Journal of Nonlinear Sciences and Applications

An efficient iterative algorithm for finding a nontrivial symmetric solution of the Yang-Baxter-like matrix equation

Check for updates

Chacha Stephen Chacha[a,b,*], Hyun-Min Kim[b,*]

[a]*Mathematics department, Mkwawa University College of Education (A constituent College of the University of Dar es salaam), P. O. Box 2513, Iringa, Tanzania.*

[b]*Mathematics department, Pusan National University, Busan, 46241, Republic of Korea.*

Abstract

This paper presents an efficient iterative method to obtain a nontrivial symmetric solution of the Yang-Baxter-like matrix equation $AXA = XAX$. Necessary conditions for the convergence of the propounded iterative method are derived. Finally, three numerical examples to illustrate the efficiency of the proposed method and the preciseness of our theoretical results are provided.

Keywords: Yang-Baxter matrix equation, iterative method, nontrivial solution, Newton's method.

2010 MSC: 15A24, 65F10, 65H10.

1. Introduction

We consider the nonlinear Yang-Baxter matrix equation (YBME)

$$AXA = XAX, \tag{1.1}$$

where A is a given $n \times n$ real matrix. The term Yang–Baxter equation was coined a result of independent work by Yang and Baxter in the late 1970s. YBME arises in real life applications in many fields such braid group, quantum field theory, statistical mechanics, quantum groups, differential equations, knot theory, and other disciplines (see [13, 20, 21]).

In some current works, YBME has been termed as the Yang-Baxter-like matrix equation ([7, 9, 10, 18]). The study of YBME has continued to be a hot topic of research for the past few decades as well as in recent years (see [1, 3, 4, 14, 16]). With some restrictions on the matrix A, many authors have been striving to exploit all nontrivial solutions (1.1). For instance, Ding and Tian [7] investigated all solutions of YBME by restricting the matrix $A = I - v^{\mathsf{T}} u$, where u and v are two n-dimensional vectors in the complex field such that $v^{\mathsf{T}} u \neq 0$. Mansour et al. [18] derived an explicit expression for obtaining all solutions of (1.1)

---

*Corresponding authors

Email addresses: `chchstephen@yahoo.com` (Chacha Stephen Chacha), `hyunmin@pusan.ac.kr` (Hyun-Min Kim)

for a given idempotent matrix. Tian [19] constructed all solutions of (1.1) for rank-one matrices. In [12] all commuting solutions of (1.1) were investigated by restricting the matrix A to be diagonalizable.

In [6, 8] the spectral solutions of (1.1) were exploited. For more details on YBME, the reader is referred to the works by Ding et al. [5, 11] (see also [15]). So far, apart from the trivial solutions of (1.1) $X = 0$ and $X = A$, most researchers have been toiling to find all the nontrivial solutions of (1.1) by imposing some restrictions on the matrix A. On the other hand, there had been very few studies on the iterative methods for obtaining a nontrivial solution of (1.1).

Therefore, this work aims at proposing a conjugate gradient based on iterative method that can be incorporated into Newton's method to find the symmetric solution of equation (1.1), and derive the necessary conditions for the existence of symmetric solution of (1.1) based on the proposed algorithm. By applying this iterative method, Newton's step can be computed even if the Fréchet derivative is singular and the existence of symmetric solution is guaranteed.

This work is outlined as follows. Some useful notations, definitions, and lemmas that will be applied in our proofs are introduced in Section 2. In Section 3, we introduce Newton's method and propose an iterative method to solve the Newton's step. Moreover, the necessary conditions for the existence of symmetric solution are derived. In Section 4, we examine our presented method experimentally to illustrate the accurateness of the established theoretical results. Finally, we present a brief conclusion in Section 5.

2. Preliminaries

We provide some important notations, definitions, and lemmas that will be exploited in our proofs.

A^T denotes the transpose of matrix A; $\|A\| = \sqrt{\text{trace}(A^\mathsf{T}A)}$ denotes the Frobenius norm of matrix A induced by the inner product; for matrices $A = [a_{ij}] \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$, then

$$A \otimes B = [a_{ij}B] = \begin{bmatrix} a_{11}B & a_{12}B & \ldots & a_{1n}B \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \in \mathbb{R}^{mp \times nq}$$

denotes the Kronecker product of matrices A and B; $\text{vec}(T) = [t_1^\mathsf{T} t_2^\mathsf{T} \cdots t_k^\mathsf{T}]$ stands for the vec operator on matrix T, where t_k is the kth column of the matrix T. For matrices A and B we have a well-known property

$$\text{vec}(AXB) = (B^\mathsf{T} \otimes A)\text{vec}(X).$$

Definition 2.1. Suppose that $f : \mathbb{R}^{m \times m} \mapsto \mathbb{R}^{m \times m}$ is a matrix function. The Fréchet derivative of a matrix function f at X in the direction H is the unique linear operator L_f that maps H to $L_f(X, H)$ such that

$$f(X + H) - f(X) - L_f(X, H) = O(\|H\|^2), \quad \forall \ X, H \in \mathbb{R}^{m \times m}.$$

