Wuhan University Journal of Natural Sciences

2021, Vol.26 No.6, 459-463

Article ID 1007-1202(2021)06-0459-05

DOI https://doi.org/10.1051/wujns/2021266459

Some Rank Formulas for the Yang-Baxter Matrix Equation $AXA = XAX$

□ **DAI Lifang, LIANG Maolin†, SHEN Yonghong**

School of Mathematics and Statistics, Tianshui Normal University, Tianshui 741001, Gansu, China

Abstract: Let A be an arbitrary square matrix, then equation $AXA = XAX$ with unknown X is called Yang-Baxter matrix equation. It is difficult to find all the solutions of this equation for any given A. In this paper, the relations between the matrices A and B are established via solving the associated rank optimization problem, where $B = AXA = XAX$, and some analytical formulas are derived, which may be useful for finding all the solutions and analyzing the structures of the solutions of the above Yang-Baxter matrix equation.

Key words: Yang-Baxter matrix equation; rank; generalized inverse

CLC number: O 124.6

0 Introduction

Throughout this paper, $\mathbf{C}^{m\times n}$ denotes the set of all $m\times n$ complex matrices. $r(A), A^+, R(A)$ stand for the rank, conjugate transpose and the range (column space) of a matrix $A\in \mathbf{C}^{m\times n}$, respectively. The Moore-Penrose inverse, denoted by A^+, is the unique solution X of the following Penrose equations

$$AXA = A,\ XAX = X,\ (AX)^* = AX,\ (XA)^* = XA$$

In addition, $E_A = I - AA^+$ and $F_A = I - A^+A$ are two oblique projectors included by $A\in \mathbf{C}^{m\times n}$.

The rank and generalized inverse of matrices are powerful tools to characterize relations among matrices, matrix equations, or matrix expressions. Two seminal references are Ref. [1] by Marsaglia and Styan, and Ref. [2] by Meyer. Some fundamental rank equalities and inequalities related to generalized inverses of block matrices were established in the two papers. Since then, the main results have widely been applied to dealing with various problems in the theory of generalized inverses of matrices and their applications, over complex field, even quaternion field, see, e.g., Refs.[3-9] and therein.

The Yang-Baxter equation has close relations to several mathematics subjects such as braid groups and knot theory, which has been widely studied, see, e.g., Refs.[10-15]. It happens that there is a similar case, the classic parameter-free Yang-Baxter equation has the same format as the following matrix equation

$$AXA = XAX \tag{1}$$

where A is any given matrix, and X is unknown to be determined. Hence, the matrix equation (1) can be referred to as Yang-Baxter matrix equation. Generally, it is difficult to obtain all the solutions of the equation because of its

Received date: 2021-07-20

Foundation item: Supported by the National Natural Science Foundation of China (11961057), the Science and Technology Project of Gansu Province (21JR1RE287 and 2021B-221), the Fuxi Scientific Research Innovation Team of Tianshui Normal University (FXD2020-03), and the Science Foundation (CXT2019-36 and CXJ2020-11) as well as the Education and Teaching Reform Project of Tianshui Normal University (JY202004 and JY203008)

Biography: DAI Lifang, female, Master, research direction: numerical linear algebra with applications. E-mail: dailf2005@163.com

† To whom correspondence should be addressed. E-mail: liangml2005@163.com; liangmaolin@tsnu.edu.cn

nonlinearity. Recently, Ding and his partners[15,16] have made several outstanding results. For instance, in Ref.[16], they have completed the existence proof based on the Brouwer fixed point via the direct iteration when the given $\boldsymbol{A}$ is a nonsingular quasi-stochastic matrix such that the inverse $\boldsymbol{A}^{-1}$ is a stochastic one. In addition, Cibotarica *et al* [17] studied equation (1) under the assumption that the given matrix $\boldsymbol{A}$ was idempotent. By means of the Jordan decomposition of $\boldsymbol{A}$, they deduced general solution of the matrix equation mentioned above.

