Journal of Applied Mathematics

Advances in Matrices, Finite and Infinite, with Applications

Guest Editors: P. N. Shivakumar, Panayiotis Psarrakos, K. C. Sivakumar, and Yang Zhang

Advances in Matrices, Finite and Infinite, with Applications

Journal of Applied Mathematics

Advances in Matrices, Finite and Infinite, with Applications

Guest Editors: P. N. Shivakumar, Panayiotis Psarrakos, K. C. Sivakumar, and Yang Zhang

Editorial Board

Contents

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 156136, 3 pages
http://dx.doi.org/10.1155/2013/156136

Editorial

Advances in Matrices, Finite and Infinite, with Applications

P. N. Shivakumar,[1] Panayiotis Psarrakos,[2] K. C. Sivakumar,[3] and Yang Zhang[1]

[1] *Department of Mathematics, University of Manitoba, Winnipeg, MB, Canada R3T 2N2*
[2] *Department of Mathematics, School of Applied Mathematical and Physical Sciences, National Technical Univeristy of Athens, Zografou Campus, 15780 Athens, Greece*
[3] *Department of Mathematics, Indian Institute of Technology Madras, Chennai 600 036, India*

Correspondence should be addressed to P. N. Shivakumar; shivaku@cc.umanitoba.ca

Received 13 June 2013; Accepted 13 June 2013

In the mathematical modeling of many problems, including, but not limited to, physical sciences, biological sciences, engineering, image processing, computer graphics, medical imaging, and even social sciences lately, solution methods most frequently involve matrix equations. In conjunction with powerful computing machines, complex numerical calculations employing matrix methods could be performed sometimes even in real time. It is well known that infinite matrices arise more naturally than finite matrices and have a colorful history in development from sequences, series and quadratic forms. Modern viewpoint considers infinite matrices more as operators defined between certain specific infinite-dimensional normed spaces, Banach spaces, or Hilbert spaces. Giant and rapid advancements in the theory and applications of finite matrices have been made over the past several decades. These developments have been well documented in many books, monographs, and research journals. The present topics of research include diagonally dominant matrices, their many extensions, inverse of matrices, their recursive computation, singular matrices, generalized inverses, inverse positive matrices with specific emphasis on their applications to economics like the Leontief input-output models, and use of finite difference and finite element methods in the solution of partial differential equations, perturbation methods, and eigenvalue problems of interest and importance to numerical analysts, statisticians, physical scientists, and engineers, to name a few.

The aim of this special issue is to announce certain recent advancements in matrices, finite and infinite, and their applications. For a review of infinite matrices and applications, see P. N. Shivakumar, and K. C. Sivakumar (Linear Algebra Appl., 2009). Specifically, an example of an application can be found in the classical problem "shape of a drum," in P. N. Shivakumar, W. Yan, and Y. Zhang (WSES transactions in Mathematics, 2011). The focal themes of this special issue are inverse positive matrices including M-matrices, applications of operator theory, matrix perturbation theory, and pseudospectra, matrix functions and generalized eigenvalue problems and inverse problems including scattering and matrices over quaternions. In the following, we present a brief overview of a few of the central topics of this special issue.

For M-matrices, let us start by mentioning the most cited books by R. Berman and R. J. Plemmons (SIAM, 1994), and by R. A. Horn and C. R. Johnson (Cambridge, 1994), as excellent sources. One of the most important properties of invertible M-matrices is that their inverses are nonnegative. There are several works that have considered generalizations of some of the important properties of M-matrices. We only point out a few of them here. The article by D. Mishra and K. C. Sivakumar (Oper. Matrices, 2012), considers generalizations of inverse nonnegativity to the Moore-Penrose inverse, while more general classes of matrices are the objects of discussion in D. Mishra and K. C. Sivakumar (Linear Algebra Appl., 2012), and A. N. Sushama, K. Premakumari and K. C. Sivakumar (Electron. J. Linear Algebra 2012).

A brief description of the papers in the issue is as follows.

In the work of Z. Zhang, the problem of solving certain optimization problems on Hermitian matrix functions is studied. Specifically, the author considers the extremal inertias and ranks of a given function $f(X,Y) : \mathbb{C}^{m\times n} \to \mathbb{C}^{m\times n}$ (which is defined in terms of matrices A, B, C, and D

in $\mathbb{C}^{m\times n}$), subject to X, Y satisfying certain matrix equations. As applications, the author presents necessary and sufficient conditions for f to be positive definite, positive semidefinite, and so forth. As another application, the author presents a characterization for $f(X, Y) = 0$, again, subject to X, Y satisfying certain matrix equations.

The task of determining parametrized splitting preconditioners for certain generalized saddle point problems is undertaken by W.-H. Luo and T.-Z. Huang in their work. The well-known and very widely applicable Sherman-Morrison-Woodbury formula for the sum of matrices is used in determining a preconditioner for linear equations where the coefficient matrix may be singular and nonsymmetric. Illustrations of the procedure are provided for Stokes and Oseen problems.

The work reported by A. González, A. Suarez and D. Garcia deals with the problem of estimating the Frobenius condition of the matrix AN, given a matrix $A \in \mathbb{R}^{n\times n}$, where N satisfies the condition that $N = \operatorname{argmin}_{X\in S}\|I - AX\|_F$. Here, S is a certain subspace of $\mathbb{R}^n$ and $\|\cdot\|_F$ denotes the Frobenius norm. The authors derive certain spectral and geometrical properties of the preconditioner N as above.

Let B, C, D and E be (not necessarily square) matrices with quaternion entries so that the matrix $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ is square. Y. Lin and Q.-W. Wang present necessary and sufficient conditions so that the top left block A of M exists satisfying the conditions that M is nonsingular and the matrix E is the top left block of M^{-1}. A general expression for such an A, when it exists, is obtained. A numerical illustration is presented.

For a complex hermitian A of order n with eigenvalues $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$, the celebrated Kantorovich inequality is $1 \le x^*Axx^*A^{-1}x \le (\lambda_n+\lambda_1)^2/4\lambda_1\lambda_n$, for all unit vectors $x \in \mathbb{C}^n$. Equivalently $0 \le x^*Axx^*A^{-1}x-1 \le (\lambda_n-\lambda_1)^2/4\lambda_1\lambda_n$, for all unit vectors $x \in \mathbb{C}^n$. F. Chen, in his work on refinements of Kantorovich inequality for hermitian matrices, obtains some refinements of the second inequality, where the proofs use only elementary ideas (http://www.math.ntua.gr/~ppsarr/).

It is pertinent to point out that, among other things, an interesting generalization of M-matrices is obtained by S. Jose and K. C. Sivakumar. The authors consider the problem of nonnegativity of the Moore-Penrose inverse of a perturbation of the form $A - XGY^T$ given that $A^\dagger \ge 0$. Using a generalized version of the Sherman-Morrison-Woodbury formula, conditions for $(A - XGY^T)^\dagger$ to be nonnegative are derived. Applications of the results are presented briefly. Iterative versions of the results are also studied.

It is well known that the concept of quaternions was introduced by Hamilton in 1843 and has played an important role in contemporary mathematics such as noncommutative algebra, analysis, and topology. Nowadays, quaternion matrices are widely and heavily used in computer science, quantum physics, altitude control, robotics, signal and colour image processing, and so on. Many questions can be reduced to quaternion matrices and solving quaternion matrix equations. This special issue has provided an excellent opportunity for researchers to report their recent results.

Let $\mathbb{H}^{n\times n}$ denote the space of all $n \times n$ matrices whose entries are quaternions. Let $Q \in \mathbb{H}^{n\times n}$ be Hermitian and self-invertible. $X \in \mathbb{H}^{n\times n}$ is called reflexive (relative to Q) if $X = QXQ$. The authors F.-L. Li, X.-Y. Hu, and L. Zhang present an efficient algorithm for finding the reflexive solution of the quaternion matrix equation: $AXB + CX^*D = F$. They also show that, given an appropriate initial matrix, their procedure yields the least reflexive solution with respect to the Frobenius norm.

In the work on the ranks of a constrained hermitian matrix expression, S. W. Yu gives formulas for the maximal and minimal ranks of the quaternion Hermitian matrix expression $C_4 - A_4XA_4^*$, where X is a Hermitian solution to quaternion matrix equations $A_1X = C_1$, $XB_1 = C_2$, and $A_3XA_3^* = C_3$, and also applies this result to some special system of quaternion matrix equations.

Y. Yao reports his findings on the optimization on ranks and inertias of a quadratic hermitian matrix function. He presents one solution for optimization problems on the ranks and inertias of the quadratic Hermitian matrix function $Q - XPX^*$ subject to a consistent system of matrix equations $AX = C$ and $XB = D$. As applications, he derives necessary and sufficient conditions for the solvability to the systems of matrix equations and matrix inequalities $AX = C$, $XB = D$, and $XPX^* = (>, <, \ge, \le)Q$.

Let R be a nontrivial real symmetric involution matrix. A complex matrix A is called R-conjugate if $\overline{A} = RAR$. In the work on the hermitian R-conjugate solution of a system of matrix equations, C.-Z. Dong, Q.-W. Wang, and Y.-P. Zhang present necessary and sufficient conditions for the existence of the hermitian R-conjugate solution to the system of complex matrix equations $AX = C$ and $XB = D$. An expression of the Hermitian R-conjugate solution is also presented.

A perturbation analysis for the matrix equation $X - \sum_{i=1}^{m} A_i^*XA_i + \sum_{j=1}^{n} B_j^*XB_j = I$ is undertaken by X.-F. Duan and Q.-W. Wang. They give a precise perturbation bound for a positive definite solution. A numerical example is presented to illustrate the sharpness of the perturbation bound.

Linear matrix equations and their optimal approximation problems have great applications in structural design, biology, statistics, control theory, and linear optimal control, and as a consequence, they have attracted the attention of several researchers for many decades. In the work of Q.-W. Wang and J. Yu, necessary and sufficient conditions of and the expressions for orthogonal solutions, symmetric orthogonal solutions, and skew-symmetric orthogonal solutions of the system of matrix equations $AX = B$ and $XC = D$ are derived. When the solvability conditions are not satisfied, the least squares symmetric orthogonal solutions and the least squares skew-symmetric orthogonal solutions are obtained. A methodology is provided to compute the least squares symmetric orthogonal solutions, and an illustrative example is given to illustrate the proposed algorithm.

Many real-world engineering systems are too complex to be defined in precise terms, and, in many matrix equations, some or all of the system parameters are vague or imprecise. Thus, solving a fuzzy matrix equation becomes important. In the work reported by X. Guo and D. Shang, the fuzzy matrix equation $\widetilde{A} \otimes \widetilde{X} \otimes \widetilde{B} = \widetilde{C}$ in which $\widetilde{A}$, $\widetilde{B}$, and $\widetilde{C}$ are $m \times m$, $n \times n$, and $m \times n$ nonnegative LR fuzzy numbers matrices, respectively, is investigated. This fuzzy matrix system, which has a wide use in control theory and control engineering, is extended into three crisp systems of linear matrix equations according to the arithmetic operations of LR fuzzy numbers. Based on the pseudoinverse matrix, a computing model to the positive fully fuzzy linear matrix equation is provided, and the fuzzy approximate solution of original fuzzy systems is obtained by solving the crisp linear matrix systems. In addition, the existence condition of nonnegative fuzzy solution is discussed, and numerical examples are given to illustrate the proposed method. Overall, the technique of LR fuzzy numbers and their operations appears to be a strong tool in investigating fully fuzzy matrix equations.

Left and right inverse eigenpairs problem arises in a natural way, when studying spectral perturbations of matrices and relative recursive problems. F.-L. Li, X.-Y. Hu, and L. Zhang consider in their work left and right inverse eigenpairs problem for k-hermitian matrices and its optimal approximate problem. The class of k-hermitian matrices contains hermitian and perhermitian matrices and has numerous applications in engineering and statistics. Based on the special properties of k-hermitian matrices and their left and right eigenpairs, an equivalent problem is obtained. Then, combining a new inner product of matrices, necessary and sufficient conditions for the solvability of the problem are given, and its general solutions are derived. Furthermore, the optimal approximate solution, a calculation procedure to compute the unique optimal approximation, and numerical experiments are provided.

Finally, let us point out a work on algebraic coding theory. Recall that a communication system, in which a receiver can verify the authenticity of a message sent by a group of senders, is referred to as a multisender authentication code. In the work reported by X. Wang, the author constructs multi-sender authentication codes using polynomials over finite fields. Here, certain important parameters and the probabilities of deceptions of these codes are also computed.

Acknowledgments

We wish to thank all the authors for their contributions and the reviewers for their valuable comments in bringing the special issue to fruition.

P. N. Shivakumar
Panayiotis Psarrakos
K. C. Sivakumar
Yang Zhang

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 680975, 7 pages
http://dx.doi.org/10.1155/2013/680975

Research Article

On Nonnegative Moore-Penrose Inverses of Perturbed Matrices

Shani Jose and K. C. Sivakumar

Department of Mathematics, Indian Institute of Technology Madras, Chennai, Tamil Nadu 600036, India

Correspondence should be addressed to K. C. Sivakumar; kcskumar@iitm.ac.in

Received 22 April 2013; Accepted 7 May 2013

Academic Editor: Yang Zhang

Nonnegativity of the Moore-Penrose inverse of a perturbation of the form $A-XGY^T$ is considered when $A^\dagger \geq 0$. Using a generalized version of the Sherman-Morrison-Woodbury formula, conditions for $(A-XGY^T)^\dagger$ to be nonnegative are derived. Applications of the results are presented briefly. Iterative versions of the results are also studied.

1. Introduction

We consider the problem of characterizing nonnegativity of the Moore-Penrose inverse for matrix perturbations of the type $A-XGY^T$, when the Moore-Penrose inverse of A is nonnegative. Here, we say that a matrix $B=(b_{ij})$ is nonnegative and denote it by $B \geq 0$ if $b_{ij} \geq 0$, $\forall i, j$. This problem was motivated by the results in [1], where the authors consider an M-matrix A and find sufficient conditions for the perturbed matrix $(A-XY^T)$ to be an M-matrix. Let us recall that a matrix $B=(b_{ij})$ is said to Z-matrix if $b_{ij} \leq 0$ for all $i, j, i \neq j$. An M-matrix is a nonsingular Z-matrix with nonnegative inverse. The authors in [1] use the well-known Sherman-Morrison-Woodbury (SMW) formula as one of the important tools to prove their main result. The SMW formula gives an expression for the inverse of $(A-XY^T)$ in terms of the inverse of A, when it exists. When A is nonsingular, $A-XY^T$ is nonsingular if and only $I-Y^TA^{-1}X$ is nonsingular. In that case,

$$\left(A-XY^T\right)^{-1}=A^{-1}-A^{-1}X\left(I-Y^TA^{-1}X\right)^{-1}Y^TA^{-1}. \quad (1)$$

The main objective of the present work is to study certain structured perturbations $A-XY^T$ of matrices A such that the Moore-Penrose inverse of the perturbation is nonnegative whenever the Moore-Penrose inverse of A is nonnegative. Clearly, this class of matrices includes the class of matrices that have nonnegative inverses, especially M-matrices. In our approach, extensions of SMW formula for singular matrices play a crucial role. Let us mention that this problem has been studied in the literature. (See, for instance [2] for matrices and [3] for operators over Hilbert spaces). We refer the reader to the references in the latter for other recent extensions.

In this paper, first we present alternative proofs of generalizations of the SMW formula for the cases of the Moore-Penrose inverse (Theorem 5) and the group inverse (Theorem 6) in Section 3. In Section 4, we characterize the nonnegativity of $(A-XGY^T)^\dagger$. This is done in Theorem 9 and is one of the main results of the present work. As a consequence, we present a result for M-matrices which seems new. We present a couple of applications of the main result in Theorems 13 and 15. In the concluding section, we study iterative versions of the results of the second section. We prove two characterizations for $(A-XY^T)^\dagger$ to be nonnegative in Theorems 18 and 21.

Before concluding this introductory section, let us give a motivation for the work that we have undertaken here. It is a well-documented fact that M-matrices arise quite often in solving sparse systems of linear equations. An extensive theory of M-matrices has been developed relative to their role in numerical analysis involving the notion of splitting in iterative methods and discretization of differential equations, in the mathematical modeling of an economy, optimization, and Markov chains [4, 5]. Specifically, the inspiration for the present study comes from the work of [1], where the authors consider a system of linear inequalities arising out of a problem in third generation wireless communication systems. The matrix defining the inequalities there is an M-matrix. In the likelihood that the matrix of this problem is singular (due to truncation or round-off errors), the earlier method becomes inapplicable. Our endeavour is to extend the applicability of these results to more general matrices, for instance, matrices with nonnegative Moore-Penrose inverses. Finally, as mentioned earlier, since matrices with nonnegative generalized inverses include in particular M-matrices, it is apparent that our results are expected to enlarge the applicability of the methods presently available for M-matrices, even in a very general framework, including the specific problem mentioned above.

2. Preliminaries

Let $\mathbb{R}$, $\mathbb{R}^n$, and $\mathbb{R}^{m\times n}$ denote the set of all real numbers, the n-dimensional real Euclidean space, and the set of all $m \times n$ matrices over $\mathbb{R}$. For $A \in \mathbb{R}^{m\times n}$, let $R(A)$, $N(A)$, $R(A)^{\perp}$, and A^T denote the range space, the null space, the orthogonal complement of the range space, and the transpose of the matrix A, respectively. For $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$, we say that $\mathbf{x}$ is nonnegative, that is, $\mathbf{x} \geq 0$ if and only if $\mathbf{x}_i \geq 0$ for all $i = 1, 2, \ldots, n$. As mentioned earlier, for a matrix B we use $B \geq 0$ to denote that all the entries of B are nonnegative. Also, we write $A \leq B$ if $B - A \geq 0$.

Let $\rho(A)$ denote the spectral radius of the matrix A. If $\rho(A) < 1$ for $A \in \mathbb{R}^{n\times n}$, then $I - A$ is invertible. The next result gives a necessary and sufficient condition for the nonnegativity of $(I - A)^{-1}$. This will be one of the results that will be used in proving the first main result.

