New Iteration Method for a Quadratic Matrix Equation Associated with an M-Matrix

GUAN Jinrui(关晋瑞)\textsuperscript{1}, SONG Ruying(宋儒瑛)\textsuperscript{1}, ZUBAIR Ahmed\textsuperscript{2}

\textsuperscript{1. Department of Mathematics, Taiyuan Normal University, Jinzhong 030619, China; 2. Institute of Mathematics and Computer Science, University of Sindh, Sindh 76080, Pakistan)

Abstract: In this paper, we consider numerical solution of a quadratic matrix equation associated with an M-matrix, which arises in the study of noisy Wiener-Hopf problems for the Markov chain. The solution of practical interest is the M-matrix solution. By a simple transformation, this quadratic matrix equation is transformed into an M-matrix algebraic Riccati equation. We propose a new iteration method for this equation and then give the convergence analysis of it. Numerical experiments are given to show that the new iteration method is feasible and effective than some existing methods in some cases.

Key words: Quadratic matrix equation; M-matrix; Algebraic Riccati equation; Iteration method

CLC Number: O241.6 AMS(2000) Subject Classification: 65F30; 65F60; 15B99

Document code: A Article ID: 1001-9847(2021)01-0001-07

1. Introduction

In this paper, we consider a quadratic matrix equation (QME)

\[ X^2 - EX - F = 0, \]  

(1.1)

where \( E, F \in \mathbb{R}^{n \times n} \), \( E \) is a diagonal matrix and \( F \) is an M-matrix. The study of equation (1.1) is motivated by noisy Wiener-Hopf problems for Markov chains. See [5, 8] for more background details.

Under some conditions, it was proved in [5] that the equation (1.1) has an M-matrix solution, which is of practical interest. In addition, by transforming it into an equivalent M-matrix algebraic Riccati equation, a fixed-point iteration method and Newton method have been developed for solving the QME in [5]. However, the fixed-point iteration method converges too slowly while the Newton method is too expensive at each iteration. So they are not very efficient for this problem. In this paper, our main aim is to develop an efficient method for solving the QME (1.1).

Received date: 2019-10-11

Foundation item: Supported by the National Natural Science Foundation of China (11401424); the Natural Science Foundation of Shanxi Province (201901D211423); the Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi (2019L0783)

Biography: GUAN Jinrui, male, Han, Shanxi, associate professor, major in numerical linear algebra.
In the following, we first review some basic results of M-matrix and M-matrix algebraic Riccati equation.

Let $A = (a_{ij}) \in \mathbb{R}^{n \times n}$. If $a_{ij} \leq 0$ for all $i \neq j$, then $A$ is called a Z-matrix. A Z-matrix $A$ is called an M-matrix if there exists a nonnegative matrix $B$ such that $A = sI - B$ and $s \geq \rho(B)$ where $\rho(B)$ is the spectral radius of $B$. In particular, $A$ is called a nonsingular M-matrix if $s > \rho(B)$ and singular M-matrix if $s = \rho(B)$.

The following lemmas can be found in [1, 11].

**Lemma 1.1** Let $A \in \mathbb{R}^{n \times n}$ be a Z-matrix. Then the following statements are equivalent:
1) $A$ is a nonsingular M-matrix;
2) $A^{-1} \geq 0$;
3) $Av > 0$ for some vectors $v > 0$;
4) All eigenvalues of $A$ have positive real part.

**Lemma 1.2** Let $A, B$ be Z-matrices. If $A$ is a nonsingular M-matrix and $A \leq B$, then $B$ is also a nonsingular M-matrix. In particular, for any nonnegative real number, $B = I + A$ is a nonsingular M-matrix.

**Lemma 1.3** Let $A$ be an M-matrix, $B \geq A$ be a Z-matrix. If $A$ is nonsingular or irreducible singular with $A \neq B$, then $B$ is also a nonsingular M-matrix.

**Lemma 1.4** Let $A, B$ be nonsingular M-matrices and $A \leq B$, then $A^{-1} \geq B^{-1}$.

M-matrix algebraic Riccati equation is of the form
\[ XCX - XD - AX + B = 0, \] (1.2)
where $A, B, C$ and $D$ are real matrices of sizes $m \times m$, $m \times n$, $n \times m$ and $n \times n$ respectively. MARE appears in many branches of applied mathematics, such as transport theory, Markov chains, stochastic process, and so on. See [2-3] and the references therein for details. For the MARE (1.2), the solution of practical interest is its minimal nonnegative solution. The following basic result is obtained in [3-4].