For a scalar γ and real square matrices U, V, W, and Z, then, the following properties holds for the trace operator

(a) $\text{trace}(U + V) = \text{trace}(U) + \text{trace}(V)$;
(b) $\text{trace}(\gamma U) = \gamma \text{trace}(U)$;
(c) $\text{trace}(U^\mathsf{T}) = \text{trace}(U)$;
(d) $\text{trace}(UV) = \text{trace}(VU)$;
(e) $\text{trace}(UVWZ) = \text{trace}(VWZU) = \text{trace}(WZUV) = \text{trace}(ZUVW)$;
(f) $\text{trace}(UVWZ) = \text{trace}(Z^\mathsf{T}W^\mathsf{T}V^\mathsf{T}U^\mathsf{T})$.

Lemma 2.2. *For any symmetric matrix* X *it holds that* $\text{trace}\left[\frac{1}{2}\left(Y + Y^\mathsf{T}\right)^\mathsf{T} X\right] = \text{trace}(Y^\mathsf{T}X)$, *where* Y *is any arbitrary* $n \times n$ *real matrix.*

Proof. The proof is straightforward from the properties of trace operator given above. Therefore, it is omitted here. □

3. Newton's method for (1.1) and an iterative method for solving a Newton's step

In this section, we derive Newton's method for (1.1) and provide an iterative algorithm to solve the Newton's step. Let define a map

$$F(X) = AXA - XAX = 0.$$

We see that the Fréchet derivative is a linear operator, $F'_X(E) : \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$, defined by

$$F'_X(E) = AEA - XAE - EAX. \tag{3.1}$$

Applying the vec operator in (3.1) we have

$$\text{vec}(F'_X(E)) = \mathcal{D}_X \text{vec}(E),$$

where $\mathcal{D}_X = A^T \otimes A - I \otimes (XA) - (XA)^T \otimes I$ is the Kronecker Fréchet derivative of $F(X)$. If the $A^T \otimes A - I \otimes (XA) - (XA)^T \otimes I$ is invertible, then, a Newton's step is calculated in the iteration

$$F'_X(E) = -F(X) \tag{3.2}$$

and the solution of (1.1) can be obtained by the Newton's iteration

$$X(i+1) = X(i) - \left[F'_{X(i)}\right]^{-1} F(X(i)), \quad \forall i = 0, 1, 2 \ldots.$$

Algorithm 3.1 (Newton's method).

(i) Input a symmetric matrix $A \in \mathbb{R}^{n \times n}$ and symmetric initial guess $X(0) \in \mathbb{R}^{n \times n}$.
(ii) Solve $E(k)$ in

$$AE(k)A - X(k)AE(k) - E(k)AX(k) = -AX(k)A + X(k)AX(k).$$

(iii) $X(k+1) = X(k) - \left[F'_{X(k)}\right]^{-1} F(X(k)) = X(k) + E(k), \quad \forall k = 0, 1, 2, \ldots.$
(iv) Check if $\|F(X(k))\| \leqslant \epsilon$, then stop, otherwise go to step (ii).
(v) Display the solution X.

Remark 3.2. We see that Newton's method for (1.1) is inapplicable if the Kronecker Fréchet derivative F'_X in step (iii) of Algorithm 3.1 is singular and the existence of the symmetric solution is not ensured. Thus, we provide an iterative Algorithm 3.3 which works even if the Kronecker Fréchet derivative F'_X is singular and ensures the existence of the symmetric solution of (1.1) when incorporated into Algorithm 3.1.

Here, we construct a conjugate gradient least squares iterative method for obtaining the symmetric solution $E(k)$ for step (ii) in Algorithm 3.1 and derive necessary condition for existence of symmetric solution.

Algorithm 3.3 (An iterative method for solving Newton's step $E(k)$).

(a) Input $A \in \mathbb{R}^{n \times n}$, symmetric matrix $X(p) \in \mathbb{R}^{n \times n}$, and symmetric initial guess $E(p0) \in \mathbb{R}^{n \times n}$. For $k = 0$, compute;
(b) $R(0) = -F(X(p)) - [AE(p0) - AX(p)AE(p0) - E(p0)AX(p)]$;
(c) $M_0 = A^T R(0) A^T - (X(p)A)^T R(0) - R(0)(AX(p))^T$;
(d) $Q(0) = \frac{1}{2} \left(M_0 + M_0^T\right)$;
(e) $\alpha(0) = \frac{\|R(0)\|^2}{\|Q(0)\|^2}$.