Lemma 1 (see [5, Lemma 2.1, Chapter 6]). *Let $A \in \mathbb{R}^{n\times n}$ be nonnegative. Then, $\rho(A) < 1$ if and only if $(I - A)^{-1}$ exists and*

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k \geq 0. \tag{2}$$

More generally, matrices having nonnegative inverses are characterized using a property called monotonicity. The notion of monotonicity was introduced by Collatz [6]. A real $n \times n$ matrix A is called *monotone* if $A\mathbf{x} \geq 0 \Rightarrow \mathbf{x} \geq 0$. It was proved by Collatz [6] that A is monotone if and only if A^{-1} exists and $A^{-1} \geq 0$.

One of the frequently used tools in studying monotone matrices is the notion of a regular splitting. We only refer the reader to the book [5] for more details on the relationship between these concepts.

The notion of monotonicity has been extended in a variety of ways to singular matrices using generalized inverses. First, let us briefly review the notion of two important generalized inverses.

For $A \in \mathbb{R}^{m\times n}$, the Moore-Penrose inverse is the unique $Z \in \mathbb{R}^{n\times m}$ satisfying the Penrose equations: $AZA = A$, $ZAZ = Z$, $(AZ)^T = AZ$, and $(ZA)^T = ZA$. The unique Moore-Penrose inverse of A is denoted by $A^{\dagger}$, and it coincides with A^{-1} when A is invertible.

The following theorem by Desoer and Whalen, which is used in the sequel, gives an equivalent definition for the Moore-Penrose inverse. Let us mention that this result was proved for operators between Hilbert spaces.

Theorem 2 (see [7]). *Let $A \in \mathbb{R}^{m\times n}$. Then $A^{\dagger} \in \mathbb{R}^{n\times m}$ is the unique matrix $X \in \mathbb{R}^{n\times m}$ satisfying*

(i) $ZAx = x$, $\forall x \in R(A^T)$,

(ii) $Zy = 0$, $\forall y \in N(A^T)$.

Now, for $A \in \mathbb{R}^{n\times n}$, any $Z \in \mathbb{R}^{n\times n}$ satisfying the equations $AZA = A$, $ZAZ = Z$, and $AZ = ZA$ is called the group inverse of A. The group inverse does not exist for every matrix. But whenever it exists, it is unique. A necessary and sufficient condition for the existence of the group inverse of A is that the index of A is 1, where the index of a matrix is the smallest positive integer k such that rank (A^{k+1}) = rank (A^k).

Some of the well-known properties of the Moore-Penrose inverse and the group inverse are given as follows: $R(A^T) = R(A^{\dagger})$, $N(A^T) = N(A^{\dagger})$, $A^{\dagger}A = P_{R(A^T)}$, and $AA^{\dagger} = P_{R(A)}$. In particular, $\mathbf{x} \in R(A^T)$ if and only if $\mathbf{x} = A^{\dagger}A\mathbf{x}$. Also, $R(A) = R(A^{\#})$, $N(A) = N(A^{\#})$, and $A^{\#}A = AA^{\#} = P_{R(A),N(A)}$. Here, for complementary subspaces L and M of $\mathbb{R}^k$, $P_{L,M}$ denotes the projection of $\mathbb{R}^k$ onto L along M. P_L denotes $P_{L,M}$ if $M = L^{\perp}$. For details, we refer the reader to the book [8].

In matrix analysis, a decomposition (splitting) of a matrix is considered in order to study the convergence of iterative schemes that are used in the solution of linear systems of algebraic equations. As mentioned earlier, regular splittings are useful in characterizing matrices with nonnegative inverses, whereas, proper splittings are used for studying singular systems of linear equations. Let us next recall this notion. For a matrix $A \in \mathbb{R}^{m\times n}$, a decomposition $A = U - V$ is called a proper splitting [9] if $R(A) = R(U)$ and $N(A) = N(U)$. It is rather well-known that a proper splitting exists for every matrix and that it can be obtained using a full-rank factorization of the matrix. For details, we refer to [10]. Certain properties of a proper splitting are collected in the next result.

Theorem 3 (see [9, Theorem 1]). *Let $A = U - V$ be a proper splitting of $A \in \mathbb{R}^{m\times n}$. Then,*

(a) $A = U(I - U^{\dagger}V)$,

(b) *$I - U^{\dagger}V$ is nonsingular and,*

(c) $A^{\dagger} = (I - U^{\dagger}V)^{-1}U^{\dagger}$.

The following result by Berman and Plemmons [9] gives a characterization for $A^{\dagger}$ to be nonnegative when A has a proper splitting. This result will be used in proving our first main result.

Theorem 4 (see [9, Corollary 4]). *Let $A = U - V$ be a proper splitting of $A \in \mathbb{R}^{m\times n}$, where $U^{\dagger} \geq 0$ and $U^{\dagger}V \geq 0$. Then $A^{\dagger} \geq 0$ if and only if $\rho(U^{\dagger}V) < 1$.*

3. Extensions of the SMW Formula for Generalized Inverses

The primary objects of consideration in this paper are generalized inverses of perturbations of certain types of a matrix

A. Naturally, extensions of the SMW formula for generalized inverses are relevant in the proofs. In what follows, we present two generalizations of the SMW formula for matrices. We would like to emphasize that our proofs also carry over to infinite dimensional spaces (the proof of the first result is verbatim and the proof of the second result with slight modifications applicable to range spaces instead of ranks of the operators concerned). However, we are confining our attention to the case of matrices. Let us also add that these results have been proved in [3] for operators over infinite dimensional spaces. We have chosen to include these results here as our proofs are different from the ones in [3] and that our intention is to provide a self-contained treatment.

Theorem 5 (see [3, Theorem 2.1]). *Let $A \in \mathbb{R}^{m\times n}$, $X \in \mathbb{R}^{m\times k}$, and $Y \in \mathbb{R}^{n\times k}$ be such that*

$$R(X) \subseteq R(A), \qquad R(Y) \subseteq R(A^T). \tag{3}$$

Let $G \in \mathbb{R}^{k\times k}$, and $\Omega = A - XGY^T$. If

$$\begin{gathered} R(X^T) \subseteq R\left(\left(G^{\dagger} - Y^T A^{\dagger} X\right)^T\right), \\ R(Y^T) \subseteq R\left(G^{\dagger} - Y^T A^{\dagger} X\right), \qquad R(X^T) \subseteq R(G), \\ R(Y^T) \subseteq R(G^T) \end{gathered} \tag{4}$$

then

$$\Omega^{\dagger} = A^{\dagger} + A^{\dagger} X\left(G^{\dagger} - Y^T A^{\dagger} X\right)^{\dagger} Y^T A^{\dagger}. \tag{5}$$

Proof. Set $Z = A^{\dagger} + A^{\dagger} X H^{\dagger} Y^T A^{\dagger}$, where $H = G^{\dagger} - Y^T A^{\dagger} X$. From the conditions (3) and (4), it follows that $AA^{\dagger}X = X$, $A^{\dagger}AY = Y$, $H^{\dagger}HX^T = X^T$, $HH^{\dagger}Y^T = Y^T$, $GG^{\dagger}X^T = X^T$, and $G^{\dagger}GY^T = Y^T$.

Now,

$$\begin{aligned} A^{\dagger}XH^{\dagger}Y^T &- A^{\dagger}XH^{\dagger}Y^T A^{\dagger}XGY^T \\ &= A^{\dagger}XH^{\dagger}\left(G^{\dagger}GY^T - Y^T A^{\dagger}XGY^T\right) \\ &= A^{\dagger}XH^{\dagger}HGY^T = A^{\dagger}XGY^T. \end{aligned} \tag{6}$$

Thus, $Z\Omega = A^{\dagger}A$. Since $R(\Omega^T) \subseteq R(A^T)$, it follows that $Z\Omega(\mathbf{x}) = x$, $\forall \mathbf{x} \in R(\Omega^T)$.

Let $\mathbf{y} \in N(\Omega^T)$. Then, $A^T\mathbf{y} - YG^TX^T\mathbf{y} = 0$ so that $X^T {A^{\dagger}}^T (A^T - YG^TX^T)\mathbf{y} = 0$. Substituting $X^T = GG^{\dagger}X^T$ and simplifying it, we get $H^TG^TX^T\mathbf{y} = 0$. Also, $A^T\mathbf{y} = YG^TX^T\mathbf{y} = YHH^{\dagger}G^TX^T\mathbf{y} = 0$ and so $A^{\dagger}\mathbf{y} = 0$. Thus, $Z\mathbf{y} = 0$ for $\mathbf{y} \in N(\Omega^T)$. Hence, by Theorem 2, $Z = \Omega^{\dagger}$. □

The result for the group inverse follows.

Theorem 6. *Let $A \in \mathbb{R}^{n\times n}$ be such that $A^{\#}$ exists. Let X, $Y \in \mathbb{R}^{n\times k}$ and G be nonsingular. Assume that $R(X) \subseteq R(A)$, $R(Y) \subseteq R(A^T)$ and $G^{-1} - Y^T A^{\#} X$ is nonsingular. Suppose that* rank $(A - XGY^T) =$ rank (A). *Then, $(A - XGY^T)^{\#}$ exists and the following formula holds:*

$$\left(A - XGY^T\right)^{\#} = A^{\#} + A^{\#}X\left(G^{-1} - Y^T A^{\#} X\right)^{-1} Y^T A^{\#}. \tag{7}$$

Conversely, if $(A-XGY^T)^{\#}$ exists, then the formula above holds, and we have rank $(A - XGY^T) =$ *rank* (A).

Proof. Since $A^{\#}$ exists, $R(A)$ and $N(A)$ are complementary subspaces of $\mathbb{R}^{n\times n}$.

Suppose that rank $(A - XGY^T) =$ rank (A). As $R(X) \subseteq R(A)$, it follows that $R(A - XGY^T) \subseteq R(A)$. Thus $R(A - XGY^T) = R(A)$. By the rank-nullity theorem, the nullity of both $A - XGY^T$ and A are the same. Again, since $R(Y) \subseteq R(A^T)$, it follows that $N(A - XGY^T) = N(A)$. Thus, $R(A - XGY^T)$ and $N(A - XGY^T)$ are complementary subspaces. This guarantees the existence of the group inverse of $A - XGY^T$.

Conversely, suppose that $(A - XGY^T)^{\#}$ exists. It can be verified by direct computation that $Z = A^{\#} + A^{\#}X(G^{-1} - Y^TA^{\#}X)^{-1}Y^TA^{\#}$ is the group inverse of $A - XGY^T$. Also, we have $(A-XGY^T)(A-XGY^T)^{\#} = AA^{\#}$, so that $R(A-XGY^T) = R(A)$, and hence the rank $(A - XGY^T) =$ rank (A). □

We conclude this section with a fairly old result [2] as a consequence of Theorem 5.

Theorem 7 (see [2, Theorem 15]). *Let $A \in \mathbb{R}^{m\times n}$ of* rank r, *$X \in \mathbb{R}^{m\times r}$ and $Y \in \mathbb{R}^{n\times r}$. Let G be an $r\times r$ nonsingular matrix. Assume that $R(X) \subseteq R(A)$, $R(Y) \subseteq R(A^T)$, and $G^{-1} - Y^T A^{\dagger} X$ is nonsingular. Let $\Omega = A - XGY^T$. Then*

$$\left(A - XGY^T\right)^{\dagger} = A^{\dagger} + A^{\dagger}X\left(G^{-1} - Y^T A^{\dagger} X\right)^{-1} Y^T A^{\dagger}. \tag{8}$$

4. Nonnegativity of $(A - XGY^T)^{\dagger}$

In this section, we consider perturbations of the form $A - XGY^T$ and derive characterizations for $(A - XGY^T)^{\dagger}$ to be nonnegative when $A^{\dagger} \geq 0$, $X \geq 0$, $Y \geq 0$ and $G \geq 0$. In order to motivate the first main result of this paper, let us recall the following well known characterization of M-matrices [5].

Theorem 8. *Let A be a Z-matrix with the representation $A = sI - B$, where $B \geq 0$ and $s \geq 0$. Then the following statements are equivalent:*

(a) *A^{-1} exists and $A^{-1} \geq 0$.*

(b) *There exists $x > 0$ such that $Ax > 0$.*

(c) *$\rho(B) < s$.*

Let us prove the first result of this article. This extends Theorem 8 to singular matrices. We will be interested in extensions of conditions (a) and (c) only.

Theorem 9. *Let $A = U - V$ be a proper splitting of $A \in \mathbb{R}^{m\times n}$ with $U^{\dagger} \geq 0$, $U^{\dagger}V \geq 0$ and $\rho(U^{\dagger}V) < 1$. Let $G \in \mathbb{R}^{k\times k}$ be nonsingular and nonnegative, $X \in \mathbb{R}^{m\times k}$, and $Y \in \mathbb{R}^{n\times k}$ be* *nonnegative such that $R(X) \subseteq R(A)$, $R(Y) \subseteq R(A^T)$ and $G^{-1} - Y^T A^\dagger X$ is nonsingular. Let $\Omega = A - XGY^T$. Then, the following are equivalent*

(a) $\Omega^\dagger \geq 0$.

(b) $(G^{-1} - Y^T A^\dagger X)^{-1} \geq 0$.

(c) $\rho(U^\dagger W) < 1$ *where* $W = V + XGY^T$.

Proof. First, we observe that since $R(X) \subseteq R(A)$, $R(Y) \subseteq R(A^T)$, and $G^{-1} - Y^T A^\dagger X$ is nonsingular, by Theorem 7, we have

$$\Omega^\dagger = A^\dagger + A^\dagger X\left(G^{-1} - Y^T A^\dagger X\right)^{-1} Y^T A^\dagger. \tag{9}$$

We thus have $\Omega\Omega^\dagger = AA^\dagger$ and $\Omega^\dagger\Omega = A^\dagger A$. Therefore,

$$R(\Omega) = R(A), \qquad N(\Omega) = N(A). \tag{10}$$

Note that the first statement also implies that $A^\dagger \geq 0$, by Theorem 4.

(a) $\Rightarrow$ (b): By taking $E = A$ and $F = XGY^T$, we get $\Omega = E - F$ as a proper splitting for Ω such that $E^\dagger = A^\dagger \geq 0$ and $E^\dagger F = A^\dagger XGY^T \geq 0$ (since $XGY^T \geq 0$). Since $\Omega^\dagger \geq 0$, by Theorem 4, we have $\rho(A^\dagger XGY^T) = \rho(E^\dagger F) < 1$. This implies that $\rho(Y^T A^\dagger XG) < 1$. We also have $Y^T A^\dagger XG \geq 0$. Thus, by Lemma 1, $(I - Y^T A^\dagger XG)^{-1}$ exists and is nonnegative. But, we have $I - Y^T A^\dagger XG = (G^{-1} - Y^T A^\dagger X)G$. Now, $0 \leq (I - Y^T A^\dagger XG)^{-1} = G^{-1}(G^{-1} - Y^T A^\dagger X)^{-1}$. This implies that $(G^{-1} - Y^T A^\dagger X)^{-1} \geq 0$ since $G \geq 0$. This proves (b).

(b) $\Rightarrow$ (c): We have $U - W = U - V - XGY^T = A - XY^T = \Omega$. Also $R(\Omega) = R(A) = R(U)$ and $N(\Omega) = N(A) = N(U)$. So, $\Omega = U - W$ is a proper splitting. Also $U^\dagger \geq 0$ and $U^\dagger W = U^\dagger(V + XGY^T) \geq U^\dagger V \geq 0$. Since $(G^{-1} - Y^T A^\dagger X)^{-1} \geq 0$, it follows from (9) that $\Omega^\dagger \geq 0$. (c) now follows from Theorem 4.

(c) $\Rightarrow$ (a): Since $A = U - V$, we have $\Omega = U - V - XGY^T = U - W$. Also we have $U^\dagger \geq 0, U^\dagger V \geq 0$. Thus, $U^\dagger W \geq 0$, since $XGY^T \geq 0$. Now, by Theorem 4, we are done if the splitting $\Omega = U - W$ is a proper splitting. Since $A = U - V$ is a proper splitting, we have $R(U) = R(A)$ and $N(U) = N(A)$. Now, from the conditions in (10), we get that $R(U) = R(\Omega)$ and $N(U) = N(\Omega)$. Hence $\Omega = U - W$ is a proper splitting, and this completes the proof. □

The following result is a special case of Theorem 9.

Theorem 10. *Let $A = U - V$ be a proper splitting of $A \in \mathbb{R}^{m\times n}$ with $U^\dagger \geq 0$, $U^\dagger V \geq 0$ and $\rho(U^\dagger V) < 1$. Let $X \in \mathbb{R}^{m\times k}$ and $Y \in \mathbb{R}^{n\times k}$ be nonnegative such that $R(X) \subseteq R(A)$, $R(Y) \subseteq R(A^T)$, and $I - Y^T A^\dagger X$ is nonsingular. Let $\Omega = A - XY^T$. Then the following are equivalent:*

(a) $\Omega^\dagger \geq 0$.

(b) $(I - Y^T A^\dagger X)^{-1} \geq 0$.

(c) $\rho(U^\dagger W) < 1$, *where* $W = V + XY^T$.

The following consequence of Theorem 10 appears to be new. This gives two characterizations for a perturbed M-matrix to be an M-matrix.

Corollary 11. *Let $A = sI - B$ where, $B \geq 0$ and $\rho(B) < s$ (i.e., A is an M-matrix). Let $X \in \mathbb{R}^{m\times k}$ and $Y \in \mathbb{R}^{n\times k}$ be nonnegative such that $I - Y^T A^\dagger X$ is nonsingular. Let $\Omega = A - XY^T$. Then the following are equivalent:*

(a) $\Omega^{-1} \geq 0$.

(b) $(I - Y^T A^\dagger X)^{-1} \geq 0$.

(c) $\rho(B(I + XY^T)) < s$.

Proof. From the proof of Theorem 10, since A is nonsingular, it follows that $\Omega\Omega^\dagger = I$ and $\Omega^\dagger\Omega = I$. This shows that Ω is invertible. The rest of the proof is omitted, as it is an easy consequence of the previous result. □

In the rest of this section, we discuss two applications of Theorem 10. First, we characterize the least element in a polyhedral set defined by a perturbed matrix. Next, we consider the following Suppose that the "endpoints" of an interval matrix satisfy a certain positivity property. Then all matrices of a particular subset of the interval also satisfy that positivity condition. The problem now is if that we are given a specific structured perturbation of these endpoints, what conditions guarantee that the positivity property for the corresponding subset remains valid.

