Journal of Applied Analysis and Computation
Volume 13, Number 5, October 2023, 2987–2994
Website:http://www.jaac-online.com
DOI:10.11948/20230115

SOLVING THE YANG-BAXTER-LIKE MATRIX EQUATION FOR A RANK-ONE MATRIX*

Mohammed Ahmed Adam Abdalrahman[1,2], Jiu Ding[3]
and Qianglian Huang[1,†]

Abstract We reduce the problem of solving the Yang-Baxter-like matrix equation $AXA = XAX$, where A is a rank-one matrix, to that of solving linear matrix equations, obtaining all solutions. We use a direct and unified approach for the both cases that A is diagonalizable or otherwise, instead of seeking the help of the Jordan canonical form or factorization of A. Based on the characterizations for the solutions, we derive a perturbation result when A is not diagonalizable.

Keywords Yang-Baxter-like matrix equation, nilpotent matrix, Moore-Penrose generalized inverse, perturbation.

MSC(2010) 15A18.

1. Introduction

For a given $n \times n$ complex matrix A with $n \geq 2$, the problem of solving the following so-called *Yang-Baxter-like matrix equation*

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

has been extensively studied in recent years (see, e.g., [1,4,7,9,11] and the references therein). This quadratic matrix equation is similar in format to the classic Yang-Baxter equation [2,15], which has close relations to knot theory, braid group theory, quantum group theory, etc. We refer to the monographs [8,16] for some historical aspects and applications of the Yang-Baxter equation.

The Yang-Baxter-like matrix equation has two trivial solutions $X = 0$ and $X = A$, but it is a difficult task to find all other solutions for a general matrix A. Solving the quadratic matrix equation for the unknown $n \times n$ matrix X is equivalent to solving a system of n^2 scalar quadratic equations of n^2 variables that are the entries of X, and determining all the solutions of the system is a challenge in general in algebraic geometry. Even for the simple case when A is a 2×2 Jordan

[†]The corresponding author.
[1]School of Mathematical Sciences, Yangzhou University, Yangzhou 225002, China
[2]Department of Mathematics, Faculty of Science and Information Technology, University of Nyala, P. O. Box 63311, Sudan
[3]School of Mathematics and Natural Sciences, University of Southern Mississippi, Hattiesburg, MS 39406
*This research is supported by National Natural Science Foundation of China (11771378).
Email: Mohammed.sudan112@gmail.com (M. A. A. Abdalrahman), Jiu.Ding@usm.edu(J. Ding), huangql@yzu.edu.cn(Q. Huang)

block with eigenvalue $\lambda \neq 0$, solving the four quadratic scalar equations takes much time via hand computation for all the non-trivial solutions.

However, if one gives the additional requirement that the solution commutes with A, that is $AX = XA$, after a series of papers in the literature, the goal of finding all the commuting solutions has been reached; see, e.g., [5–7, 13]. Thus, the remaining problem is to find all the non-commuting solutions of (1.1). For some special classes of matrices of A, such as elementary matrices and low ranked matrices, all the non-commuting solutions have been obtained [4, 12, 14], by means of the Jordan canonical form decomposition. However, even for a nilpotent matrix A with a couple of 2×2 Jordan blocks, finding all non-commuting solutions of the corresponding Yang-Baxter-like matrix equation is still an open problem.

So far, the approach to solving the Yang-Baxter-like matrix equation is by means of the Jordan canonical form J of the given matrix A, and then solving the "simplified" Yang-Baxter-like matrix equation

$$JYJ = YJY. \tag{1.2}$$

As is well known (Lemma 1 in [5]), the two equations (1.1) and (1.2) are equivalent in the sense that X solves (1.1) if and only if $Y = S^{-1}XS$ solves (1.2), where $S^{-1}AS = J$. Although in theory the reduction of the matrix equation with the original matrix A to the one with the simplest matrix J in the "similarity class" of A can help us find some or all solutions by using the structural analysis of the Jordan blocks in the canonical form of A, finding the Jordan canonical form of the given matrix and the corresponding matrix S, whose columns are eigenvectors and generalized eigenvectors associated to all the eigenvalues of A, is not a simple task, especially the numerical computation of J and S. It turns out that obtaining partial or all solutions of (1.1) directly without resort to the costly reduction process of the Jordan canonical form decomposition is important and desired in applications.

A recent paper [10] proposed an alternative approach that avoids using the Jordan canonical form of the known matrix A. For the case that $A = pq^T$ with p and q given nonzero vectors, a necessary and sufficient condition is given for a matrix X of the same order as A to be a solution of (1.1), based on which some expressions of all commuting or non-commuting solutions are obtained.