While $R(k) \neq 0$ or $Q(k) \neq 0$, evaluate

(i) $\alpha_k = \frac{\|R(k)\|^2}{\|Q(k)\|^2}$;
(ii) $E(pk+1) = E(pk) + \alpha_k Q(k)$;

(iii) $R(k+1) = -F(X(p)) - [AE(pk+1)A - X(p)AE(pk+1) - E(k)AX(p)]$;

(iv) $M_{k+1} = A^T R(k+1)A^T - (X(p)A)^T R(k+1) - R(k+1)(AX(p))^T$;

 (v) $\beta(k) = \frac{\|R(k+1)\|^2}{\|R(k)\|^2}$;

(vi) $Q(k+1) = \frac{1}{2}\left(M(k+1) + M(k+1)^T\right) + \beta(k)Q(k)$.

Remark 3.4 (Conditions for the existence of symmetric solution in Algorithm 3.3).

1. The sequence of matrices $Q(k)$ and $E(pk)$ are symmetric $\forall k = 0, 1, \ldots$.
2. When $R(k) = 0$, then the kth iterate is the approximate symmetric solution.

Now, we prove the following lemmas and theorems based on Algorithm 3.3.

Lemma 3.5. *Presume that* $E(p)$ *is a symmetric solution of pth Newton's iteration* (3.2), *and the sequences* $\{M(k)\}$, $\{R(k)\}$, $\{E(pk)\}$ *are generated by Algorithm 3.3. Then,*

$$\text{trace}\left[M(k)^T(E(p) - E(pk))\right] = \|R(k)\|^2, \quad \forall k = 0, 1, 2, \ldots.$$

Proof. From Algorithm 3.3, we have

$$\begin{aligned}
&\text{trace}\left[M(k)^T(E(p) - E(pk))\right] \\
&= \text{trace}\left\{\left[A^T R(k)A^T - (X(p)A)^T R(k) - R(k)(AX(p))^T\right]^T(E(p) - E(pk))\right\} \\
&= \text{trace}\left\{R(k)^T\left[A^T(E(p) - E(pk))A^T - X(p)A(E(p) - E(pk)) - (E(p) - E(pk))AX(p)\right]\right\} \\
&= \text{trace}\left\{R(k)^T\left[-F(X(p)) - [AE(pk)A - X(p)AE(pk) - E(pk)AX(p)]\right]\right\} \\
&= \text{trace}\left\{R(k)^T R(k)\right\} = \|R(k)\|^2.
\end{aligned}$$

Thus, the proof is completed. □

Lemma 3.6. *Assume that* $E(p)$ *is a symmetric solution of pth Newton's iteration* (3.2). *Then,*

$$\text{trace}\left[Q(k)^T(E(p) - E(pk))\right] = \|R(k)\|^2, \quad \forall k = 0, 1, 2, \ldots. \tag{3.3}$$

Proof. We prove via mathematical induction. Suppose $k = 0$, it follows from Algorithm 3.3, Lemma 2.2, and Lemma 3.5 that

$$\begin{aligned}
\text{trace}\left[Q(0)^T(E(p) - E(p0))\right] &= \text{trace}\left[\frac{1}{2}\left(M_0 + M_0^T\right)^T(E(p) - E(p0))\right] \\
&= \text{trace}\left[M_0^T(E(p) - E(p0))\right] = \|R(0)\|^2.
\end{aligned}$$

Now presume that statement (3.3) holds true for $k = h \in \mathbb{N}$, we need to show that the statement (3.3) also holds for $k = h + 1 \in \mathbb{N}$. From Algorithm 3.3, Lemma 2.2, and Lemma 3.5, we have

$$\begin{aligned}
&\text{trace}\left[Q(h+1)^T(E(p) - E(ph+1))\right] \\
&= \text{trace}\left\{\left[\frac{1}{2}\left(M(h+1) + M(h+1)^T\right)^T + \beta(h)Q(h)\right]^T(E(p) - E(ph+1))\right\} \\
&= \text{trace}\left[M(h+1)^T(E(p) - E(ph+1))\right] + \beta(h)\text{trace}\left[Q(h)^T(E(p) - E(ph+1))\right] \\
&= \|R(h+1)\|^2 + \beta(h)\text{trace}\left[Q(h)^T(E(p) - E(ph) - \alpha(h)Q(h))\right] \\
&= \|R(h+1)\|^2 + \beta(h)\text{trace}\left[Q(h)^T(E(p) - E(ph))\right] - \beta(h)\alpha(h)\|Q(h)\|^2 \\
&= \|R(h+1)\|^2 + \beta(h)\|R_h\|^2 - \beta(h)\|R_h\|^2 \\
&= \|R(h+1)\|^2 + \|R(h+1)\|^2 - \|R(h+1)\|^2 = \|R(h+1)\|^2.
\end{aligned}$$

The proof is completed as required. □

Remark 3.7. From Lemma 3.6, for the Newton's iteration (3.2) to have a symmetric solution, then the sequences $\{R(k)\}$ and $\{Q(k)\}$ generated by Algorithm 3.3 should be nonzero.