The first result is motivated by Theorem 12 below. Let us recall that with respect to the usual order, an element $\mathbf{x}^* \in \mathcal{X} \subseteq \mathbb{R}^n$ is called a *least element* of $\mathcal{X}$ if it satisfies $\mathbf{x}^* \leq \mathbf{x}$ for all $\mathbf{x} \in \mathcal{X}$. Note that a nonempty set may not have the least element, and if it exists, then it is unique. In this connection, the following result is known.

Theorem 12 (see [11, Theorem 3.2]). *For $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$, let*

$$\begin{aligned}\mathcal{X}_{\mathbf{b}} = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} + \mathbf{y} \geq \mathbf{b},\ P_{N(A)}\mathbf{x} = 0,\\ P_{R(A)}\mathbf{y} = 0,\ \text{for some } \mathbf{y} \in \mathbb{R}^m\}.\end{aligned} \tag{11}$$

Then, a vector $\mathbf{x}^*$ is the least element of $\mathcal{X}_{\mathbf{b}}$ if and only if $\mathbf{x}^* = A^\dagger\mathbf{b} \in \mathcal{X}_{\mathbf{b}}$ with $A^\dagger \geq 0$.

Now, we obtain the nonnegative least element of a polyhedral set defined by a perturbed matrix. This is an immediate application of Theorem 10.

Theorem 13. *Let $A \in \mathbb{R}^{m\times n}$ be such that $A^\dagger \geq 0$. Let $X \in \mathbb{R}^{m\times k}$, and let $Y \in \mathbb{R}^{n\times k}$ be nonnegative, such that $R(X) \subseteq R(A)$, $R(Y) \subseteq R(A^T)$, and $I - Y^T A^\dagger X$ is nonsingular. Suppose that $(I - Y^T A^\dagger X)^{-1} \geq 0$. For $\mathbf{b} \in \mathbb{R}^m$, $b \geq 0$, let*

$$\begin{aligned}\mathcal{S}_{\mathbf{b}} = \{\mathbf{x} \in \mathbb{R}^n : \Omega\mathbf{x} + \mathbf{y} \geq \mathbf{b},\ P_{N(\Omega)}\mathbf{x} = 0,\\ P_{R(\Omega)}\mathbf{y} = 0,\ \text{for some } \mathbf{y} \in \mathbb{R}^m\},\end{aligned} \tag{12}$$

where $\Omega = A - XY^T$. Then, $\mathbf{x}^* = (A - XY^T)^\dagger\mathbf{b}$ is the least element of $\mathcal{S}_b$.

Proof. From the assumptions, using Theorem 10, it follows that $\Omega^\dagger \geq 0$. The conclusion now follows from Theorem 12. □

To state and prove the result for interval matrices, let us first recall the notion of interval matrices. For $A, B \in \mathbb{R}^m \times n$, an interval (matrix) $J = [A, B]$ is defined as $J = [A, B] = \{C : A \leq C \leq B\}$. The interval $J = [A, B]$ is said to be range-kernel regular if $R(A) = R(B)$ and $N(A) = N(B)$. The following result [12] provides necessary and sufficient conditions for $C^\dagger \geq 0$ for $C \in K$, where $K = \{C \in J : R(C) = R(A) = R(B), N(C) = N(A) = N(B)\}$.

Theorem 14. *Let $J = [A, B]$ be range-kernel regular. Then, the following are equivalent:*

(a) $C^\dagger \geq 0$ *whenever* $C \in K$,

(b) $A^\dagger \geq 0$ *and* $B^\dagger \geq 0$.

In such a case, we have $C^\dagger = \sum_{j=0}^{\infty} (B^\dagger (B - C))^j B^\dagger$.

Now, we present a result for the perturbation.

Theorem 15. *Let $J = [A, B]$ be range-kernel regular with $A^\dagger \geq 0$ and $B^\dagger \geq 0$. Let X_1, X_2, Y_1, and Y_2 be nonnegative matrices such that $R(X_1) \subseteq R(A)$, $R(Y_1) \subseteq R(A^T)$, $R(X_2) \subseteq R(B)$, and $R(Y_2) \subseteq R(B^T)$. Suppose that $I - Y_1^T A^\dagger X_1$ and $I - Y_2^T A^\dagger X_2$ are nonsingular with nonnegative inverses. Suppose further that $X_2 Y_2^T \leq X_1 Y_1^T$. Let $\widetilde{J} = [E, F]$, where $E = A - X_1 Y_1^T$ and $F = B - X_2 Y_2^T$. Finally, let $\widetilde{K} = \{Z \in \widetilde{J} : R(Z) = R(E) = R(F)$ and $N(Z) = N(E) = N(F)\}$. Then,*

(a) $E^\dagger \geq 0$ *and* $F^\dagger \geq 0$.

(b) $Z^\dagger \geq 0$ *whenever* $Z \in \widetilde{K}$.

In that case, $Z^\dagger = \sum_{j=0}^{\infty} (B^\dagger B - F^\dagger Z)^j F^\dagger$.

Proof. It follows from Theorem 7 that

$$E^\dagger = A^\dagger + A^\dagger X_1 \left(I - Y_1^T A^\dagger X_1\right)^{-1} Y_1^T A^\dagger,$$
$$F^\dagger = B^\dagger + B^\dagger X_2 \left(I - Y_2^T B^\dagger X_2\right)^{-1} Y_2^T B^\dagger. \quad (13)$$

Also, we have $EE^\dagger = AA^\dagger$, $E^\dagger E = A^\dagger A$, $F^\dagger F = B^\dagger B$, and $FF^\dagger = BB^\dagger$. Hence, $R(E) = R(A)$, $N(E) = N(A)$, $R(F) = R(B)$, and $N(F) = N(B)$. This implies that the interval $\widetilde{J}$ is range-kernel regular. Now, since E and F satisfy the conditions of Theorem 10, we have $E^\dagger \geq 0$ and $F^\dagger \geq 0$ proving (a). Hence, by Theorem 14, $Z^\dagger \geq 0$ whenever $Z \in \widetilde{K}$. Again, by Theorem 14, we have $Z^\dagger = \sum_{j=0}^{\infty} (F^\dagger (F - Z))^j F^\dagger = \sum_{j=0}^{\infty} (B^\dagger B - F^\dagger Z)^j F^\dagger$. □

5. Iterations That Preserve Nonnegativity of the Moore-Penrose Inverse

In this section, we present results that typically provide conditions for iteratively defined matrices to have nonnegative Moore-Penrose inverses given that the matrices that we start with have this property. We start with the following result about the rank-one perturbation case, which is a direct consequence of Theorem 10.

Theorem 16. *Let $A \in \mathbb{R}^{n \times n}$ be such that $A^\dagger \geq 0$. Let $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^n$ be nonnegative vectors such that $\mathbf{x} \in R(A)$, $\mathbf{y} \in R(A^T)$, and $1 - \mathbf{y}^T A^\dagger \mathbf{x} \neq 0$. Then $(A - \mathbf{x}\mathbf{y}^T)^\dagger \geq 0$ if and only if $\mathbf{y}^T A^\dagger \mathbf{x} < 1$.*

Example 17. Let us consider $A = (1/3)\begin{pmatrix} 1 & -1 & 1 \\ -1 & 4 & -1 \\ 1 & -1 & 1 \end{pmatrix}$. It can be verified that A can be written in the form $A = \begin{pmatrix} I_2 \\ e_1^T \end{pmatrix}(I_2 + \mathbf{e}_1 e_1^T)^{-1} C^{-1} (I_2 + \mathbf{e}_1 e_1^T)^{-1} \begin{pmatrix} I_2 & \mathbf{e}_1 \end{pmatrix}$ where $\mathbf{e}_1 = (1, 0)^T$ and C is the 2×2 circulant matrix generated by the row $(1, 1/2)$. We have $A^\dagger = (1/2)\begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \geq 0$. Also, the decomposition $A = U - V$, where $U = (1/3)\begin{pmatrix} 1 & -1 & 2 \\ -1 & 4 & -2 \\ 1 & -1 & 2 \end{pmatrix}$, and $V = (1/3)\begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix}$ is a proper splitting of A with $U^\dagger \geq 0$, $U^\dagger V \geq 0$, and $\rho(U^\dagger V) < 1$.

Let $\mathbf{x} = \mathbf{y} = (1/3, 1/6, 1/3)^T$. Then, $\mathbf{x}$ and let $\mathbf{y}$ are nonnegative, $\mathbf{x} \in R(A)$ and $\mathbf{y} \in R(A^T)$. We have $1 - \mathbf{y}^T A^\dagger \mathbf{x} = 5/12 > 0$ and $\rho(U^\dagger W) < 1$ for $W = V + \mathbf{x}\mathbf{y}^T$. Also, it can be seen that $(A - \mathbf{x}\mathbf{y}^T)^\dagger = (1/20)\begin{pmatrix} 47 & 28 & 47 \\ 28 & 32 & 28 \\ 47 & 28 & 47 \end{pmatrix} \geq 0$. This illustrates Theorem 10.

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in \mathbb{R}^m$ and $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k \in \mathbb{R}^n$ be nonnegative. Denote $X_i = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i)$ and $Y_i = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_i)$ for $i = 1, 2, \ldots, k$. Then $A - X_k Y_k^T = A - \sum_{i=1}^{k} \mathbf{x}_i \mathbf{y}_i^T$. The following theorem is obtained by a recurring application of the rank-one result of Theorem 16.

Theorem 18. *Let $A \in \mathbb{R}^{m \times n}$ and let X_i, Y_i be as above. Further, suppose that $\mathbf{x}_i \in R(A - X_{i-1} Y_{i-1}^T)$, $\mathbf{y}_i \in R(A - X_{i-1} Y_{i-1}^T)^T$ and $1 - \mathbf{y}_i^T (A - X_{i-1} Y_{i-1}^T)^\dagger \mathbf{x}_i$ be nonzero. Let $(A - X_{i-1} Y_{i-1}^T)^\dagger \geq 0$, for all $i = 1, 2, \ldots, k$. Then $(A - X_k Y_k^T)^\dagger \geq 0$ if and only if $\mathbf{y}_i^T (A - X_{i-1} Y_{i-1}^T)^\dagger \mathbf{x}_i < 1$, where $A - X_0 Y_0^T$ is taken as A.*

Proof. Set $B_i = A - X_i Y_i^T$ for $i = 0, 1, \ldots, k$, where B_0 is identified as A. The conditions in the theorem can be written as:

$$\mathbf{x}_i \in R(B_{i-1}), \quad \mathbf{y}_i \in R\left(B_{i-1}^T\right),$$
$$1 - \mathbf{y}_i^T B_{i-1}^\dagger \mathbf{x}_i \neq 0, \quad (14)$$
$$\forall i = 1, 2 \ldots, k.$$

Also, we have $B_i^\dagger \geq 0$, for all $i = 0, 1, \ldots, k - 1$.

Now, assume that $B_k^\dagger \geq 0$. Then by Theorem 16 and the conditions in (14) for $i = k$, we have $\mathbf{y}_k B_{k-1}^\dagger \mathbf{x}_k < 1$. Also since by assumption $B_i^\dagger \geq 0$ for all $i = 0, 1 \ldots, k - 1$, we have $\mathbf{y}_i^T B_{i-1}^\dagger \mathbf{x}_i < 1$ for all $i = 1, 2 \ldots, k$.

The converse part can be proved iteratively. Then condition $\mathbf{y}_1^T {B_0}^\dagger \mathbf{x}_1 < 1$ and the conditions in (14) for $i = 1$ imply that ${B_1}^\dagger \geq 0$. Repeating the argument for $i = 2$ to k proves the result. □

The following result is an extension of Lemma 2.3 in [1], which is in turn obtained as a corollary. This corollary will be used in proving another characterization for $(A - X_k Y_k^T)^\dagger \geq 0$.

Theorem 19. *Let $A \in \mathbb{R}^{n\times n}$, $\mathbf{b},\mathbf{c} \in \mathbb{R}^n$ be such that $-\mathbf{b} \geq 0$, $-\mathbf{c} \geq 0$ and $\alpha (\in \mathbb{R}) > 0$. Further, suppose that $\mathbf{b} \in R(A)$ and $\mathbf{c} \in R(A^T)$. Let $\widehat{A} = \begin{pmatrix} A & \mathbf{b} \\ \mathbf{c}^T & \alpha \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$. Then, $\widehat{A}^\dagger \geq 0$ if and only if $A^\dagger \geq 0$ and $\mathbf{c}^T A^\dagger \mathbf{b} < \alpha$.*

Proof. Let us first observe that $AA^\dagger \mathbf{b} = \mathbf{b}$ and $\mathbf{c}^T A^\dagger A = \mathbf{c}^T$. Set

$$X = \begin{pmatrix} A^\dagger + \dfrac{A^\dagger \mathbf{b}\mathbf{c}^T A^\dagger}{\alpha - \mathbf{c}^T A^\dagger \mathbf{b}} & \dfrac{-A^\dagger \mathbf{b}}{\alpha - \mathbf{c}^T A^\dagger \mathbf{b}} \\ -\dfrac{\mathbf{c}^T A^\dagger}{\alpha - \mathbf{c}^T A^\dagger \mathbf{b}} & \dfrac{1}{\alpha - \mathbf{c}^T A^\dagger \mathbf{b}} \end{pmatrix}. \tag{15}$$

It then follows that $\widehat{A}X = \begin{pmatrix} AA^\dagger & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}$ and $X\widehat{A} = \begin{pmatrix} A^\dagger A & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}$. Using these two equations, it can easily be shown that $X = \widehat{A}^\dagger$.

Suppose that $A^\dagger \geq 0$ and $\beta = \alpha - \mathbf{c}^T A^\dagger \mathbf{b} > 0$. Then, $\widehat{A}^\dagger \geq 0$. Conversely suppose that $\widehat{A}^\dagger \geq 0$. Then, we must have $\beta > 0$. Let $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n+1}\}$ denote the standard basis of $\mathbb{R}^{n+1}$. Then, for $i = 1, 2, \ldots, n$, we have $0 \leq \widehat{A}^\dagger(\mathbf{e}_i) = (A^\dagger \mathbf{e}_i + (A^\dagger \mathbf{b}\mathbf{c}^T A^\dagger \mathbf{e}_i/\beta), -(\mathbf{c}^T A^\dagger \mathbf{e}_i/\beta))^T$. Since $\mathbf{c} \leq 0$, it follows that $A^\dagger \mathbf{e}_i \geq 0$ for $i = 1, 2, \ldots, n$. Thus, $A^\dagger \geq 0$. □

Corollary 20 (see [1, Lemma 2.3]). *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular. Let $\mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ be such that $-\mathbf{b} \geq 0$, $-\mathbf{c} \geq 0$ and $\alpha (\in \mathbb{R}) > 0$. Let $\widehat{A} = \begin{pmatrix} A & \mathbf{b} \\ \mathbf{c}^T & \alpha \end{pmatrix}$. Then, $\widehat{A}^{-1} \in \mathbb{R}^{(n+1)\times(n+1)}$ is nonsingular with $\widehat{A}^{-1} \geq 0$ if and only if $A^{-1} \geq 0$ and $\mathbf{c}^T A^{-1} \mathbf{b} < \alpha$.*

Now, we obtain another necessary and sufficient condition for $(A - X_k Y_k^T)^\dagger$ to be nonnegative.

Theorem 21. *Let $A \in \mathbb{R}^{m\times n}$ be such that $A^\dagger \geq 0$. Let $\mathbf{x}_i$, $\mathbf{y}_i$ be nonnegative vectors in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, for every $i = 1, \ldots, k$, such that*

$$\mathbf{x}_i \in R\left(A - X_{i-1} Y_{i-1}^T\right), \qquad \mathbf{y}_i \in R\left(A - X_{i-1} Y_{i-1}^T\right)^T, \tag{16}$$

where $X_i = (\mathbf{x}_1, \ldots, \mathbf{x}_i)$, $Y_i = (\mathbf{y}_1, \ldots, \mathbf{y}_i)$ and $A - X_0 Y_0^T$ is taken as A. If $H_i = I_i - Y_i^T A^\dagger X_i$ is nonsingular and $1 - \mathbf{y}_i^T A^\dagger \mathbf{x}_i$ is positive for all $i = 1, 2, \ldots, k$, then $(A - X_k Y_k^T)^\dagger \geq 0$ if and only if

$$\mathbf{y}_i^T A^\dagger \mathbf{x}_i < 1 - \mathbf{y}_i^T A^\dagger X_{i-1} H_{i-1}^{-1} Y_{i-1}^T A^\dagger \mathbf{x}_i, \quad \forall i = 1, \ldots, k. \tag{17}$$

Proof. The range conditions in (16) imply that $R(X_k) \subseteq R(A)$ and $R(Y_k) \subseteq R(A^T)$. Also, from the assumptions, it follows that H_k is nonsingular. By Theorem 10, $(A - X_k Y_k^T)^\dagger \geq 0$ if and only if $H_k^{-1} \geq 0$. Now,

$$H_k = \begin{pmatrix} H_{k-1} & -Y_{k-1}^T A^\dagger \mathbf{x}_k \\ -\mathbf{y}_k^T A^\dagger X_{k-1} & 1 - \mathbf{y}_k^T A^\dagger \mathbf{x}_k \end{pmatrix}. \tag{18}$$

Since $1 - \mathbf{y}_k^T A^\dagger \mathbf{x}_k > 0$, using Corollary 20, it follows that $H_k^{-1} \geq 0$ if and only if $\mathbf{y}_k^T A^\dagger X_{k-1} H_{k-1}^{-1} Y_{k-1}^T A^\dagger \mathbf{x}_k < 1 - \mathbf{y}_k^T A^\dagger \mathbf{x}_k$ and $H_{k-1}^{-1} \geq 0$.