**Lemma 1.5** For the MARE (1.2), if $K = \begin{pmatrix} D & -C \\ -B & A \end{pmatrix}$ is a nonsingular M-matrix or an irreducible singular M-matrix, then (1.2) has a minimal nonnegative solution $S$. If $K$ is a nonsingular M-matrix, then $A - SC$ and $D - CS$ are also nonsingular M-matrices. If $K$ is irreducible M-matrix, then $S > 0$ and $A - SC$ and $D - CS$ are also irreducible M-matrices.

When $K$ is an irreducible singular M-matrix, there exist unique, up to a multiplicative constant, $u > 0$ and $v > 0$ such that $u^T K = 0$, $K v = 0$ and $u^T v = 1$. Partition the vectors $u$ and $v$ according to the block structure of the matrix $M$ as $u^T = [u_1^T, u_2^T]$, $v^T = [v_1^T, v_2^T]$. Let $\mu = u_2^T v_2 - u_1^T v_1$, we have the following result.[3]

**Lemma 1.6** If $K$ is an irreducible singular M-matrix and $S$ is the minimal nonnegative solution of the MARE (1.2). Then
(i) when $\mu < 0$, $D - CS$ is singular and $A - SC$ is nonsingular;
(ii) when $\mu > 0$, $D - CS$ is nonsingular and $A - SC$ is singular;
(iii) when $\mu = 0$, both $D - CS$ and $A - SC$ are singular.

Efficient methods for solving the MARE (1.2) include the Schur method, the fixed-point iteration, the Newton iteration, the doubling algorithms and etc.\[2-3, 6-7, 9-10\]
2. A New Iteration Method

In this section, we first briefly introduce the fixed-point iteration method in [5] for solving the QME (1.1), and then propose a new iteration method for solving (1.1).

By introducing $X = I - Y$, the equation (1.1) can be transformed into

$$Y^2 - (\alpha I - E)Y - Y\alpha + (\alpha^2 I - \alpha E - F) = 0.$$  

(2.1)

It was proved in [5] that when

$$\alpha \geq \max_{1 \leq i, j \leq n} (e_{ij} + \sqrt{e_{ii}^2 + 4f_{ii}})/2, \quad (2.2)$$

$\alpha^2 I - \alpha E - F$ is a nonnegative matrix. In addition, if $F$ is a nonsingular M-matrix, then

$$K = \begin{pmatrix} \alpha I & -I \\ -\alpha^2 I + \alpha E + F & \alpha I - E \end{pmatrix} \quad (2.3)$$

is a nonsingular M-matrix, and if $F$ is an irreducible singular M-matrix, then $K$ is an irreducible singular M-matrix.

By the above analysis and the theory of MARE, the following results were obtained in [5].

**Theorem 2.1** If $F$ is a nonsingular M-matrix, then (1.1) has exactly one M-matrix as its solution and the M-matrix is nonsingular. If $F$ is an irreducible singular M-matrix, then (1.1) has M-matrix solutions and all elements of each M-matrix solution are nonzero. In addition, let $u, v$ be positive vectors such that $Fv = 0$ and $u^TF = 0$, then

1) if $u^TEv \leq 0$, then (1.1) has exactly one M-matrix as its solution and the M-matrix is singular.

2) if $u^TEv > 0$, then (1.1) has exactly one nonsingular M-matrix as its solution but may also have singular M-matrices as its solutions.

For solving the QME (1.1), a fixed-point iteration method was proposed in [5] as follows. Fixed-point iteration method:

$$Y_{k+1} = (2\alpha I - E)^{-1}(Y_k^2 + \alpha^2 I - \alpha E - F), \quad Y_0 = 0.$$  

(2.4)

Convergence analysis showed that the sequence $\{Y_k\}$ in (2.4) is monotonically increasing and converges to the minimal nonnegative solution of the equation (2.1). In addition, the convergence rate of the fixed-point iteration method is linear for the noncritical case, and is sublinear for the critical case. However, experiments in [5] showed that the fixed-point iteration needs a large number of iterations to converge, though at each iteration it is very cheap.