Lemma 3.8. *Suppose that the matrix sequences $\{R(k)\}$ and $\{Q(k)\}$ are generated by Algorithm 3.3. Then, it holds that*

$$\text{trace}(R(k)^\mathsf{T} R(j)) = 0 \text{ and } \text{trace}(Q(k)^\mathsf{T} Q(j)) = 0, \ \forall j = 0, 1, \dots, l, \text{ where } 1 \leqslant l \text{ and } k > j. \quad (3.4)$$

Proof. We prove by mathematical induction.

Step 1: Let $l = 1$, it follows that

$$
\begin{aligned}
\text{trace}\left[R(1)^\mathsf{T} R(0)\right] &= \text{trace}\left\{\left[-F(X(p)) - [AE(p1)A - X(p)AE(p1) - E(p1)AX(p)]\right]^\mathsf{T} R(0)\right\} \\
&= \text{trace}\Big\{\left[-F(X(p)) - [AE(p0)A - X(p)AE(p0) - E(p0)AX(p)\right. \\
&\quad + \left. \alpha(0)(AQ(0)A - X(p)AQ(0) - Q(0)AX(p))]\right]^\mathsf{T} R(0)\Big\} \\
&= \text{trace}\left\{[R(0) - \alpha(0)\left(AQ(0)A - X(p)AQ(0) - Q(0)AX(p))\right]^\mathsf{T} R(0)\right\} \\
&= \|R(0)\|^2 - \text{trace}\left\{\alpha(0)\left(Q(0)^\mathsf{T}\left[A^\mathsf{T} R(0)A^\mathsf{T} - (X(p)A)^\mathsf{T} R(0) - R(0)(AX(p))^\mathsf{T}\right]\right)\right\} \\
&= \|R(0)\|^2 - \alpha(0)\text{trace}\left[Q(0)^\mathsf{T} M_0\right] \\
&= \|R(0)\|^2 - \alpha(0)\text{trace}\left[Q(0)^\mathsf{T}\frac{1}{2}\left(M_0 + M_0^\mathsf{T}\right)\right] \\
&= \|R(0)\|^2 - \alpha(0)\text{trace}\left[Q(0)^\mathsf{T} Q(0)\right] = 0,
\end{aligned}
$$

and

$$
\begin{aligned}
\text{trace}\left[Q(1)^\mathsf{T} Q(0)\right] &= \text{trace}\left[\left[\frac{1}{2}\left(M_1 + M_1^\mathsf{T}\right) + \beta(0)Q(0)\right]^\mathsf{T} Q(0)\right] \\
&= \text{trace}\left(M_1^\mathsf{T} Q(0)\right) + \beta(0)\text{trace}\left(Q(0)^\mathsf{T} Q(0)\right) \\
&= \text{trace}\left[\left[A^\mathsf{T} R(1)A^\mathsf{T} - (X(p)A)^\mathsf{T} R(1) - R(1)(AX(p))^\mathsf{T}\right]^\mathsf{T} Q(0)\right] + \beta(0)\|Q(0)\|^2 \\
&= \text{trace}\left[R(1)^\mathsf{T}\left[AQ(0)A - X(p)AQ(0) - Q(0)AX(p)\right]\right] + \frac{\|R(1)\|^2}{\|R(0)\|^2}\|Q(0)\|^2 \\
&= \text{trace}\left[R(1)^\mathsf{T}\left[A\frac{1}{\alpha(0)}(E(p1) - E(p0))A - X(p)A\frac{1}{\alpha(0)}(E(p1) - E(p0))\right.\right. \\
&\quad \left.\left. - \frac{1}{\alpha(0)}(E(p1) - E(p0))AX(p)\right]\right] + \frac{\|R(1)\|^2}{\|R(0)\|^2}\|Q(0)\|^2 \\
&= \frac{1}{\alpha(0)}\text{trace}\left[R(1)^\mathsf{T}(R(0) - R(1))\right] + \frac{\|R(1)\|^2}{\|R(0)\|^2}\|Q(0)\|^2 \\
&= \frac{1}{\alpha(0)}\left(\text{trace}\left[R(1)^\mathsf{T} R(0)\right] - \text{trace}\left[R(1)^\mathsf{T} R(1)\right]\right) + \frac{\|R(1)\|^2}{\|R(0)\|^2}\|Q(0)\|^2 \\
&= -\frac{1}{\alpha(0)}\text{trace}\left[R(1)^\mathsf{T} R(1)\right] + \frac{\|R(1)\|^2}{\|R(0)\|^2}\|Q(0)\|^2 \\
&= -\frac{\|R(1)\|^2}{\|R(0)\|^2}\|Q(0)\|^2 + \frac{\|R(1)\|^2}{\|R(0)\|^2}\|Q(0)\|^2.
\end{aligned}
$$

