Journal of Computational Mathematics, Vol.25, No.5, 2007, 498–511.

MODIFIED BERNOULLI ITERATION METHODS FOR QUADRATIC MATRIX EQUATION *

Zhong-Zhi Bai and Yong-Hua Gao
(*State Key Laboratory of Scientific/Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100080, P.R. China*
Email: bzz@lsec.cc.ac.cn)

Abstract

We construct a modified Bernoulli iteration method for solving the quadratic matrix equation $AX^2 + BX + C = 0$, where A, B and C are square matrices. This method is motivated from the Gauss-Seidel iteration for solving linear systems and the Sherman-Morrison-Woodbury formula for updating matrices. Under suitable conditions, we prove the local linear convergence of the new method. An algorithm is presented to find the solution of the quadratic matrix equation and some numerical results are given to show the feasibility and the effectiveness of the algorithm. In addition, we also describe and analyze the block version of the modified Bernoulli iteration method.

Mathematics subject classification: 65F10, 65F15, 65N30.
Key words: Quadratic matrix equation, Quadratic eigenvalue problem, Solvent, Bernoulli's iteration, Newton's method, Local convergence.

1. Introduction

Quadratic matrix equations (**QME)** appear in many areas of scientific computing and engineering applications. Besides the famous *algebraic Riccati equation* (**ARE**) in control theory [5, 12, 18], the quadratic matrix equation

$$\mathcal{Q}(X) = AX^2 + BX + C = 0, \quad \text{with} \quad A,\ B,\ C \in \mathbb{R}^{n\times n}, \tag{1.1}$$

occurs in a variety of problems. For example, the *quadratic eigenvalue problem* (**QEP**) arising in the analysis of damped structural systems and vibration problems [6, 7, 15, 17, 26, 27], the *Quasi-Birth-Death* (**QBD**) problem used as stochastic models in telecommunication computer performance and inventory control [4, 19], and the noisy Wiener-Hopf problems coming from Markov chains [10, 16, 24, 25].

A lot of work has been done for computing the numerical solution of the QME (1.1). Davis [6] considered Newton's method, Higham and Kim [15] studied Newton's method with exact line search, and He, Beini and Rhee [13] discussed a cyclic reduction algorithm for the QBD problem; see also [3]. Besides, as special cases of the quadratic matrix equation (1.1), several linear matrix equations have been studied in [20, 21, 22, 8] where existence conditions and direct algorithms about the solutions were presented. Other different techniques for analyzing and computing the solutions of some general matrix equations can be found in [1, 11].

From the discussion in [14] we know that the *Bernoulli iteration* (**BI**) method is more efficient for over-damped quadratic eigenvalue problems and the QBD problem. This motivates

* Received December 15, 2005; final revised September 1, 2006; accepted October 1, 2006.

us to further study this iteration method for the QME (1.1). By making use of techniques of Gauss-Seidel relaxation sweep [9] for solving linear systems and the Sherman-Morrison-Woodbury formula [2, 9] for updating matrix, we establish a new iteration method and its block version for solving the QME (1.1). These methods are technical modifications of the Bernoulli iteration method in [14], and are of lower computing costs in actual applications. Convergence analyses show that these new methods have local linear convergence rates, and numerical implementations show that they are feasible and effective solvers for the QME (1.1).

The organization of the paper is as follows. After introducing some basic notations and concepts in Section 2, we establish the general *modified Bernoulli iteration* (**MBI**) method in Section 3. In Section 4, we derive a block version of the MBI method, which is called as *block modified Bernoulli iteration* (**BMBI**) method. In Section 5, we study the local convergence properties of both MBI and BMBI methods under suitable conditions. In Section 6, some numerical results are given to show the feasibility and effectiveness of our new methods. Finally, in Section 7, we use some remarks to end this paper.

2. Notations and Concepts

We introduce some notations and concepts, which are necessary for our subsequent statements.

A solution of the QME (1.1) is also called a solvent. The QEP corresponding to the QME (1.1) is defined as

$$\mathcal{Q}(\lambda)x = (\lambda^2 A + \lambda B + C)x = 0. \tag{2.1}$$

As is known, if A is nonsingular, then $\mathcal{Q}(\lambda)$ has exactly $2n$ eigenvalues λ_j $(j = 1, 2, \cdots, 2n)$, which can be ordered with respect to their absolute values in the following form:

$$|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_{2n}|. \tag{2.2}$$

We use $\lambda(A)$ to denote the eigenvalue set of the matrix A.

Definition 2.1. ([14]) *Let the eigenvalues λ_j $(j = 1, 2, \ldots, 2n)$ of the QEP* (2.1) *be ordered in the form of* (2.2). *Then a solvent S_1 of $Q(X)$ is called a dominant solvent if $\lambda(S_1) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ and $|\lambda_n| > |\lambda_{n+1}|$; and a solvent S_2 of $Q(X)$ is called a minimal solvent if $\lambda(S_2) = \{\lambda_{n+1}, \lambda_{n+2}, \ldots, \lambda_{2n}\}$ and $|\lambda_n| > |\lambda_{n+1}|$.*

One type of Bernoulli iteration methods for finding the dominant and the minimal solvents are as follows:

$$A + (B + CW_{k-1})W_k = 0, \qquad W_0 = 0,\ k = 1, 2, \ldots \tag{2.3}$$

and

$$(AX_{k-1} + B)X_k + C = 0, \qquad X_0 = 0,\ k = 1, 2, \ldots. \tag{2.4}$$

From [14, Theorem 10], we see that

$$\lim_{k\to\infty} W_k = S_1^{-1} \quad \text{and} \quad \lim_{k\to\infty} X_k = S_2$$

hold under the condition that the QME (1.1) has a dominant solvent S_1 and a minimal solvent S_2. Moreover, the asymptotic convergence rates of these two iterations are linear, with the convergence factor $\sigma = |\lambda_{n+1}|/|\lambda_n|$.