The purpose of this paper is to establish concise characterizations in terms of the given matrix A for a matrix X to solve (1.1) when A is a general rank one matrix without being factorized. We show that solving the quadratic matrix equation (1.1) with a rank-one matrix A is equivalent to solving several linear matrix equations. Thus, we can construct a numerical scheme to find all solutions of the matrix equation, so robust linear solvers can be applied to solve (1.1) numerically. Our approach and analysis are more direct than before and the results are simpler to express. Such results together with the concept of generalized inverses also lead to a perturbation analysis for solving (1.1) when A is not diagonalizable.

In the next section we solve (1.1) directly for the first case that A is diagonalizable, and we present the solution result in Section 3 for the case that A is not diagonalizable. We discuss numerical methods for finding all the solutions and develop a perturbation result in Section 4. We conclude in Section 5.

2. Solutions When A is Diagonalizable

Let A be an $n \times n$ complex matrix of rank one. As usual, by $r(A), R(A)$, and $N(A)$ we denote the rank, range, and null space of A, respectively. Suppose that X is a solution of the corresponding Yang-Baxter-like matrix equation (1.1). We shall characterize X in terms of linear matrix equations. In this section we assume that A is diagonalizable.

Since $r(A) = 1$, it is easy to see that A is similar to a diagonal matrix of diagonal entries $0, \ldots, 0, c$ with $c \neq 0$, so c is a nonzero eigenvalue of A. Since the other eigenvalue of A is 0 of multiplicity $n-1$, it follows that the minimal polynomial of A is $\lambda(\lambda - c)$, from which

$$A^2 = cA. \tag{2.1}$$

Theorem 2.1. *Suppose that the given rank-one matrix A is diagonalizable. Then an $n \times n$ matrix X solves* (1.1) *if and only if one of the following conditions is satisfied:*

(i) Either $AX = 0$ or $XA = 0$;

(ii) Both $AX = cA$ and $XA = cA$ hold, where c is a nonzero eigenvalue of A.

Moreover, X is a commuting solution if and only if either $AX = XA = 0$ or $AX = XA = cA$.

Proof. The matrix A has a unique nonzero eigenvalue c and satisfies the equation (2.1). For the sufficiency part, first assume that $AX = 0$ or $XA = 0$. Then $AXA = XAX = 0$. Now assume that X satisfies $AX = cA$ and $XA = cA$. Then

$$AXA = cA \cdot A = cA^2 = c \cdot cA = c^2 A$$

and

$$XAX = cA \cdot X = c \cdot cA = c^2 A,$$

so X solves (1.1).

For the necessity part, let X satisfy $AXA = XAX$. First suppose $XAX = 0$. If $AX \neq 0$ and $XA \neq 0$, then $X : R(A) \to \{0\}$ from $R(AX) = R(A)$, that is $XA = 0$, a contradiction to the assumption that $XA \neq 0$. This gives (i). Now suppose $XAX \neq 0$. Then from $XAXA = AXA^2 = cAXA$, we have

$$(X - cI)AXA = 0, \tag{2.2}$$

where I is the identity matrix. Since $AXA \neq 0$, the rank of AXA is 1, so $R(AXA) = R(A)$. Thus the equality (2.2) implies that $X - cI$ maps $R(A)$ into $\{0\}$, and consequently $XA - cA = (X - cI)A = 0$.

Now $XAX = AXA \neq 0$ implies $X^T A^T X^T = A^T X^T A^T \neq 0$. Then from $X^T A^T X^T A^T = A^T X^T (A^T)^2 = cA^T X^T A^T$,

$$(X^T - cI)A^T X^T A^T = 0,$$

so $(X^T - cI)A^T = 0$ since $R(A^T X^T A^T) = R(A^T)$, from which $AX = cA$.

The last statement of the theorem is obvious. □

Remark 2.1. When the rank-one matrix A has one nonzero eigenvalue c, there are nonsingular matrix solutions to the system $AX = cA$ and $XA = cA$, so there are nonsingular matrix solutions to (1.1); for example, $X = cI$ is one of such solutions. This fact can also be seen from the following illustrative example.

Example 2.1. Let A be a 3×3 matrix with all entries 1. Then $r(A) = 1$ and $A^2 = 3A$.