However, as is pointed out above, there may be many other solutions of this matrix equation which are not found. In this paper, letting $\boldsymbol{A}$ be an $n\times n$ matrix, we shall make analysis on the relation between the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ via considering an associated rank optimization, where $\boldsymbol{B}=\boldsymbol{AXA}=\boldsymbol{XAX}$, and some rank formulas will be established, which may provide ideas for finding all the solutions and analyzing the structures of the solutions of the Yang-Baxter matrix equation (1).

Actually, for some $\boldsymbol{B}\in\mathbf{C}^{n\times n}$, suppose that matrix equation $\boldsymbol{B}=\boldsymbol{AXA}$ is consistent, then we have the following equality-constrained rank optimization problem

$$\min_{A\times A=B} r(\boldsymbol{B}-\boldsymbol{XAX}) \tag{2}$$

Obviously, if the minimal rank arrives at zero, then it is equivalent to (1).

The outline of this paper is as follows. In Section 1, some basic formulas on the rank of matrices or matrix expressions will be given. In Section 2, the rank optimization (2) will be studied, after that some rank formulas with respect to the Yang-Baxter matrix equation (1) will be shown. As applications, the general solution of (1) will be expressed under some constrained conditions. Finally, we conclude this paper with some remarks.

1 Preliminary Knowledge

In this section, we introduce some necessary results with respect to the ranks of block matrix and matrix expressions.

Lemma 1[1,2] Let $\boldsymbol{A}\in\mathbf{C}^{m\times n}$, $\boldsymbol{B}\in\mathbf{C}^{m\times k}$ and $\boldsymbol{C}\in\mathbf{C}^{l\times n}$. Then

(a) $r[\boldsymbol{A},\boldsymbol{B}]=r(\boldsymbol{A})+r(\boldsymbol{E}_{\boldsymbol{A}}\boldsymbol{B})=r(\boldsymbol{B})+r(\boldsymbol{E}_{\boldsymbol{B}}\boldsymbol{A})$.

(b) $r\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{C}\end{bmatrix}=r(\boldsymbol{A})+r(\boldsymbol{CF}_{\boldsymbol{A}})=r(\boldsymbol{C})+r(\boldsymbol{AF}_{\boldsymbol{C}})$.

(c) $r\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}=r\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{C}\end{bmatrix}+r[\boldsymbol{A},\boldsymbol{B}]-r(\boldsymbol{A})$
$+r[\boldsymbol{E}_{A_2}(\boldsymbol{D}-\boldsymbol{CA}^{+}\boldsymbol{B})\boldsymbol{F}_{A_1}]$,

where $\boldsymbol{A}_1=\boldsymbol{E}_{\boldsymbol{A}}\boldsymbol{B},\boldsymbol{A}_2=\boldsymbol{CF}_{\boldsymbol{A}}$.

(d) In particular, if $R(\boldsymbol{B})\subseteq R(\boldsymbol{A})$ and $R(\boldsymbol{C}^*)\subseteq R(\boldsymbol{A}^*)$, then $r(\boldsymbol{D}-\boldsymbol{CA}^{+}\boldsymbol{B})=r(\boldsymbol{A})+r\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}$.

Lemma 2[3] Given matrices $\boldsymbol{A}\in\mathbf{C}^{m\times n}$, $\boldsymbol{B}_i\in\mathbf{C}^{m\times p_i}$ and $\boldsymbol{C}_i\in\mathbf{C}^{q_i\times n}$ ($i=1,2$), and let

$$\mathcal{P}(\boldsymbol{X}_1,\boldsymbol{X}_2)=\boldsymbol{A}-\boldsymbol{B}_1\boldsymbol{X}_1\boldsymbol{C}_1-\boldsymbol{B}_2\boldsymbol{X}_2\boldsymbol{C}_2$$

