Journal of Nonlinear Modeling and Analysis http://jnma.ca; http://jnma.ijournal.cn
Volume 2, Number 1, March 2020, 111–123 DOI:10.12150/jnma.2020.111

All Commuting Solutions of a Quadratic Matrix Equation for General Matrices*

Qixiang Dong[1] and Jiu Ding[2,†]

Abstract Using the Jordan canonical form and the theory of Sylvester's equation, we find all the commuting solutions of the quadratic matrix equation $AXA = XAX$ for an arbitrary given matrix A.

Keywords Jordan canonical form, Sylvester's equation.

MSC(2010) 15A24, 15A18.

1. Introduction

The purpose of this paper is to determine all the commuting solutions of the quadratic matrix equation

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

where A is a given $n \times n$ complex matrix. This equation is called the *Yang-Baxter-like matrix equation* since it has a similar pattern to the classical Yang-Baxter equation introduced independently by Yang in [11] and Baxter in [1], which is famous in statistical physics with close relations to knot theory, braid groups, and quantum groups [8, 12].

Finding all the solutions of (1.1) is difficult for general A, and so far it is only possible for some special matrices as in [10]. This is due to the fact that if we multiple out the both sides of the equation, solving it is equivalent to solving a system of n^2 quadratic polynomial equations in n^2 variables, which is a challenging task in general. Thus, the current research on solving (1.1) is mainly focused on finding commuting solutions, namely the solutions that commute with A. Some recent papers have been devoted to finding various commuting solutions of (1.1) with different assumptions on A. In particular, corresponding to each eigenvalue of A, a spectral projection solution was obtained in [3]. When all the eigenvalues of A are semi-simple, the whole set of the commuting solutions of (1.1) has been successfully constructed in [6] with the help of a result on unique solutions of the Sylvester equation.

[†]the corresponding author.
Email address: qxdong@yzu.edu.cn(Q. Dong), jiu.ding@usm.edu, jiudin@gmail.com(J. Ding)
[1]School of Mathematical Sciences, Yangzhou University, Yangzhou 225002, China
[2]School of Mathematics and Natural Sciences, University of Southern Mississippi, Hattiesburg, MS 39406, USA
[*]The first author was supported by the National Natural Science Foundation of China (11571300 and 11871064).

A natural question arises: can we find all the commuting solutions of (1.1) if A is not diagonalizable? A serious study about it began with the paper [7] in which all the commuting solutions have been described when A is a general nilpotent matrix, based on the above mentioned result on the Sylvester equation and the structure theorem of [2, 13] on matrices that commute with a Jordan block with eigenvalue zero.

In this paper, based on the ideas developed in the above works, we want to extend the main result of [7] from a nilpotent matrix to an arbitrary one. We shall give a general solution structure theorem on all the commuting solutions of (1.1), thus giving an answer to the question of finding all commuting solutions of a general Yang-Baxter-like matrix equation. After the paper was written up, we learnt that the same problem was also studied in a recent paper [9] with a different approach. In the next section we present some key lemmas for our purpose, and the main result will be given in Section 3. Some concrete examples constitute in Section 4 to illustrate our theorem, and we conclude with Section 5.

2. Preliminaries

Let A be an arbitrary $n \times n$ complex matrix. The following lemma provides an equivalent way to solve (1.1) for commuting solutions, which was proved in [7].

Lemma 2.1. *A matrix X satisfies $AX = XA$ and $AXA = XAX$ if and only if $AX = XA$ and $X(X - A)A = 0$.*

As proved in [4] (Lemma 3.1), solving (1.1) for a given matrix A is equivalent to solving a simpler Yang-Baxter-like matrix equation

$$JYJ = YJY, \tag{2.1}$$

where $J = U^{-1}AU$ is the Jordan form of A, and the solutions X to (1.1) and the solutions Y to (2.1) satisfy the relation $X = UYU^{-1}$. So from Lemma 2.1, we just need to solve the system

$$JY = YJ,\ Y(Y - J)J = 0$$

to find all the commuting solutions of (2.1). Then all the commuting solutions to (1.1) are given by $X = UYU^{-1}$.

Denote

$$J_j(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & \lambda & 1 \\ 0 & 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix}$$

the $j \times j$ Jordan block with eigenvalue λ. In particular, the Jordan block $J_j(0)$ corresponding to eigenvalue 0 satisfies $J_j(0)^j = 0$. The following lemma is a generalization of Theorem 5.15 of [2] from eigenvalue zero to any eigenvalue, but its proof is basically the same and is included for reader's convenience.