A basis of $N(A) = N(A^T)$ is given by

$$u_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad u_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

From the above theorem, all solutions of (1.1) are:

1) $$X = \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \end{bmatrix}, \ \forall\, r_i \in \mathbb{C},$$

2) $$X = \begin{bmatrix} s_1 & s_4 \\ s_2 & s_5 \\ s_3 & s_6 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \ \forall\, s_i \in \mathbb{C},$$

3) $$X = 3I + \begin{bmatrix} x+y+z+w & -x-z & -y-w \\ -x-y & x & y \\ -z-w & z & w \end{bmatrix}, \quad \forall\, x, y, z, w \in \mathbb{C}.$$

3. Solutions When A is not Diagonalizable

Again let $r(A) = 1$ for the given $n \times n$ complex matrix A. The remaining situation for us to discuss is that A is not diagonalizable. Then 0 is the only eigenvalue of A, so $A^2 = 0$ since $r(A) = 1$. In other words, A is a nilpotent matrix of order 2. Consequently, $R(A) \subset N(A)$.

Theorem 3.1. *Suppose that the given rank-one matrix A is not diagonalizable. Then an $n \times n$ matrix X solves* (1.1) *if and only if $AX = 0$ or $XA = 0$. Moreover, X is a commuting solution if and only if $AX = XA = 0$.*

Proof. The assumption implies that $A^2 = 0$, so $R(A) \subset N(A)$.

For the sufficiency part, suppose $AX = 0$ or $XA = 0$. Then clearly $AXA = XAX = 0$, so X solves (1.1).

For the necessity part, let X be a solution of (1.1). Suppose $AX \neq 0$. Then $0 < r(AX) \leq r(A) = 1$, from which $r(AX) = 1$. So $R(AX) = R(A)$ since $R(AX) \subset R(A)$ and $\dim R(A) = 1$. From $AXA = XAX$, we see that $(AX)^2 = AXAX = A \cdot AXA = A^2XA = 0$ since $A^2 = 0$. Thus, AX is a nilpotent matrix of rank one and order 2, and it follows that $R(A) = R(AX) \subset N(AX)$. Therefore, $AXA = 0$.

Then $X \cdot AX = AXA = 0$, which implies that $X : R(AX) = R(A) \to \{0\}$. In other words, $XA : \mathbb{C}^n \to \{0\}$. Hence $XA = 0$.

Now, suppose that $XA \neq 0$. Then $A^T X^T \neq 0$. By the same reason as above, $R(A^T) = R(A^T X^T) \subset N(A^T X^T)$, so $A^T X^T A^T = 0$. Thus, $X^T A^T X^T = (AXA)^T = A^T X^T A^T = 0$, which means that $X^T : R(A^T X^T) = R(A^T) \to \{0\}$, namely $X^T A^T = 0$. Hence $AX = 0$.

The last conclusion is obvious. □

Remark 3.1. The proof of Theorem 3.1 indicates that under its condition, all solutions X of (1.1) must make $AXA = XAX = 0$, a consequence of which is that X cannot be nonsingular, since otherwise $A = 0$ from multiplying X^{-1} to the both sides of the equality $XAX = 0$. This is unlike the situation in the previous section.

4. Computation of Solutions and Perturbations

From the characterizations proved in the previous two sections, we know that, for the class of rank-one matrices A, solving the nonlinear matrix equation (1.1) is equivalent to solving the linear matrix equations $AX = 0$ and $XA = 0$ separately if $A^2 = 0$, and the linear matrix equations $AX = 0$ and $XA = 0$ separately or $AX = cA$ and $XA = cA$ simultaneously if the equation (2.1) true for some $c \neq 0$. Hence numerical linear algebra can be effectively used to find such solutions.

For the case that A is nilpotent, solving the system of $AX = XA = 0$ will give all the commuting solutions, and when A is not nilpotent, solving the system $AX = XA = 0$ or that of $AX = XA = cA$ will give all the commuting solutions. The equation $AX = 0$ means that all columns of X belong to $N(A)$, so if one finds a basis $\{u_1, \ldots, u_{n-1}\}$ of $N(A)$, then all columns of X are linear combinations of $u_1, \ldots, u_{n-1}$. It follows that all solutions of $AX = 0$ can be expressed as $X = UR$ with R an arbitrary $(n-1) \times n$ matrix, where $U = [u_1, \ldots, u_{n-1}]$.

Similarly, all solutions to the equation $XA = 0$ can be written as $X = LV$ with L an arbitrary $n \times (n-1)$ matrix, where $V^T = [v_1, \ldots, v_{n-1}]$ whose columns constitute a basis of $N(A^T)$.

As for the consistent nonhomogeneous equation $AX = cA$, since $X = cI$ is one particular solution, all its solutions are $X = UR + cI$ with the same U and R as above. Similarly, all solutions to $XA = cA$ have the general expression $X = LV + cI$ with the same V and L as above. Their common solutions are given by those R and L such that $UR = LV$. Since c is nonzero, if we make any norm $\|UR\| = \|LV\| < c$, then X is nonsingular by Banach's lemma.