3. The MBI Method

Without loss of generality, throughout the paper we will only focus on the MBI method based on the iteration scheme (2.4), as discussions with respect to the iteration scheme (2.3) can proceed in an analogous fashion.

To this end, we rewrite (2.4) as:

$$(AX_k + B)X_{k+1} + C = 0, \qquad X_0 = 0,\ k = 0, 1, 2, \ldots. \tag{3.1}$$

Let the matrices X_k and C be decomposed according to their columns, i.e.,

$$X_k = (X_{k,1}, X_{k,2}, \ldots, X_{k,n}) \quad \text{and} \quad C = (C_1, C_2, \ldots, C_n), \tag{3.2}$$

where $X_{k,i}, C_i \in \mathbb{R}^n$, $i = 1, 2, \ldots, n$, are the columns of X_k and C, respectively. Let $X_k^{(i)}$ be the updated matrix of X_k, whose first i columns are replaced correspondingly by the already computed columns of X_{k+1}, i.e.,

$$X_k^{(i)} = (X_{k+1,1}, X_{k+1,2}, \ldots, X_{k+1,i}, X_{k,i+1}, \ldots, X_{k,n}).$$

If the columns of X_{k+1} are computed successively by solving (3.1), and each time the currently obtained information is used promptly like the Gauss-Seidel iteration employed to a linear system, we can establish the following *modified Bernoulli iteration* (**MBI**) method for solving the QME (1.1):

The MBI method.

1. Given a starting guess X_0.
2. For $k = 0, 1, 2, \ldots$ until $\{X_k\}$ is convergent.
   - Set $X_k^{(0)} := X_k$.
   - Successively solve

$$(AX_k^{(i)} + B)X_{k+1,i+1} + C_{i+1} = 0, \quad i = 0, 1, \ldots, n-1. \tag{3.3}$$

   - Set $X_{k+1} := (X_{k+1,1}, X_{k+1,2}, \ldots, X_{k+1,n})$.

In order to avoid to solve n different linear systems at each iteration step of the MBI method, we need to consider and utilize the special structures and the internal relationships among the matrices $AX_k^{(i)}, i = 0, 1, \ldots, n-1$, in the iteration scheme (3.3). As $X_k^{(i+1)}$ and $X_k^{(i)}$ only differ at one column, or in other words, $X_k^{(i+1)}$ is a rank-one correction of $X_k^{(i)}$, we can use the Sherman-Morrison formula for updating matrix to further reduce the computing cost of the MBI method.

Assume that $A \in \mathbb{R}^{n\times n}$ is a nonsingular matrix, and $u, v \in \mathbb{R}^n$ are two vectors. If $1 + v^T A^{-1} u \neq 0$, then the matrix $A + uv^T$ is nonsingular and its inverse is given by

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^T A^{-1}}{1 + v^T A^{-1} u}.$$

This is the well-known Sherman-Morrison formula, see [2, 9].

Because

$$\begin{aligned} AX_k^{(i)} + B &= AX_k^{(i-1)} + B + A(X_k^{(i)} - X_k^{(i-1)}) \\ &= AX_k^{(i-1)} + B + A(X_{k+1,i} - X_{k,i})e_i^T, \end{aligned}$$

where e_i denotes the i-th column of the n-by-n identity matrix, by using the above Sherman-Morrison formula, we can obtain

$$\begin{aligned} (AX_k^{(i)} + B)^{-1} \;=\; & (AX_k^{(i-1)} + B)^{-1} \\ & - \frac{(AX_k^{(i-1)} + B)^{-1}A(X_{k+1,i} - X_{k,i})e_i^T(AX_k^{(i-1)} + B)^{-1}}{1 + e_i^T(AX_k^{(i-1)} + B)^{-1}A(X_{k+1,i} - X_{k,i})}. \end{aligned}$$

Here, we have assumed that

$$1 + e_i^T(AX_k^{(i-1)} + B)^{-1}A(X_{k+1,i} - X_{k,i}) \neq 0$$

holds; see Theorem 5.2 for conditions that guarantee the validity of this assumption. Hence, from (3.3) we can explicitly express $X_{k+1,i+1}$ as

$$\begin{aligned} X_{k+1,i+1} &= -(AX_k^{(i)} + B)^{-1}C_{i+1} \\ &= -(AX_k^{(i-1)} + B)^{-1}C_{i+1} \\ &\quad + \frac{(AX_k^{(i-1)} + B)^{-1}A(X_{k+1,i} - X_{k,i})e_i^T(AX_k^{(i-1)} + B)^{-1}C_{i+1}}{1 + e_i^T(AX_k^{(i-1)} + B)^{-1}A(X_{k+1,i} - X_{k,i})}. \end{aligned} \tag{3.4}$$

