East Asian Math. J.
Vol. 35 (2019), No. 3, pp. 341–349
http://dx.doi.org/10.7858/eamj.2019.030

ON NEWTON'S METHOD FOR SOLVING A SYSTEM OF NONLINEAR MATRIX EQUATIONS

Taehyeong Kim, Sang-Hyup Seo, and Hyun-Min Kim*

Abstract. In this paper, we are concerned with the minimal positive solution to system of the nonlinear matrix equations $A_1X^2+B_1Y+C_1=0$ and $A_2Y^2+B_2X+C_2=0$, where A_i is a positive matrix or a nonnegative irreducible matrix, C_i is a nonnegative matrix and $-B_i$ is a nonsingular M-matrix for $i=1,2$. We apply Newton's method to system and present a modified Newton's iteration which is validated to be efficient in the numerical experiments. We prove that the sequences generated by the modified Newton's iteration converge to the minimal positive solution to system of nonlinear matrix equations.

1. Introduction

In this paper, we consider the following system of nonlinear matrix equations that can be expressed in the form

$$\begin{cases} F_1(X,Y)=A_1X^2+B_1Y+C_1=0, \\ F_2(X,Y)=A_2Y^2+B_2X+C_2=0, \end{cases} \tag{1.1}$$

where $X,Y\in\mathbb{R}^{n\times n}$ are unknown matrices, A_i is a positive matrix or a nonnegative irreducible matrix, C_i is a nonnegative matrix for $i=1,2$, $-B_1$ and $-B_2$ are nonsingular M-matrices.

Set $A=\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$, $B=\begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}$, $C=\begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}$, and $Z=\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix}$, then system (1.1) can be equivalently reformulated as

$$F(Z)=AZ^2+BP^TZP+C=0, \tag{1.2}$$

where $P=\begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix}$ is a permutation matrix.

Received May 1, 2019; Accepted May 21, 2019.
2010 *Mathematics Subject Classification.* 15A24, 65F10, 65H10.
Key words and phrases. Matrix equation, minimal positive solution, Newton's method, M-matrix .
*Corresponding author.
This work was supported by a 2-Year Research Grant of Pusan National University.

(pISSN 1226-6973, eISSN 2287-2833)

In the past several decades, solving matrix equations has been a hot topic in linear algebra field. Among various matrix equations, the quadratic matrix equation

$$AZ^2 + BZ + C = 0 \tag{1.3}$$

has draw much attention. This kind of matrix equation has important applications in Quasi-birth-death process, random walk, tandem Jackson queue, see [1, 7, 11, 12] and the references there in.

The quadratic matrix equation of the form (1.2) is a generalization of equation (1.3), which has been widely studied. In [2, 3], Newton's method was applied to equation (1.3). Supporting theory and implication details were also showed there. Higham and Kim [5] incorporated exact line search into Newton's method which improves the global convergence properties of Newton's method. Moreover, there are many results about the perturbation analysis on equation (1.3), see [5, 9, 10] for details about the normwise, mixed and componentwise condition numbers. Similar results can be obtained for the equation of the form (1.2).

An efficient way to solve system (1.1) is to apply Newton's method to equation (1.2). Note that the size of the matrices in equation (1.2) is $2n \times 2n$, the computational cost of Newton's method may be very expensive if n is large. We present a modified Newton's iteration where the matrices dealt with are of size $n \times n$. We prove that the sequences generated by the modified iteration method converge to the minimal positive solution of system (1.1). It is validated in the numerical experiments that the modified Newton's method works much more efficiently than the Newton's method that is directly applied to equation (1.2).

This paper is organized as follows. In Section 2, we present a modified Newton's method and prove that the sequences generated by the iteration converge to the minimal positive solution of equation (1.1). In Section 3, some numerical experiments are given to show the efficiency of the modified Newton's iteration.

We begin with the notations used throughout this paper. $\mathbb{R}^{n\times n}$ stands for the set of $n \times n$ matrices with elements in field $\mathbb{R}$. For $X = (x_{ij}) \in \mathbb{R}^{n\times n}$, we write $X \geq 0$ ($X > 0$) and say that X is nonnegative (positive) if $x_{ij} \geq 0$ ($x_{ij} > 0$) holds for all i, j, and $X \geq Y$ ($X > Y$) is used as a different notation for $X - Y \geq 0$ ($X - Y > 0$). For a matrix $A = (a_{ij}) \in \mathbb{R}^{n\times n}$, $\mathrm{vec}(A)$ is a vector defined by $\mathrm{vec}(A) = (a_1^T, \ldots, a_n^T)^T$ with a_i as the i-th column of A. For matrices $B = (b_{ij}) \in \mathbb{R}^{m\times n}$ and $C \in \mathbb{R}^{p\times q}$, $B \otimes C = (b_{ij}C)$ is the Kronecker product of matrices B and C, which is a $mp \times nq$ block matrix. I_n is the identity matrix of size $n \times n$.