In the following, we will propose a new iteration method for computing the minimal nonnegative solution of the equation (2.1).

First, write the equation (2.1) as

$$(2\alpha I - E - Y)Y = \alpha^2 I - \alpha E - F.$$  

Then we can get the following iterations

$$(2\alpha I - E - Y_k)Y_{k+1} = \alpha^2 I - \alpha E - F.$$  

Thus the new iteration method can be stated as follows.

**New iteration method:**

$$Y_{k+1} = (2\alpha I - E - Y_k)^{-1}(\alpha^2 I - \alpha E - F), \quad Y_0 = 0.$$  

(2.5)
Compared with the fixed-point iteration method (2.4), the new iteration method (2.5) is a little expensive, since at each iteration a matrix inverse is required to compute. However, the new iteration method may need less iterations than the fixed-point iteration method, which will be confirmed by the numerical experiments. So it is feasible.

3. Convergence Analysis

In the following, we give convergence analysis of the new iteration method (2.5).

The following lemma can be concluded from Theorem 2.1, Lemma 1.5 and Lemma 1.6.

**Lemma 3.1** Let the parameter \( \alpha \) satisfy (2.2). If \( F \) in (1.1) is a nonsingular M-matrix, then the equation (2.1) has a unique minimal nonnegative solution \( S_\alpha \), and \( \alpha I - S_\alpha \) are nonsingular M-matrices. If \( F \) in (1.1) is an irreducible singular M-matrix, then the equation (2.1) has a unique minimal nonnegative solution \( S_\alpha \), and \( \alpha I - S_\alpha \) are irreducible M-matrices. In addition, let \( u, v \) be positive vectors such that \( Fv = 0, u^TF = 0 \), then

1) If \( u^TFv = 0 \), then both \( \alpha I - S_\alpha \) are singular M-matrices;
2) If \( u^TFv < 0 \), then \( \alpha I - S_\alpha \) is nonsingular and \( \alpha I - E - S_\alpha \) is singular;
3) If \( u^TFv > 0 \), then \( \alpha I - S_\alpha \) is singular and \( \alpha I - E - S_\alpha \) is nonsingular.

**Proof** When \( F \) is a nonsingular M-matrix or an irreducible singular M-matrix, equation (2.1) has a unique minimal nonnegative solution \( S_\alpha \) by Lemma 1.5. When \( F \) is nonsingular, then both \( \alpha I - S_\alpha \) and \( \alpha I - E - S_\alpha \) are nonsingular M-matrices.

When \( F \) is irreducible singular, take \( \bar{u} = (u^T(\alpha I - E), u^T)^T \) and \( \bar{v} = (v^T, \alpha v^T)^T \), then we have \( \bar{u}^TK = 0, K\bar{v} = 0 \) and \( \mu = -u^TFv \). By Lemma 1.6, the conclusion follows.

**Theorem 3.1** Let \( F \) in (1.1) be a nonsingular M-matrix or an irreducible singular M-matrix and the parameter \( \alpha \) satisfy (2.2). Then the sequence \( \{Y_k\} \) generated by (2.5) is well defined, converges to \( S_\alpha \), and satisfy

\[
0 \leq Y_k \leq Y_{k+1}, \quad Y_{k+1} \leq S_\alpha, \tag{3.1}
\]

where \( S_\alpha \) is the unique minimal nonnegative solution of the equation (2.1).

**Proof** We first prove (3.1) by induction.

When \( k = 0 \), we have \( Y_1 = (2\alpha I - E)^{-1}(\alpha^2 I - \alpha E - F) \). Since \( \alpha \) satisfies (2.2), we have \( 2\alpha I > E \) and \( \alpha^2 I - \alpha E - F \geq 0 \). It is clear that \( 0 = Y_0 \leq Y_1 \). Since \( S_\alpha \) is the minimal nonnegative solution of (2.1), it holds that \( (2\alpha I - E - S_\alpha)S_\alpha = \alpha^2 I - \alpha E - F \). By Lemma 3.1, we know \( \alpha I - E - S_\alpha \) is an M-matrix, and hence, by Lemma 1.2, \( 2\alpha I - E - S_\alpha \) is a nonsingular M-matrix. Thus \( S_\alpha = (2\alpha I - E - S_\alpha)^{-1}(\alpha^2 I - \alpha E - F) \). It is evident \( (2\alpha I - E)^{-1} \leq (2\alpha I - E - S_\alpha)^{-1} \), thus \( Y_1 \leq S_\alpha \).