Denote by

$$\begin{cases} u_{k,i} = A(X_{k+1,i} - X_{k,i}), \\ p_{k,i} = (AX_k^{(i-1)} + B)^{-1}u_{k,i}, \\ q_{k,i} = (AX_k^{(i-1)} + B)^{-1}C_{i+1}. \end{cases} \tag{3.5}$$

Then, it follows from (3.4) that

$$X_{k+1,i+1} = -(AX_k^{(i-1)} + B)^{-1}C_{i+1} + \frac{e_i^T q_{k,i}}{1 + e_i^T p_{k,i}} p_{k,i}.$$

By recursively operating this equality we have

$$\begin{aligned} X_{k+1,i+1} &= -(AX_k^{(i-1)} + B)^{-1}C_{i+1} + \frac{e_i^T q_{k,i}}{1 + e_i^T p_{k,i}} p_{k,i} \\ &= -(AX_k^{(i-2)} + B)^{-1}C_{i+1} + \frac{e_{i-1}^T q_{k,i-1}}{1 + e_{i-1}^T p_{k,i-1}} p_{k,i-1} + \frac{e_i^T q_{k,i}}{1 + e_i^T p_{k,i}} p_{k,i} \\ &= \cdots \\ &= -(AX_k + B)^{-1}C_{i+1} + \sum_{l=1}^{i} \frac{e_l^T q_{k,l}}{1 + e_l^T p_{k,l}} p_{k,l}. \end{aligned} \tag{3.6}$$

From (3.6), we see that the MBI method can be considered as a modification of the Bernoulli iteration method. Evidently, the extra cost introduced by such a modification is involved in the

computations of $u_{k,i}, p_{k,i}$ and $q_{k,i}$. Hence, to reduce the computing cost, in actual applications we can replace (3.5) by the following approximated formulas:

$$\begin{cases} p_{k,i} = (AX_k + B)^{-1} u_{k,i}, \\ q_{k,i} = (AX_k + B)^{-1} C_{i+1}. \end{cases} \tag{3.7}$$

Consequently, the resulting method for finding the minimal solvent of QME (1.1) can be described as follows.

Method 3.1. (The MBI method).

1. *Input matrices A, B, C and a tolerance ϵ. Set $X_0 := 0$ and $k := 0$.*
2. *For $i = 1 : n$, compute $X_{k+1,i}$ by solving*

$$(AX_k + B)X_{k+1,i} = -C_i.$$

3. *For $i = 1 : n - 1$,*

   3.1 compute $u_{k,i} = A(X_{k+1,i} - X_{k,i})$,

   3.2 compute $p_{k,i}$ by solving $(AX_k + B)p_{k,i} = u_{k,i}$,

   3.3 set $q_{k,i} := X_{k+1,i+1}$,

   3.4 for $j = i + 1 : n$, compute

$$X_{k+1,j} := X_{k+1,j} + \frac{e_i^T q_{k,i}}{1 + e_i^T p_{k,i}} p_{k,i}.$$

4. *Compute $\rho = \frac{\|X_{k+1} - X_k\|_1}{\|X_k\|_1}$.*
5. *If $\rho \leqslant \epsilon$, then stop and set $X := X_{k+1}$.*
6. *Otherwise, set $k := k + 1$ and return to Step 2.*

We observe that in the execution of Method 3.1, at each iteration step we need to solve $2n-1$ systems of linear equations with the same coefficient matrix $AX_k + B$ and compute $n-1$ matrix-vector products with the matrix A. This is the main cost in the implementation of the MBI method.

4. The BMBI Method

To derive a block version of the MBI method, we first review the Sherman-Morrison-Woodbury formula for updating a matrix corrected by a low-rank matrix; see [9, 2].

Assume that $A \in \mathbb{R}^{n\times n}$ is a nonsingular matrix, and $U, V \in \mathbb{R}^{n\times m}$ are two matrices. If $I + V^T A^{-1} U$ is nonsingular, then the matrix $A + UV^T$ is nonsingular and it holds that

$$(A + UV^T)^{-1} = A^{-1} - A^{-1}U(I + V^T A^{-1} U)^{-1} V^T A^{-1}.$$

This formula is known as the Sherman-Morrison-Woodbury formula, and is a block version of the above-described Sherman-Morrison formula.

Analogous to (3.2), we decompose X_k and C into column blocks as

$$X_k = (X_{k,1}, X_{k,2}, \ldots, X_{k,m}) \quad \text{and} \quad C = (C_1, C_2, \ldots, C_m), \tag{4.1}$$

where $X_{k,i}, C_i \in \mathbb{R}^{n\times n_i}$, $i = 1, 2, \ldots, m$, are the column blocks of X_k and C, respectively, with n_i $(i = 1, 2, \ldots, m)$ being m positive integers satisfying $\sum_{i=1}^{m} n_i = n$.

Similarly to the MBI method, we can describe the following *blockwise modified Bernoulli iteration* (**BMBI**) method for solving the QME (1.1):

The BMBI method.

1. Given a starting guess X_0.
2. For $k = 0, 1, 2, \ldots$ until $\{X_k\}$ is convergent,
   - Set $X_k^{(0)} := X_k$.
   - Successively solve
   $$(AX_k^{(i)} + B)X_{k+1,i+1} + C_{i+1} = 0, \quad i = 0, 1, \ldots, m-1. \tag{4.2}$$
   - Set $X_{k+1} := (X_{k+1,1}, X_{k+1,2}, \ldots, X_{k+1,m})$.

