*, 428:2791–2804, 2008.

[13] 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.

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

[15] Nicholas J. Higham and Hyun-Min Kim. Solving a quadratic matrix equation by Newton's method with exact line searches. *SIAM J. Matrix Anal. Appl.*, 23:303–316, 2001.

[16] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 2nd edition, 1995.

[17] Hyun-Min Kim. Convergence of Newton's method for solving a class of quadratic matrix equations. *Honam Math. J.*, 30(2):399–409, 2008.

[18] Peter Lancaster. *Lambda-matrices and Vibrating Systems*. Pergamon Press, 1966.

[19] Peter Lancaster and Miron Tismenetsky. *The Theory of Matrices with Applications*. Academic Press, 2 edition, 1985.

[20] Guy Latouche and Vaidyanathan Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modeling*. ASA-SIAM, 1999.

[21] George Poole and Thomas Boullion. A survey on M-matrices. *SIAM Rev.*, 16(4):419–427, 1974.

[22] 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.

[23] 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.

[24] 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.

Sang-hyup Seo, Where
Email address: `saibie1677@gmail.com`