For a diagonalizable rank-one matrix A so that the equation (2.1) holds for some $c \neq 0$, a small rank-preserving perturbation $\tilde{A}$ of A is still diagonalizable, from which $\tilde{A}^2 = \tilde{c}\tilde{A}$ with $\tilde{c}$ a small perturbation of c. However, if a rank-one matrix A is not diagonalizable so that $A^2 = 0$, most small rank-preserving perturbations of A become diagonalizable. Since numerically solving the Yang-Baxter-like matrix equation involves errors, a perturbation analysis of (1.1) is of practical importance, which has not appeared in the literature. From the characterizations for all solutions of (1.1) in Sections 2 and 3 for the two different cases of A, we can deduce a perturbation result when a given rank-one nilpotent matrix A is perturbed to a rank-one matrix $\tilde{A}$. We shall use $\| \;\|$ to denote the vector 2-norm and the induced matrix 2-norm. For any $m \times n$ matrix $B = [b_{ij}]$, let

$$\|B\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |b_{ij}|^2 \right)^{1/2}$$

be the Frobenius norm of B, whose square is the trace of the matrix $B^H B$, and let $B^\dagger$ be the Moore-Penrose generalized inverse of B Campbell [3]. It is well-known that for any two matrices B and C such that BC is defined, $\|BC\|_F \leq \|B\|_F \|C\|_F$.

Theorem 4.1. *Suppose that the given rank-one matrix A is not diagonalizable so that $A^2 = 0$. Let $\tilde{A} = A + E$ be an $n \times n$ complex matrix such that $r(\tilde{A}) = 1$ and $\tilde{A}^2 = \epsilon \tilde{A}$, where ϵ is a complex number. Then for any solution $\tilde{X}$ to the perturbed Yang-Baxter-like matrix equation $\tilde{A}\tilde{X}\tilde{A} = \tilde{X}\tilde{A}\tilde{X}$, there is a solution X to* (1.1) *such that*

$$\|\tilde{X} - X\|_F \leq \sqrt{2} \left[|\epsilon| + (|\epsilon| + \|\tilde{X}\|_F) \|A^\dagger E\|_F \right].$$

Proof. Write $\tilde{A} = [\tilde{a}_1 \ \cdots \ \tilde{a}_n]$ and $\tilde{X} = [\tilde{x}_1 \ \cdots \ \tilde{x}_n]$. Then $\tilde{A}\tilde{x}_j = \epsilon \tilde{a}_j$ for $j = 1, \ldots, n$. Let x_j be the orthogonal projection of $\tilde{x}_j$ onto $N(A)$ for each j. Then $AX = 0$ with $X = [x_1 \ \cdots \ x_n]$, and $\tilde{x}_j - x_j \in N(A)^\perp$ for each j, where $N(A)^\perp$ is the orthogonal complement of $N(A)$ in $\mathbb{C}^n$. Now fix $j = 1, \ldots, n$.

Subtracting $Ax_j = 0$ from $\tilde{A}\tilde{x}_j = \epsilon \tilde{a}_j$ gives

$$A(\tilde{x}_j - x_j) = \epsilon \tilde{a}_j - E\tilde{x}_j.$$

By the definition of the generalized inverse, $A^\dagger A(\tilde{x}_j - x_j) = \tilde{x}_j - x_j$, so

$$\tilde{x}_j - x_j = A^\dagger(\epsilon \tilde{a}_j - E\tilde{x}_j),$$

from which

$$\|\tilde{x}_j - x_j\|^2 \leq 2 \left(|\epsilon|^2 \|A^\dagger \tilde{a}_j\|^2 + \|A^\dagger E \tilde{x}_j\|^2 \right).$$

Consequently,

$$\begin{aligned} \|\tilde{X} - X\|_F^2 &= \sum_{j=1}^{n} \|\tilde{x}_j - x_j\|^2 \\ &\leq 2 \left(|\epsilon|^2 \sum_{j=1}^{n} \|A^\dagger \tilde{a}_j\|^2 + \sum_{j=1}^{n} \|A^\dagger E \tilde{x}_j\|^2 \right) \\ &= 2 \left(|\epsilon|^2 \|A^\dagger \tilde{A}\|_F^2 + \|A^\dagger E \tilde{X}\|_F^2 \right). \end{aligned}$$