In order to avoid to solve m different linear systems at each iteration step of the BMBI method, we need to consider and utilize the special structures and the internal relationships among the matrices $AX_k^{(i)}, i = 0, 1, \ldots, m-1$, in the iteration scheme (4.2). As $X_k^{(i+1)}$ and $X_k^{(i)}$ only differ at one column block, or in other words, $X_k^{(i+1)}$ is a low-rank correction of $X_k^{(i)}$, we can use the Sherman-Morrison-Woodbury formula for updating the matrix to further reduce the computing cost of the BMBI method. We have

$$\begin{aligned} AX_k^{(i)} + B &= AX_k^{(i-1)} + B + A(X_k^{(i)} - X_k^{(i-1)}) \\ &= AX_k^{(i-1)} + B + A(X_{k+1,i} - X_{k,i})E_i, \end{aligned}$$

where

$$E_i = (0, \ldots, 0, I_{n_i}, 0, \ldots, 0) \in \mathbb{R}^{n_i\times n}$$

and $I_{n_i} \in \mathbb{R}^{n_i\times n_i}$ is the n_i-by-n_i identity matrix. By making use of the Sherman-Morrison-Woodbury formula, we obtain

$$\begin{aligned} (AX_k^{(i)} + B)^{-1} = (AX_k^{(i-1)} + B)^{-1} &- (AX_k^{(i-1)} + B)^{-1}A(X_{k+1,i} - X_{k,i}) \\ &\cdot[I + E_i(AX_k^{(i-1)} + B)^{-1}A(X_{k+1,i} - X_{k,i})]^{-1}E_i(AX_k^{(i-1)} + B)^{-1}. \end{aligned}$$

Here, we have assumed that the matrix

$$I + E_i(AX_k^{(i-1)} + B)^{-1}A(X_{k+1,i} - X_{k,i})$$

is nonsingular; see Theorem 5.2 for conditions that guarantee the validity of this assumption. Hence, from (4.2) we can explicitly express $X_{k+1,i+1}$ as

$$\begin{aligned} & X_{k+1,i+1} = -(AX_k^{(i)} + B)^{-1}C_{i+1} \\ = \; & -(AX_k^{(i-1)} + B)^{-1}C_{i+1} + (AX_k^{(i-1)} + B)^{-1}A(X_{k+1,i} - X_{k,i}) \\ & \cdot[I + E_i(AX_k^{(i-1)} + B)^{-1}A(X_{k+1,i} - X_{k,i})]^{-1}E_i(AX_k^{(i-1)} + B)^{-1}C_{i+1}. \end{aligned} \tag{4.3}$$

Denote by

$$\begin{cases} U_{k,i} = A(X_{k+1,i} - X_{k,i}), \\ P_{k,i} = (AX_k^{(i-1)} + B)^{-1}U_{k,i}, \\ Q_{k,i} = (AX_k^{(i-1)} + B)^{-1}C_{i+1}. \end{cases} \tag{4.4}$$

Then (4.3) can be rewritten as

$$X_{k+1,i+1} = -(AX_k^{(i-1)} + B)^{-1}C_{i+1} + P_{k,i}(I + E_iP_{k,i})^{-1}E_iQ_{k,i}.$$

By recursively employing this equality, we have

$$\begin{aligned} X_{k+1,i+1} &= -(AX_k^{(i-1)} + B)^{-1}C_{i+1} + P_{k,i}(I + E_iP_{k,i})^{-1}E_iQ_{k,i} \\ &= -(AX_k^{(i-1)} + B)^{-1}C_{i+1} + P_{k,i-1}(I + E_{i-1}P_{k,i-1})^{-1}E_{i-1}Q_{k,i-1} \\ &\quad + P_{k,i}(I + E_iP_{k,i})^{-1}E_iQ_{k,i} \\ &= \cdots \\ &= -(AX_k + B)^{-1}C_{i+1} + \sum_{l=1}^{i} P_{k,l}(I + E_lP_{k,l})^{-1}E_lQ_{k,l}. \end{aligned}$$

Similar to (3.7), in actual applications we adopt the following approximations to replace the items in (4.4):

$$\begin{cases} P_{k,i} = (AX_k + B)^{-1}U_{k,i}, \\ Q_{k,i} = (AX_k + B)^{-1}C_{i+1}. \end{cases}$$

Then, we can construct the BMBI method for finding the minimal solvent of the quadratic matrix equation (1.1).

Method 4.1. (The BMBI method).