2. A modified Newton's iteration

In this section, we propose a modified Newton's method for obtaining the solution of system (1.1). We show that the sequences generated by the modified method converge to the minimal positive solution of (1.1).

Definition 2.1. ([4]) *A matrix $A \in \mathbb{R}^{n\times n}$ is an M-matrix if $A = sI - B$ for some nonnegative matrix B and s with $s \geq \rho(B)$ where ρ is the spectral radius; it is a singular M-matrix if $s = \rho(B)$ and a nonsingular M-matrix if $s > \rho(B)$.*

Theorem 2.2. ([6]) *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.*
(5) *$Av \geq 0$ implies $v \geq 0$.*

One of the numerical methods to solve system (1.1) is to apply Newton's method to equation (1.2). According to the ideas developed in [2, 3, 5], Newton's iteration for solving equation (1.2) can be stated as

$$\begin{cases} AZ_iH_i + AH_iZ_i + BP^TH_iP = -F(Z_i), \\ Z_{i+1} = Z_i + H_i. \end{cases} \qquad i = 1, 2, \ldots \tag{2.1}$$

Note that the matrices in equation (2.1) is of size $2n \times 2n$, which implies that the computation cost by iteration (2.1) is very expensive if n is very large. Hence, it is not so practical. To reduce the computation cost, we propose a modified Newton's iteration where the matrices dealt with are of size $n \times n$.

Consider the following iteration

$$\begin{cases} A_1X_iH_{i,1} + A_1H_{i,1}X_i + B_1H_{i,2} = -F_1(X_i, Y_i), \\ A_2Y_iH_{i,2} + A_2H_{i,2}Y_i + B_2H_{i,1} = -F_2(X_i, Y_i), \\ X_{i+1} = X_i + H_{i,1}, \\ Y_{i+1} = Y_i + H_{i,2}. \end{cases} \tag{2.2}$$

Given $X_0 = Y_0 = 0$, we will prove that the sequences $\{X_i\}$ and $\{Y_i\}$ generated by (2.2) converge to the minimal positive solution of system (1.1). By compatibility with Kronecker products, the first two equations of (2.2) are equivalent to

$$\begin{aligned} (I_n \otimes A_1X_i + X_i^T \otimes A_1)\mathrm{vec}(H_{i,1}) + (I_n \otimes B_1)\mathrm{vec}(H_{i,2}) &= \mathrm{vec}(-F_1(X_1, Y_1)), \\ (I_n \otimes A_2Y_i + Y_i^T \otimes A_2)\mathrm{vec}(H_{i,2}) + (I_n \otimes B_2)\mathrm{vec}(H_{i,1}) &= \mathrm{vec}(-F_2(X_1, Y_1)). \end{aligned} \tag{2.3}$$

For convenience of notation, we let

$$M_i = -\begin{bmatrix} I_n \otimes B_1 & I_n \otimes A_1X_i + X_i^T \otimes A_1 \\ I_n \otimes A_2Y_i + Y_i^T \otimes A_2 & I_n \otimes B_2 \end{bmatrix}.$$

Then, (2.3) can be rewritten as

$$M_i \begin{bmatrix} \mathrm{vec}(H_{i,2}) \\ \mathrm{vec}(H_{i,1}) \end{bmatrix} = \begin{bmatrix} \mathrm{vec}(F_1(X_i, Y_i)) \\ \mathrm{vec}(F_2(X_i, Y_i)) \end{bmatrix}. \tag{2.4}$$

Suppose that M_i is nonsingular, it follows from (2.2) that

$$\begin{bmatrix} \mathrm{vec}(Y_{i+1}) \\ \mathrm{vec}(X_{i+1}) \end{bmatrix} = \begin{bmatrix} \mathrm{vec}(Y_i) \\ \mathrm{vec}(X_i) \end{bmatrix} + (M_i)^{-1} \begin{bmatrix} \mathrm{vec}(F_1(X_i, Y_i)) \\ \mathrm{vec}(F_2(X_i, Y_i)) \end{bmatrix},$$

which leads to

$$M_i \begin{bmatrix} \mathrm{vec}(Y_{i+1}) \\ \mathrm{vec}(X_{i+1}) \end{bmatrix} = \begin{bmatrix} \mathrm{vec}(-A_1 X_i^2 + C_1) \\ \mathrm{vec}(-A_2 Y_i^2 + C_2) \end{bmatrix}. \tag{2.5}$$