Suppose that the assertions (3.1) hold for \( k = l - 1 \), i.e. \( 0 \leq Y_{l-1} \leq Y_l \), \( Y_l \leq S_\alpha \).

Since \( Y_l \leq S_\alpha \), we know \( 2\alpha I - E - Y_l \geq 2\alpha I - E - S_\alpha \) is a nonsingular M-matrix by Lemma 1.3. Thus \( Y_{l+1} \) is well defined. From \( 0 \leq Y_{l-1} \leq Y_l \leq S_\alpha \) and Lemma 1.4, we know \( Y_{l+1} - Y_l \geq Y_{l-1} - Y_l \geq 0 \). We have

\[
Y_{l+1} - Y_l = (2\alpha I - E - Y_l)^{-1}(\alpha^2 I - \alpha E - F) - (2\alpha I - E - Y_{l-1})^{-1}(\alpha^2 I - \alpha E - F)
\]

\[
\geq 0,
\]

and

\[
Y_{l+1} - S_\alpha = (2\alpha I - E - Y_l)^{-1}(\alpha^2 I - \alpha E - F) - (2\alpha I - E - S_\alpha)^{-1}(\alpha^2 I - \alpha E - F)
\]
\[ \begin{align*}
&= [(2\alpha I - E - Y_1)^{-1} - (2\alpha I - E - S_\alpha)^{-1}] (\alpha^2 I - \alpha E - F) \\
&\leq 0.
\end{align*} \]

Hence the assertions (3.1) hold for \( k = l + 1 \). Thus we have proved by induction that the assertions (3.1) hold for all \( k \geq 0 \).

Since \( \{Y_k\} \) is nonnegative, monotonically increasing and bounded from above, there is a nonnegative matrix \( Y \) such that \( \lim_{k \to \infty} Y_k = Y \). From (3.1) we know \( Y \leq S_\alpha \). On the other hand, taking the limit in (2.5), we know \( Y \) is a solution of (2.1), thus \( S_\alpha \leq Y \). Hence \( S_\alpha = Y \).

**Theorem 3.2** Let \( F \) in (1.1) be a nonsingular M-matrix or an irreducible singular M-matrix and the parameter \( \alpha \) satisfy (2.2). Then the convergence rate of (2.5) is given by

\[ \limsup_{k \to \infty} \|S_\alpha - Y_k\|_2^{1/k} \leq \frac{\lambda}{\alpha + \delta}, \tag{3.2} \]

where \( \lambda = \rho(S_\alpha) \) and \( \delta \) is the minimum eigenvalue of the M-matrix \( \alpha I - E - S_\alpha \).

**Proof** We have

\[ S_\alpha - Y_k = (2\alpha I - E - S_\alpha)^{-1} (\alpha^2 I - \alpha E - F) - (2\alpha I - E - Y_{k-1})^{-1} (\alpha^2 I - \alpha E - F) \]

\[ = [(2\alpha I - E - S_\alpha)^{-1} - (2\alpha I - E - Y_{k-1})^{-1}] (\alpha^2 I - \alpha E - F) \]

\[ = (2\alpha I - E - S_\alpha)^{-1} (S_\alpha - Y_{k-1}) (2\alpha I - E - Y_{k-1})^{-1} (\alpha^2 I - \alpha E - F) \]

\[ \leq (2\alpha I - E - S_\alpha)^{-1} (S_\alpha - Y_{k-1}) (2\alpha I - E - S_\alpha)^{-1} (\alpha^2 I - \alpha E - F) \]

\[ = (2\alpha I - E - S_\alpha)^{-1} (S_\alpha - Y_{k-1}) S_\alpha \]

\[ \leq \cdots \]

\[ \leq (2\alpha I - E - S_\alpha)^{-k} (S_\alpha - Y_0) S_\alpha^k. \]