1. *Input matrices A, B, C and a tolerance ϵ. Set $X_0 := 0$ and k:=0.*
2. *For $i = 1 : m$, compute $X_{k+1,i}$ by solving*

   $$(AX_k + B)X_{k+1,i} = -C_i.$$

3. *For $i = 1 : m - 1$,*
   - *3.1 compute $U_{k,i} = A(X_{k+1,i} - X_{k,i})$,*
   - *3.2 compute $P_{k,i}$ by solving $(AX_k + B)P_{k,i} = U_{k,i}$,*
   - *3.3 set $Q_{k,i} := X_{k+1,i+1}$,*
   - *3.4 compute $V_{k,i}$ by solving $(I + E_iP_{k,i})V_{k,i} = E_i$,*
   - *3.5 for $j = i + 1 : m$, compute*

     $$X_{k+1,j} := X_{k+1,j} + P_{k,i}V_{k,i}Q_{k,i}.$$

4. *Compute $\rho = \frac{\|X_{k+1} - X_k\|_1}{\|X_k\|_1}$.*
5. *If $\rho \leqslant \epsilon$, then stop and set $X := X_{k+1}$.*
6. *Otherwise, set $k =: k + 1$ and return to Step 2.*

We observe that in the execution of Method 4.1, at each iteration step we need to solve $2m-1$ systems of linear equations with the same coefficient matrix AX_k+B and compute $m-1$ matrix-matrix products with the matrix A. This is the main cost in the implementation of the BMBI method.

As a special case, we consider the BMBI method when $m=2$. Now,

$$X_k=(X_{k,1},X_{k,2}) \quad \text{and} \quad C=(C_1,C_2),$$

where $X_{k,i},C_i\in\mathbb{R}^{n\times n_i}$, $i=1,2$, are the two column blocks of X_k and C, respectively, with $n_i(i=1,2)$ being two positive integers satisfying $n_1+n_2=n$. The BMBI method then becomes to the following:

The BMBI method.

1. Given a starting guess X_0.
2. For $k=0,1,2,\ldots$ until $\{X_k\}$ is convergent.
   - Set $X_k^{(0)}:=X_k$.
   - Successively solve
   $$(AX_k^{(i)}+B)X_{k+1,i+1}+C_{i+1}=0, \quad i=0,1.$$
   - Set $X_{k+1}:=(X_{k+1,1},X_{k+1,2})$.

By applying the Sherman-Morrison-Woodbury formula and using the same argument as for the BMBI method, we obtain

$$X_{k+1,2}=-Q+P(I+EP)^{-1}EQ,$$

where

$$\begin{cases} U=A(X_{k+1,1}-X_{k,1}), & P=(AX_k+B)^{-1}U,\\ Q=(AX_k+B)^{-1}C_2, & E=(I_{n_1},0). \end{cases}$$

Hence, the BMBI method reduces to the following two-block case:

Method 4.2. (The BMBI method (Two-Block Case)).

1. *Input matrices A,B,C and a tolerance ϵ. Set $X_0:=0$ and k:=0.*
2. *For $i=1:2$, compute $X_{k+1,i}$ by solving*
   $$(AX_k+B)X_{k+1,i}=-C_i.$$
3. *Update $X_{k+1,2}$ according to the following:*
   - *3.1 compute $U=A(X_{k+1,1}-X_{k,1})$,*
   - *3.2 compute P by solving $(AX_k+B)P=U$,*
   - *3.3 set $Q:=X_{k+1,2}$,*
   - *3.4 compute V by solving $(I+EP)V=E$,*
   - *3.5 compute $X_{k+1,2}:=X_{k+1,2}+PVQ$.*
4. *Compute $\rho=\frac{\|X_{k+1}-X_k\|_1}{\|X_k\|_1}$.*
5. *If $\rho\leqslant\epsilon$, then stop and set $X:=X_{k+1}$.*
6. *Otherwise, set $k=:k+1$ and return to Step 2.*

5. Convergence Theorem

In this section, under suitable conditions we will establish the local convergence theorem for the BMBI method (4.2). The local convergence theorems for the MBI method can be analogously demonstrated.

Theorem 5.1. *Assume that the QME* (1.1) *has a minimal and a dominant solvent. Let* X^* *be the minimal solvent such that*

$$\theta := 2\|(AX^* + B)^{-1}\|_F^2\|A\|_F\|C\|_F < 2^{\frac{1}{m}} - 1.$$

Then, there exists an open ball $\mathcal{S} = \mathcal{S}(X^*, r)$ *such that, for* $\forall X_0 \in \mathcal{S}$*, the iteration sequence* $\{X_k\}$ *generated by the BMBI method satisfies*

$$\|X_{k+1} - X^*\|_F \leqslant \gamma\|X_k - X^*\|_F, \quad k = 0, 1, 2, \ldots,$$

where $\gamma = (1+\theta)^m - 1 \in (0,1)$ *and* $\|\cdot\|_F$ *is the Frobenius matrix norm.*

Proof. From the definition of the iteration sequence $\{X_k\}$ we obtain

$$\begin{aligned} X_{k+1,l} - X_l^* &= [-(AX_k^{(l-1)} + B)^{-1}C_l] - [-(AX^* + B)^{-1}C_l] \\ &= [(AX^* + B)^{-1} - (AX_k^{(l-1)} + B)^{-1}]C_l \\ &= (AX_k^{(l-1)} + B)^{-1}A(X_k^{(l-1)} - X^*)(AX^* + B)^{-1}C_l. \end{aligned}$$

Here, we have used the equality

$$(AX_k^{(l-1)} + B) - (AX^* + B) = A(X_k^{(l-1)} - X^*). \tag{5.1}$$

Denote by

$$\alpha = \|(AX^* + B)^{-1}\|_F.$$

When

$$\|A(X_k - X^*)\|_F \leqslant \|A\|_F\|X_k - X^*\|_F < \frac{1}{2\alpha(1+\theta)^m} \tag{5.2}$$

or

$$\|X_k - X^*\|_F < r := \frac{1}{2\alpha(1+\theta)^m\|A\|_F},$$

we can proceed the proof by considering each block of the matrix X_{k+1}.