Theorem 2.3. *Suppose A_i is a positive matrix or a nonnegative irreducible matrix, C_i is a nonnegative matrix for $i = 1, 2$, $-B_1$ and $-B_2$ are nonsingular M-matrices in (1.1). Suppose that there is a pair of positive matrices (U, V) such that $F_1(U, V) \leq 0$ and $F_2(U, V) \leq 0$. Set $X_0 = Y_0 = 0$, then the sequences $\{X_i\}$ and $\{Y_i\}$ generated by iteration (2.2) converge to the minimal positive solution of system (1.1), that is, there is a pair of matrices (S_X, S_Y) which is the minimal positive solution of (1.1) such that*

$$\lim_{i\to\infty} X_i = S_X,$$
$$\lim_{i\to\infty} Y_i = S_Y.$$

Moreover,

$$M_i = -\begin{bmatrix} I_n \otimes B_1 & I_n \otimes A_1 X_i + X_i^T \otimes A_1 \\ I_n \otimes A_2 Y_i + Y_i^T \otimes A_2 & I_n \otimes B_2 \end{bmatrix}$$

is a nonsingular M-matrix for each X_i, Y_i.

Proof. We use mathematical induction. Let U and V be positive matrices such that

$$\begin{cases} F_1(U, V) = A_1 U^2 + B_1 V + C_1 \leq 0, \\ F_2(U, V) = A_2 V^2 + B_2 U + C_2 \leq 0. \end{cases} \tag{2.6}$$

Since $X_0 = Y_0 = 0, (-B_1)^{-1} \geq 0$, $(-B_2)^{-1} \geq 0$ and $H_1 = P(-B)^{-1} C P^T \geq 0$. Therefore, the following statements

$$\begin{cases} X_i < U, \\ Y_i < V, \end{cases} \tag{2.7}$$

$$M_i \text{ is a nonsingular } M\text{-matrix}, \tag{2.8}$$

$$\begin{cases} X_i \leq X_{i+1}, \\ Y_i \leq Y_{i+1} \end{cases} \tag{2.9}$$

are true for $i = 0$. Suppose that the statements (2.7)−(2.9) are true for $i = k \in \mathbb{N}$, we can obtain from (2.5) and (2.6),

$$\begin{aligned}
&M_k \begin{bmatrix} \mathrm{vec}(V - Y_{k+1}) \\ \mathrm{vec}(U - X_{k+1}) \end{bmatrix} \\
&= M_k \begin{bmatrix} \mathrm{vec}(V) \\ \mathrm{vec}(U) \end{bmatrix} - M_k \begin{bmatrix} \mathrm{vec}(Y_{k+1}) \\ \mathrm{vec}(X_{k+1}) \end{bmatrix} \\
&= \begin{bmatrix} \mathrm{vec}(-B_1 V - A_1 U X_k - A_1 X_k U + A_1 X_k^2 - C_1) \\ \mathrm{vec}(-B_2 U - A_2 V Y_k - A_2 Y_k V + A_2 Y_k^2 - C_2) \end{bmatrix} \\
&\geq \begin{bmatrix} \mathrm{vec}(A_1 U^2 - A_1 U X_k - A_1 X_k U + A_1 X_k^2) \\ \mathrm{vec}(A_2 V^2 - A_2 V Y_k - A_2 Y_k V + A_2 Y_k^2) \end{bmatrix} \\
&= \begin{bmatrix} \mathrm{vec}(A_1 (U - X_k)^2) \\ \mathrm{vec}(A_2 (V - Y_k)^2) \end{bmatrix} > 0.
\end{aligned} \tag{2.10}$$

Since (2.8) is true for $i = k$, (2.10) implies $X_{k+1} < U$ and $Y_{k+1} < V$. For showing the statement (2.8) for $i = k + 1$, we consider an equation

$$M_{k+1} \begin{bmatrix} \mathrm{vec}(V - Y_{k+1}) \\ \mathrm{vec}(U - X_{k+1}) \end{bmatrix}. \tag{2.11}$$