Hence

\[ \|S_\alpha - Y_k\|_2^{1/k} \leq \|(2\alpha I - E - S_\alpha)^{-k}\|_2^{1/k} \cdot \|(S_\alpha - Y_0)\|_2^{1/k} \cdot \|S_\alpha^k\|_2^{1/k}. \]

Taking limit on both side and noting that \( \rho(A) = \lim_{k \to \infty} \|A^k\|_2^{1/k} \), we have

\[ \limsup_{k \to \infty} \|S_\alpha - Y_k\|_2^{1/k} \leq \rho((2\alpha I - E - S_\alpha)^{-1}) \cdot \rho(S_\alpha). \]

Since \( \alpha I - E - S_\alpha \) is an M-matrix and \( S_\alpha \) is a nonnegative matrix, we can easily verify that

\[ \rho((2\alpha I - E - S_\alpha)^{-1}) = \frac{1}{\alpha + \delta}, \quad \rho(S_\alpha) = \lambda, \]

where \( \delta \) is the minimum nonnegative eigenvalue of \( \alpha I - E - S_\alpha \) and \( \lambda \) is the Perron eigenvalue of \( S_\alpha \). Thus the conclusion (3.2) holds.

**Corollary 3.1** If \( F \) in (1.1) is a nonsingular M-matrix, then the convergence rate of (2.5) is linear. If \( F \) is an irreducible singular M-matrix, then for the cases 2) and 3) in Lemma 3.1, the convergence rate of (2.5) is linear; for the case 1) the convergence rate of (2.5) is sublinear.

**Proof** When \( F \) is a nonsingular M-matrix, \( \alpha I - S_\alpha \) and \( \alpha I - E - S_\alpha \) are nonsingular M-matrices. Hence \( \alpha > \lambda \) and \( \delta > 0 \). Thus \( \frac{\alpha}{\alpha + \delta} < 1 \), the convergence rate of (2.5) is linear.

If \( F \) is an irreducible singular M-matrix, and the cases 2) and 3) happen, then one of \( \alpha I - S_\alpha \) and \( \alpha I - E - S_\alpha \) is a nonsingular M-matrix. We have \( \frac{\alpha}{\alpha + \delta} < 1 \), the convergence rate of (2.5) in this case is still linear.

If \( F \) is an irreducible singular M-matrix and the case 1) happens, then \( \alpha I - S_\alpha \) and \( \alpha I - E - S_\alpha \) are both singular M-matrices. Thus we have \( \alpha = \rho(S_\alpha) \) and \( \delta = 0 \). Hence \( \frac{\alpha}{\alpha + \delta} = 1 \), the convergence rate of (2.5) is sublinear in this case.
Corollary 3.2  The optimal parameter of the new iteration method (2.5) is given by
\[
\alpha_{\text{opt}} = \max_{1 \leq i \leq n} e_{ii} + \sqrt{e_{ii}^2 + 4f_{ii}}.
\]

Proof Let \( \alpha_1 > 0, \alpha_2 > 0 \) be two parameters that satisfy (2.2) and \( \alpha_1 > \alpha_2 \). Then the convergence factor of the new iteration method (2.5) are respectively
\[
r_1 = \frac{\lambda_1}{2\alpha_1 - \lambda_1}, \quad r_2 = \frac{\lambda_2}{2\alpha_2 - \lambda_2},
\]
where \( \lambda_1 = \rho(S_{\alpha_1}), \lambda_2 = \rho(S_{\alpha_2}) \) and \( S_{\alpha_1}, S_{\alpha_2} \) are the minimal nonnegative solutions of (2.1) respectively. Since \( \alpha_1 I - S_1 = \alpha_2 I - S_2 \), we have \( S_1 = (\alpha_1 - \alpha_2)I + S_2 \) and \( \lambda_1 = (\alpha_1 - \alpha_2) + \lambda_2 \).
Thus
\[
r_1 = \frac{\lambda_1}{2\alpha_1 - \lambda_1} = \frac{(\alpha_1 - \alpha_2) + \lambda_2}{2\alpha_1 - (\alpha_1 - \alpha_2) - \lambda_2} = \frac{(\alpha_1 - \alpha_2) + \lambda_2}{2\alpha_2 + (\alpha_1 - \alpha_2) - \lambda_2} > \frac{\lambda_2}{2\alpha_2 - \lambda_2} = r_2,
\]
and the conclusion follows.