Now, applying the above argument to the matrix H_{k-1}, we have that $H_{k-1}^{-1} \geq 0$ holds if and only if $H_{k-2}^{-1} \geq 0$ holds and $\mathbf{y}_{k-1}^T A^\dagger X_{k-2}(I_{k-2} - Y_{k-2}^T A^\dagger X_{k-2})^{-1} Y_{k-2}^T A^\dagger \mathbf{x}_{k-1} < 1 - \mathbf{y}_{k-1}^T A^\dagger \mathbf{x}_{k-1}$. Continuing the above argument, we get, $(A - X_k \mathbf{y}_k^T)^\dagger \geq 0$ if and only if $\mathbf{y}_i^T A^\dagger X_{i-1} H_{i-1}^{-1} Y_{i-1}^T A^\dagger \mathbf{x}_i < 1 - \mathbf{y}_i^T A^\dagger \mathbf{x}_i$, $\forall i = 1, \ldots, k$. This is condition (17). □

We conclude the paper by considering an extension of Example 17.

Example 22. For a fixed n, let C be the circulant matrix generated by the row vector $(1, (1/2), (1/3), \ldots (1/(n-1)))$. Consider

$$A = \begin{pmatrix} I \\ e_1^T \end{pmatrix} \left(I + \mathbf{e}_1 e_1^T\right)^{-1} C^{-1} \left(I + \mathbf{e}_1 e_1^T\right)^{-1} \begin{pmatrix} I & \mathbf{e}_1 \end{pmatrix}, \tag{19}$$

where I is the identity matrix of order $n-1$, $\mathbf{e}_1$ is $(n-1)\times 1$ vector with 1 as the first entry and 0 elsewhere. Then, $A \in \mathbb{R}^{n\times n}$. $A^\dagger$ is the $n\times n$ nonnegative Toeplitz matrix

$$A^\dagger = \begin{pmatrix} 1 & \dfrac{1}{n-1} & \dfrac{1}{n-2} & \cdots & \dfrac{1}{2} & 1 \\ \dfrac{1}{2} & 1 & \dfrac{1}{n-1} & \cdots & \dfrac{1}{3} & \dfrac{1}{2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dfrac{1}{n-1} & \dfrac{1}{n-2} & \dfrac{1}{n-3} & \cdots & 1 & \dfrac{1}{n-1} \\ 1 & \dfrac{1}{n-1} & \dfrac{1}{n-2} & \cdots & \dfrac{1}{2} & 1 \end{pmatrix}. \tag{20}$$

Let x_i be the first row of $B_{i-1}^\dagger$ written as a column vector multiplied by $1/n^i$ and let y_i be the first column of $B_{i-1}^\dagger$ multiplied by $1/n^i$, where $B_i = A - X_i Y_i^T$, with B_0 being identified with A, $i = 0, 2, \ldots, k$. We then have $\mathbf{x}_i \geq 0$, $\mathbf{y}_i \geq 0$, $x_1 \in R(A)$ and $y_1 \in R(A^T)$. Now, if $B_{i-1}^\dagger \geq 0$ for each iteration, then the vectors x_i and y_i will be nonnegative. Hence, it is enough to check that $1 - y_i^T B_{i-1}^\dagger x_i > 0$ in order to get $B_i^\dagger \geq 0$. Experimentally, it has been observed that for $n > 3$, the condition $1 - y_i^T B_{i-1}^\dagger x_i > 0$ holds true for any iteration. However, for $n = 3$, it is observed that $B_1^\dagger \geq 0$, $B_2^\dagger \geq 0$, $1 - y_1^T A^\dagger x_1 > 0$, and $1 - y_2^\dagger B_1^\dagger x_2 > 0$. But, $1 - y_3^T B_2^\dagger x_3 < 0$, and in that case, we observe that $B_3^\dagger$ is not nonnegative.

References

[1] J. Ding, W. Pye, and L. Zhao, "Some results on structured M-matrices with an application to wireless communications," *Linear Algebra and its Applications*, vol. 416, no. 2-3, pp. 608–614, 2006.

[2] S. K. Mitra and P. Bhimasankaram, "Generalized inverses of partitioned matrices and recalculation of least squares estimates for data or model changes," *Sankhyā A*, vol. 33, pp. 395–410, 1971.

[3] C. Y. Deng, "A generalization of the Sherman-Morrison-Woodbury formula," *Applied Mathematics Letters*, vol. 24, no. 9, pp. 1561–1564, 2011.

[4] A. Berman, M. Neumann, and R. J. Stern, *Nonnegative Matrices in Dynamical Systems*, John Wiley & Sons, New York, NY, USA, 1989.

[5] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Society for Industrial and Applied Mathematics (SIAM), 1994.

[6] L. Collatz, *Functional Analysis and Numerical Mathematics*, Academic Press, New York, NY, USA, 1966.

[7] C. A. Desoer and B. H. Whalen, "A note on pseudoinverses," *Society for Industrial and Applied Mathematics*, vol. 11, pp. 442–447, 1963.

[8] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses:Theory and Applications*, Springer, New York, NY, USA, 2003.

[9] A. Berman and R. J. Plemmons, "Cones and iterative methods for best least squares solutions of linear systems," *SIAM Journal on Numerical Analysis*, vol. 11, pp. 145–154, 1974.

[10] D. Mishra and K. C. Sivakumar, "On splittings of matrices and nonnegative generalized inverses," *Operators and Matrices*, vol. 6, no. 1, pp. 85–95, 2012.

[11] D. Mishra and K. C. Sivakumar, "Nonnegative generalized inverses and least elements of polyhedral sets," *Linear Algebra and its Applications*, vol. 434, no. 12, pp. 2448–2455, 2011.

[12] M. R. Kannan and K. C. Sivakumar, "Moore-Penrose inverse positivity of interval matrices," *Linear Algebra and its Applications*, vol. 436, no. 3, pp. 571–578, 2012.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 320392, 7 pages
http://dx.doi.org/10.1155/2013/320392

Research Article

A New Construction of Multisender Authentication Codes from Polynomials over Finite Fields

Xiuli Wang

College of Science, Civil Aviation University of China, Tianjin 300300, China

Correspondence should be addressed to Xiuli Wang; xlwang@cauc.edu.cn

Received 3 February 2013; Accepted 7 April 2013

Academic Editor: Yang Zhang

Multisender authentication codes allow a group of senders to construct an authenticated message for a receiver such that the receiver can verify the authenticity of the received message. In this paper, we construct one multisender authentication code from polynomials over finite fields. Some parameters and the probabilities of deceptions of this code are also computed.

1. Introduction

Multisender authentication code was firstly constructed by Gilbert et al. [1] in 1974. Multisender authentication system refers to who a group of senders, cooperatively send a message to a receiver; then the receiver should be able to ascertain that the message is authentic. About this case, many scholars and researchers had made great contributions to multisender authentication codes, such as [2–6].

In the actual computer network communications, multisender authentication codes include sequential model and simultaneous model. Sequential model is that each sender uses his own encoding rules to encode a source state orderly, the last sender sends the encoded message to the receiver, and the receiver receives the message and verifies whether the message is legal or not. Simultaneous model is that all senders use their own encoding rules to encode a source state, and each sender sends the encoded message to the synthesizer, respectively; then the synthesizer forms an authenticated message and verifies whether the message is legal or not. In this paper, we will adopt the second model.

In a simultaneous model, there are four participants: a group of senders $U = \{U_1, U_2, \ldots, U_n\}$, the key distribution center, he is responsible for the key distribution to senders and receiver, including solving the disputes between them, a receiver R, and a synthesizer, where he only runs the trusted synthesis algorithm. The code works as follows: each sender and receiver has their own Cartesian authentication code, respectively. Let $(S, E_i, T_i; f_i)$ $(i = 1, 2, \ldots, n)$ be the senders' Cartesian authentication code, $(S, E_R, T; g)$ be the receiver's Cartesian authentication code, $h : T_1 \times T_2 \times \cdots \times T_n \rightarrow T$ be the synthesis algorithm, and $\pi_i : E \rightarrow E_i$ be a subkey generation algorithm, where E is the key set of the key distribution center. When authenticating a message, the senders and the receiver should comply with the protocol. The key distribution center randomly selects an encoding rule $e \in E$ and sends $e_i = \pi_i(e)$ to the ith sender U_i $(i = 1, 2, \ldots, n)$, secretly; then he calculates e_R by e according to an effective algorithm and secretly sends e_R to the receiver R. If the senders would like to send a source state s to the receiver R, U_i computes $t_i = f_i(s, e_i)$ $(i = 1, 2, \ldots, n)$ and sends $m_i = (s, t_i)$ $(i = 1, 2, \ldots, n)$ to the synthesizer through an open channel. The synthesizer receives the message $m_i = (s, t_i)$ $(i = 1, 2, \ldots, n)$ and calculates $t = h(t_1, t_2, \ldots, t_n)$ by the synthesis algorithm h and then sends message $m = (s, t)$ to the receiver; he checks the authenticity by verifying whether $t = g(s, e_R)$ or not. If the equality holds, the message is authentic and is accepted. Otherwise, the message is rejected.

We assume that the key distribution center is credible, and though he know the senders' and receiver's encoding rules, he will not participate in any communication activities. When transmitters and receiver are disputing, the key distribution center settles it. At the same time, we assume that the system follows the Kerckhoff principle in which, except the actual used keys, the other information of the whole system is public.

In a multisender authentication system, we assume that the whole senders are cooperative to form a valid message; that is, all senders as a whole and receiver are reliable. But there are some malicious senders who together cheat the receiver; the part of senders and receiver are not credible, and they can take impersonation attack and substitution attack. In the whole system, we assume that $\{U_1, U_2, \ldots, U_n\}$ are senders, R is a receiver, E_i is the encoding rules set of the sender U_i, and E_R is the decoding rules set of the receiver R. If the source state space S and the key space E_R of receiver R are according to a uniform distribution, then the message space M and the tag space T are determined by the probability distribution of S and E_R. $L = \{i_1, i_2, \ldots, i_l\} \subset \{1, 2, \ldots, n\}$, $l < n$, $U_L = \{U_{i_1}, U_{i_2}, \ldots, U_{i_l}\}$, $E_L = \{E_{U_{i_1}}, E_{U_{i_2}}, \ldots, E_{U_{i_l}}\}$. Now consider that let us consider the attacks from malicious groups of senders. Here, there are two kinds of attack.

The opponent's impersonation attack to receiver: U_L, after receiving their secret keys, encode a message and send it to the receiver. U_L are successful if the receiver accepts it as legitimate message. Denote by P_I the largest probability of some opponent's successful impersonation attack to receiver; it can be expressed as

$$P_I = \max_{m \in M} \left\{ \frac{\left| \{ e_R \in E_R \mid e_R \subset m \} \right|}{\left| E_R \right|} \right\}. \tag{1}$$

The opponent's substitution attack to the receiver: U_L replace m with another message m', after they observe a legitimate message m. U_L are successful if the receiver accepts it as legitimate message; it can be expressed as

$$P_S = \max_{m \in M} \left\{ \frac{\max_{m' \neq m \in M} \left| \{ e_R \in E_R \mid e_R \subset m, m' \} \right|}{\left| \{ e_R \in E_R \mid e_R \subset m \} \right|} \right\}. \tag{2}$$

There might be l malicious senders who together cheat the receiver; that is, the part of senders and the receiver are not credible, and they can take impersonation attack. Let $L = \{i_1, i_2, \ldots, i_l\} \subset \{1, 2, \ldots, n\}$, $l < n$ and $E_L = \{E_{U i_1}, E_{U i_2}, \ldots, E_{U i_l}\}$. Assume that $U_L = \{U_{i_1}, U_{i_2}, \ldots, U_{i_l}\}$, after receiving their secret keys, send a message m to the receiver R; U_L are successful if the receiver accepts it as legitimate message. Denote by $P_U(L)$ the maximum probability of success of the impersonation attack to the receiver. It can be expressed as

$$P_U(L) = \max_{e_L \in E_L} \max_{e_L \in e_U} \left\{ \frac{\max_{m \in M} \left| \{ e_R \in E_R \mid e_R \subset m, p(e_R, e_P) \neq 0 \} \right|}{\left| \{ e_R \in E_R \mid p(e_R, e_P) \neq 0 \} \right|} \right\}. \tag{3}$$

Notes. $p(e_R, e_P) \neq 0$ implies that any information s encoded by e_T can be authenticated by e_R.

In [2], Desmedt et al. gave two constructions for MRA-codes based on polynomials and finite geometries, respectively. To construct multisender or multireceiver authentication by polynomials over finite fields, many researchers have done much work, for example, [7–9]. There are other constructions of multisender authentication codes that are given in [3–6]. The construction of authentication codes is combinational design in its nature. We know that the polynomial over finite fields can provide a better algebra structure and is easy to count. In this paper, we construct one multisender authentication code from the polynomial over finite fields. Some parameters and the probabilities of deceptions of this code are also computed. We realize the generalization and the application of the similar idea and method of the paper [7–9].

2. Some Results about Finite Field

Let F_q be the finite field with q elements, where q is a power of a prime p and F is a field containing F_q; denote by F_q^* be the nonzero elements set of F_q. In this paper, we will use the following conclusions over finite fields.

Conclusion 1. A generator α of F_q^* is called a primitive element of F_q.

Conclusion 2. Let $\alpha \in F_q$; if some polynomials contain α as their root and their leading coefficient are 1 over F_q, then the polynomial having least degree among all such polynomials is called a minimal polynomial over F_q.

Conclusion 3. Let $|F| = q^n$, then F is an n-dimensional vector space over F_q. Let α be a primitive element of F_q and $g(x)$ the minimal polynomial about α over F_q; then $\dim g(x) = n$ and $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ is a basis of F. Furthermore, $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ is linear independent, and it is equal to $\alpha, \alpha^2, \ldots, \alpha^{n-1}, \alpha^n$ (α is a primitive element, $\alpha \neq 0$) is also linear independent; moreover, $\alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{n-1}}, \alpha^{p^n}$ is also linear independent.

Conclusion 4. Consider $(x_1 + x_2 + \cdots + x_n)^m = (x_1)^m + (x_2)^m + \cdots + (x_n)^m$, where $x_i \in F_q$, $(1 \leq i \leq n)$ and m is a nonnegative power of character p of F_q.

Conclusion 5. Let $m \leq n$. Then, the number of $m \times n$ matrices of rank m over F_q is $q^{m(m-1)/2} \prod_{i=n-m+1}^{n} (q^i - 1)$.

More results about finite fields can be found in [10–12].

3. Construction

Let the polynomial $p_j(x) = a_{j1} x^{p^n} + a_{j2} x^{p^{(n-1)}} + \cdots + a_{jn} x^p$ $(1 \leq j \leq k)$, where the coefficient $a_{il} \in F_q$, $(1 \leq l \leq n)$, and these vectors by the composition of their coefficient are linearly independent. The set of source states $S = F_q$; the set of ith transmitter's encoding rules $E_{U_i} = \{p_1(x_i), p_2(x_i), \ldots, p_k(x_i), x_i \in F_q^*\}$ $(1 \leq i \leq n)$; the set of receiver's encoding rules $E_R = \{p_1(\alpha), p_2(\alpha), \ldots, p_k(\alpha)$, where α is a primitive element of $F_q\}$; the set of ith transmitter's tags $T_i = \{t_i \mid t_i \in F_q\}$ $(1 \leq i \leq n)$; the set of receiver's tags $T = \{t \mid t \in F_q\}$.

Define the encoding map $f_i : S \times E_{U_i} \to T_i$, $f_i(s, e_{U_i}) = sp_1(x_i) + s^2 p_2(x_i) + \cdots + s^k p_k(x_i)$, $1 \le i \le n$.

The decoding map $f : S \times E_R \to T$, $f(s, e_R) = sp_1(\alpha) + s^2 p_2(\alpha) + \cdots + s^k p_k(\alpha)$.

The synthesizing map $h : T_1 \times T_2 \times \cdots \times T_n \to T$, $h(t_1, t_2, \ldots, t_n) = t_1 + t_2 + \cdots + t_n$.

The code works as follows.

Assume that q is larger than, or equal to, the number of the possible message and $n \le q$.

3.1. Key Distribution. The key distribution center randomly generates k ($k \le n$) polynomials $p_1(x), p_2(x), \ldots, p_k(x)$, where $p_j(x) = a_{j1}x^{p^n} + a_{j2}x^{p^{(n-1)}} + \cdots + a_{jn}x^p$ ($1 \le j \le k$), and make these vectors by composed of their coefficient is linearly independent, it is equivalent to the column vectors of the matrix $\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} \end{pmatrix}$ is linearly independent. He selects n distinct nonzero elements $x_1, x_2, \ldots, x_n \in F_q$ again and makes x_i ($1 \le i \le n$) secret; then he sends privately $p_1(x_i), p_2(x_i), \ldots, p_k(x_i)$ to the sender U_i ($1 \le i \le n$). The key distribution center also randomly chooses a primitive element α of F_q satisfying $x_1 + x_2 + \cdots + x_n = \alpha$ and sends $p_1(\alpha), p_2(\alpha), \ldots, p_k(\alpha)$ to the receiver R.

3.2. Broadcast. If the senders want to send a source state $s \in S$ to the receiver R, the sender U_i calculates $t_i = f_i(s, e_{U_i}) = A_s(x_i) = sp_1(x_i) + s^2 p_2(x_i) + \cdots + s^k p_k(x_i)$, $1 \le i \le n$ and then sends $A_s(x_i) = t_i$ to the synthesizer.

3.3. Synthesis. After the synthesizer receives $t_1, t_2, \ldots, t_n$, he calculates $h(t_1, t_2, \ldots, t_n) = t_1 + t_2 + \cdots + t_n$ and then sends $m = (s, t)$ to the receiver R.

3.4. Verification. When the receiver R receives $m = (s, t)$, he calculates $t' = g(s, e_R) = A_s(\alpha) = sp_1(\alpha) + s^2 p_2(\alpha) + \cdots + s^k p_k(\alpha)$. If $t = t'$, he accepts t; otherwise, he rejects it.

Next, we will show that the above construction is a well defined multisender authentication code with arbitration.