For $l = 1$, based on (5.1) and (5.2), by making use of the perturbation lemma [23] we have

$$\|(AX_k^{(0)} + B)^{-1}\|_F < \frac{\alpha}{1 - 1/(2(1+\theta)^m)} < 2\alpha.$$

It then follows that

$$\begin{aligned} \|X_{k+1,1} - X_1^*\|_F &\leqslant 2\alpha\|A\|_F\|X_k - X^*\|_F\alpha\|C_1\|_F \\ &\leqslant 2\alpha^2\|A\|_F\|C\|_F\|X_k - X^*\|_F. \end{aligned}$$

Hence,

$$\begin{aligned} \|X_k^{(1)} - X^*\|_F &\leqslant \|X_k^{(0)} - X^*\|_F + \|X_{k+1,1} - X_1^*\|_F \\ &\leqslant (1+\theta)\|X_k - X^*\|_F \end{aligned}$$

and

$$\|A(X_k^{(1)} - X^*)\|_F \leqslant \|A\|_F \|X_k^{(1)} - X^*\|_F < \frac{1}{2\alpha(1+\theta)^{m-1}}.$$

Again, by making use of the perturbation lemma, we have

$$\|(AX_k^{(1)} + B)^{-1}\|_F \leqslant \frac{\alpha}{1 - 1/(2(1+\theta)^{m-1})} < 2\alpha.$$

Hence, for $l = 2$, it holds that

$$\begin{aligned}\|X_{k+1,2} - X_2^*\|_F &\leqslant 2\alpha\|A\|_F(1+\theta)\|X_k - X^*\|_F \alpha\|C_2\|_F \\ &\leqslant \theta(1+\theta)\|X_k - X^*\|_F.\end{aligned}$$

By recursively continuing this process, we can obtain

$$\|X_k^{(l)} - X^*\|_F \leqslant \|X_k - X^*\|_F + \sum_{i=1}^{l} \|X_{k+1,i} - X_i^*\|_F$$

and

$$\|X_{k+1,l} - X_l^*\|_F \leqslant \theta(1+\theta)^{l-1}\|X_k - X^*\|_F, \quad l = 3, \ldots, m.$$

Consequently, we get

$$\begin{aligned}\|X_{k+1} - X^*\|_F &\leqslant \sum_{l=1}^{m} \|X_{k+1,l} - X_l^*\|_F \\ &\leqslant \sum_{l=1}^{m} \theta(1+\theta)^{l-1}\|X_k - X^*\|_F \\ &= [(1+\theta)^m - 1]\|X_k - X^*\|_F \\ &= \gamma\|X_k - X^*\|_F.\end{aligned}$$

This completes the proof of the theorem. ∎

The convergence conditions of Theorem 5.1 readily guarantee the well-definiteness of the MBI and the BMBI iteration sequences. This fact is precisely described in the following theorem.

Theorem 5.2. *Assume that the QME* (1.1) *has a minimal and a dominant solvent. Let* X^* *be the minimal solvent such that*

$$\theta := 2\|(AX^* + B)^{-1}\|_F^2 \|A\|_F \|C\|_F < 2^{\frac{1}{m}} - 1.$$

Then, there exists an open ball $\mathcal{S} = \mathcal{S}(X^*, r)$ *such that, for* $\forall X_0 \in \mathcal{S}$*, it holds that*

(i) $d_{k,i} := 1 + e_i^T(AX_k^{(i-1)} + B)^{-1}A(X_{k+1,i} - X_{k,i}) \neq 0$, $i = 1, \ldots, n$,

(ii) $D_{k,i} := I + E_i(AX_k^{(i-1)} + B)^{-1}A(X_{k+1,i} - X_{k,i})$ *is nonsingular,* $i = 1, \ldots, m$.

That is to say, the iteration sequences $\{X_k\}$ *generated by both MBI and BMBI methods are well defined.*

Proof. Theorem 5.1 and its proof show that the following estimates hold true:

(F_1) $\|X_{k+1} - X^*\|_F \leqslant \gamma\|X_k - X^*\|_F$, $k = 0, 1, \ldots$, and

(F_2) $\|(AX_k^{(i-1)} + B)^{-1}\|_F \leqslant 2\alpha$, $i = 1, \ldots, n$, $k = 0, 1, \ldots$,

where $\gamma \in (0, 1)$ is a constant and $\alpha = \|(AX^* + B)^{-1}\|_F$. Because

$$\begin{aligned}
d_{k,i} &\geqslant 1 - |e_i^T (AX_k^{(i-1)} + B)^{-1} A(X_{k+1,i} - X_{k,i})| \\
&\geqslant 1 - \|e_i^T (AX_k^{(i-1)} + B)^{-1}\|_F \cdot \|A\|_F \cdot \|X_{k+1,i} - X_{k,i}\|_F \\
&\geqslant 1 - \|(AX_k^{(i-1)} + B)^{-1}\|_F \cdot \|A\|_F \cdot \|X_{k+1,i} - X_{k,i}\|_F \\
&\geqslant 1 - \|(AX_k^{(i-1)} + B)^{-1}\|_F \cdot \|A\|_F \cdot (\|X_{k+1,i} - X^*\|_F + \|X_{k,i} - X^*\|_F) \\
&> 1 - 2\alpha \cdot \|A\|_F \cdot \gamma^k (1 + \gamma) \cdot \|X_0 - X^*\|_F \\
&= 1 - 2\alpha(1 + \gamma)\|A\|_F \cdot \gamma^k \|X_0 - X^*\|_F,
\end{aligned}$$