Lemma 2.2. *Let j and k be two natural numbers and let Y be a $j \times k$ matrix. Then $J_j(\lambda)Y = YJ_k(\lambda)$ if and only if*

$$Y = [0\ \hat{Y}] \ \text{ or } \ Y = \begin{bmatrix} \hat{Y} \\ 0 \end{bmatrix},$$

depending on whether $j \le k$ or $j \ge k$, where $\hat{Y}$ is an upper triangular Toeplitz matrix

$$Y = \begin{bmatrix} y_1 & y_2 & \cdots & \cdots & y_{l-1} & y_l \\ 0 & y_1 & y_2 & \cdots & \cdots & y_{l-1} \\ 0 & 0 & \ddots & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & y_2 \\ 0 & 0 & \cdots & \cdots & 0 & y_1 \end{bmatrix}, \quad l = \min\{j,k\}, \tag{2.2}$$

with $y_1, \ldots, y_l$ arbitrary complex numbers.

Proof Since $J_j(\lambda) = \lambda I_j + J_j(0)$ and $J_k(\lambda) = \lambda I_k + J_k(0)$, where I_j and I_k are the $j \times j$ and $k \times k$ identity matrices respectively, the condition $J_j(\lambda)Y = YJ_k(\lambda)$ is satisfied if and only if $J_j(0)Y = YJ_k(0)$. Let y_{pq} be the (p,q)-entry of Y. Then the (p,q)-entry of $J_j(0)Y$ is $y_{p+1,q}$ for $p = 1, \ldots, j-1$ and $q = 1, \ldots, k$, and all the (j,q)-entries are 0. Similarly, the (p,q)-entry of $YJ_k(0)$ is $y_{p,q-1}$ for $p = 1, \ldots, j$ and $q = 2, \ldots, k$, and all the $(p,1)$-entries are 0. Consequently $J_j(0)Y = YJ_k(0)$ if and only if

$$y_{p+1,q} = y_{p,q-1},\ y_{p0} = y_{j+1,q} = 0, \ \ p = 1, \ldots, j,\ q = 1, \ldots, k,$$

Hence K is given by (2.2) with $y_q \equiv y_{1q}$ arbitrary numbers for $q = 1, \ldots l$. □

We also need a general result from matrix theory, which was proved in [5] (Lemma 2.3), so that the problem of finding all the commuting solutions of (2.1) can be reduced to that of finding all the commuting solutions of (2.1) with J replaced by its diagonal blocks associated with distinct eigenvalues.

Lemma 2.3. *Let a square matrix $H = \operatorname{diag}(H_1, \ldots, H_d)$ be block diagonal with square diagonal blocks. Suppose that H_j and H_k have no common eigenvalues whenever $j \ne k$. If K is a square matrix such that $HK = KH$, then $K = \operatorname{diag}(K_1, \ldots, K_d)$, where K_j has the same size as H_j for all j.*

As an application of Lemmas 2.1 and 2.2, we show the following result, which was obtained in [4] as Theorem 3.1 via another argument based on projections.

Theorem 2.1. *If J is a single Jordan block $J_n(\lambda)$ with $\lambda \ne 0$, then all the commuting solutions of (2.1) are the trivial ones $Y = 0$ and $Y = J$.*

Proof By Lemma 2.2, all the solutions Y of the equation $J_n(\lambda)Y = YJ_n(\lambda)$ are given by (2.2) with $l = n$, so y_1 is the only eigenvalue of Y. Now the additional equation $Y(Y - J_n(\lambda))J_n(\lambda) = 0$ is reduced to $Y(Y - J_n(\lambda)) = 0$ since $J_n(\lambda)$ is nonsingular. Thus from $J_n(\lambda) = \lambda I + J_n(0)$ we have $YJ_n(0) = (Y - \lambda I)Y$. The fact $J_n(0)^n = 0$ implies that

$$0 = YJ_n(0)^n = YJ_n(0)J_n(0)^{n-1} = (Y - \lambda I)YJ_n(0)^{n-1} = \cdots = (Y - \lambda I)^n Y.$$

Let u be an eigenvector of Y associated with eigenvalue y_1. Then

$$0 = (Y - \lambda I)^n Y u = [(y_1 - \lambda)^n y_1]u,$$

from which $(y_1 - \lambda)^n y_1 = 0$. Hence $y_1 = \lambda$ or $y_1 = 0$.

If $y_1 = \lambda$, then Y is nonsingular and so $Y(Y - J_n(\lambda)) = 0$ implies that $Y = J_n(\lambda)$. If $y_1 = 0$, then the only eigenvalue $-\lambda$ of $Y - \lambda I$ is nonzero, hence the matrix $(Y - \lambda I)^n$ is nonsingular, resulting in $Y = 0$. □

Remark 2.1. If $\lambda = 0$, then all the commuting solutions of (2.1) have been obtained in [7] (Theorem 2.1).