Lemma 1. *Let $C_i = (S, E_{P_i}, T_i, f_i)$; then the code is an A-code, $1 \le i \le n$.*

Proof. (1) For any $e_{U_i} \in E_{U_i}$, $s \in S$, because $E_{U_i} = \{p_1(x_i), p_2(x_i), \ldots, p_k(x_i),\ x_i \in F_q^*\}$, so $t_i = sp_1(x_i) + s^2 p_2(x_i) + \cdots + s^k p_k(x_i) \in T_i = F_q$. Conversely, for any $t_i \in T_i$, choose $e_{U_i} = \{p_1(x_i), p_2(x_i), \ldots, p_k(x_i),\ x_i \in F_q^*\}$, where $p_j(x) = a_{j1}x^{p^n} + a_{j2}x^{p^{(n-1)}} + \cdots + a_{jn}x^p$ ($1 \le j \le k$), and let $t_i = f_i(s, e_{U_i}) = sp_1(x_i) + s^2 p_2(x_i) + \cdots + s^k p_k(x_i)$; it is equivalent to

$$\left(x_i^{p^n}, x_i^{p^{n-1}}, \ldots, x_i^p\right) \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} \end{pmatrix} \begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix} = t_i. \tag{4}$$

It follows that

$$\begin{pmatrix} x_1^{p^n} & x_1^{p^{n-1}} & \cdots & x_1^p \\ x_2^{p^n} & x_2^{p^{n-1}} & \cdots & x_2^p \\ \vdots & \vdots & \vdots & \vdots \\ x_n^{p^n} & x_n^{p^{n-1}} & \cdots & x_n^p \end{pmatrix} \times \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} \end{pmatrix} \begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix}. \tag{5}$$

Denote

$$A = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} \end{pmatrix},$$
$$X = \begin{pmatrix} x_1^{p^n} & x_1^{p^{n-1}} & \cdots & x_1^p \\ x_2^{p^n} & x_2^{p^{n-1}} & \cdots & x_2^p \\ \vdots & \vdots & \vdots & \vdots \\ x_n^{p^n} & x_n^{p^{n-1}} & \cdots & x_n^p \end{pmatrix}, \tag{6}$$
$$S = \begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix}, \qquad t = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix}.$$

The above linear equation is equivalent to $XAS = t$, because the column vectors of A are linearly independent, X is equivalent to a Vandermonde matrix, and X is inverse; therefore, the above linear equation has a unique solution, so s is only defined; that is, f_i ($1 \le i \le n$) is a surjection.

(2) If $s' \in S$ is another source state satisfying $sp_1(x_i) + s^2 p_2(x_i) + \cdots + s^k p_k(x_i) = s' p_1(x_i) + s'^2 p_2(x_i) + \cdots + s'^k p_k(x_i) = t_i$, and it is equivalent to $(s - s')p_1(x_i) + (s^2 - s'^2)p_2(x_i) + \cdots + (s^k - s'^k)p_k(x_i) = 0$, then

$$\left(x_i^{p^n}, x_i^{p^{n-1}}, \ldots, x_i^p\right) \times \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} \end{pmatrix} \begin{pmatrix} s - s' \\ s^2 - s'^2 \\ \vdots \\ s^k - s'^k \end{pmatrix} = 0. \tag{7}$$

Thus

$$\begin{pmatrix} x_1^{p^n} & x_1^{p^{n-1}} & \cdots & x_1^{p} \\ x_2^{p^n} & x_2^{p^{n-1}} & \cdots & x_2^{p} \\ \vdots & \vdots & \vdots & \vdots \\ x_n^{p^n} & x_n^{p^{n-1}} & \cdots & x_n^{p} \end{pmatrix} \times \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} \end{pmatrix} \begin{pmatrix} s-s' \\ s^2-s'^2 \\ \vdots \\ s^k-s'^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (8)$$

Similar to (1), we know that the homogeneous linear equation $XAS = 0$ has a unique solution; that is, there is only zero solution, so $s = s'$. So, s is the unique source state determined by e_{U_i} and t_i; thus, C_i $(1 \le i \le n)$ is an A-code. □

Lemma 2. *Let $C = (S, E_R, T, g)$; then the code is an A-code.*

Proof. (1) For any $s \in S$, $e_R \in E_R$, from the definition of e_R, we assume that $E_R = \{p_1(\alpha), p_2(\alpha), \ldots, p_k(\alpha)$, where α is a primitive element of $F_q\}$, $g(s, e_R) = sp_1(\alpha) + s^2 p_2(\alpha) + \cdots + s^k p_k(\alpha) \in T = F_q$; on the other hand, for any $t \in T$, choose $e_R = \{p_1(\alpha), p_2(\alpha), \ldots, p_k(\alpha)$, where α is a primitive element of $F_q\}$, $g(s, e_R) = sp_1(\alpha) + s^2 p_2(\alpha) + \cdots + s^k p_k(\alpha) = t$; it is equivalent to

$$\left(\alpha^{p^n}, \alpha^{p^{n-1}}, \cdots, \alpha^{p}\right) \times \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} \end{pmatrix} \begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix} = t, \quad (9)$$

$$A = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} \end{pmatrix};$$

that is, $(\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^{p}) A \begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix} = t$. From Conclusion 3, we know that $(\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^{p})$ is linearly independent and the column vectors of A are also linearly independent; therefore, the above linear equation has unique solution, so s is only defined; that is, g is a surjection.

(2) If s' is another source state satisfying $t = g(s', e_R)$, then

$$\left(\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^{p}\right) A \begin{pmatrix} s' \\ s'^2 \\ \vdots \\ s'^k \end{pmatrix} = \left(\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^{p}\right) A \begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix}; \quad (10)$$

that is, $(\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^{p}) A \left(\begin{bmatrix} s' \\ s'^2 \\ \vdots \\ s'^k \end{bmatrix} - \begin{bmatrix} s \\ s^2 \\ \vdots \\ s^k \end{bmatrix} \right) = 0$. Similar to (1), we get that the homogeneous linear equation $(\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^{p}) A (S' - S) = 0$ has a unique solution; that is, there is only zero solution, so $S = S'$; that is, $s = s'$. So, s is the unique source state determined by e_R and t; thus, $C = (S, E_R, T, g)$ is an A-code.

At the same time, for any valid $m = (s, t)$, we have known that $\alpha = x_1 + x_2 + \cdots + x_n$, and it follows that $t' = sp_1(\alpha) + s^2 p_2(\alpha) + \cdots + s^k p_k(\alpha) = sp_1(x_1 + x_2 + \cdots + x_n) + s^2 p_2(x_1 + x_2 + \cdots + x_n) + \cdots + s^k p_k(x_1 + x_2 + \cdots + x_n)$. We also have known that $p_j(x) = a_{j1} x^{p^n} + a_{j2} x^{p^{(n-1)}} + \cdots + a_{jn} x^{p}$ $(1 \le j \le k)$; from Conclusion 4, $(x_1 + x_2 + \cdots + x_n)^{p^m} = (x_1)^{p^m} + (x_2)^{p^m} + \cdots + (x_n)^{p^m}$, where m is a nonnegative power of character p of F_q, and we get $p_j(x_1 + x_2 + \cdots + x_n) = p_j(x_1) + p_j(x_2) + \cdots + p_j(x_n)$; therefore, $t' = sp_1(\alpha) + s^2 p_2(\alpha) + \cdots + s^k p_k(\alpha) = (sp_1(x_1) + s^2 p_2(x_1) + \cdots + s^k p_k(x_1)) + (sp_1(x_2) + s^2 p_2(x_2) + \cdots + s^k p_k(x_2)) + \cdots + (sp_1(x_n) + s^2 p_2(x_n) + \cdots + s^k p_k(x_n)) = t_1 + t_2 + \cdots + t_n = t$, and the receiver R accepts m. □

From Lemmas 1 and 2, we know that such construction of multisender authentication codes is reasonable and there are n senders in this system. Next, we compute the parameters of this code and the maximum probability of success in impersonation attack and substitution attack by the group of senders.

Theorem 3. *Some parameters of this construction are* $|S| = q$, $|E_{U_i}| = [q^{k(k-1)/2} \prod_{i=n-k+1}^{n} (q^i - 1)] \binom{q-1}{1} = [q^{k(k-1)/2} \prod_{i=n-k+1}^{n} (q^i - 1)](q-1)$ $(1 \le i \le n)$, $|T_i| = q$ $(1 \le i \le n)$, $|E_R| = [q^{k(k-1)/2} \prod_{i=n-k+1}^{n} (q^i - 1)] \varphi(q-1)$, $|T| = q$. *Where $\varphi(q-1)$ is the Euler function of $q-1$, it represents the number of primitive element of F_q here.*

Proof. For $|S| = q$, $|T_i| = q$, and $|T| = q$, the results are straightforward. For E_{U_i}, because $E_{U_i} = \{p_1(x_i), p_2(x_i), \ldots, p_k(x_i)$, $x_i \in F_q^*\}$, where $p_j(x) = a_{j1} x^{p^n} + a_{j2} x^{p^{(n-1)}} + \cdots + a_{jn} x^{p}$ $(1 \le j \le k)$, and these vectors by the composition of their coefficient are linearly independent, it is equivalent to the columns of $A = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} \end{pmatrix}$ is linear independent.

From Conclusion 5, we can conclude that the number of A satisfying the condition is $q^{k(k-1)/2}\prod_{i=n-k+1}^{n}(q^i-1)$. On the other hand, the number of distinct nonzero elements x_i $(1 \le i \le n)$ in F_q is $\binom{q-1}{1} = q-1$, so $|E_{U_i}| = [q^{k(k-1)/2}\prod_{i=n-k+1}^{n}(q^i-1)](q-1)$. For E_R, $E_R = \{p_1(\alpha), p_2(\alpha), \ldots, p_k(\alpha)$, where α is a primitive element of $F_q\}$. For α, from Conclusion 1, a generator of F_q^* is called a primitive element of F_q, $|F_q^*| = q-1$; by the theory of the group, we know that the number of generator of F_q^* is $\varphi(q-1)$; that is, the number of α is $\varphi(q-1)$. For $p_1(x), p_2(x), \ldots, p_k(x)$. From above, we have confirmed that the number of these polynomials is $q^{k(k-1)/2}\prod_{i=n-k+1}^{n}(q^i-1)$; therefore, $|E_R| = [q^{k(k-1)/2}\prod_{i=n-k+1}^{n}(q^i-1)]\varphi(q-1)$. □

Lemma 4. *For any $m \in M$, the number of e_R contained m is $\varphi(q-1)$.*

Proof. Let $m = (s,t) \in M$, $e_R = \{p_1(\alpha), p_2(\alpha), \ldots, p_k(\alpha)$, where α is a primitive element of $F_q\} \in E_R$. If $e_R \subset m$, then $sp_1(\alpha) + s^2p_2(\alpha) + \cdots + s^kp_k(\alpha) = t \Leftrightarrow (\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^p)A\begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix} = t$. For any α, suppose that there is another A' such that $(\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^p)A'\begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix} = t$, then $(\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^p)(A - A')\begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix} = 0$, because $\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^p$ is linearly independent, so $(A-A')\begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix} = 0$, but $\begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix}$ is arbitrarily; therefore, $A - A' = 0$; that is, $A = A'$, and it follows that A is only determined by α. Therefore, as $\alpha \in E_R$, for any given s and t, the number of e_R contained in m is $\varphi(q-1)$. □

Lemma 5. *For any $m = (s,t) \in M$ and $m' = (s',t') \in M$ with $s \ne s'$, the number of e_R contained m and m' is 1.*

Proof. Assume that $e_R = \{p_1(\alpha), p_2(\alpha), \ldots, p_k(\alpha)$, where α is a primitive element of $F_q\} \in E_R$. If $e_R \subset m$ and $e_R \subset m'$, then $sp_1(\alpha) + s^2p_2(\alpha) + \cdots + s^kp_k(\alpha) = t \Leftrightarrow (\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^p)A\begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix} = t$, $s'p_1(\alpha) + s'^2p_2(\alpha) + \cdots + s'^kp_k(\alpha) = t \Leftrightarrow (\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^p)A\begin{pmatrix} s' \\ s'^2 \\ \vdots \\ s'^k \end{pmatrix} = t$. It is equivalent to $(\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^p)A\begin{pmatrix} s-s' \\ s^2-s'^2 \\ \vdots \\ s^k-s'^k \end{pmatrix} = t-t'$ because $s \ne s'$, so $t \ne t'$; otherwise, we assume that $t = t'$ and since $\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^p$ and the column vectors of A both are linearly independent, it forces that $s = s'$; this is a contradiction. Therefore, we get

$$(t-t')^{-1}\left[\left(\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^p\right)A\begin{pmatrix} s-s' \\ s^2-s'^2 \\ \vdots \\ s^k-s'^k \end{pmatrix}\right] = 1, \qquad (*)$$

since t, t' is given, $(t-t')^{-1}$ is unique, by equation $(*)$, for any given s, s' and t, t', we obtain that $(\alpha^{p^n}, \alpha^{p^{n-1}}, \ldots, \alpha^p)A$ is only determined; thus, the number of e_R contained m and m' is 1. □

Lemma 6. *For any fixed $e_U = \{p_1(x_i), p_2(x_i), \ldots, p_k(x_i)$, $x_i \in F_q^*\}$ $(1 \le i \le n)$ containing a given e_L, then the number of e_R which is incidence with e_U is $\varphi(q-1)$.*

Proof. For any fixed $e_U = \{p_1(x_i), p_2(x_i), \ldots, p_k(x_i)$, $x_i \in F_q^*\}$ $(1 \le i \le n)$ containing a given e_L, we assume that $p_j(x_i) = a_{j1}x_i^{p^n} + a_{j2}x_i^{p^{(n-1)}} + \cdots + a_{jn}x_i^p$ $(1 \le j \le k,\ 1 \le i \le n)$, $e_R = \{p_1(\alpha), p_2(\alpha), \ldots, p_k(\alpha)$, where α is a primitive element of $F_q\}$. From the definitions of e_R and e_U and Conclusion 4, we can conclude that e_R is incidence with e_U if and only if $x_1 + x_2 + \cdots + x_n = \alpha$. For any α, since $\mathrm{Rank}(1,1,\ldots,1) = \mathrm{Rank}(1,1,\ldots,1,\alpha) = 1 < n$, so the equation $x_1+x_2+\cdots+x_n = \alpha$ always has a solution. From the proof of Theorem 3, we know the number of e_R which is incident with e_U (i.e., the number of all E_R) is $[q^{k(k-1)/2}\prod_{i=n-k+1}^{n}(q^i-1)]\ \varphi(q-1)$. □

Lemma 7. *For any fixed $e_U = \{p_1(x_i), p_2(x_i), \ldots, p_k(x_i)$, $x_i \in F_q^*\}$ $(1 \le i \le n)$ containing a given e_L and $m = (s,t)$, the number of e_R which is incidence with e_U and contained in m is 1.*

Proof. For any $s \in S$, $e_R \in E_R$, we assume that $e_R = \{p_1(\alpha), p_2(\alpha), \ldots, p_k(\alpha)$, where α is a primitive element of $F_q\}$. Similar to Lemma 6, for any fixed $e_U = \{p_1(x_i), p_2(x_i), \ldots, p_k(x_i)$, $x_i \in F_q^*\}$, $(1 \le i \le n)$ containing a given e_L, we have known that e_R is incident with e_U if and only if

$$x_1 + x_2 + \cdots + x_n = \alpha. \qquad (11)$$

Again, with $e_R \subset m$, we can get

$$sp_1(\alpha) + s^2p_2(\alpha) + \cdots + s^kp_k(\alpha) = t. \qquad (12)$$

By (11) and (12) and the property of $p_j(x)$ $(1 \le j \le k)$, we have the following conclusion:

$$sp_1\left(\sum_{i=1}^{n}x_i\right) + s^2p_2\left(\sum_{i=1}^{n}x_i\right) + \cdots + s^kp_k\left(\sum_{i=1}^{n}x_i\right)$$
$$= t \Longleftrightarrow \left(p_1\left(\sum_{i=1}^{n}x_i\right), p_2\left(\sum_{i=1}^{n}x_i\right), \ldots, p_k\left(\sum_{i=1}^{n}x_i\right)\right)$$

$$\times \begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix}$$

$$= t \Longleftrightarrow \left(\sum_{i=1}^{n} p_1(x_i), \sum_{i=1}^{n} p_2(x_i), \dots, \sum_{i=1}^{n} p_k(x_i) \right) \begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix}$$

$$= t \Longleftrightarrow \left(\left(\sum_{i=1}^{n} x_i \right)^n, \left(\sum_{i=1}^{n} x_i \right)^{n-1}, \dots, \left(\sum_{i=1}^{n} x_i \right) \right) A \begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix}$$

$$= t \Longleftrightarrow \left[\left(\sum_{i=1}^{n} p_1(x_i), \sum_{i=1}^{n} p_2(x_i), \dots, \sum_{i=1}^{n} p_k(x_i) \right) - \left(\left(\sum_{i=1}^{n} x_i \right)^n, \left(\sum_{i=1}^{n} x_i \right)^{n-1}, \dots, \left(\sum_{i=1}^{n} x_i \right) \right) A \right] \times \begin{pmatrix} s \\ s^2 \\ \vdots \\ s^k \end{pmatrix} = 0, \tag{13}$$

because s is any given. Similar to the proof of Lemma 4, we can get $(\sum_{i=1}^{n} p_1(x_i), \sum_{i=1}^{n} p_2(x_i), \dots, \sum_{i=1}^{n} p_k(x_i)) - ((\sum_{i=1}^{n} x_i)^n, (\sum_{i=1}^{n} x_i)^{n-1}, \dots, (\sum_{i=1}^{n} x_i))A = 0$; that is, $((\sum_{i=1}^{n} x_i)^n, (\sum_{i=1}^{n} x_i)^{n-1}, \dots, (\sum_{i=1}^{n} x_i))A = (\sum_{i=1}^{n} p_1(x_i), \sum_{i=1}^{n} p_2(x_i), \dots, \sum_{i=1}^{n} p_k(x_i))$, but $p_1(x_i), p_2(x_i), \dots, p_k(x_i)$ and x_i $(1 \le i \le n)$ also are fixed; thus, α and $A$are only determined, so the number of e_R which is incident with e_U and contained in m is 1. □

Theorem 8. *In the constructed multisender authentication codes, if the senders' encoding rules and the receiver's decoding rules are chosen according to a uniform probability distribution, then the largest probabilities of success for different types of deceptions, respectively, are*

$$P_I = \frac{1}{q^{k(k-1)/2} \prod_{i=n-k+1}^{n} (q^i - 1)},$$
$$P_S = \frac{1}{\varphi(q-1)}, \tag{14}$$
$$P_U(L) = \frac{1}{[q^{k(k-1)/2} \prod_{i=n-k+1}^{n} (q^i - 1)] \varphi(q-1)}.$$

Proof. By Theorem 3 and Lemma 4, we get

$$P_I = \max_{m \in M} \left\{ \frac{|\{e_R \in E_R \mid e_R \subset m\}|}{|E_R|} \right\}$$
$$= \frac{\varphi(q-1)}{[q^{k(k-1)/2} \prod_{i=n-k+1}^{n} (q^i - 1)] \; \varphi(q-1)}$$
$$= \frac{1}{q^{k(k-1)/2} \prod_{i=n-k+1}^{n} (q^i - 1)}. \tag{15}$$

By Lemmas 4 and 5, we get

$$P_S = \max_{m \in M} \left\{ \frac{\max_{m' \neq m \in M} |\{e_R \in E_R \mid e_R \subset m, m'\}|}{|\{e_R \in E_R \mid e_R \subset m\}|} \right\} = \frac{1}{\varphi(q-1)}. \tag{16}$$

By Lemmas 6 and 7, we get

$$P_U(L) = \max_{e_L \in E_L} \max_{e_L \in e_U} \left\{ \frac{\max_{m \in M} |\{e_R \in E_R \mid e_R \subset m, p(e_R, e_P) \neq 0\}|}{|\{e_R \in E_R \mid p(e_R, e_P) \neq 0\}|} \right\}$$
$$= \frac{1}{[q^{k(k-1)/2} \prod_{i=n-k+1}^{n} (q^i - 1)] \varphi(q-1)}. \tag{17}$$

□

Acknowledgments

This paper is supported by the NSF of China (61179026) and the Fundamental Research of the Central Universities of China Civil Aviation University of Science special (ZXH2012k003).