4. Numerical Experiments

In this section, numerical experiments are given to show the feasibility and effectiveness of the new iteration method. We compare the new iteration method (NM) with the fixed-point iteration method (FP) in [5], and present numerical results of each experiment in terms of iteration numbers (IT), CPU time (CPU) in seconds and residue (Res), where the residue is defined to be
\[
\text{Res} := \|X^2 - EX - F\|_{\infty}/\|F\|_{\infty}.
\]
In our implementations, all iterations are run in MATLAB (R2012a) on a personal computer and are terminated when the current iterate satisfies \( \|X - \hat{X}\|_{\infty}/\|\hat{X}\|_{\infty} < 10^{-6} \).

Example 4.1[5] Consider the equation (1.1) with
\[
E = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}, \quad F = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.
\]
The numerical result is summarized in Tab. 4.1.

<table>
<thead>
<tr>
<th>Method</th>
<th>IT</th>
<th>CPU</th>
<th>Res</th>
</tr>
</thead>
<tbody>
<tr>
<td>FP</td>
<td>133</td>
<td>0.0046</td>
<td>2.4479e-06</td>
</tr>
<tr>
<td>NM</td>
<td>98</td>
<td>0.0021</td>
<td>2.3049e-06</td>
</tr>
</tbody>
</table>

Example 4.2[5] Consider the equation (1.1) with
\[
E = \begin{pmatrix} aI_{10} \\ bI_{10} \end{pmatrix}, \quad F = \begin{pmatrix} 1 & -1 \\ 1 & \ddots \\ \ddots & 1 \\ -1 & 1 \end{pmatrix} \in \mathbb{R}^{20 \times 20},
\]
where we take \( a = 2, b = 1 \). The numerical result is summarized in Tab. 4.2.

<table>
<thead>
<tr>
<th>Method</th>
<th>IT</th>
<th>CPU</th>
<th>Res</th>
</tr>
</thead>
<tbody>
<tr>
<td>FP</td>
<td>29</td>
<td>0.0028</td>
<td>8.1216e-07</td>
</tr>
<tr>
<td>NM</td>
<td>20</td>
<td>0.0019</td>
<td>5.4632e-07</td>
</tr>
</tbody>
</table>
Example 4.3 Consider the equation (1.1) with coefficient matrices defined as
\[ E = \text{diag}(1 : n); \quad F = \text{rand}(n, n); \quad F = \text{diag}(Fe) - F, \]
where \( e = (1, 1, \cdots, 1)^T \). For different sizes of \( n \), we list the numerical results in Tab. 4.3.

<table>
<thead>
<tr>
<th>( n )</th>
<th>Method</th>
<th>IT</th>
<th>CPU</th>
<th>Res</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>FP</td>
<td>127</td>
<td>0.0711</td>
<td>1.4847e-04</td>
</tr>
<tr>
<td></td>
<td>NM</td>
<td>82</td>
<td>0.0559</td>
<td>6.9822e-05</td>
</tr>
<tr>
<td>200</td>
<td>FP</td>
<td>168</td>
<td>0.3505</td>
<td>3.2440e-04</td>
</tr>
<tr>
<td></td>
<td>NM</td>
<td>100</td>
<td>0.2324</td>
<td>1.5371e-04</td>
</tr>
<tr>
<td>500</td>
<td>FP</td>
<td>238</td>
<td>6.2241</td>
<td>8.4503e-04</td>
</tr>
<tr>
<td></td>
<td>NM</td>
<td>133</td>
<td>2.2135</td>
<td>4.3075e-04</td>
</tr>
</tbody>
</table>

From Tabs. 4.1-4.3, we can conclude that the new iteration method needs less iteration number and CPU time than the fixed-point iteration method. Hence it is feasible and effective.

References:

与M-矩阵相关的一类二次矩阵方程的新迭代法

关晋瑞¹，宋儒瑛²，Zubair Ahmed²
(1. 太原师范学院数学系, 山西 晋中 030619; 2. 信德大学数学与计算机系, 信德 76080, 巴基斯坦)


关键词: 二次矩阵方程; M-矩阵; 代数Riccati方程; 迭代法