3. Commuting Solutions for a General Matrix

We solve (1.1) for commuting solutions with an arbitrary matrix A that has d distinct eigenvalues $\lambda_1, \ldots, \lambda_d$. Let

$$J = \begin{bmatrix} D_1 & & & \\ & D_2 & & \\ & & \ddots & \\ & & & D_d \end{bmatrix} \tag{3.1}$$

be the Jordan form of A, where for $t = 1, \ldots, d$, each D_t is itself a block matrix consisting of all the Jordan blocks associated with eigenvalue λ_t in the increasing order of block sizes. Let U denote a nonsingular matrix such that $A = UJU^{-1}$.

For each $t = 1, \ldots, d$, without loss of generality, we can write D_t as

$$D_t = \begin{bmatrix} \lambda_t I_{m_t} & & & \\ & J_{2,\lambda_t} & & \\ & & \ddots & \\ & & & J_{r_t,\lambda_t} \end{bmatrix}, \quad t = 1, \ldots, d,$$

where I_{m_t} is the $m_t \times m_t$ identity matrix with a possibility that $m_t = 0$, and r_t is the maximum size of the Jordan blocks corresponding to eigenvalue λ_t. For $j = 2, \ldots, r_t$, the block diagonal matrix J_{j,λ_t} has the structure

$$J_{j,\lambda_t} = \begin{bmatrix} J_j(\lambda_t) & & \\ & \ddots & \\ & & J_j(\lambda_t) \end{bmatrix},$$

in which the $j \times j$ Jordan block $J_j(\lambda_t)$ appears $s_j(\lambda_t)$ times. Of course, if a Jordan block of some size $j < r_t$ does not exist, then the corresponding J_{j,λ_t} will not be present inside D_t.

Since the eigenvalues of $D_1, \ldots, D_d$ in (3.1) are distinct, Lemma 2.3 immediately implies the following result.

Proposition 3.1. *All the commuting solutions of (2.1) are of the form $Y = diag\{Y_1, \ldots, Y_d\}$, where each Y_t is a general commuting solution of the Yang-Baxter-like matrix equation*

$$D_t Y_t D_t = Y_t D_t Y_t, \quad t = 1, \ldots, d.$$

Hence, we are lead to solving the above equations directly for all $t = 1, \ldots, d$. Since such equations are of the same type, we can suppress the subscript t to simplify the notation in our further analysis. This way, we restate our subproblem as follows.

Given a block diagonal matrix $J \equiv J(\lambda)$ of the form

$$\begin{bmatrix} \lambda I_m & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_r \end{bmatrix}, \quad J_j = \begin{bmatrix} J_j(\lambda) & & \\ & \ddots & \\ & & J_j(\lambda) \end{bmatrix}, \quad j = 2, \ldots, r, \tag{3.2}$$

where the $j \times j$ Jordan block $J_j(\lambda)$ appears s_j times in J_j for $j = 2, \ldots, r$. And we solve the corresponding Yang-Baxter-like matrix equation (2.1) for all the commuting solutions.

First we determine all matrices Y that commute with J. Since $J = \lambda I + J(0)$, the commutability equation $JY = YJ$ is equivalent to

$$J(0)Y = YJ(0).$$

Partition Y as an $r \times r$ block matrix

$$Y = \begin{bmatrix} Y_{11} & \cdots & Y_{1r} \\ \vdots & \vdots & \vdots \\ Y_{r1} & \cdots & Y_{rr} \end{bmatrix} \tag{3.3}$$

with the same sizes of the diagonal blocks of $J(0)$, so Y_{11} is $m \times m$ and Y_{ii} is $is_i \times is_i$ for $i = 2, \ldots, r$. Then according to the analysis of [7] for obtaining the structure of (3.3), we have the following proposition.

Proposition 3.2. *Let J be defined by (3.2). Then all the solutions of the equation $JY = YJ$ are given by (3.3) such that*

(i) Y_{11} is any $m \times m$ complex matrix;

(ii) for $i = 2, \ldots, r$, the $s_i \times 1$ block matrix $Y_{i1} = [(A_1^{(i)})^T \cdots (A_{s_j}^{(i)})^T]^T$ with the blocks $A_k^{(i)}$ of size $i \times m$ given by

$$A_k^{(i)} = \begin{bmatrix} a_1^{(i,k)} & \cdots\cdots & a_m^{(i,k)} \\ 0 & \cdots\cdots & 0 \\ \vdots & \cdots\cdots & \vdots \\ 0 & \cdots\cdots & 0 \end{bmatrix}, \quad \forall\, a_1^{(i,k)}, \ldots, a_m^{(i,k)}; \quad k = 1, \ldots, s_i; \tag{3.4}$$