with variables $\boldsymbol{X}_1,\boldsymbol{X}_2$. Then

$$\max_{X_1,X_2} r[\mathcal{P}(\boldsymbol{X}_1,\boldsymbol{X}_2)]$$
$$=\min\left\{r[\boldsymbol{A}\ \ \boldsymbol{B}_1\ \ \boldsymbol{B}_2],\ r\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{C}_1\\ \boldsymbol{C}_2\end{bmatrix},\ r\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}_1\\ \boldsymbol{C}_2 & \boldsymbol{0}\end{bmatrix},\ r\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}_2\\ \boldsymbol{C}_1 & \boldsymbol{0}\end{bmatrix}\right\}$$

$$\min_{X_1,X_2} r[\mathcal{P}(\boldsymbol{X}_1,\boldsymbol{X}_2)]$$
$$=r[\boldsymbol{A}\ \ \boldsymbol{B}_1\ \ \boldsymbol{B}_2]+r\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{C}_1\\ \boldsymbol{C}_2\end{bmatrix}+\max\left\{r\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}_1\\ \boldsymbol{C}_2 & \boldsymbol{0}\end{bmatrix}\right.$$
$$\left.-r\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}_1 & \boldsymbol{B}_2\\ \boldsymbol{C}_2 & \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}-r\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}_1\\ \boldsymbol{C}_1 & \boldsymbol{0}\\ \boldsymbol{C}_2 & \boldsymbol{0}\end{bmatrix},\ r\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}_2\\ \boldsymbol{C}_1 & \boldsymbol{0}\end{bmatrix}-r\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}_1 & \boldsymbol{B}_2\\ \boldsymbol{C}_1 & \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}-r\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}_2\\ \boldsymbol{C}_1 & \boldsymbol{0}\\ \boldsymbol{C}_2 & \boldsymbol{0}\end{bmatrix}\right\}$$

The following lemma is well-known (see, e.g., Ref. [18]).

Lemma 3 Let $\boldsymbol{A},\boldsymbol{B}$ and $\boldsymbol{C}$ be given matrices with appropriate sizes. Then matrix equation $\boldsymbol{AXB}=\boldsymbol{C}$ is consistent, if and only if one of the following conditions holds:

(a) $\boldsymbol{AA}^{+}\boldsymbol{CB}^{+}\boldsymbol{B}=\boldsymbol{C}$.

(b) $\boldsymbol{AA}^{+}\boldsymbol{C}=\boldsymbol{C}$ and $\boldsymbol{CB}^{+}\boldsymbol{B}=\boldsymbol{C}$.

(c) $R(\boldsymbol{C})\subseteq R(\boldsymbol{A})$ and $R(\boldsymbol{C}^*)\subseteq R(\boldsymbol{A}^*)$.

In this case, the general solution of the matrix equation is $\boldsymbol{X}=\boldsymbol{A}^{+}\boldsymbol{CB}^{+}+\boldsymbol{F}_{\boldsymbol{A}}\boldsymbol{U}+\boldsymbol{VE}_{\boldsymbol{B}}$, where $\boldsymbol{U}$ and $\boldsymbol{V}$ with appropriate sizes are arbitrary.

2 Main Results

The optimization problem (2) is indeed the extremal ranks problem upon the nonlinear matrix expression under equality constraint. Tian[5,19] developed an algebraic linearization method to solve the rank and inertia problems of some nonlinear (Hermitian) matrix expressions. The key technique of this method derives from the following block matrix operation

$$\begin{bmatrix}\boldsymbol{I}_m & -\boldsymbol{X}\\ \boldsymbol{0} & \boldsymbol{I}_k\end{bmatrix}\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{XP}\\ \boldsymbol{PY} & \boldsymbol{P}\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_n & \boldsymbol{0}\\ -\boldsymbol{Y} & \boldsymbol{I}_l\end{bmatrix}=\begin{bmatrix}\boldsymbol{Q}-\boldsymbol{XPY} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{P}\end{bmatrix},$$

where $Q \in \mathbf{C}^{m\times n}$ and $P \in \mathbf{C}^{k\times l}$ are given. This equality implies that

$$r(Q - XPY) = r\begin{bmatrix} Q & XP \\ PY & P \end{bmatrix} - r(P) \quad (3)$$

On the other hand, for the given matrices A and B as in (1) and (2), applying Lemma 3 to matrix equation $AXA = B$, we know that this equation is consistent if and only if

$$AA^+BA^+A = B\text{, or } AA^+B = B \text{ and } BA^+A = B \quad (4)$$

In that case, the general solution of which is

$$X = A^+BA^+ + F_AU + VE_A \quad (5)$$

with $U, V \in \mathbf{C}^{n\times n}$.