By compatibility with Kronecker products, equation (2.11) can be calculated as

$$\begin{aligned}
&M_{k+1} \begin{bmatrix} \mathrm{vec}(V - Y_{k+1}) \\ \mathrm{vec}(U - X_{k+1}) \end{bmatrix} \\
&= M_{k+1} \begin{bmatrix} \mathrm{vec}(V) \\ \mathrm{vec}(U) \end{bmatrix} - M_{k+1} \begin{bmatrix} \mathrm{vec}(Y_{k+1}) \\ \mathrm{vec}(X_{k+1}) \end{bmatrix} \\
&= \begin{bmatrix} \mathrm{vec}(B_1 Y_{k+1} + A_1 X_{k+1}^2 + A_1 X_{k+1}^2) \\ \mathrm{vec}(B_2 X_{k+1} + A_2 Y_{k+1}^2 + A_2 Y_{k+1}^2) \end{bmatrix} \\
&\quad - \begin{bmatrix} \mathrm{vec}(B_1 V + A_1 X_{k+1} U + A_1 U X_{k+1}) \\ \mathrm{vec}(B_2 U + A_2 Y_{k+1} V + A_2 V Y_{k+1}) \end{bmatrix} \\
&\geq \begin{bmatrix} \mathrm{vec}(A_1 X_{k+1}^2 - A_1 X_k X_{k+1} - A_1 X_{k+1} X_k + A_1 X_k^2) \\ \mathrm{vec}(A_2 Y_{k+1}^2 - A_2 Y_k Y_{k+1} - A_2 Y_{k+1} Y_k + A_2 Y_k^2) \end{bmatrix} \\
&\quad + \begin{bmatrix} \mathrm{vec}(A_1 U^2 - A_1 X_{k+1} U - A_1 U X_{k+1} + A_1 X_{k+1}^2) \\ \mathrm{vec}(A_2 V^2 - A_2 Y_{k+1} V - A_2 V Y_{k+1} + A_2 Y_{k+1}^2) \end{bmatrix} \\
&= \begin{bmatrix} \mathrm{vec}(A_1 (X_{k+1} - X_k)^2) \\ \mathrm{vec}(A_2 (Y_{k+1} - Y_k)^2) \end{bmatrix} + \begin{bmatrix} \mathrm{vec}(A_1 (U - X_{k+1})^2) \\ \mathrm{vec}(A_2 (V - Y_{k+1})^2) \end{bmatrix} > 0.
\end{aligned}$$

Since the statement (2.7) is true for $i = k + 1$, M_{k+1} is a nonsingular M-matrix. It implies that the statement (2.8) is true with $i = k + 1$. To show the statement (2.9) is true for $i = k+1$, we consider following equation

$$\begin{aligned} & M_{k+1} \begin{bmatrix} \mathrm{vec}(Y_{k+2} - Y_{k+1}) \\ \mathrm{vec}(X_{k+2} - X_{k+1}) \end{bmatrix} \\ & = \begin{bmatrix} \mathrm{vec}(B_1 Y_{k+1} + A_1 X_{k+1}^2 + A_1 X_{k+1}^2) \\ \mathrm{vec}(B_2 X_{k+1} + A_2 Y_{k+1}^2 + A_2 Y_{k+1}^2) \end{bmatrix} \\ & \quad - \begin{bmatrix} \mathrm{vec}(B_1 Y_{k+2} + A_1 X_{k+1} X_{k+2} + A_1 X_{k+2} X_{k+1}) \\ \mathrm{vec}(B_2 X_{k+2} + A_2 Y_{k+1} Y_{k+2} + A_2 Y_{k+2} Y_{k+1}) \end{bmatrix} \\ & = \begin{bmatrix} \mathrm{vec}(A_1 X_{k+1}^2 + B_1 Y_{k+1} + C_1) \\ \mathrm{vec}(A_2 Y_{k+1}^2 + B_2 X_{k+1} + C_2) \end{bmatrix} \\ & \geq \begin{bmatrix} \mathrm{vec}(A_1 (X_{k+1} - X_k)^2) \\ \mathrm{vec}(A_2 (Y_{k+1} - Y_k)^2) \end{bmatrix} \geq 0. \end{aligned}$$

Since M_{k+1} is a nonsingular M-matrix, we get $X_{k+1} \leq X_{k+2}$ and $Y_{k+1} \leq Y_{k+2}$. Finally, the statement (2.9) is true for $i = k+1$.