(iii) for $j = 2, \ldots, r$, the $1 \times s_j$ block matrix $Y_{1j} = [B_1^{(j)} \cdots B_{s_j}^{(j)}]$ with the blocks $B_k^{(j)}$ of size $m \times j$ given by

$$B_k^{(j)} = \begin{bmatrix} 0 \cdots 0 \; b_1^{(j,k)} \\ \vdots \ldots \vdots \quad \vdots \\ 0 \cdots 0 \; b_m^{(j,k)} \end{bmatrix}, \quad \forall\, b_1^{(j,k)}, \ldots, b_m^{(j,k)}; \quad k = 1, \ldots, s_j; \tag{3.5}$$

(iv) for $i, j = 2, \ldots, r$, the $s_i \times s_j$ block matrix $Y_{ij} = [A_{uv}^{(i,j)}]$ with the blocks $A_{uv}^{(i,j)}$ of size $i \times j$ given by

$$A_{uv}^{(i,j)} = \begin{bmatrix} a_{uv1}^{(i,j)} & a_{uv2}^{(i,j)} & \cdots & a_{uvj}^{(i,j)} \\ 0 & a_{uv1}^{(i,j)} & \ddots & \vdots \\ \vdots & 0 & \ddots & a_{uv2}^{(i,j)} \\ 0 & 0 & \cdots & a_{uv1}^{(i,j)} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \forall\, a_{uv1}^{(i,j)}, \ldots, a_{uvj}^{(i,j)} \tag{3.6}$$

if $i \geq j$, and if $i < j$,

$$A_{uv}^{(i,j)} = \begin{bmatrix} 0 \cdots 0 & a_{uv1}^{(i,j)} & a_{uv2}^{(i,j)} & \cdots & a_{uvi}^{(i,j)} \\ \vdots \quad\; \vdots & 0 & a_{uv1}^{(i,j)} & \ddots & \vdots \\ \vdots \quad\; \vdots & 0 & 0 & \ddots & a_{uv2}^{(i,j)} \\ 0 \cdots 0 & 0 & 0 & \cdots & a_{uv1}^{(i,j)} \end{bmatrix}, \quad \forall\, a_{uv1}^{(i,j)}, \ldots, a_{uvi}^{(i,j)}. \tag{3.7}$$

After solving the commutability equation $JY = YJ$ as above, we turn to solving the second equation $Y(Y - J)J = 0$ with Y already given by Proposition 3.1. The case that $\lambda = 0$ was already investigated in [7] and all the commuting solutions of (1.1) were given in Theorem 3.1 therein, since J is a nilpotent matrix. So in what follows we assume that $\lambda \neq 0$. Then, since J is nonsingular, it is enough to solve the simpler equation

$$Y(Y - J) = 0.$$

With the block matrix structure of Y and J, we have that

$$Y(Y - J) = \begin{bmatrix} Y_{11} & \cdots & Y_{1r} \\ Y_{21} & \cdots & Y_{2r} \\ \vdots & \vdots & \vdots \\ Y_{r1} & \cdots & Y_{rr} \end{bmatrix} \begin{bmatrix} Y_{11} - \lambda I_m & \cdots & Y_{1r} \\ Y_{21} & \cdots & Y_{2s} \\ \vdots & \ddots & \vdots \\ Y_{r1} & \cdots & Y_{rr} - J_r \end{bmatrix}$$

$$= \begin{bmatrix} Y_{11}(Y_{11}-\lambda I_m)+\sum_{j=2}^{r} Y_{1j}Y_{j1} & \cdots & \sum_{j=2}^{r-1} Y_{1j}Y_{jr}+Y_{1r}(Y_{rr}-J_r) \\ Y_{21}(Y_{11}-\lambda I_m)+\sum_{j=2}^{r} Y_{2j}Y_{j1} & \cdots & \sum_{j=2}^{r-1} Y_{2j}Y_{jr}+Y_{2r}(Y_{rr}-J_r) \\ \vdots & \vdots & \vdots \\ Y_{r1}(Y_{11}-\lambda I_m)+\sum_{j=2}^{r} Y_{rj}Y_{j1} & \cdots & \sum_{j=2}^{r-1} Y_{rj}Y_{jr}+Y_{rr}(Y_{rr}-J_r) \end{bmatrix}.$$

To simplify the above expressions, we need the following lemma.