On the basis of the above analysis, we obtain the following theorem.

Theorem 1 Let matrices A and B be given as in (1) and (2), then

$$\max_{AXA=B} (B - XAX) = \max\left\{ r\begin{bmatrix} A & B \\ B & ABA \end{bmatrix} + r\begin{bmatrix} BA \\ B \end{bmatrix} + r[AB, B] - r[ABA, B] - r\begin{bmatrix} ABA \\ B \end{bmatrix} - r(A) \right.$$
$$\left. \cdot r\begin{bmatrix} BA \\ B \end{bmatrix} + r[AB, B] + r(B) - r(AB) - r(BA) - r(A) \right\} \quad (6)$$

Proof Substituting (5) into $B - XAX$, then it follows from (3) and (4) that

$$\begin{aligned} &r(B - XAX) \\ &= r[B - (A^+BA^+ + F_AU + VE_A)A(A^+BA^+ + F_AU + VE_A)] \\ &= r\begin{bmatrix} B & A^+BA^+A + F_AUA \\ AA^+BA^+ + AVE_A & A \end{bmatrix} - r(A) \\ &= r\begin{bmatrix} B & A^+B + F_AUA \\ BA^+ + AVE_A & A \end{bmatrix} - r(A) \\ &= r\left(\begin{bmatrix} B & A^+B \\ BA^+ & A \end{bmatrix} + \begin{bmatrix} F_A \\ 0 \end{bmatrix} U[0, A] + \begin{bmatrix} 0 \\ A \end{bmatrix} V[E_A, 0] \right) - r(A) \\ &:= r[\mathcal{L}(U,V)] - r(A) \end{aligned}$$

This identity implies that the rank optimization problem (2) is equivalent to find the minimal rank of the linear matrix-valued function $\mathcal{L}(U,V)$, namely,

$$\min_{AXA=B} r(B - XAX) = \min_{U,V} r[\mathcal{L}(U,V)] - r(A) \quad (7)$$

Applying Lemma 2 to $\mathcal{L}(U,V)$ yields

$$\min_{U,V} r[\mathcal{L}(U,V)] = r\begin{bmatrix} B & A^+B & F_A & 0 \\ BA^+ & A & 0 & A \end{bmatrix} + r\begin{bmatrix} B & A^+B \\ BA^+ & A \\ 0 & A \\ E_A & 0 \end{bmatrix} + \max\{s_1, s_2\} \quad (8)$$

where

$$s_1 = r\begin{bmatrix} B & A^+B & F_A \\ BA^+ & A & 0 \\ E_A & 0 & 0 \end{bmatrix} - r\begin{bmatrix} B & A^+B & F_A & 0 \\ BA^+ & A & 0 & A \\ E_A & 0 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} B & A^+B & F_A \\ BA^+ & A & 0 \\ 0 & A & 0 \\ E_A & 0 & 0 \end{bmatrix}$$

$$s_2 = r\begin{bmatrix} B & A^+B & 0 \\ BA^+ & A & A \\ 0 & A & 0 \end{bmatrix} - r\begin{bmatrix} B & A^+B & F_A & 0 \\ BA^+ & A & 0 & A \\ 0 & A & 0 & 0 \end{bmatrix} - r\begin{bmatrix} B & A^+B & 0 \\ BA^+ & A & A \\ 0 & A & 0 \\ E_A & 0 & 0 \end{bmatrix}$$