Presume that the statement (3.4) holds for $l = s \in \mathbb{N}$. We show that it holds for $l = s + 1 \in \mathbb{N}$. From Algorithm 3.3, we have

$$\text{trace}\left[R(s+1)^\mathsf{T} R(s)\right] = \text{trace}\left[[R(s) - \alpha(s)\left(AQ(s)A - X(p)AQ(s) - Q(s)AX(p))\right]^\mathsf{T} R(s)\right]$$

$$= \text{trace} \left[R(s)^\mathsf{T} R(s) \right] - \alpha(s)\text{trace} \left[[AQ(s)A - X(p)AQ(s) - Q(s)AX(p)]^\mathsf{T} R(s) \right]$$

$$= \|R(s)\|^2 - \alpha(s)\text{trace} \left[Q(s)^\mathsf{T} \left(A^\mathsf{T} R(s)A^\mathsf{T} - (X(p)A)^\mathsf{T} R(s) - R(s)(AX(p)^\mathsf{T}) \right) \right]$$

$$= \|R(s)\|^2 - \alpha(s)\text{trace} \left[Q(s)^\mathsf{T} M_s \right]$$

$$= \|R(s)\|^2 - \alpha(s)\text{trace} \left[Q(s)^\mathsf{T} \frac{1}{2}(M_s + M_s^\mathsf{T}) \right]$$

$$= \|R(s)\|^2 - \alpha(s)\text{trace} \left[Q(s)^\mathsf{T} (Q(s) - \beta_{s-1}Q(s-1)) \right]$$

$$= \|R(s)\|^2 - \alpha(s)\|Q(s)\|^2 + \alpha(s)\beta_{s-1}\text{trace} \left[Q(s)^\mathsf{T} Q(s-1) \right]$$

$$= \|R(s)\|^2 - \|R(s)\|^2 + 0 = 0.$$

Likewise, we have

$$\text{trace} \left[Q(s+1)^\mathsf{T} Q(s) \right] = \text{trace} \left[\left[\frac{1}{2}\left(M_{s+1} + M(s+1)^\mathsf{T} \right) + \beta(s)Q(s) \right]^\mathsf{T} Q(s) \right]$$

$$= \text{trace} \left[M(s+1)^\mathsf{T} Q(s) \right] + \beta(s)\|Q(s)\|^2$$

$$= \text{trace} \left[\left[A^\mathsf{T} R(s+1)A^\mathsf{T} - (X(p)A)^\mathsf{T} R(s+1) - R(s+1)(AX(p)^\mathsf{T} \right]^\mathsf{T} Q(s) \right] + \beta(s)\|Q(s)\|^2$$

$$= \text{trace} \left[R(s+1)^\mathsf{T} [AQ(s)A - X(p)AQ(s) - Q(s)AX(p)] \right] + \beta(s)\|Q(s)\|^2$$

$$= \text{trace} \left[R(s+1)^\mathsf{T} \frac{1}{\alpha(s)}(R(s) - R(s+1)) \right] + \beta(s)\|Q(s)\|^2$$

$$= -\frac{1}{\alpha(s)}\|R(s+1)\|^2 + \beta(s)\|Q(s)\|^2$$

$$= -\frac{\|Q(s)\|^2}{\|R(s)\|^2}\|R(s+1)\|^2 + \frac{\|R(s+1)\|^2}{\|R(s)\|^2}\|Q(s)\|^2 = 0.$$

Consequently, we have $\text{trace} \left[R(k)^\mathsf{T} R(k-1) \right] = 0$ and $\text{trace} \left[Q(k)^\mathsf{T} Q(k-1) \right] = 0$, $\forall k = 0, 1, \ldots, l$.