Lemma 3.1. *$Y_{ij}Y_{jk}=0$ if $i+k\le j$.*

Proof Suppose $i+k\le j$. Let $k=1$. If $i=1$, then $j\ge 2$ and from the zero structure of (3.4) and (3.5), $B_u^{(j)}A_u^{(j)}=0$ for $u=1,\dots,s_j$, so

$$Y_{1j}Y_{j1}=[B_1^{(j)}\cdots B_{s_j}^{(j)}][(A_1^{(j)})^T\cdots(A_{s_j}^{(j)})^T]^T=\sum_{u=1}^{s_j}B_u^{(j)}A_u^{(j)}=0.$$

If $2\le i<j$, then $A_{uv}^{(i,j)}A_v^{(j)}=0$ for all u,v from (3.7) and (3.4). Thus

$$Y_{ij}Y_{j1}=\begin{bmatrix} A_{11}^{(i,j)} & \cdots & A_{1s_j}^{(i,j)} \\ \vdots & \vdots & \vdots \\ A_{s_i1}^{(i,j)} & \cdots & A_{s_is_j}^{(i,j)} \end{bmatrix}\begin{bmatrix} A_1^{(j)} \\ \vdots \\ A_{s_j}^{(j)} \end{bmatrix}=\begin{bmatrix} \sum_{v=1}^{s_j}A_{1v}^{(i,j)}A_v^{(j)} \\ \vdots \\ \sum_{v=1}^{s_j}A_{s_iv}^{(i,j)}A_v^{(j)} \end{bmatrix}=\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Now let $i=1$. If $2\le k<j$, then since $B_u^{(j)}A_{uv}^{(j,k)}=0$ for all u,v,

$$\begin{aligned} Y_{1j}Y_{jk} &= [B_1^{(j)}\cdots B_{s_j}^{(j)}]\begin{bmatrix} A_{11}^{(j,k)} & \cdots & A_{1s_k}^{(j,k)} \\ \vdots & \vdots & \vdots \\ A_{s_j1}^{(j,k)} & \cdots & A_{s_js_k}^{(j,k)} \end{bmatrix} \\ &= \left[\sum_{u=1}^{s_j}B_u^{(j)}A_{u1}^{(j,k)}\cdots\sum_{u=1}^{s_j}B_u^{(j)}A_{us_k}^{(j,k)}\right]=[0\cdots 0]. \end{aligned}$$

Finally, Let $i\ge 2$ and $k\ge 2$. Then $i<j$ and $k<j$, so from (3.6) and (3.7) we have $A_{uv}^{(i,j)}A_{vl}^{(j,k)}=0$ for all u,v,l. Therefore,

$$\begin{aligned} Y_{ij}Y_{jk} &= \begin{bmatrix} A_{11}^{(i,j)} & \cdots & A_{1s_j}^{(i,j)} \\ \vdots & \vdots & \vdots \\ A_{s_i1}^{(i,j)} & \cdots & A_{s_is_j}^{(i,j)} \end{bmatrix}\begin{bmatrix} A_{11}^{(j,k)} & \cdots & A_{1s_k}^{(j,k)} \\ \vdots & \vdots & \vdots \\ A_{s_j1}^{(j,k)} & \cdots & A_{s_js_k}^{(j,k)} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{v=1}^{s_j}A_{1v}^{(i,j)}A_{v1}^{(j,k)} & \cdots & \sum_{v=1}^{s_j}A_{1v}^{(i,j)}A_{vs_k}^{(j,k)} \\ \vdots & \vdots & \vdots \\ \sum_{v=1}^{s_j}A_{s_iv}^{(i,j)}A_{v1}^{(j,k)} & \cdots & \sum_{v=1}^{s_j}A_{s_iv}^{(i,j)}A_{vs_k}^{(j,k)} \end{bmatrix}=\begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}.\square \end{aligned}$$

Let H_{ij} denote the (i,j)-block of $Y(Y-J)$. Then by Lemmas 3.1,

$$H_{ij}=\sum_{k=1,k\ne j}^{\min\{i+j-1,r\}}Y_{ik}Y_{kj}+Y_{ij}(Y_{jj}-J_j),\ \ \forall\, i,j=1,\dots,r,$$

where $J_1 \equiv \lambda I_m$. In particular, $H_{11} = Y_{11}(Y_{11} - \lambda I_m)$, and all the solutions of $H_{11} = 0$ are $Y_{11} = \lambda P$ with P any $m \times m$ projection matrix [6].

It is time to state the main result of this paper after summarizing the above computation and the main result of [7].