References

[1] E. N. Gilbert, F. J. MacWilliams, and N. J. A. Sloane, "Codes which detect deception," *The Bell System Technical Journal*, vol. 53, pp. 405–424, 1974.

[2] Y. Desmedt, Y. Frankel, and M. Yung, "Multi-receiver/multi-sender network security: efficient authenticated multicast/feedback," in *Proceedings of the the 11th Annual Conference of the IEEE Computer and Communications Societies (Infocom '92)*, pp. 2045–2054, May 1992.

[3] K. Martin and R. Safavi-Naini, "Multisender authentication schemes with unconditional security," in *Information and Communications Security*, vol. 1334 of *Lecture Notes in Computer Science*, pp. 130–143, Springer, Berlin, Germany, 1997.

[4] W. Ma and X. Wang, "Several new constructions of multi trasmitters authentication codes," *Acta Electronica Sinica*, vol. 28, no. 4, pp. 117–119, 2000.

[5] G. J. Simmons, "Message authentication with arbitration of transmitter/receiver disputes," in *Advances in Cryptology—EUROCRYPT '87, Workshop on the Theory and Application of of Cryptographic Techniques*, vol. 304 of *Lecture Notes in Computer Science*, pp. 151–165, Springer, 1988.

[6] S. Cheng and L. Chang, "Two constructions of multi-sender authentication codes with arbitration based linear codes to be published in ," *WSEAS Transactions on Mathematics*, vol. 11, no. 12, 2012.

[7] R. Safavi-Naini and H. Wang, "New results on multi-receiver authentication codes," in *Advances in Cryptology—EUROCRYPT '98 (Espoo)*, vol. 1403 of *Lecture Notes in Comput. Sci.*, pp. 527–541, Springer, Berlin, Germany, 1998.

[8] R. Aparna and B. B. Amberker, "Multi-sender multi-receiver authentication for dynamic secure group communication," *International Journal of Computer Science and Network Security*, vol. 7, no. 10, pp. 47–63, 2007.

[9] R. Safavi-Naini and H. Wang, "Bounds and constructions for multireceiver authentication codes," in *Advances in cryptology—ASIACRYPT'98 (Beijing)*, vol. 1514 of *Lecture Notes in Comput. Sci.*, pp. 242–256, Springer, Berlin, Germany, 1998.

[10] S. Shen and L. Chen, *Information and Coding Theory*, Science press in China, 2002.

[11] J. J. Rotman, *Advanced Modern Algebra*, High Education Press in China, 2004.

[12] Z. Wan, *Geometry of Classical Group over Finite Field*, Science Press in Beijing, New York, NY, USA, 2002.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 435730, 8 pages
http://dx.doi.org/10.1155/2013/435730

Research Article

Geometrical and Spectral Properties of the Orthogonal Projections of the Identity

Luis González, Antonio Suárez, and Dolores García

Department of Mathematics, Research Institute SIANI, University of Las Palmas de Gran Canaria, Campus de Tafira, 35017 Las Palmas de Gran Canaria, Spain

Correspondence should be addressed to Luis González; luisglez@dma.ulpgc.es

Received 28 December 2012; Accepted 10 April 2013

Academic Editor: K. Sivakumar

We analyze the best approximation AN (in the Frobenius sense) to the identity matrix in an arbitrary matrix subspace AS ($A \in \mathbb{R}^{n\times n}$ nonsingular, S being any fixed subspace of $\mathbb{R}^{n\times n}$). Some new geometrical and spectral properties of the orthogonal projection AN are derived. In particular, new inequalities for the trace and for the eigenvalues of matrix AN are presented for the special case that AN is symmetric and positive definite.

1. Introduction

The set of all $n \times n$ real matrices is denoted by $\mathbb{R}^{n\times n}$, and I denotes the identity matrix of order n. In the following, A^T and $\operatorname{tr}(A)$ denote, as usual, the transpose and the trace of matrix $A \in \mathbb{R}^{n\times n}$. The notations $\langle\cdot,\cdot\rangle_F$ and $\|\cdot\|_F$ stand for the Frobenius inner product and matrix norm, defined on the matrix space $\mathbb{R}^{n\times n}$. Throughout this paper, the terms orthogonality, angle, and cosine will be used in the sense of the Frobenius inner product.

Our starting point is the linear system

$$Ax = b, \quad A \in \mathbb{R}^{n\times n}, \ x, b \in \mathbb{R}^n, \tag{1}$$

where A is a large, nonsingular, and sparse matrix. The resolution of this system is usually performed by iterative methods based on Krylov subspaces (see, e.g., [1, 2]). The coefficient matrix A of the system (1) is often extremely ill-conditioned and highly indefinite, so that in this case, Krylov subspace methods are not competitive without a good preconditioner (see, e.g., [2, 3]). Then, to improve the convergence of these Krylov methods, the system (1) can be preconditioned with an adequate nonsingular preconditioning matrix N, transforming it into any of the equivalent systems

$$\begin{gathered} NAx = Nb, \\ ANy = b, \qquad x = Ny, \end{gathered} \tag{2}$$

the so-called left and right preconditioned systems, respectively. In this paper, we address only the case of the right-hand side preconditioned matrices AN, but analogous results can be obtained for the left-hand side preconditioned matrices NA.

The preconditioning of the system (1) is often performed in order to get a preconditioned matrix AN as close as possible to the identity in some sense, and the preconditioner N is called an approximate inverse of A. The closeness of AN to I may be measured by using a suitable matrix norm like, for instance, the Frobenius norm [4]. In this way, the problem of obtaining the best preconditioner N (with respect to the Frobenius norm) of the system (1) in an arbitrary subspace S of $\mathbb{R}^{n\times n}$ is equivalent to the minimization problem; see, for example, [5]

$$\min_{M\in S}\|AM - I\|_F = \|AN - I\|_F. \tag{3}$$

The solution N to the problem (3) will be referred to as the "optimal" or the "best" approximate inverse of matrix A in the subspace S. Since matrix AN is the best approximation to the identity in subspace AS, it will be also referred to as the orthogonal projection of the identity matrix onto the subspace AS. Although many of the results presented in this paper are also valid for the case that matrix N is singular, from now on, we assume that the optimal approximate inverse N (and thus also the orthogonal projection AN) is a nonsingular matrix. The solution N to the problem (3) has been studied as a natural generalization of the classical Moore-Penrose inverse in [6], where it has been referred to as the S-Moore-Penrose inverse of matrix A.

The main goal of this paper is to derive new geometrical and spectral properties of the best approximations AN (in the sense of formula (3)) to the identity matrix. Such properties could be used to analyze the quality and theoretical effectiveness of the optimal approximate inverse N as preconditioner of the system (1). However, it is important to highlight that the purpose of this paper is purely theoretical, and we are not looking for immediate numerical or computational approaches (although our theoretical results could be potentially applied to the preconditioning problem). In particular, the term "optimal (or best) approximate inverse" is used in the sense of formula (3) and not in any other sense of this expression.

Among the many different works dealing with practical algorithms that can be used to compute approximate inverses, we refer the reader to for example, [4, 7–9] and to the references therein. In [4], the author presents an exhaustive survey of preconditioning techniques and, in particular, describes several algorithms for computing sparse approximate inverses based on Frobenius norm minimization like, for instance, the well-known SPAI and FSAI algorithms. A different approach (which is also focused on approximate inverses based on minimizing $\|AM - I\|_F$) can be found in [7], where an iterative descent-type method is used to approximate each column of the inverse, and the iteration is done with "sparse matrix by sparse vector" operations. When the system matrix is expressed in block-partitioned form, some preconditioning options are explored in [8]. In [9], the idea of "target" matrix is introduced, in the context of sparse approximate inverse preconditioners, and the generalized Frobenius norms $\|B\|_{F,H}^2 = \operatorname{tr}(BHB^T)$ (H symmetric positive definite) are used, for minimization purposes, as an alternative to the classical Frobenius norm.

The last results of our work are devoted to the special case that matrix AN is symmetric and positive definite. In this sense, let us recall that the cone of symmetric and positive definite matrices has a rich geometrical structure and, in this context, the angle that any symmetric and positive definite matrix forms with the identity plays a very important role [10]. In this paper, the authors extend this geometrical point of view and analyze the geometrical structure of the subspace of symmetric matrices of order n, including the location of all orthogonal matrices not only the identity matrix.

This paper has been organized as follows. In Section 2, we present some preliminary results required to make the paper self-contained. Sections 3 and 4 are devoted to obtain new geometrical and spectral relations, respectively, for the orthogonal projections AN of the identity matrix. Finally, Section 5 closes the paper with its main conclusions.

2. Some Preliminaries

Now, we present some preliminary results concerning the orthogonal projection AN of the identity onto the matrix subspace $AS \subset \mathbb{R}^{n\times n}$. For more details about these results and for their proofs, we refer the reader to [5, 6, 11].

Taking advantage of the prehilbertian character of the matrix Frobenius norm, the solution N to the problem (3) can be obtained using the orthogonal projection theorem. More precisely, the matrix product AN is the orthogonal projection of the identity onto the subspace AS, and it satisfies the conditions stated by the following lemmas; see [5, 11].

Lemma 1. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Let N be the solution to the problem* (3). *Then,*

$$0 \le \|AN\|_F^2 = \operatorname{tr}(AN) \le n, \tag{4}$$

$$0 \le \|AN - I\|_F^2 = n - \operatorname{tr}(AN) \le n. \tag{5}$$

An explicit formula for matrix N can be obtained by expressing the orthogonal projection AN of the identity matrix onto the subspace AS by its expansion with respect to an orthonormal basis of AS [5]. This is the idea of the following lemma.

Lemma 2. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular. Let S be a linear subspace of $\mathbb{R}^{n\times n}$ of dimension d and $\{M_1, \ldots, M_d\}$ a basis of S such that $\{AM_1, \ldots, AM_d\}$ is an orthogonal basis of AS. Then, the solution N to the problem* (3) *is*

$$N = \sum_{i=1}^{d} \frac{\operatorname{tr}(AM_i)}{\|AM_i\|_F^2} M_i, \tag{6}$$

and the minimum (*residual*) *Frobenius norm is*

$$\|AN - I\|_F^2 = n - \sum_{i=1}^{d} \frac{[\operatorname{tr}(AM_i)]^2}{\|AM_i\|_F^2}. \tag{7}$$

Let us mention two possible options, both taken from [5], for choosing in practice the subspace S and its corresponding basis $\{M_i\}_{i=1}^d$. The first example consists of considering the subspace S of $n\times n$ matrices with a prescribed sparsity pattern, that is,

$$S = \left\{M \in \mathbb{R}^{n\times n} : m_{ij} = 0 \ \forall (i,j) \notin K\right\}, \\ K \subset \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}. \tag{8}$$

Then, denoting by $M_{i,j}$ the $n\times n$ matrix whose only nonzero entry is $m_{ij} = 1$, a basis of subspace S is clearly $\{M_{i,j} : (i,j) \in K\}$, and then $\{AM_{i,j} : (i,j) \in K\}$ will be a basis of subspace AS (since we have assumed that matrix A is nonsingular). In general, this basis of AS is not orthogonal, so that we only need to use the Gram-Schmidt procedure to obtain an orthogonal basis of AS, in order to apply the orthogonal expansion (6).

For the second example, consider a linearly independent set of $n \times n$ real symmetric matrices $\{P_1, \ldots, P_d\}$ and the corresponding subspace

$$S' = \operatorname{span}\left\{P_1 A^T, \ldots, P_d A^T\right\}, \tag{9}$$

which clearly satisfies

$$\begin{aligned} S' &\subseteq \left\{ M = PA^T : P \in \mathbb{R}^{n\times n}\ P^T = P \right\} \\ &= \left\{ M \in \mathbb{R}^{n\times n} : (AM)^T = AM \right\}. \end{aligned} \tag{10}$$

Hence, we can explicitly obtain the solution N to the problem (3) for subspace S', from its basis $\{P_1A^T, \ldots, P_dA^T\}$, as follows. If $\{AP_1A^T, \ldots, AP_dA^T\}$ is an orthogonal basis of subspace AS', then we just use the orthogonal expansion (6) for obtaining N. Otherwise, we use again the Gram-Schmidt procedure to obtain an orthogonal basis of subspace AS', and then we apply formula (6). The interest of this second example stands in the possibility of using the conjugate gradient method for solving the preconditioned linear system, when the symmetric matrix AN is positive definite. For a more detailed exposition of the computational aspects related to these two examples, we refer the reader to [5].

Now, we present some spectral properties of the orthogonal projection AN. From now on, we denote by $\{\lambda_i\}_{i=1}^n$ and $\{\sigma_i\}_{i=1}^n$ the sets of eigenvalues and singular values, respectively, of matrix AN arranged, as usual, in nonincreasing order, that is,

$$\begin{aligned} |\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n| > 0, \\ \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0. \end{aligned} \tag{11}$$

The following lemma [11] provides some inequalities involving the eigenvalues and singular values of the preconditioned matrix AN.

Lemma 3. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Then,*

$$\begin{aligned} \sum_{i=1}^n \lambda_i^2 &\leq \sum_{i=1}^n |\lambda_i|^2 \leq \sum_{i=1}^n \sigma_i^2 = \|AN\|_F^2 \\ &= \operatorname{tr}(AN) = \sum_{i=1}^n \lambda_i \leq \sum_{i=1}^n |\lambda_i| \leq \sum_{i=1}^n \sigma_i. \end{aligned} \tag{12}$$

The following fact [11] is a direct consequence of Lemma 3.

Lemma 4. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Then, the smallest singular value and the smallest eigenvalue's modulus of the orthogonal projection AN of the identity onto the subspace AS are never greater than 1. That is,*

$$0 < \sigma_n \leq |\lambda_n| \leq 1. \tag{13}$$

The following theorem [11] establishes a tight connection between the closeness of matrix AN to the identity matrix and the closeness of $\sigma_n(|\lambda_n|)$ to the unity.

Theorem 5. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Let N be the solution to the problem* (3). *Then,*

$$\begin{aligned} (1 - |\lambda_n|)^2 \leq (1 - \sigma_n)^2 \leq \|AN - I\|_F^2 \\ \leq n\left(1 - |\lambda_n|^2\right) \leq n\left(1 - \sigma_n^2\right). \end{aligned} \tag{14}$$

Remark 6. Theorem 5 states that the closer the smallest singular value σ_n of matrix AN is to the unity, the closer matrix AN will be to the identity, that is, the smaller $\|AN - I\|_F$ will be, and conversely. The same happens with the smallest eigenvalue's modulus $|\lambda_n|$ of matrix AN. In other words, we get a good approximate inverse N of A when σ_n ($|\lambda_n|$) is sufficiently close to 1.

To finish this section, let us mention that, recently, lower and upper bounds on the normalized Frobenius condition number of the orthogonal projection AN of the identity onto the subspace AS have been derived in [12]. In addition, this work proposes a natural generalization (related to an arbitrary matrix subspace S of $\mathbb{R}^{n\times n}$) of the normalized Frobenius condition number of the nonsingular matrix A.

3. Geometrical Properties

In this section, we present some new geometrical properties for matrix AN, N being the optimal approximate inverse of matrix A, defined by (3). Our first lemma states some basic properties involving the cosine of the angle between matrix AN and the identity, that is,

$$\cos(AN, I) = \frac{\langle AN, I\rangle_F}{\|AN\|_F \|I\|_F} = \frac{\operatorname{tr}(AN)}{\|AN\|_F \sqrt{n}}. \tag{15}$$

Lemma 7. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Let N be the solution to the problem* (3). *Then,*

$$\cos(AN, I) = \frac{\operatorname{tr}(AN)}{\|AN\|_F \sqrt{n}} = \frac{\|AN\|_F}{\sqrt{n}} = \frac{\sqrt{\operatorname{tr}(AN)}}{\sqrt{n}}, \tag{16}$$

$$0 \leq \cos(AN, I) \leq 1, \tag{17}$$

$$\|AN - I\|_F^2 = n\left(1 - \cos^2(AN, I)\right). \tag{18}$$

Proof. First, using (15) and (4) we immediately obtain (16). As a direct consequence of (16), we derive that $\cos(AN, I)$ is always nonnegative. Finally, using (5) and (16), we get

$$\|AN - I\|_F^2 = n - \operatorname{tr}(AN) = n\left(1 - \cos^2(AN, I)\right) \tag{19}$$

and the proof is concluded. □

Remark 8. In [13], the authors consider an arbitrary approximate inverse Q of matrix A and derive the following equality:

$$\begin{aligned} \|AQ - I\|_F^2 &= \left(\|AQ\|_F - \|I\|_F\right)^2 \\ &\quad + 2\left(1 - \cos(AQ, I)\right)\|AQ\|_F \|I\|_F, \end{aligned} \tag{20}$$

that is, the typical decomposition (valid in any inner product space) of the strong convergence into the convergence of the norms $(\|AQ\|_F - \|I\|_F)^2$ and the weak convergence $(1 - \cos(AQ, I))\|AQ\|_F\|I\|_F$. Note that for the special case that Q is the optimal approximate inverse N defined by (3), formula (18) has stated that the strong convergence is reduced just to the weak convergence and, indeed, just to the cosine $\cos(AN, I)$.