Step 2: We assume that $\text{trace} \left[R(s)^\mathsf{T} R(j) \right] = 0$ and $\text{trace} \left[Q(s)^\mathsf{T} Q(j) \right] = 0$, $\forall j = 0, 1, \ldots, l-1$. By Algorithm 3.3 and Lemma 2.2, along with the assumptions made, it follows that

$$\text{trace} \left[R(s+1)^\mathsf{T} R(j) \right] = \text{trace} \left[[R(s) - \alpha(s)(AQ(s)A - X(p)AQ(s) - Q(s)AX(p))]^\mathsf{T} R(j) \right]$$

$$= \text{trace} \left[R(s)^\mathsf{T} R(j) \right] - \alpha(s)\text{trace} \left[Q(s)^\mathsf{T} \left(A^\mathsf{T} R(j)A^\mathsf{T} - (X(p)A)^\mathsf{T} R(j) - R(j)(AX(p))^\mathsf{T} \right) \right]$$

$$= \text{trace} \left[R(s)^\mathsf{T} R(j) \right] - \alpha(s)\text{trace} \left[Q(s)^\mathsf{T} M(j) \right]$$

$$= 0 - \alpha(s)\text{trace} \left[Q(s)^\mathsf{T} \frac{1}{2}(M(j) + M(j)^\mathsf{T}) \right]$$

$$= -\alpha(s)\text{trace} \left[Q(s)^\mathsf{T} (Q(j) - \beta(j-1)Q(j-1)) \right] = 0.$$

Lastly, we verify that $\text{trace} \left[Q(s+1)^\mathsf{T} Q(j) \right] = 0$.

$$\text{trace} \left[Q(s+1)^\mathsf{T} Q(j) \right] = \text{trace} \left[\left[\frac{1}{2}\left(M_{s+1} + M(s+1)^\mathsf{T} \right) + \beta(s)Q(s) \right]^\mathsf{T} Q(j) \right]$$

$$= \text{trace} \left[M(s+1)^\mathsf{T} Q(j) \right]$$

$$= \text{trace} \left[\left[A^\mathsf{T} R(s+1)A^\mathsf{T} - (X(p)A)^\mathsf{T} R(s+1) - R(s+1)(AX(p))^\mathsf{T} \right]^\mathsf{T} Q(j) \right]$$

$$= \text{trace} \left[R(s+1)^\mathsf{T} [AQ(j)A - X(k)AQ(j) - Q(j)AX(p)] \right]$$

$$= \text{trace} \left[R(s+1)^\mathsf{T} \frac{1}{\alpha(j)}(R(j) - R(j+1)) \right]$$

$$= \frac{1}{\alpha(j)} \text{trace} \left[R(s+1)^\mathsf{T} R(j) \right] - \frac{1}{\alpha(j)} \text{trace} \left[R(s+1)^\mathsf{T} R(j+1) \right] = 0,$$

$\forall j = 0, 1, \ldots, s - 1$. The proof is done.   □

From Lemma 3.8, it is apparent that if $k > 0$, and $R(i) \neq 0$, $\forall i = 0, 1, \ldots, k$, then, the sequences $R(i)$, $R(j)$ generated by Algorithm 3.3 are orthogonal $\forall j \neq i$. We have the following remark.

Remark 3.9. According to Lemma 3.8, if there exists a nonnegative integer k($k > 0$) such that $R(i) \neq 0$, $\forall i = 0, 1, \ldots, k$ in Algorithm 3.3, consequently the sequence of matrices $R(i)$ and $R(j)$ are orthogonal $\forall i \neq j$.

Theorem 3.10. *Presume that the* p*th Newton's iteration* (3.2) *has an approximate symmetric solution. Then, for an arbitrary symmetric initial guess* $E(p0)$, *its symmetric solution can be obtained within finite iterative steps.*

Proof. From Lemma 3.8, assume that $R(k) \neq 0$, $\forall k = 0, 1, \ldots, n^2 - 1$. Since the pth Newton's iteration (3.2) has a symmetric solution, then based on Remark 3.9, it is certain that there exists a nonnegative integer k such that $Q(k) \neq 0$. Consequently, we can compute $E(pn^2)$ and R_{n^2} by Algorithm 3.3. Furthermore, from Lemma 3.8, we notice that $\text{trace}(R(n^2)^\mathsf{T} R(k)) = 0$, $\forall k = 0, 1, 2, \ldots, n^2 - 1$ and $\text{trace}(R(i)^\mathsf{T} R(j)) = 0$, $\forall i, j = 0, 1, \ldots, n^2 - 1$ with $i \neq j$. That is for an orthogonal basis $\{R(k)\}$, it follows that

$$\text{trace}(R(i)^\mathsf{T}, R(k)) = \begin{cases} \|R(k)\|^2, & \forall i = k, \\ 0, & \forall i \neq k. \end{cases}$$

However, we understand that $\text{trace}(R(n^2)^\mathsf{T} R(k)) = 0$ holds true if $R(n^2) = 0$. Thus, $E(pn^2)$ is the solution of the pth Newton's iteration (3.2).   □

Now, we are in a position to demonstrate the convergence of Algorithm 3.1 to a nontrivial symmetric solution of (1.1) when incorporated with Algorithm 3.3.