Theorem 3.1. *Let A be an $n\times n$ complex matrix with distinct eigenvalues $\lambda_1, \dots, \lambda_d$, and let U be a nonsingular matrix such that the Jordan form $J = U^{-1}AU$ of A is given by (3.1). Then all the commuting solutions of the corresponding Yang-Baxter-like matrix equation (1.1) are $X = UYU^{-1}$ with $Y = diag\{Y_1, \dots, Y_d\}$, where for $t = 1, \dots, d$, the matrix Y_t is an $r_t \times r_t$ block matrix as given by (3.3) with $r = r_t$ and satisfies the following:*

I. If $\lambda_t = 0$, then

(i) Y_{11} is any $m \times m$ complex matrix;

(ii) for $i = 2, \dots, r_t$, the $s_i \times 1$ block matrix $Y_{i1} = [(A_1^{(i)})^T \cdots (A_{s_i}^{(i)})^T]^T$ with the blocks $A_k^{(i)}$ of size $i \times m$ given by (3.4);

(iii) for $j = 2, \dots, r_t$, the $1 \times s_j$ block matrix $Y_{1j} = [B_1^{(j)} \cdots B_{s_j}^{(j)}]$ with the blocks $B_k^{(j)}$ of size $m \times j$ given by (3.5);

(iv) for $i, j = 2, \dots, r_t$, the $s_i \times s_j$ block matrix $Y_{ij} = [A_{uv}^{(i,j)}]$ with the blocks $A_{uv}^{(i,j)}$ of size $i \times j$ given by (3.6) if $i \geq j$ or (3.7) if $i < j$ that satisfy

$$\sum_{k=2,k\neq j}^{\min\{i+j-2,r\}} Y_{ik}Y_{kj}J_j + Y_{ij}(Y_{jj} - J_j)J_j = 0,\ i, j = 2, \dots, r_t.$$

II. If $\lambda_t \neq 0$, then

(i) $Y_{11} = \lambda_t P$, where P is any $m \times m$ projection matrix;

(ii) for $i = 2, \dots, r_t$, the $s_i \times 1$ block matrix $Y_{i1} = [(A_1^{(i)})^T \cdots (A_{s_i}^{(i)})^T]^T$ with the blocks $A_k^{(i)}$ of size $i\times m$ given by (3.4) and the $1\times s_j$ block matrix $Y_{1i} = [B_1^{(i)} \cdots B_{s_i}^{(i)}]$ with the blocks $B_k^{(i)}$ of size $m \times i$ given by (3.5), and for $i, j = 2, \dots, r_t$, the $s_i \times s_j$ block matrix $Y_{ij} = [A_{uv}^{(i,j)}]$ with the blocks $A_{uv}^{(i,j)}$ of size $i \times j$ given by (3.6) if $i \geq j$ or (3.7) if $i < j$ satisfy the matrix equations

$$\sum_{k=1,k\neq j}^{\min\{i+j-1,r\}} Y_{ik}Y_{kj} + Y_{ij}(Y_{jj} - J_j) = 0,\quad \forall\, i, j = 1, \dots, r_t, \tag{3.8}$$

where $J_1 \equiv \lambda I_m$.

4. A Concrete Example

We study in more details a special case in this section to illustrate our general result. More specifically, we assume that A has only one eigenvalue λ and $r = 2$. Since the case of $\lambda = 0$ was studied in detail by [7], we assume that $\lambda \neq 0$. With s and I denoting s_2 and I_m, system (3.8) in Theorem 3.1 becomes

$$\begin{cases} Y_{11}(Y_{11} - \lambda I) = 0, \\ Y_{11}Y_{12} + Y_{12}(Y_{22} - J_2) = 0, \\ Y_{22}Y_{21} + Y_{21}(Y_{11} - \lambda I) = 0, \\ Y_{21}Y_{12} + Y_{22}(Y_{22} - J_2) = 0. \end{cases}$$

Since the solutions of the first equation are $Y_{11} = \lambda P$ with P any $m \times m$ projection matrix, the above system is simplified to

$$\begin{cases} \lambda P Y_{12} + Y_{12}(Y_{22} - J_2) = 0, \\ Y_{22}Y_{21} - \lambda Y_{21}(I - P) = 0, \\ Y_{21}Y_{12} + Y_{22}(Y_{22} - J_2) = 0. \end{cases} \tag{4.1}$$

Here

$$Y_{21} = \begin{bmatrix} A_1 \\ \vdots \\ A_s \end{bmatrix}, \ Y_{12} = [B_1 \cdots B_s], \ Y_{22} = \begin{bmatrix} A_{11} & \cdots & A_{1s} \\ \vdots & \dots & \vdots \\ A_{s1} & \cdots & A_{ss} \end{bmatrix},$$

where

$$A_i = \begin{bmatrix} a_1^{(i)} & \cdots & a_s^{(i)} \\ 0 & \cdots & 0 \end{bmatrix}, B_j = \begin{bmatrix} 0 & b_1^{(j)} \\ \vdots & \vdots \\ 0 & b_s^{(j)} \end{bmatrix}, A_{ij} = \begin{bmatrix} a_{ij} & b_{ij} \\ 0 & a_{ij} \end{bmatrix}, i, j = 1, \dots, s.$$