Remark 9. More precisely, formula (18) states that the closer $\cos(AN, I)$ is to the unity (i.e., the smaller the angle $\angle(AN, I)$ is), the smaller $\|AN - I\|_F$ will be, and conversely. This gives us a new measure of the quality (in the Frobenius sense) of the approximate inverse N of matrix A, by comparing the minimum residual norm $\|AN - I\|_F$ with the cosine of the angle between AN and the identity, instead of with $\mathrm{tr}(AN)$, $\|AN\|_F$ (Lemma 1), or σ_n, $|\lambda_n|$ (Theorem 5). So for a fixed nonsingular matrix $A \in \mathbb{R}^{n\times n}$ and for different subspaces $S \subset \mathbb{R}^{n\times n}$, we have

$$\begin{aligned} \mathrm{tr}(AN) \nearrow n &\Longleftrightarrow \|AN\|_F \nearrow \sqrt{n} \Longleftrightarrow \sigma_n \nearrow 1 \Longleftrightarrow |\lambda_n| \nearrow 1 \\ &\Longleftrightarrow \cos(AN, I) \nearrow 1 \Longleftrightarrow \|AN - I\|_F \searrow 0. \end{aligned} \tag{21}$$

Obviously, the optimal theoretical situation corresponds to the case

$$\begin{aligned} \mathrm{tr}(AN) = n &\Longleftrightarrow \|AN\|_F = \sqrt{n} \Longleftrightarrow \sigma_n = 1 \Longleftrightarrow \lambda_n = 1 \\ &\Longleftrightarrow \cos(AN, I) = 1 \Longleftrightarrow \|AN - I\|_F = 0 \\ &\Longleftrightarrow N = A^{-1} \Longleftrightarrow A^{-1} \in S. \end{aligned} \tag{22}$$

Remark 10. Note that the ratio between $\cos(AN, I)$ and $\cos(A, I)$ is independent of the order n of matrix A. Indeed, assuming that $\mathrm{tr}(A) \neq 0$ and using (16), we immediately obtain

$$\frac{\cos(AN, I)}{\cos(A, I)} = \frac{\|AN\|_F : \sqrt{n}}{\mathrm{tr}(A) : \|A\|_F\sqrt{n}} = \|AN\|_F \frac{\|A\|_F}{\mathrm{tr}(A)}. \tag{23}$$

The following lemma compares the trace and the Frobenius norm of the orthogonal projection AN.

Lemma 11. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Let N be the solution to the problem* (3). *Then,*

$$\begin{aligned} \mathrm{tr}(AN) \leq \|AN\|_F &\Longleftrightarrow \|AN\|_F \leq 1 \Longleftrightarrow \mathrm{tr}(AN) \leq 1 \\ &\Longleftrightarrow \cos(AN, I) \leq \frac{1}{\sqrt{n}}, \end{aligned} \tag{24}$$

$$\begin{aligned} \mathrm{tr}(AN) \geq \|AN\|_F &\Longleftrightarrow \|AN\|_F \geq 1 \Longleftrightarrow \mathrm{tr}(AN) \geq 1 \\ &\Longleftrightarrow \cos(AN, I) \geq \frac{1}{\sqrt{n}}. \end{aligned} \tag{25}$$

Proof. Using (4), we immediately obtain the four leftmost equivalences. Using (16), we immediately obtain the two rightmost equivalences. □

The next lemma provides us with a relationship between the Frobenius norms of the inverses of matrices A and its best approximate inverse N in subspace S.

Lemma 12. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Let N be the solution to the problem* (3). *Then,*

$$\left\|A^{-1}\right\|_F \left\|N^{-1}\right\|_F \geq 1. \tag{26}$$

Proof. Using (4), we get

$$\begin{aligned} \|AN\|_F \leq \sqrt{n} = \|I\|_F &= \left\|(AN)(AN)^{-1}\right\|_F \\ &\leq \|AN\|_F \left\|(AN)^{-1}\right\|_F \Longrightarrow \left\|(AN)^{-1}\right\|_F \geq 1, \end{aligned} \tag{27}$$

and hence

$$\left\|A^{-1}\right\|_F \left\|N^{-1}\right\|_F \geq \left\|N^{-1}A^{-1}\right\|_F = \left\|(AN)^{-1}\right\|_F \geq 1, \tag{28}$$

and the proof is concluded. □

The following lemma compares the minimum residual norm $\|AN - I\|_F$ with the distance (with respect to the Frobenius norm) $\|A^{-1} - N\|_F$ between the inverse of A and the optimal approximate inverse N of A in any subspace $S \subset \mathbb{R}^{n\times n}$. First, note that for any two matrices $A, B \in \mathbb{R}^{n\times n}$ (A nonsingular), from the submultiplicative property of the Frobenius norm, we immediately get

$$\begin{aligned} \|AB - I\|_F^2 &= \left\|A\left(B - A^{-1}\right)\right\|_F^2 \\ &\leq \|A\|_F^2 \left\|B - A^{-1}\right\|_F^2. \end{aligned} \tag{29}$$

However, for the special case that $B = N$ (the solution to the problem (3)), we also get the following inequality.

Lemma 13. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Let N be the solution to the problem* (3). *Then,*

$$\|AN - I\|_F^2 \leq \|A\|_F \left\|N - A^{-1}\right\|_F. \tag{30}$$

Proof. Using the Cauchy-Schwarz inequality and (5), we get

$$\begin{aligned} &\left|\left\langle A^{-1} - N, A^T\right\rangle_F\right| \leq \left\|A^{-1} - N\right\|_F \left\|A^T\right\|_F \\ &\Longrightarrow \left|\mathrm{tr}\left(\left(A^{-1} - N\right)A\right)\right| \leq \left\|A^{-1} - N\right\|_F \left\|A^T\right\|_F \\ &\Longrightarrow \left|\mathrm{tr}\left(A\left(A^{-1} - N\right)\right)\right| \leq \left\|N - A^{-1}\right\|_F \|A\|_F \\ &\Longrightarrow |\mathrm{tr}(I - AN)| \leq \left\|N - A^{-1}\right\|_F \|A\|_F \\ &\Longrightarrow n - \mathrm{tr}(AN) \leq \left\|N - A^{-1}\right\|_F \|A\|_F \\ &\Longrightarrow \|AN - I\|_F^2 \leq \|A\|_F \left\|N - A^{-1}\right\|_F. \end{aligned} \tag{31}$$

□

The following extension of the Cauchy-Schwarz inequality, in a real or complex inner product space $(H, \langle \cdot, \cdot \rangle)$, was obtained by Buzano [14]. For all $a, x, b \in H$, we have

$$|\langle a, x\rangle \cdot \langle x, b\rangle| \le \frac{1}{2}\left(\|a\|\,\|b\| + |\langle a, b\rangle|\right)\|x\|^2. \tag{32}$$

The next lemma provides us with lower and upper bounds on the inner product $\langle AN, B\rangle_F$, for any $n \times n$ real matrix B.

Lemma 14. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Let N be the solution to the problem (3). Then, for every $B \in \mathbb{R}^{n\times n}$, we have*

$$\left|\langle AN, B\rangle_F\right| \le \frac{1}{2}\left(\sqrt{n}\|B\|_F + |\operatorname{tr}(B)|\right). \tag{33}$$

Proof. Using (32) for $a = I$, $x = AN$, $b = B$, and (4), we get

$$\begin{aligned}
&\left|\langle I, AN\rangle_F \cdot \langle AN, B\rangle_F\right| \le \frac{1}{2}\left(\|I\|_F\|B\|_F + \left|\langle I, B\rangle_F\right|\right)\|AN\|_F^2\\
&\Longrightarrow \left|\operatorname{tr}(AN) \cdot \langle AN, B\rangle_F\right|\\
&\le \frac{1}{2}\left(\sqrt{n}\|B\|_F + |\operatorname{tr}(B)|\right)\|AN\|_F^2\\
&\Longrightarrow \left|\langle AN, B\rangle_F\right| \le \frac{1}{2}\left(\sqrt{n}\|B\|_F + |\operatorname{tr}(B)|\right).
\end{aligned} \tag{34}$$

□

The next lemma provides an upper bound on the arithmetic mean of the squares of the n^2 terms in the orthogonal projection AN. By the way, it also provides us with an upper bound on the arithmetic mean of the n diagonal terms in the orthogonal projection AN. These upper bounds (valid for any matrix subspace S) are independent of the optimal approximate inverse N, and thus they are independent of the subspace S and only depend on matrix A.

Lemma 15. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular with $\operatorname{tr}(A) \neq 0$ and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Let N be the solution to the problem (3). Then,*

$$\begin{aligned}
\frac{\|AN\|_F^2}{n^2} &\le \frac{\|A\|_F^2}{[\operatorname{tr}(A)]^2},\\
\frac{\operatorname{tr}(AN)}{n} &\le \frac{n\|A\|_F^2}{[\operatorname{tr}(A)]^2}.
\end{aligned} \tag{35}$$

Proof. Using (32) for $a = A^T$, $x = I$, and $b = AN$, and the Cauchy-Schwarz inequality for $\langle A^T, AN\rangle_F$ and (4), we get

$$\begin{aligned}
&\left|\left\langle A^T, I\right\rangle_F \cdot \langle I, AN\rangle_F\right|\\
&\le \frac{1}{2}\left(\left\|A^T\right\|_F\|AN\|_F + \left|\left\langle A^T, AN\right\rangle_F\right|\right)\|I\|_F^2\\
&\Longrightarrow |\operatorname{tr}(A) \cdot \operatorname{tr}(AN)|\\
&\le \frac{n}{2}\left(\|A\|_F\|AN\|_F + \left|\left\langle A^T, AN\right\rangle_F\right|\right)\\
&\le \frac{n}{2}\left(\|A\|_F\|AN\|_F + \left\|A^T\right\|_F\|AN\|_F\right)\\
&= n\|A\|_F\|AN\|_F\\
&\Longrightarrow |\operatorname{tr}(A)|\,\|AN\|_F^2 \le n\|A\|_F\|AN\|_F\\
&\Longrightarrow \frac{\|AN\|_F}{n} \le \frac{\|A\|_F}{|\operatorname{tr}(A)|}\\
&\Longrightarrow \frac{\|AN\|_F^2}{n^2} \le \frac{\|A\|_F^2}{[\operatorname{tr}(A)]^2} \Longrightarrow \frac{\operatorname{tr}(AN)}{n} \le \frac{n\|A\|_F^2}{[\operatorname{tr}(A)]^2}.
\end{aligned} \tag{36}$$

□

Remark 16. Lemma 15 has the following interpretation in terms of the quality of the optimal approximate inverse N of matrix A in subspace S. The closer the ratio $n\|A\|_F/|\operatorname{tr}(A)|$ is to zero, the smaller $\operatorname{tr}(AN)$ will be, and thus, due to (5), the larger $\|AN - I\|_F$ will be, and this happens for any matrix subspace S.

Remark 17. By the way, from Lemma 15, we obtain the following inequality for any nonsingular matrix $A \in \mathbb{R}^{n\times n}$. Consider any matrix subspace S s.t. $A^{-1} \in S$. Then, $N = A^{-1}$, and using Lemma 15, we get

$$\begin{aligned}
\frac{\|AN\|_F^2}{n^2} = \frac{\|I\|_F^2}{n^2} = \frac{1}{n} &\le \frac{\|A\|_F^2}{[\operatorname{tr}(A)]^2}\\
&\Longrightarrow |\operatorname{tr}(A)| \le \sqrt{n}\|A\|_F.
\end{aligned} \tag{37}$$

4. Spectral Properties

In this section, we present some new spectral properties for matrix AN, N being the optimal approximate inverse of matrix A, defined by (3). Mainly, we focus on the case that matrix AN is symmetric and positive definite. This has been motivated by the following reason. When solving a large nonsymmetric linear system (1) by using Krylov methods, a possible strategy consists of searching for an adequate optimal preconditioner N such that the preconditioned matrix AN is symmetric positive definite [5]. This enables one to use the conjugate gradient method (CG-method), which is, in general, a computationally efficient method for solving the new preconditioned system [2, 15].

Our starting point is Lemma 3, which has established that the sets of eigenvalues and singular values of any orthogonal projection AN satisfy

$$\begin{aligned}
\sum_{i=1}^n \lambda_i^2 \le \sum_{i=1}^n |\lambda_i|^2 &\le \sum_{i=1}^n \sigma_i^2\\
= \sum_{i=1}^n \lambda_i \le \sum_{i=1}^n |\lambda_i| &\le \sum_{i=1}^n \sigma_i.
\end{aligned} \tag{38}$$

Let us particularize (38) for some special cases.

First, note that if AN is normal (i.e., for all $1 \leq i \leq n$: $\sigma_i = |\lambda_i|$ [16]), then (38) becomes

$$\sum_{i=1}^{n}\lambda_i^2 \leq \sum_{i=1}^{n}|\lambda_i|^2 = \sum_{i=1}^{n}\sigma_i^2 = \sum_{i=1}^{n}\lambda_i \leq \sum_{i=1}^{n}|\lambda_i| = \sum_{i=1}^{n}\sigma_i. \tag{39}$$

In particular, if AN is symmetric ($\sigma_i = |\lambda_i| = \pm\lambda_i \in \mathbb{R}$), then (38) becomes

$$\sum_{i=1}^{n}\lambda_i^2 = \sum_{i=1}^{n}|\lambda_i|^2 = \sum_{i=1}^{n}\sigma_i^2 = \sum_{i=1}^{n}\lambda_i \leq \sum_{i=1}^{n}|\lambda_i| = \sum_{i=1}^{n}\sigma_i. \tag{40}$$

In particular, if AN is symmetric and positive definite ($\sigma_i = |\lambda_i| = \lambda_i \in \mathbb{R}^+$), then the equality holds in all (38), that is,

$$\sum_{i=1}^{n}\lambda_i^2 = \sum_{i=1}^{n}|\lambda_i|^2 = \sum_{i=1}^{n}\sigma_i^2 = \sum_{i=1}^{n}\lambda_i = \sum_{i=1}^{n}|\lambda_i| = \sum_{i=1}^{n}\sigma_i. \tag{41}$$

The next lemma compares the traces of matrices AN and $(AN)^2$.

Lemma 18. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Let N be the solution to the problem* (3). *Then,*

(i) *for any orthogonal projection AN*

$$\operatorname{tr}\left((AN)^2\right) \leq \|AN\|_F^2 = \operatorname{tr}(AN), \tag{42}$$

(ii) *for any symmetric orthogonal projection AN*

$$\left\|(AN)^2\right\|_F \leq \operatorname{tr}\left((AN)^2\right) = \|AN\|_F^2 = \operatorname{tr}(AN), \tag{43}$$

(iii) *for any symmetric positive definite orthogonal projection AN*

$$\left\|(AN)^2\right\|_F \leq \operatorname{tr}\left((AN)^2\right) = \|AN\|_F^2 = \operatorname{tr}(AN) \leq [\operatorname{tr}(AN)]^2. \tag{44}$$

Proof. (i) Using (38), we get

$$\sum_{i=1}^{n}\lambda_i^2 \leq \sum_{i=1}^{n}\sigma_i^2 = \sum_{i=1}^{n}\lambda_i. \tag{45}$$

(ii) It suffices to use the obvious fact that $\|(AN)^2\|_F \leq \|AN\|_F^2$ and the following equalities taken from (40):

$$\sum_{i=1}^{n}\lambda_i^2 = \sum_{i=1}^{n}\sigma_i^2 = \sum_{i=1}^{n}\lambda_i. \tag{46}$$

(iii) It suffices to use (43) and the fact that (see, e.g., [17, 18]) if P and Q are symmetric positive definite matrices then $\operatorname{tr}(PQ) \leq \operatorname{tr}(P)\operatorname{tr}(Q)$ for $P = Q = AN$. □

The rest of the paper is devoted to obtain new properties about the eigenvalues of the orthogonal projection AN for the special case that this matrix is symmetric positive definite.

First, let us recall that the smallest singular value and the smallest eigenvalue's modulus of the orthogonal projection AN are never greater than 1 (see Lemma 4). The following theorem establishes the dual result for the largest eigenvalue of matrix AN (symmetric positive definite).

Theorem 19. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Let N be the solution to the problem* (3). *Suppose that matrix AN is symmetric and positive definite. Then, the largest eigenvalue of the orthogonal projection AN of the identity onto the subspace AS is never less than* 1. *That is,*

$$\sigma_1 = \lambda_1 \geq 1. \tag{47}$$

Proof. Using (41), we get

$$\sum_{i=1}^{n}\lambda_i^2 = \sum_{i=1}^{n}\lambda_i \Longrightarrow \lambda_n^2 - \lambda_n = \sum_{i=1}^{n-1}\left(\lambda_i - \lambda_i^2\right). \tag{48}$$

Now, since $\lambda_n \leq 1$ (Lemma 4), then $\lambda_n^2 - \lambda_n \leq 0$. This implies that at least one summand in the rightmost sum in (48) must be less than or equal to zero. Suppose that such summand is the kth one ($1 \leq k \leq n-1$). Since AN is positive definite, then $\lambda_k > 0$, and thus

$$\lambda_k - \lambda_k^2 \leq 0 \Longrightarrow \lambda_k \leq \lambda_k^2 \Longrightarrow \lambda_k \geq 1 \Longrightarrow \lambda_1 \geq 1 \tag{49}$$

and the proof is concluded. □

In Theorem 19, the assumption that matrix AN is positive definite is essential for assuring that $|\lambda_1| \geq 1$, as the following simple counterexample shows. Moreover, from Lemma 4 and Theorem 19, respectively, we have that the smallest and largest eigenvalues of AN (symmetric positive definite) satisfy $\lambda_n \leq 1$ and $\lambda_1 \geq 1$, respectively. Nothing can be asserted about the remaining eigenvalues of the symmetric positive definite matrix AN, which can be greater than, equal to, or less than the unity, as the same counterexample also shows.