Therefore, the statements (2.7), (2.8) and (2.9) are true for all $i \in \mathbb{N}$. Therefore, the matrix sequences $\{X_i\}$ and $\{Y_i\}$ are monotonically increasing and bounded above. By Monotone convergence theorem, there are positive matrices S_X and S_Y such that $\lim_{i\to\infty} X_i = S_X$ and $\lim_{i\to\infty} Y_i = S_Y$. Moreover, for any other pair of positive solutions (S_U, S_V), since it holds that $F_1(S_U, S_V) \leq 0$ and $F_2(S_U, S_V) \leq 0$, we get from statement (2.7) that $S_X \leq S_U$ and $S_Y \leq S_V$. Hence, $\{S_X, S_Y\}$ is the pair of minimal positive solution of system (1.1).

□

3. Numerical example

In this section, we compare the modified iteration (2.2) with Newton's iteration (2.1) through numerical examples. The comparison criterion is the time that takes to find the solution of the same system of nonlinear matrix equations. All of computations are performed by using MATLAB/version 2019a with 1.90GHz Intel Core i7 CPU and 8GB memory.

Example 3.1. *Let $n = 10, 15, \ldots, 50$, and A_i, B_i, C_i for $i = 1, 2$ are $n \times n$ matrices which, in MATLAB code, are defined as*

$$\begin{aligned} A_1 &= \textsf{rand(n)}, \\ A_2 &= \textsf{rand(n)}, \\ B_1 &= \textsf{rand(n)*n-eye(n)*n\^{}2}, \\ B_2 &= \textsf{rand(n)*n-eye(n)*n\^{}2}, \\ C_1 &= \textsf{rand(n)}, \\ C_2 &= \textsf{rand(n)}. \end{aligned}$$

And set $Z_0 = 0$ in (2.1), then Theorem 2.5 implies that (1.1) has the minimal positive solution (X, Y). We set tolerance to 10^{-16} for break each iteration. The relative residue of (2.1) and (2.2) is defined as

$$\delta = \frac{||AZ_{i+1}^2 + BP^T Z_{i+1} P + C||_F}{||A||_F ||Z_{i+1}||_F^2 + ||B||_F ||P^T||_F |Z_{i+1}||_F ||P||_F + ||C||_F} < 10^{-16}.$$

If each relative residue is less than tolerance, terminate each iteration. We repeat this process for 30 times and compare the average time taken for each iteration.

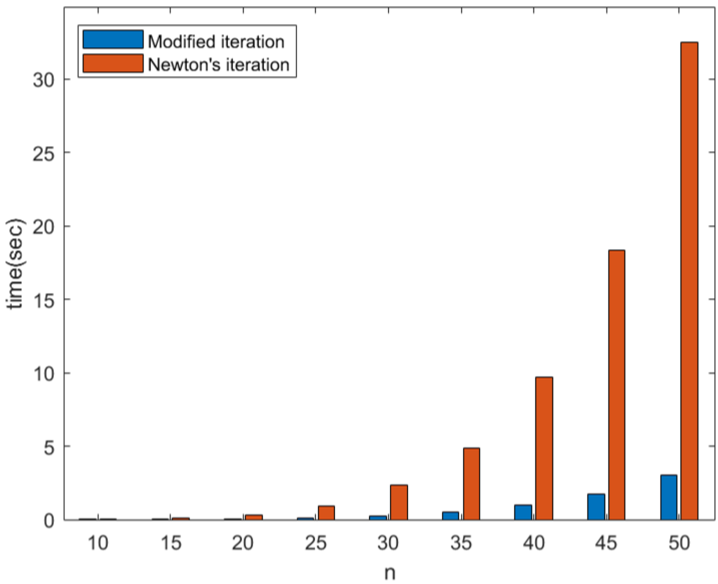

FIGURE 1. Comparision of CPU time

Figure 1 shows that when n is increasing, the difference of CPU time to find the solution of system (1.1) between the modified iteration and the classical Newton's iteration becomes larger.

Example 3.2. *Let $n = 10, 30, 50$ and A_i, B_i, C_i for $i = 1, 2$ are $n \times n$ matrices which are defined as*

$$\begin{aligned}
A_1 &= \textit{rand(n)\^{}2}, \\
A_2 &= \textit{rand(n)}, \\
B_1 &= \textit{rand(n)-eye(n)*n\^{}2}, \\
B_2 &= \textit{rand(n)-eye(n)*n\^{}2}, \\
C_1 &= \textit{rand(n)\^{}2}, \\
C_2 &= \textit{rand(n)}.
\end{aligned}$$