We give a simple numerical example before ending the section. Let $n = 3$ and let A be its Jordan form $J = \mathrm{diag}(1, J_2(1))$. Then $\lambda = 1, m = 1, r = 2$, and $s = 1$. The 1×1 projection matrix P is either 1 or 0. Let

$$Y_{21} = \begin{bmatrix} x \\ 0 \end{bmatrix}, \ Y_{12} = [0 \ y], \ Y_{22} = \begin{bmatrix} z & w \\ 0 & z \end{bmatrix}.$$

Then for $P = 1$, the corresponding system (4.1) is

$$\begin{cases} [0 \ y] + [0 \ y] \begin{bmatrix} z-1 & w-1 \\ 0 & z-1 \end{bmatrix} & = 0, \\ \begin{bmatrix} z & w \\ 0 & z \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} & = 0, \\ \begin{bmatrix} x \\ 0 \end{bmatrix} [0 \ y] + \begin{bmatrix} z & w \\ 0 & z \end{bmatrix} \begin{bmatrix} z-1 & w-1 \\ 0 & z-1 \end{bmatrix} & = 0, \end{cases}$$

which is reduced to

$$\begin{cases} yz & = 0, \\ xz & = 0, \\ z(z-1) & = 0, \\ xy + 2zw - z - w & = 0. \end{cases}$$

Thus, $z = 0$, from which

$$Y = \begin{bmatrix} 1 & 0 & y \\ x & 0 & xy \\ 0 & 0 & 0 \end{bmatrix}, \ \forall\, x, y,$$

and $z = 1$, from which

$$Y = J = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Similarly, when $P = 0$, the system (4.1) becomes

$$\begin{cases} [0\ y] \begin{bmatrix} z-1 & w-1 \\ 0 & z-1 \end{bmatrix} & = 0, \\ \begin{bmatrix} z & w \\ 0 & z \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} - \begin{bmatrix} x \\ 0 \end{bmatrix} & = 0, \\ \begin{bmatrix} x \\ 0 \end{bmatrix} [0\ y] + \begin{bmatrix} z & w \\ 0 & z \end{bmatrix} \begin{bmatrix} z-1 & w-1 \\ 0 & z-1 \end{bmatrix} & = 0, \end{cases}$$

which is just

$$\begin{cases} y(z-1) & = 0, \\ x(z-1) & = 0, \\ z(z-1) & = 0, \\ xy + 2zw - z - w & = 0. \end{cases}$$

Hence $z = 0$ so $Y = 0$, and $z = 1$ so

$$Y = \begin{bmatrix} 0 & 0 & y \\ x & 1 & 1-xy \\ 0 & 0 & 1 \end{bmatrix}, \quad \forall\, x, y.$$

We generalize the example as follows. Let $n \geq 4$ and let A be its Jordan form $J = \mathrm{diag}(\lambda, J_{n-1}(\lambda))$ with $\lambda \neq 0$. Partition the solution matrix Y as

$$Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix},$$

where Y_{11} is the number λ or 0,

$$Y_{21} = \begin{bmatrix} x \\ 0 \end{bmatrix}, \quad Y_{12} = [0\ \ y], \quad Y_{22} = \begin{bmatrix} y_1 & y_2 & \cdots & y_{n-2} & y_{n-1} \\ 0 & y_1 & \cdots & \cdots & y_{n-2} \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \cdots & y_2 \\ 0 & 0 & \cdots & 0 & y_1 \end{bmatrix}.$$

Here x and y are numbers and 0 is the $(n-1)$-dimensional column or row zero vector.

Then for $Y_{11} = \lambda$, the corresponding system (4.1) is

$$\begin{cases} [0 \;\; y]\,[Y_{22} - J_{n-1}(0)] & = 0, \\ Y_{22} \begin{bmatrix} x \\ 0 \end{bmatrix} & = 0, \\ \begin{bmatrix} x \\ 0 \end{bmatrix} [0 \;\; y] + Y_{22}\,[Y_{22} - J_{n-1}(\lambda)] & = 0, \end{cases}$$

which is equivalent to

$$\begin{cases} yy_1 & = 0, \\ xy_1 & = 0, \\ y_1(y_1 - \lambda) & = 0, \\ y_1(y_2 - 1) + y_2(y_1 - \lambda) & = 0, \\ y_1 y_3 + y_2(y_2 - 1) + y_3(y_1 - \lambda) & = 0, \\ \vdots & \vdots \;\vdots \\ y_1 y_{n-2} + \cdots + y_{n-4} y_3 + y_{n-3}(y_2 - 1) + y_{n-2}(y_1 - \lambda) & = 0, \\ xy + y_1 y_{n-1} + \cdots + y_{n-3} y_3 + y_{n-2}(y_2 - 1) + y_{n-1}(y_1 - \lambda) & = 0. \end{cases}$$