we see that $d_{k,i} > 0$ holds true for all k when X_0 is sufficiently close to X^*. This shows the validity of (i). Analogously, as

$$\begin{aligned}
f_{k,i} &= \|E_i (AX_k^{(i-1)} + B)^{-1} A(X_{k+1,i} - X_{k,i})\|_F \\
&\leqslant \|E_i (AX_k^{(i-1)} + B)^{-1}\|_F \|A\|_F \|X_{k+1,i} - X_{k,i}\|_F \\
&\leqslant \|(AX_k^{(i-1)} + B)^{-1}\|_F \|A\|_F \|X_{k+1,i} - X_{k,i}\|_F \\
&\leqslant \|(AX_k^{(i-1)} + B)^{-1}\|_F \|A\|_F (\|X_{k+1,i} - X^*\|_F + \|X_{k,i} - X^*\|_F) \\
&\leqslant 2\alpha \cdot \|A\|_F \cdot \gamma^k (1 + \gamma) \cdot \|X_0 - X^*\|_F \\
&= 2\alpha(1 + \gamma)\|A\| \cdot \gamma^k \|X_0 - X^*\|_F,
\end{aligned}$$

we see that $f_{k,i} < 1$ holds true for all k when X_0 is sufficiently close to X^*. By making use of the perturbation lemma, we immediately know that (ii) is valid. ∎

6. Numerical Results

We use two examples to illustrate the numerical feasibility and effectiveness of the two-block version of the BMBI method.

All implementations are run in MATLAB (Version 6.5) with machine precision 10^{-16} on a Pentium IV personal computer, started from $X_0 = 0$, and terminated when the current iteration satisfies

$$\frac{\|X_i - X_{i-1}\|_1}{\|X_i\|_1} \leqslant n \cdot eps,$$

where eps $= 2^{-52}$ is the floating-point relative accuracy, and is roughly equal to 2.22×10^{-16}.

Example 6.1. *Consider a QME* (1.1) *for which*

$$A = \begin{pmatrix} 15 & -5 & & & \\ -5 & 15 & -5 & & \\ & \ddots & \ddots & \ddots & \\ & & -5 & 15 & -5 \\ & & & -5 & 15 \end{pmatrix} + \alpha e e^T, \quad B = \begin{pmatrix} 20 & -10 & & & \\ -10 & 30 & -10 & & \\ & \ddots & \ddots & \ddots & \\ & & -10 & 30 & -10 \\ & & & -10 & 20 \end{pmatrix},$$

and $C = I_n$, where $e \in \mathbb{R}^n$ is the vector with all elements being equal to one and $I_n \in \mathbb{R}^{n \times n}$ is the n-by-n identity matrix.

Table 6.1: Numerical Results for Example 6.1

Method	α	0.10	0.15	0.19	0.195	0.198
BI	IT	22	34	78	115	231
	ERR	2.20e-14	1.04e-14	1.93e-14	2.12e-14	2.06e-14
BMBI	IT	22	32	72	105	207
	ERR	1.45e-14	2.15e-14	1.81e-14	1.89e-14	2.13e-14

Table 6.2: Numerical Results for Example 6.2

Method	n	20	40	60	80	100
BI	IT	104	189	269	346	420
	ERR	3.41e-15	7.89e-15	1.24e-14	1.70e-14	2.21e-14
BMBI	IT	98	182	261	338	412
	ERR	4.44e-15	8.62e-15	1.33e-14	1.70e-14	2.19e-14

This example is essentially similar to the one used in [14]. We take $n = 100$. For different α, we list the iteration step (**IT**) and the relative error (**ERR**) for the BI and the BMBI methods in Table 6.1. We remark that both methods converge to the same solution of the QME (1.1), and the α selected here satisfy the over-damped condition (see [14]). From Table 6.1 we see that the BMBI method outperforms the BI method in both iteration steps and approximation accuracy.

Example 6.2. *Consider a QME* (1.1) *for which* $A = C = I_n$ *and*

$$B = \begin{pmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 4 & -1 \\ & & & -1 & 4 \end{pmatrix}.$$

This is an over-damped problem, too, due to $\lambda_{min}(B)^2 - 4\|A\|_2\|C\|_2 > 0$, where $\lambda_{min}(B)$ denotes the smallest eigenvalue of the matrix B. For different n, we list the iteration steps and the relative errors for the BI and the BMBI methods in Table 6.2. Again, we remark that both methods can converge to the same solution of the QME (1.1). The numerical results in Table 6.2 further confirm that the BMBI method outperforms the BI method in iteration steps.

Consequently, the BMBI method may be effective and practical than the BI method for solving the quadratic matrix equation (1.1).