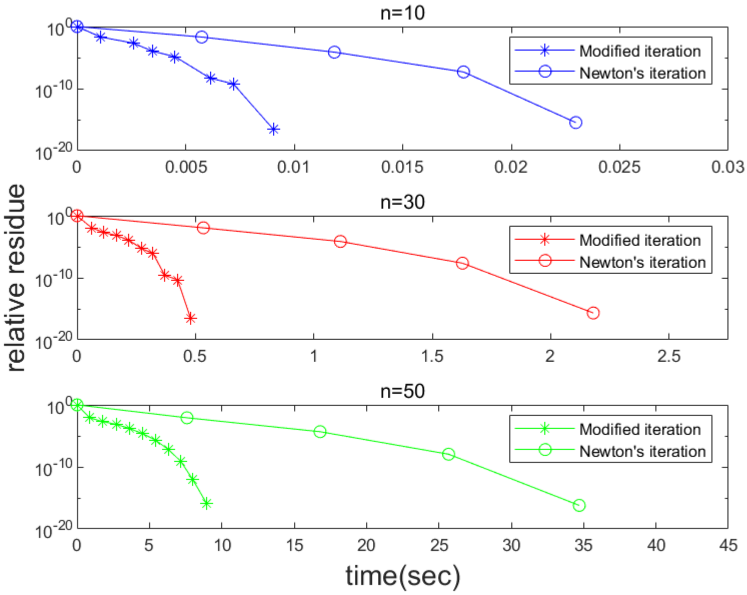

FIGURE 2. Comparision of CPU time and relative residue

We set same tolerance and relative residue as in Example 3.1.

Figure 2 shows that the modified Newton's method (2.2) requires more iteration numbers than iteration (2.1), but it takes less CPU time for convergence. The modified Newton's method (2.2) terminates before Newton's iteration repeating twice.

4. Conclusion

In this paper, to obtain the minimal positive solution of system (1.1), if Newton's method is applied to equation (1.2), it has the disadvantage that the size of the matrices used for computation in the iteration is $2n \times 2n$, which may lead to expensive computational cost. Therefore, we propose a modified Newton's method (2.2) which reduces the size of the matrices used for computation to $n \times n$. And from several numerical examples, we show that our modified Newton's method is much faster than Newton's method which is directly applied to equation (1.2).

References

[1] D.A. Bini & B. Meini, The cyclic reduction algorithm: from Poisson equation to stochastic processes and beyond, Numerical Algorithms, 51 (2009) 23–60.

[2] G. J. Davis, Algorithm 598: An algorithm to compute solvents of the matrix equation $AX^2 + BX + C = 0$, ACM Trans. Math. Software, 9 (1983) 246–254.
[3] J. E. Dennis & Jr. and R. B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice-Hall, Englewood Cliffs, NJ, 1983.
[4] C.-H. Guo & N.J. Higham, Iterative solution of a nonsymmetric algebraic Riccati equation, SIAM J. Matrix Anal. Appl., 29 (2007) 396–412.
[5] N.J. Higham & H.-M. Kim, Solving a duadratic matrix equation by Newton's method with exact line searches, SIAM J. Matrix. Anal. Appl., 23 (2006) 303–316.
[6] R.A. Horn & C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1994.
[7] JR. Jackson, Networks of waiting lines, Operations Research, 5 (1957) 518–521.
[8] H.-M. Kim Convergence of Newton's method for solving a class of quadratic matrix equations, Honam Mathematic J.,30 (2008) 399–409.
[9] L.-D. Liu & X. Lu, Two kinds of condition numbers for the quadratic matrix equation, Appl. Math. Comput., 219 (2013) 8759–8769.
[10] J. Meng & S.-H. Seo & H.-M. Kim, Condition numbers and backward error of a matrix polynomial equation arising in stochastic models, J. Sci. Comput., 76 (2018) 759–776.
[11] M. Miyazawa, Light tail asymptotics in multidimensional rflecting processes for queueing networks, Top, 19 (2011) 233–299.
[12] M.F. Neuts, Structured stochastic matrices of M/G/1 type and their applications, volume 5 of Probability: Pure and Applied. Marcel Dekker, Inc., New York, 1989.

Taehyeong Kim
Department of Mathematics, Pusan National University, Busan 46241, Republic of Korea
E-mail address: xogud7936@pusan.ac.kr

Sang-Hyup Seo
Department of Mathematics, Pusan National University, Busan 46241, Republic of Korea
E-mail address: saibie1677@gmail.com

Hyun-Min Kim
Department of Mathematics, Pusan National University, Busan 46241, Republic of Korea
E-mail address: hyunmin@pusan.ac.kr