It follows that $y_1 = 0$ or $y_1 = \lambda$. If $y_1 = 0$, then from the above system, $y_2 = \cdots = y_{n-2} = 0$ and $y_{n-1} = xy/\lambda$ with x and y arbitrary numbers. Thus the corresponding solutions are

$$Y = \begin{bmatrix} \lambda & 0 & \cdots & 0 & y \\ x & 0 & \cdots & 0 & \frac{xy}{\lambda} \\ 0 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & \vdots & 0 \end{bmatrix}.$$

If $y_1 = \lambda$, then $y_2 = 1$ and $x = y = y_3 = \cdots = y_{n-1} = 0$, so $Y = J$.

On the other hand, for $Y_{11} = 0$, we have the system

$$\begin{cases} [0 \;\; y]\,[Y_{22} - J_{n-1}(\lambda)] & = 0, \\ (Y_{22} - \lambda I) \begin{bmatrix} x \\ 0 \end{bmatrix} & = 0, \\ \begin{bmatrix} x \\ 0 \end{bmatrix} [0 \;\; y] + Y_{22}\,[Y_{22} - J_{n-1}(\lambda)] & = 0, \end{cases}$$

which is the same as

$$\begin{cases} y(y_1-\lambda) & = 0, \\ x(y_1-\lambda) & = 0, \\ y_1(y_1-\lambda) & = 0, \\ y_1(y_2-1)+y_2(y_1-\lambda) & = 0, \\ y_1y_3+y_2(y_2-1)+y_3(y_1-\lambda) & = 0, \\ \vdots & \vdots \ \vdots \\ y_1y_{n-2}+\cdots+y_{n-4}y_3+y_{n-3}(y_2-1)+y_{n-2}(y_1-\lambda) & = 0, \\ xy+y_1y_{n-1}+\cdots+y_{n-3}y_3+y_{n-2}(y_2-1)+y_{n-1}(y_1-\lambda) & = 0. \end{cases}$$

Again, $y_1 = 0$ or λ. In the case of $y_1 = 0$, we have $y = x = 0$ since $\lambda \neq 0$. Then $y_2 = \cdots y_{n-1} = 0$, so $Y = 0$. If $y_1 = \lambda$, then $y_2 = 1$, from which $y_3 = \cdots y_{n-2} = 0$ and $y_{n-1} = -xy/\lambda$ with x and y arbitrary numbers. Therefore, the corresponding solutions are

$$Y = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & y \\ x & \lambda & 1 & 0 & -\frac{xy}{\lambda} & \\ 0 & 0 & \ddots & \ddots & 0 & \\ \vdots & \vdots & \ddots & \vdots & 1 & \\ 0 & 0 & \cdots & \vdots & \lambda & \end{bmatrix}.$$

As an application of our main result, we have obtained the same result as Theorem 3.4 of [4] for all the commuting solutions of $JYJ = YJY$ for the class of matrices.

5. Conclusions

We have found all the commuting solutions of the Yang-Baxter-like matrix equation (1.1) for an arbitrary matrix A, based on a reduction of the problem to the case with A having only one eigenvalue that is nonzero. By solving the commutability equation $JY = YJ$ and a simpler matrix product homogeneous equation $Y(Y-J) = 0$ in succession with J the Jordan form of A, we obtained a system of matrix equations of smaller size, the solutions of which constitute all the commuting solutions of the original equation.

It is much more difficult to find all the non-commuting solutions of (1.1) for general A, and we hope to be able to explore this challenging problem in the future.

References

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

[2] C. Cullen, *Matrices and Linear Transformations*, Dover, 1990.

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

[4] J. Ding, N. Rhee, and C. Zhang, *Further solutions of a Yang-Baxter-like matrix equation*, East Asian J. Applied Math., 2013, 2, 352–362.

[5] J. Ding, N. Rhee, and C. Zhang, *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, J. Ding, and Q. Huang, *All commuting solutions of a quadratic matrix equation for nilpotent matrices*, Algb. Collo., 2018, 25(1), 31–44.

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

[9] 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.

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

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

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

[13] X. Zhan, *Matrix Theory*, Graduate Studies in Mathematics, AMS, 2013.