Now, we simplify (8) by the elementary block matrix operations. Noting that the consistency conditions in (4), it is clear that

$$A^+B = A^+BA^+A \text{ and } BA^+ = AA^+BA^+$$

i.e.,

$$A^+BF_A = 0 \text{ and } E_ABA^+ = 0.$$

These equalities as well as (4) will be frequently used in the sequel. Thus, by the elementary transformations of block matrices, Lemma 1 (a)-(c) follow that

$$r\begin{bmatrix} B & A^+B & F_A & 0 \\ BA^+ & A & 0 & A \end{bmatrix} = r(A) + r(F_A) + r[A^+AB, A^+B] \quad (9)$$

$$r\begin{bmatrix} B & A^+B \\ BA^+ & A \\ 0 & A \\ E_A & 0 \end{bmatrix} = r(A) + r(E_A) + r\begin{bmatrix} BAA^+ \\ BA^+ \end{bmatrix} = r(A) + r(E_A) + r\begin{bmatrix} BA \\ B \end{bmatrix} \quad (10)$$

$$r\begin{bmatrix} B & A^+B & F_A \\ BA^+ & A & 0 \\ E_A & 0 & 0 \end{bmatrix} = r(E_A) + r(F_A) + r\begin{bmatrix} A^+ABAA^+ & A^+B \\ BA^+ & A \end{bmatrix}$$
$$= r(E_A) + r(F_A) + r(A^+ABAA^+ - A^+BA^+BA^+) + r(A) \quad (11)$$

Furthermore, since

$$r(A^+ABAA^+ - A^+BA^+BA^+) = r(ABA - BA^+B),$$

it then follows from (4) and Lemma 1(d) that

$$r(A^+ABAA^+ - A^+BA^+BA^+) = r\begin{bmatrix} A & B \\ B & ABA \end{bmatrix} - r(A) \quad (12)$$

Putting (12) into (11) deduces that

$$r\begin{bmatrix} B & A^+B & F_A \\ BA^+ & A & 0 \\ E_A & 0 & 0 \end{bmatrix} = r(E_A) + r(F_A) + r\begin{bmatrix} A & B \\ B & ABA \end{bmatrix} \quad (13)$$

Similarly, we have

$$r\begin{bmatrix} B & A^+B & F_A & 0 \\ BA^+ & A & 0 & A \\ E_A & 0 & 0 & 0 \end{bmatrix}$$

$$= r(E_A) + r(F_A) + r(A) + r[A^+ABAA^+, A^+B]$$

$$= r(E_A) + r(F_A) + r(A) + r[ABA, B] \quad (14)$$

$$r\begin{bmatrix} B & A^+B & F_A \\ BA^+ & A & 0 \\ 0 & A & 0 \\ E_A & 0 & 0 \end{bmatrix} = r(E_A) + r(F_A) + r(A) + r\begin{bmatrix} A^+ABAA^+ \\ BA^+ \end{bmatrix}$$

$$= r(E_A) + r(F_A) + r(A) + r\begin{bmatrix} ABA \\ B \end{bmatrix} \quad (15)$$

$$r\begin{bmatrix} B & A^+B & 0 \\ BA^+ & A & A \\ 0 & A & 0 \end{bmatrix} = 2r(A) + r(B) \quad (16)$$

$$r\begin{bmatrix} B & A^+B & F_A & 0 \\ BA^+ & A & 0 & A \\ 0 & A & 0 & 0 \end{bmatrix} = r(F_A) + 2r(A) + r(A^+AB)$$

$$= r(F_A) + 2r(A) + r(AB) \quad (17)$$

$$r\begin{bmatrix} B & A^+B & 0 \\ BA^+ & A & A \\ 0 & A & 0 \\ E_A & 0 & 0 \end{bmatrix} = r(E_A) + 2r(A) + r(BAA^+)$$

$$= r(E_A) + 2r(A) + r(BA) \quad (18)$$

Hence, substituting equalities (9), (10), (13)-(18) into (8), from (7) we know that (6) holds.