Theorem 3.11. *Presume that* (1.1) *has a symmetric solution and each Newton's iteration is consistent for symmetric initial guess* $X(0)$. *The sequence* $\{X(k)\}$ *is generated by Algorithm 3.1 with* $X(0)$ *such that* $\lim_{k \to \infty} X(k) = X_*$, *and the matrix* X_* *satisfies* $F(X_*) = 0$, *then,* X_* *is a symmetric solution of* (1.1).

Proof. Since each Newton's iteration has a symmetric solution, from Theorem 3.10 and Newton's method we can obtain the sequence $\{X(k)\}$ which is the set of symmetric matrices. Furthermore, the Newton's sequence converges to the solution X_* which is a symmetric solution of (1.1).   □

4. Numerical examples

In this section, some numerical tests are provided to demonstrate our theoretical results. All the tests are performed by MATLAB R2015a. Due to the influence of round off error, we consider a matrix R as zero if $\|R\| < \epsilon = 10^{-06}$. We give three applications of our theoretical results. The first example considers a real application to a well-known matrix [2] in the classical Yang-Baxter equation. In the second and third examples, we consider artificial matrices.

Example 4.1. In this example, we consider (1.1) with $A = \begin{pmatrix} a & 0 & 0 & d \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ d & 0 & 0 & a \end{pmatrix}$ for the completely integrable

system studied in [17] (see also [2]). As assumed in [7], $b = 1, c = 0$, and $d = a - 1$ with $a \neq 1$, therefore $A = I - uv^\mathsf{T}$, where $u = (1 - a, 0, 0, 1 - a)^\mathsf{T}$ and $v = (1, 0, 0, 1)^\mathsf{T}$. Setting $a = 1/2$, $X(0) = 0.0002 * I_4$, $E(0) = \text{zeros}(4)$ and applying Algorithms 3.1 and 3.3, we get our symmetric solution

$$X = 10^{-03} \begin{pmatrix} 0.099979991996799 & 0 & 0 & 0.100020008003201 \\ 0 & -0.000040016006403 & 0 & 0 \\ 0 & 0 & -0.000040016006403 & 0 \\ 0.100020008003201 & 0 & 0 & 0.099979991996799 \end{pmatrix}$$

with a corresponding residual 4.0×10^{-08}.

Example 4.2. We consider (1.1) with $A = \frac{W_g}{40}$, where

$$
W_g = \begin{cases} \frac{1}{g^{i-1}}, & \text{if } i = j, \\ \frac{1}{i+j+1}, & \text{if } i \neq j, \end{cases} \quad i, j = 1, 2, 3, \ldots, g,
$$

and g is the size of matrix A. Then employing Algorithms 3.1 and 3.3 with $g = 4$, $X_0 = 0.02I$, and $E(0) = 0$, and iterating one step, we have the approximate symmetric solution of (1.1)

$$
X = \begin{pmatrix} 0.018508011860414 & -0.000927806052941 & -0.000756902508248 & -0.000684082682123 \\ -0.000927806052941 & 0.019097274532220 & -0.000504815622338 & -0.000566077052292 \\ -0.000756902508248 & -0.000504815622338 & 0.019220439996582 & -0.000392552347481 \\ -0.000684082682123 & -0.000566077052292 & -0.000392552347481 & 0.019445356064656 \end{pmatrix}
$$

with a corresponding residual 7.22×10^{-07}.

Example 4.3. We consider (1.1) with $A = \begin{pmatrix} 0.000382 & 0.000157 & 0.000395 \\ 0.000157 & 0 & 0.000478 \\ 0.000395 & 0.000478 & 0.001065 \end{pmatrix}$. Applying Algorithms 3.1 and 3.3 with $X(0) = 0.0025I$ and $E(0)$ and iterating one step we obtain a symmetric solution of (1.1)

$$
X(1) = \begin{pmatrix} 0.002352708907435 & -0.000088923544160 & -0.000231206116464 \\ -0.000088923544160 & 0.002376752258012 & -0.000203348192003 \\ -0.000231206116464 & -0.000203348192003 & 0.001926142918804 \end{pmatrix}
$$

with a corresponding residual $\|AX(1)A - X(1)AX(1)\| = 1.91 \times 10^{-09}$.