It follows that

$$\begin{aligned} \|\tilde{X} - X\|_F &\leq \sqrt{2} \left(|\epsilon|^2 \|A^\dagger \tilde{A}\|_F^2 + \|A^\dagger E\|_F^2 \|\tilde{X}\|_F^2 \right)^{\frac{1}{2}} \\ &\leq \sqrt{2}(|\epsilon| \|A^\dagger \tilde{A}\|_F + \|A^\dagger E\|_F \|\tilde{X}\|_F). \end{aligned}$$

Since $A^\dagger A$ is the orthogonal projection onto $N(A)^\perp$, which is one dimensional, $\|A^\dagger A\|_F$ equals 1, which is the trace of $(A^\dagger A)^H A^\dagger A = A^\dagger A$, so $\|A^\dagger \tilde{A}\|_F = \|A^\dagger A + A^\dagger E\|_F \leq 1 + \|A^\dagger E\|_F$. Therefore,

$$\|\tilde{X} - X\|_F \leq \sqrt{2} \left[|\epsilon|(1 + \|A^\dagger E\|_F) + \|\tilde{X}\|_F \|A^\dagger E\|_F \right]$$

$$= \sqrt{2}\left[|\epsilon| + (|\epsilon| + \|\tilde{X}\|_F)\|A^\dagger E\|_F\right].$$

□

Remark 4.1. Since the solution $\tilde{X}$ in Theorem 4.1 also satisfies $\tilde{X}\tilde{A} = \epsilon\tilde{A}$, by considering the orthogonal projection of each row vector of $\tilde{X}$ onto $N(A^T)$, we obtain another solution X to (1.1) that satisfies $XA = 0$.

5. Conclusions

We have given a direct approach to finding all the solutions of (1.1) without taking the initial step of reducing the given matrix A to its Jordan canonical form as done before in the literature. For a general $n \times n$ complex matrix A of rank-one in the Yang-Baxter-like matrix equation (1.1), we have derived two necessary and sufficient conditions for the solutions of the equation, depending on whether A is diagonalizable or not. The first case implies that A has a nonzero eigenvalue and the second case happens when A is also a nilpotent matrix of order 2. In both cases, the problem of solving the nonlinear matrix equation is simplified to solving linear matrix equations. Based on such results, a natural numerical scheme to find all the solutions of the matrix equation is proposed, and a perturbation bound has also been obtained for the structurally unstable case when A is nilpotent.

It will be interesting and desired to develop a usable perturbation upper bound for the case when the given rank-one matrix A is diagonalizable. This is possible after exploring the linear structure of the solution set for the system $AX = XA = cA$ and applying a minimal distance argument such as the least squares technique.

Acknowledgements

This manuscript has benefited greatly from the constructive comments and helpful suggestions of the anonymous referees, the authors would like to express their deep gratitude to them.

References

[1] M. Adam, J. Ding, Q. Huang and L. Zhu, *Solving a class of quadratic matrix equations*, Applied Math. Lett., 2018, 82, 58–63.

[2] R. J. Baxter, *Partition function of the eight-vertex lattice model*, Ann. Phys., 1972, 70, 193–228.

[3] S. L. Campbell and C. D. Meyer, *Generalized Inverses of Linear Transformations*, Pitman, 1979.

[4] J. Ding and N. Rhee, *Spectral solutions of the Yang-Baxter matrix equation*, J. Math. Anal. Appl., 2013, 402, 567–573.

[5] J. Ding, C. Zhang and N. Rhee, *Commuting solutions of the Yang-Baxter matrix equation*, Applied Math. Lett., 2015, 44, 1–4.

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

[7] Q. Dong and J. Ding, *All commuting solutions of a quadratic matrix equation for general matrices*, J. Nonlinear Modeling Anal., 2020, 2(1), 111–123.

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

[9] Q. Huang, M. Adam, J. Ding and L. Zhu, *All non-commuting solutions of the Yang-Baxter matrix equation for a class of diagonalizable matrices*, Operators and Matrices, 2019, 1, 187–195.

[10] L. Lu, *Manifold expressions of all solutions of the Yang-Baxter-like matrix equation for rank-one matrices*, Appl. Math. Lett., 2022, 108175.

[11] H. Mukherjee and A. M, *Solutions to the matrix Yang-Baxter equation*, arXiv: 2209. 04605, 2022.

[12] D. Shen and M. Wei, *All solutions of the Yang-Baxter-like matrix equation for diagonalizable coefficient matrix with two different eigenvalues*, Applied Math. Lett., 2020, 101, 106048.

[13] D. Shen, M. Wei and Z. Jia, *On commuting solutions of the Yang-Baxter-like matrix equation*, J. Math. Anal. Appl., 2018, 462(1), 665–696.

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

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

[16] C. Yang and M. Ge, *Braid Group, Knot Theory, and Statistical Mechanics*, World Scientific, 1989.