From Theorem 1, we obtain another main result of this paper on the Yang-Baxter matrix equation (1) and rank optimization problem (2), which reveals the relation between the matrices A and B included in Yang-Baxter matrix equation (1).

Theorem 2 Let A and B be given matrices satisfying (4). Then the Yang-Baxter matrix equation $AXA = B = XAX$ is consistent, if and only if

(a) $r\begin{bmatrix} A & B \\ B & ABA \end{bmatrix} = r(A)$, $r[ABA, B] = r[AB, B]$,

$r\begin{bmatrix} ABA \\ B \end{bmatrix} = r\begin{bmatrix} BA \\ B \end{bmatrix}$, $r(A) = r(B)$.

(b) $r\begin{bmatrix} BA \\ B \end{bmatrix} + r[AB, B] + r(B) = r(AB) + r(BA) + r(A)$.

Proof According to (6), Yang-Baxter matrix equation (1) is consistent, if and only if

$$r\begin{bmatrix} A & B \\ B & ABA \end{bmatrix} + r\begin{bmatrix} BA \\ B \end{bmatrix} + r[AB, B] - r[ABA, B]$$

$$- r\begin{bmatrix} ABA \\ B \end{bmatrix} - r(A) = 0,$$

and

$$r\begin{bmatrix} BA \\ B \end{bmatrix} + r[AB, B] + r(B) - r(AB) - r(BA) - r(A) = 0.$$

Furthermore, the first equality can be rewritten as

$$\left(r\begin{bmatrix} A & B \\ B & ABA \end{bmatrix} - r(A)\right) + \left(r\begin{bmatrix} BA \\ B \end{bmatrix} - r\begin{bmatrix} ABA \\ B \end{bmatrix}\right)$$

$$+ \left(r[AB, B] - r[ABA, B]\right) = 0 \quad (19)$$

and notice that the left items of (19) are nonnegative, which implies that (a) and (b) hold true.

From this theorem, one can obtain the following conclusion.

Corollary 1 Suppose that the matrix A in Theorem 1 is idempotent, i.e., $A^2 = A$, then

$$\min_{AXA=B} r(B - XAX) = r\begin{bmatrix} A & B \\ B & B \end{bmatrix} - r(A) = r(B - BA^+B) \quad (20)$$

Moreover, matrix equations $AXA = B = XAX$ is consistent, if and only if $B = BA^+B$. In this case, the general solution of (1) can be expressed as

$$X = AWA + AU_1(I_n - A) + (I_n - A)U_2A + (I_n - A)U_3(I_n - A) \quad (21)$$

where idempotent W commuting with A is arbitrary, U_3 is an arbitrary matrix with order n, and U_1, U_2, satisfy

$$[AWAU_1(I_n - A)] + (I_n - A)U_2[AWA + AU_1(I_n - A)] = 0.$$

In particular, let $B = AXA$, then the general solution commuting with A of the Yang-Baxter matrix equation (1) is

$$X = AWA + (I_n - A)Z(I_n - A) \quad (22)$$

where any $Z \in \mathbf{C}^{n\times n}$.

Proof Combining (4) and the above assumptions, we get $\boldsymbol{ABA}=\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{B}$ and $r(\boldsymbol{B})\leqslant r(\boldsymbol{A})$, which implies from (6) that (20) is true. It is easy to verify that (21) and (22) are the required general solutions.

Remark 1 Let $\boldsymbol{A}=\boldsymbol{S}\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{S}^{-1}$ be the Jordan form of $\boldsymbol{A}$ with rank r, then we can verify that (21) is indeed the same as (6) in Ref. [17], while (22) is the special case of (21) when the matrices $\boldsymbol{C}$ and $\boldsymbol{D}$ are null matrices.