5. Conclusion

Recently, many researchers have been toiling to find all solutions of (1.1) by imposing some restrictions on matrix A. However, in the literature, there is an inconsiderable work on the algorithms to solve equation (1.1).

In this paper, we incorporated an efficient conjugate gradient based on iterative method into Newton's method and exploited it to obtain the approximate symmetric solution of (1.1). Moreover, we derived some necessary conditions for the existence of the symmetric solution for the suggested iterative method. Numerical experiments reveal the validity of our theoretical results for both elementary and some general matrices considered. Our proposed Algorithm converges very slowly when the initial guess of not sufficiently close to the solution. On the other hand, it converges very fast if the initial guess is sufficiently very close to the solution of our equation. The choice of $E(0)$ and $X(0)$ is very crucial in our proposed method. Thus, this work acts as the benchmark for future research.

Acknowledgment

References

[1] D. Bachiller, F. Cedó, *A family of solutions of the Yang-Baxter equation*, J. Algebra, **412** (2014), 218–229. 1
[2] R. J. Baxter, *Partition function of the eight–vertex lattice model*, Ann. Physics, **70** (1972), 193–228. 4, 4.1
[3] F. Cedó, E. Jespers, J. Okniski, *Retractability of set theoretic solutions of the Yang-Baxter equation*, Adv. Math., **224** (2010), 2472–2484. 1

[4] J. Ding, N. H. Rhee, *A nontrivial solution to a stochastic matrix equation*, East Asian J. Appl. Math., **2** (2012), 277–284. 1

[5] J. Ding, N. H. Rhee, *Computing solutions of the Yang-Baxter-like matrix equation for diagonalisable matrices*, East Asian J. Appl. Math., **5** (2015), 75–84. 1

[6] J. Ding, N. H. Rhee, *Spectral solutions of the Yang–Baxter matrix equation*, J. Math. Anal. Appl., **402** (2013), 567–573. 1

[7] J. Ding, H. Tian, *Solving the YangBaxter–like matrix equation for a class of elementary matrices*, Comput. Math. Appl., **72** (2016), 1541–1548. 1, 4.1

[8] J. Ding, C. Zhang, *On the structure of the spectral solutions of the Yang–Baxter matrix equation*, Appl. Math. Lett., **35** (2014), 86–89. 1

[9] J. Ding, C. Zhang, N. H. Rhee, *Further solutions of a Yang-Baxter–like matrix equation*, East Asian J. Appl. Math., **2** (2013), 352–362. 1

[10] J. Ding, C. Zhang, N. H. Rhee, *Commuting solutions of the Yang-Baxter matrix equation*, Appl. Math. Lett., **44** (2015), 1–4. 1

[11] J. Ding, A. Zhou, *Eigenvalues of rank-one updated matrices with some applications*, Appl. Math. Lett., **20** (2007), 1223–1226. 1

[12] Q. Dong, J. Ding, *Complete commuting solutions of the Yang-Baxter-like matrix equation for diagonalizable matrices*, Comput. Math. Appl., **72** (2016), 194–201. 1

[13] F. Felix, *Nonlinear Equations, Quantum Groups and Duality Theorems: A Primer on the Yang-Baxter Equation*, VDM Verlag, Germany, (2009). 1

[14] T. Gateva–Ivanova, *Quadratic algebras, Yang-Baxter equation, and Artin–Schelter regularity*, Adv. Math., **230** (2012), 2153–2175. 1

[15] J. Hietarinta, *All solutions to the constant quantum Yang-Baxter equation in two dimensions*, Phys. Lett. A, **165** (1992), 245–251. 1

[16] M. Jimbo, *Yang-Baxter Equation in Integrable Systems*, World Scientific Publishing Co., Teaneck, (1989). 1

[17] P. P. Kulish, E. K. Sklyanin, *Solutions of the Yang-Baxter equation, in: Yang-Baxter Equation in Integrable Systems*, Adv. Series in Math. Phys., **10** (1990), 172–198. 4.1

[18] S. I. A. Mansour, J. Ding, Q. L. Huang, *Explicit solutions of the Yang–Baxter–like matrix equation for an idempotent matrix*, Appl. Math. Lett., **63** (2017), 71–76. 1

[19] H. Tian, *All solutions of the Yang–Baxter–like matrix equation for rank-one matrices*, Appl. Math. Lett., **51** (2016), 55–59. 1

[20] C. N. Yang, *Some exact results for the many–body problem in one dimension with repulsive delta–function interaction*, Phys. Rev. Lett., **19** (1967), 1312–1315. 1

[21] C. Yang, M. Ge, *Braid Group, Knot Theory, and Statistical Mechanics*, World Scientific Publishing Co., Teaneck, (1989). 1