7. Concluding Remarks

We have constructed a modified Bernoulli iteration method and its block variant for solving the quadratic matrix equation (1.1), and proved the local convergence of these new methods under suitable conditions. If we neglect the sparsity of the involved matrices in the QME (1.1), then the computational complexity of these new methods is $\mathcal{O}(n^3)$, i.e., at each iteration step their costs are $\mathcal{O}(n^3)$ operations. However, the computational complexity of the Newton iteration method is $\mathcal{O}(n^4)$ and the Bernoulli iteration method is $\mathcal{O}(n^3)$. Therefore, at each iteration

step the costs of our new methods are less than the Newton but are comparable to the Bernoulli iteration methods. Because our new methods use the currently available information promptly like the Gauss-Seidel iteration applied to the system of linear equations, they may show faster convergence speeds than the Bernoulli and the Newton iteration methods in actual applications.

Acknowledgments. Supported by The Special Funds For Major State Basic Research Projects (No. G1999032803), The China NNSF Outstanding Young Scientist Foundation (No. 10525102), The National Natural Science Foundation (No. 10471146), and The National Basic Research Program (No. 2005CB321702), P.R. China.

References

[1] Z.-Z. Bai, A class of iteration methods based on the Moser formula for nonlinear equations in Markov chains, *Linear Algebra Appl.*, **266** (1997), 219-241.

[2] Z.-Z. Bai, G.-Q. Li and L.-Z. Lu, Combinative preconditioners of modified incomplete Cholesky factorization and Sherman-Morrison-Woodbury update for self-adjoint elliptic Dirichlet-periodic boundary value problems, *J. Comput. Math.*, **22**:6 (2004), 833-856.

[3] Z.-Z. Bai, X.-X. Guo and J.-F. Yin, On two iteration methods for the quadratic matrix equations, *Int. J. Numer. Anal. Model.*, **2**:Supp (2005), 114-122.

[4] 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-and-death processes, *Ann. Appl. Probab.*, **7** (1997), 134-155.

[5] S. Bittanti, A.J. Laub and J.C. Willems, The Riccati Equation, Springer-Verlag, Berlin, 1991.

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

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

[8] Y.-B. Deng and X.-Y. Hu, On solutions of matrix equation $AXA^T+BYB^T=C$, *J. Comput. Math.*, **23**:1 (2005), 17-26.

[9] G.H. Golub and C.F. Van Loan, Matrix Computations, 3rd Edition, The Johns Hopkins University Press, Baltimore, 1996.

[10] C.-H. Guo, On a quadratic matrix equation associated with an M-matrix, *IMA J. Numer. Anal.*, **23** (2003), 11-27.

[11] X.-X. Guo, On Hermitian positive definite solution of nonlinear matrix equation, *J. Comput. Math.*, **23**:5 (2005), 513-526.

[12] X.-X. Guo and Z.-Z. Bai, On the minimal nonnegative solution of nonsymmetric algebraic Riccati equation, *J. Comput. Math.*, **23**:3 (2005), 305-320.

[13] C. He, B. Meini and N.H. Rhee, A shifted cyclic reduction algorithm for quasi-birth-death problems, *SIAM J. Matrix Anal. Appl.*, **23** (2001/2002), 673-691.

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

[15] N.J. Higham and H.-M. Kim, Solving a quadratic matrix equation by Newton's method with exact line searches, *SIAM J. Matrix Anal. Appl.*, **23** (2001), 303-316.

[16] J. Kennedy and D. Williams, Probabilistic factorization of a quadratic matrix polynomial, *Math. Proc. Cambridge Philos. Soc.*, **107** (1990), 591-600.

[17] P. Lancaster, Lambda-Matrices and Vibrating Systems, Pergamon Press, Oxford, 1966.

[18] P. Lancaster and L. Rodman, Algebraic Riccati Equations, The Clarendon Press, Oxford, 1995.

[19] G. Latouche and V. Ramaswami, Introduction to Matrix Analytic Methods in Stochastic Modeling, SIAM, Philadelphia, 1999.

[20] A.-P. Liao and Z.-Z. Bai, Least-squares solutions of the matrix equation $A^*XA=D$ in bisymmetric matrix set, *Math. Numer. Sinica*, **24**:1 (2002), 9-20.

[21] A.-P. Liao and Z.-Z. Bai, Least-squares solution of $AXB = D$ over symmetric positive semidefinite matrices X, *J. Comput. Math.*, **21**:2 (2003), 175-182.

[22] A.-P. Liao and Z.-Z. Bai, Least squares symmetric and skew-symmetric solutions of the matrix equation $AXA^T+BYB^T=C$ with the least norm, *Math. Numer. Sinica*, **27**:1 (2005), 81-95.

[23] J.M. Ortega and W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, SIAM, Philadelphia, 2000.

[24] L.C.G. Rogers, Fluid models in queueing theory and Wiener-Hopf factorization of Markov chains, *Ann. Appl. Probab.*, **4** (1994), 390-413.

[25] L.C.G. Rogers and Z. Shi, Computing the invariant law of a fluid model, *J. Appl. Probab.*, **31** (1994), 885-896.

[26] H.A. Smith, R.K. Singh and D.C. Sorensen, Formulation and solution of the non-linear, damped eigenvalue problem for skeletal systems, *Int. J. Numer. Meth. Engrg.*, **38** (1995), 3071-3085.

[27] Z.C. Zheng, G.X. Ren and W.J. Wang, A reduction method for large scale unsymmetric eigenvalue problems in structural dynamics, *J. Sound Vib.*, **199** (1997), 253-268.