Corollary 2 Let $\boldsymbol{B}=\boldsymbol{A}$ in Theorem 1, then

$$\min_{\boldsymbol{AXA}=\boldsymbol{A}} r(\boldsymbol{A}-\boldsymbol{XAX})=\max\{r(\boldsymbol{A}-\boldsymbol{A}^3),2[r(\boldsymbol{A})-r(\boldsymbol{A}^2)]\}$$

That is to say, the Yang-Baxter matrix equation (1) satisfies $\boldsymbol{AXA}=\boldsymbol{XAX}=\boldsymbol{A}$ if and only if $\boldsymbol{A}^3=\boldsymbol{A}$.

Proof From (6) we know that the result holds true.

3 Conclusion

In this paper, we have studied the Yang-Baxter matrix equation (1). By using the rank of matrices, this equation is converted into solving an associated minimal rank optimization problem, i.e., (2). After that, some analytical formulas with respect to the known matrices were given. As applications, the general solutions of (1) were derived under some assumptions. These rank formulas may be useful for finding all the solutions and analyzing the structures of the solutions of the Yang-Baxter matrix equation. How to express the general solution of (1) is an interesting but challenging problem.

References

[1] Marsaglia G, Styan G. Equalities and inequalities for ranks of matrices [J]. *Linear Multilinear Algebra*, 1974, **2**: 269-292.

[2] Meyer Jr C. Generalized inverses and ranks of block matrices [J]. *SIAM J Appl Math*, 1973, **25**:597-602.

[3] Tian Y G. Upper and lower bounds for ranks of matrix expressions using generalized inverses [J]. *Linear Algebra Appl*, 2002, **355**(1-3): 187-214.

[4] Tian Y G. On properties of BLUEs under general linear regression models [J]. *J Stat Plan Inf*, 2013, **143**(4): 771-782.

[5] Tian Y G. Extremal ranks of a quadratic matrix expression with applications [J]. *Linear Multilinear Algebra*, 2011, **59**(6): 627-644.

[6] Xiong Z P, Qin Y Y. On the inverse of a special Schur complement [J]. *Appl Math Comput*, 2012, **218**(14): 7679-7684.

[7] Wang Q W, Yu S W, Zhang Q. The real solutions to a system of quaternion matrix equations with applications [J]. *Comm Algebra*, 2009, **37**(1): 2060-2079.

[8] Wang Q W, Zhang X, He Z H. On the Hermitian structures of the solution to a pair of matrix equations [J]. *Linear Multilinear Algebra*, 2012, **61**(1): 73-90.

[9] Rashedi S, Ebadi G, Biswas A. The maximal and minimal ranks of a quaternion matrix expression with applications [J]. *J Egypt Math Soc*, 2013, **21**(3): 175-183.

[10] Drinfeld V. On some unsolved problems in quantum group theory [J]. *Lecture Notes in Mathematics*, 1992, **1510**: 1-8.

[11] Etingof P, Schedler T, Soloviev A. Set-theoretical solutions of the quantum Yang-Baxter equation [J]. *Duke Math J*, 1999, **100**: 169-178.

[12] Gateva-Ivanova T, Bergh M. Semi-groups of I-type [J]. *J Algebra*, 1998, **206**: 97-112.

[13] Felix F. *Nonlinear Equations*, *Quantum Groups And Duality Theorems*: *A Primer on the Yang-Baxter Equation* [D]. Buffalo: The State University of New York at Buffalo, 2001.

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

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

[16] Ding J, Rhee N. A nontrivial solution to a stochastic matrix equation [J]. *East Asian J Appl Math*, 2012, **2**: 277-284.

[17] Cibotarica A, Ding J, Kolibal J, *et al*. Solutions of the Yang-Baxter matrix equation for an idempotent [J]. *Numer Algebra Control Opt*, 2013, **3**(2): 347-352.

[18] Ben-Israel A, Greville T. *Generalized Inverses*: *Theory and Applications* [M]. New York: R.E. Krieger Publishing Company, 1980.

[19] Tian Y G. Solving optimization problems on ranks and inertias of some constrained nonlinear matrix functions via an algebraic linearization method [J]. *Nonlinear Anal*, 2012, **75**(2): 717-734.

□