Example 20. For $n = 3$, let

$$A_k = \begin{pmatrix} 3 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad k \in \mathbb{R}, \tag{50}$$

let I_3 be identity matrix of order 3, and let S be the subspace of all 3×3 scalar matrices; that is, $S = \operatorname{span}\{I_3\}$. Then the solution N_k to the problem (3) for subspace S can be immediately obtained by using formula (6) as follows:

$$N_k = \frac{\operatorname{tr}(A_k)}{\|A_k\|_F^2} I_3 = \frac{k+4}{k^2+10} I_3 \tag{51}$$

and then we get

$$A_k N_k = \frac{k+4}{k^2+10} A_k = \frac{k+4}{k^2+10}\begin{pmatrix} 3 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{52}$$

Let us arrange the eigenvalues and singular values of matrix A_kN_k, as usual, in nonincreasing order (as shown in (11)).

On one hand, for $k = -2$, we have

$$A_{-2}N_{-2} = \frac{1}{7}\begin{pmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{53}$$

and then

$$\begin{aligned} \sigma_1 &= |\lambda_1| = \frac{3}{7} \\ &\geq \sigma_2 = |\lambda_2| = \frac{2}{7} \\ &\geq \sigma_3 = |\lambda_3| = \frac{1}{7}. \end{aligned} \tag{54}$$

Hence, $A_{-2}N_{-2}$ is indefinite and $\sigma_1 = |\lambda_1| = 3/7 < 1$.

On the other hand, for $1 < k < 3$, we have (see matrix (52))

$$\begin{aligned} \sigma_1 &= \lambda_1 = 3\frac{k+4}{k^2+10} \\ &\geq \sigma_2 = \lambda_2 = k\frac{k+4}{k^2+10} \\ &\geq \sigma_3 = \lambda_3 = \frac{k+4}{k^2+10}, \end{aligned} \tag{55}$$

and then

$$\begin{aligned} k = 2: \quad & \sigma_1 = \lambda_1 = \frac{9}{7} > 1, \\ & \sigma_2 = \lambda_2 = \frac{6}{7} < 1, \\ & \sigma_3 = \lambda_3 = \frac{3}{7} < 1, \\ k = \frac{5}{2}: \quad & \sigma_1 = \lambda_1 = \frac{6}{5} > 1, \\ & \sigma_2 = \lambda_2 = 1, \\ & \sigma_3 = \lambda_3 = \frac{2}{5} < 1, \\ k = \frac{8}{3}: \quad & \sigma_1 = \lambda_1 = \frac{90}{77} > 1, \\ & \sigma_2 = \lambda_2 = \frac{80}{77} > 1, \\ & \sigma_3 = \lambda_3 = \frac{30}{77} < 1. \end{aligned} \tag{56}$$

Hence, for A_kN_k positive definite, we have (depending on k) $\lambda_2 < 1$, $\lambda_2 = 1$, or $\lambda_2 > 1$.

The following corollary improves the lower bound zero on both $\mathrm{tr}(AN)$, given in (4), and $\cos(AN, I)$, given in (17).

Corollary 21. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Let N be the solution to the problem* (3). *Suppose that matrix AN is symmetric and positive definite. Then,*

$$1 \leq \|AN\|_F \leq \mathrm{tr}(AN) = \|AN\|_F^2 \leq n, \tag{57}$$

$$\cos(AN, I) \geq \frac{1}{\sqrt{n}}. \tag{58}$$

Proof. Denote by $\|\cdot\|_2$ the spectral norm. Using the well-known inequality $\|\cdot\|_2 \leq \|\cdot\|_F$ [19], Theorem 19, and (4), we get

$$\begin{aligned} &\|AN\|_F \geq \|AN\|_2 = \sigma_1 = \lambda_1 \geq 1 \\ &\quad\Longrightarrow 1 \leq \|AN\|_F \leq \mathrm{tr}(AN) = \|AN\|_F^2 \leq n. \end{aligned} \tag{59}$$

Finally, (58) follows immediately from (57) and (25). □

Let us mention that an upper bound on all the eigenvalues moduli and on all singular values of any orthogonal projection AN can be immediately obtained from (38) and (4) as follows:

$$\begin{aligned} &\sum_{i=1}^{n}|\lambda_i|^2 \leq \sum_{i=1}^{n}\sigma_i^2 = \|AN\|_F^2 \leq n \\ &\quad\Longrightarrow |\lambda_i|, \quad \sigma_i \leq \sqrt{n}, \ \forall i = 1, 2, \ldots, n. \end{aligned} \tag{60}$$

Our last theorem improves the upper bound given in (60) for the special case that the orthogonal projection AN is symmetric positive definite.

Theorem 22. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n\times n}$. Let N be the solution to the problem* (3). *Suppose that matrix AN is symmetric and positive definite. Then, all the eigenvalues of matrix AN satisfy*

$$\sigma_i = \lambda_i \leq \frac{1+\sqrt{n}}{2} \quad \forall i = 1, 2, \ldots, n. \tag{61}$$

Proof. First, note that the assertion is obvious for the smallest singular value since $|\lambda_n| \leq 1$ for any orthogonal projection AN (Lemma 4). For any eigenvalue of AN, we use the fact that $x - x^2 \leq 1/4$ for all $x > 0$. Then from (41), we get

$$\begin{aligned} &\sum_{i=1}^{n}\lambda_i^2 = \sum_{i=1}^{n}\lambda_i \\ &\quad\Longrightarrow \lambda_1^2 - \lambda_1 = \sum_{i=2}^{n}\left(\lambda_i - \lambda_i^2\right) \leq \frac{n-1}{4} \\ &\quad\Longrightarrow \lambda_1 \leq \frac{1+\sqrt{n}}{2} \Longrightarrow \lambda_i \leq \frac{1+\sqrt{n}}{2} \\ &\qquad \forall i = 1, 2, \ldots, n. \end{aligned} \tag{62}$$

□

5. Conclusion

In this paper, we have considered the orthogonal projection AN (in the Frobenius sense) of the identity matrix onto an arbitrary matrix subspace AS ($A \in \mathbb{R}^{n\times n}$ nonsingular, $S \subset \mathbb{R}^{n\times n}$). Among other geometrical properties of matrix AN, we have established a strong relation between the quality of the approximation $AN \approx I$ and the cosine of the angle $\angle(AN, I)$. Also, the distance between AN and the identity has been related to the ratio $n\|A\|_F/|\operatorname{tr}(A)|$ (which is independent of the subspace S). The spectral analysis has provided lower and upper bounds on the largest eigenvalue of the symmetric positive definite orthogonal projections of the identity.

Acknowledgments

The authors are grateful to the anonymous referee for valuable comments and suggestions, which have improved the earlier version of this paper. This work was partially supported by the "Ministerio de Economía y Competitividad" (Spanish Government), and FEDER, through Grant Contract CGL2011-29396-C03-01.

References

[1] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, Cambridge, Mass, USA, 1994.

[2] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publishing, Boston, Mass, USA, 1996.

[3] S.-L. Wu, F. Chen, and X.-Q. Niu, "A note on the eigenvalue analysis of the SIMPLE preconditioning for incompressible flow," *Journal of Applied Mathematics*, vol. 2012, Article ID 564132, 7 pages, 2012.

[4] M. Benzi, "Preconditioning techniques for large linear systems: a survey," *Journal of Computational Physics*, vol. 182, no. 2, pp. 418–477, 2002.

[5] G. Montero, L. González, E. Flórez, M. D. García, and A. Suárez, "Approximate inverse computation using Frobenius inner product," *Numerical Linear Algebra with Applications*, vol. 9, no. 3, pp. 239–247, 2002.

[6] A. Suárez and L. González, "A generalization of the Moore-Penrose inverse related to matrix subspaces of $C^{n\times m}$," *Applied Mathematics and Computation*, vol. 216, no. 2, pp. 514–522, 2010.

[7] E. Chow and Y. Saad, "Approximate inverse preconditioners via sparse-sparse iterations," *SIAM Journal on Scientific Computing*, vol. 19, no. 3, pp. 995–1023, 1998.

[8] E. Chow and Y. Saad, "Approximate inverse techniques for block-partitioned matrices," *SIAM Journal on Scientific Computing*, vol. 18, no. 6, pp. 1657–1675, 1997.

[9] R. M. Holland, A. J. Wathen, and G. J. Shaw, "Sparse approximate inverses and target matrices," *SIAM Journal on Scientific Computing*, vol. 26, no. 3, pp. 1000–1011, 2005.

[10] R. Andreani, M. Raydan, and P. Tarazaga, "On the geometrical structure of symmetric matrices," *Linear Algebra and Its Applications*, vol. 438, no. 3, pp. 1201–1214, 2013.

[11] L. González, "Orthogonal projections of the identity: spectral analysis and applications to approximate inverse preconditioning," *SIAM Review*, vol. 48, no. 1, pp. 66–75, 2006.

[12] A. Suárez and L. González, "Normalized Frobenius condition number of the orthogonal projections of the identity," *Journal of Mathematical Analysis and Applications*, vol. 400, no. 2, pp. 510–516, 2013.

[13] J.-P. Chehab and M. Raydan, "Geometrical properties of the Frobenius condition number for positive definite matrices," *Linear Algebra and Its Applications*, vol. 429, no. 8-9, pp. 2089–2097, 2008.

[14] M. L. Buzano, "Generalizzazione della diseguaglianza di Cauchy-Schwarz," *Rendiconti del Seminario Matematico Università e Politecnico di Torino*, vol. 31, pp. 405–409, 1974 (Italian).

[15] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, 1952.

[16] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.

[17] E. H. Lieb and W. Thirring, "Inequalities for the moments of the eigenvalues of the Schrödinger Hamiltonian and their relation to Sobolev inequalities," in *Studies in Mathematical Physics: Essays in Honor of Valentine Bargmann*, E. Lieb, B. Simon, and A. Wightman, Eds., pp. 269–303, Princeton University Press, Princeton, NJ, USA, 1976.

[18] Z. Ulukök and R. Türkmen, "On some matrix trace inequalities," *Journal of Inequalities and Applications*, vol. 2010, Article ID 201486, 8 pages, 2010.

[19] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 489295, 6 pages
http://dx.doi.org/10.1155/2013/489295

Research Article

A Parameterized Splitting Preconditioner for Generalized Saddle Point Problems

Wei-Hua Luo[1,2] and Ting-Zhu Huang[1]

[1] *School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China*
[2] *Key Laboratory of Numerical Simulation of Sichuan Province University, Neijiang Normal University, Neijiang, Sichuan 641112, China*

Correspondence should be addressed to Ting-Zhu Huang; tingzhuhuang@126.com

Received 4 January 2013; Revised 31 March 2013; Accepted 1 April 2013

Academic Editor: P. N. Shivakumar

By using Sherman-Morrison-Woodbury formula, we introduce a preconditioner based on parameterized splitting idea for generalized saddle point problems which may be singular and nonsymmetric. By analyzing the eigenvalues of the preconditioned matrix, we find that when α is big enough, it has an eigenvalue at 1 with multiplicity at least n, and the remaining eigenvalues are all located in a unit circle centered at 1. Particularly, when the preconditioner is used in general saddle point problems, it guarantees eigenvalue at 1 with the same multiplicity, and the remaining eigenvalues will tend to 1 as the parameter $\alpha \to 0$. Consequently, this can lead to a good convergence when some GMRES iterative methods are used in Krylov subspace. Numerical results of Stokes problems and Oseen problems are presented to illustrate the behavior of the preconditioner.

1. Introduction

In some scientific and engineering applications, such as finite element methods for solving partial differential equations [1, 2], and computational fluid dynamics [3, 4], we often consider solutions of the generalized saddle point problems of the form

$$\begin{pmatrix} A & B^T \\ -B & C \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times n}$ $(m \le n)$, and $C \in \mathbb{R}^{m\times m}$ are positive semidefinite, $x, f \in \mathbb{R}^n$, and $y, g \in \mathbb{R}^m$. When $C = 0$, (1) is a general saddle point problem which is also a researching object for many authors.

It is well known that when the matrices A, B, and C are large and sparse, the iterative methods are more efficient and attractive than direct methods assuming that (1) has a good preconditioner. In recent years, a lot of preconditioning techniques have arisen for solving linear system; for example, Saad [5] and Chen [6] have comprehensively surveyed some classical preconditioning techniques, including ILU preconditioner, triangular preconditioner, SPAI preconditioner, multilevel recursive Schur complements preconditioner, and sparse wavelet preconditioner. Particularly, many preconditioning methods for saddle problems have been presented recently, such as dimensional splitting (DS) [7], relaxed dimensional factorization (RDF) [8], splitting preconditioner [9], and Hermitian and skew-Hermitian splitting preconditioner [10].

Among these results, Cao et al. [9] have used splitting idea to give a preconditioner for saddle point problems where the matrix A is symmetric and positive definite and B is of full row rank. According to his preconditioner, the eigenvalues of the preconditioned matrix would tend to 1 when the parameter $t \to \infty$. Consequently, just as we have seen from those examples of [9], preconditioner has guaranteed a good convergence when some iterative methods were used.

In this paper, being motivated by [9], we use the splitting idea to present a preconditioner for the system (1), where A may be nonsymmetric and singular (when $\mathrm{rank}(B) < m$). We find that, when the parameter is big enough, the preconditioned matrix has the eigenvalue at 1 with multiplicity at least n, and the remaining eigenvalues are all located in a unit circle centered at 1. Particularly, when the precondidtioner is used in some general saddle point problems (namely, $C = 0$), we see that the multiplicity of the eigenvalue at 1 is also at least n, but the remaining eigenvalues will tend to 1 as the parameter $\alpha \to 0$.

The remainder of the paper is organized as follows. In Section 2, we present our preconditioner based on the splitting idea and analyze the bound of eigenvalues of the preconditioned matrix. In Section 3, we use some numerical examples to show the behavior of the new preconditioner. Finally, we draw some conclusions and outline our future work in Section 4.

2. A Parameterized Splitting Preconditioner

Now we consider using splitting idea with a variable parameter to present a preconditioner for the system (1).

Firstly, it is evident that when $\alpha \neq 0$, the system (1) is equivalent to

$$\begin{pmatrix} A & B^T \\ \frac{-B}{\alpha} & \frac{C}{\alpha} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ \frac{g}{\alpha} \end{pmatrix}. \tag{2}$$

Let

$$M = \begin{pmatrix} A & B^T \\ \frac{-B}{\alpha} & I \end{pmatrix}, \qquad N = \begin{pmatrix} 0 & 0 \\ 0 & I - \frac{C}{\alpha} \end{pmatrix},$$
$$X = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad F = \begin{pmatrix} f \\ \frac{g}{\alpha} \end{pmatrix}. \tag{3}$$

Then the coefficient matrix of (2) can be expressed by $M - N$. Multiplying both sides of system (2) from the left with matrix M^{-1}, we have

$$\left(I - M^{-1}N\right) X = M^{-1}F. \tag{4}$$

Hence, we obtain a preconditioned linear system from (1) using the idea of splitting and the corresponding preconditioner is

$$H = \begin{pmatrix} A & B^T \\ \frac{-B}{\alpha} & I \end{pmatrix}^{-1} \begin{pmatrix} I & 0 \\ 0 & \frac{I}{\alpha} \end{pmatrix}. \tag{5}$$

Now we analyze the eigenvalues of the preconditioned system (4).

Theorem 1. *The preconditioned matrix $I - M^{-1}N$ has an eigenvalue at 1 with multiplicity at least n. The remaining eigenvalues λ satisfy*

$$\lambda = \frac{s_1 + s_2}{s_1 + \alpha}, \tag{6}$$

where $s_1 = \omega^T BA^{-1}B^T\omega$, $s_2 = \omega^T C\omega$, and $\omega \in \mathbb{C}^m$ satisfies

$$\left(\frac{BA^{-1}B^T}{\alpha} + \frac{C}{\alpha}\right)\omega = \lambda\left(I + \frac{BA^{-1}B^T}{\alpha}\right)\omega, \quad \|\omega\| = 1. \tag{7}$$

Proof. Because

$$M^{-1} = \begin{pmatrix} \left(A + \frac{B^TB}{\alpha}\right)^{-1} & -A^{-1}B^T\left(I + \frac{BA^{-1}B^T}{\alpha}\right)^{-1} \\ \frac{B}{\alpha}\left(A + \frac{B^TB}{\alpha}\right)^{-1} & \left(I + \frac{BA^{-1}B^T}{\alpha}\right)^{-1} \end{pmatrix}, \tag{8}$$

we can easily get

$$I - M^{-1}N = \begin{pmatrix} I & A^{-1}B^T\left(I + \frac{BA^{-1}B^T}{\alpha}\right)^{-1}\left(I - \frac{C}{\alpha}\right) \\ 0 & \left(I + \frac{BA^{-1}B^T}{\alpha}\right)^{-1}\left(\frac{BA^{-1}B^T}{\alpha} + \frac{C}{\alpha}\right) \end{pmatrix}, \tag{9}$$

which implies the preconditioned matrix $I - M^{-1}N$ has an eigenvalue at 1 with multiplicity at least n.

For the remaining eigenvalues, let

$$\left(I + \frac{BA^{-1}B^T}{\alpha}\right)^{-1}\left(\frac{BA^{-1}B^T}{\alpha} + \frac{C}{\alpha}\right)\omega = \lambda\omega \tag{10}$$

with $\|\omega\| = 1$; then we have

$$\left(\frac{BA^{-1}B^T}{\alpha} + \frac{C}{\alpha}\right)\omega = \lambda\left(I + \frac{BA^{-1}B^T}{\alpha}\right)\omega. \tag{11}$$

By multiplying both sides of this equality from the left with ω^T, we can get

$$\lambda = \frac{s_1 + s_2}{s_1 + \alpha}. \tag{12}$$

This completes the proof of Theorem 1. □

Remark 2. From Theorem 1, we can get that when parameter α is big enough, the modulus of nonnil eigenvalues λ will be located in interval $(0, 1)$.

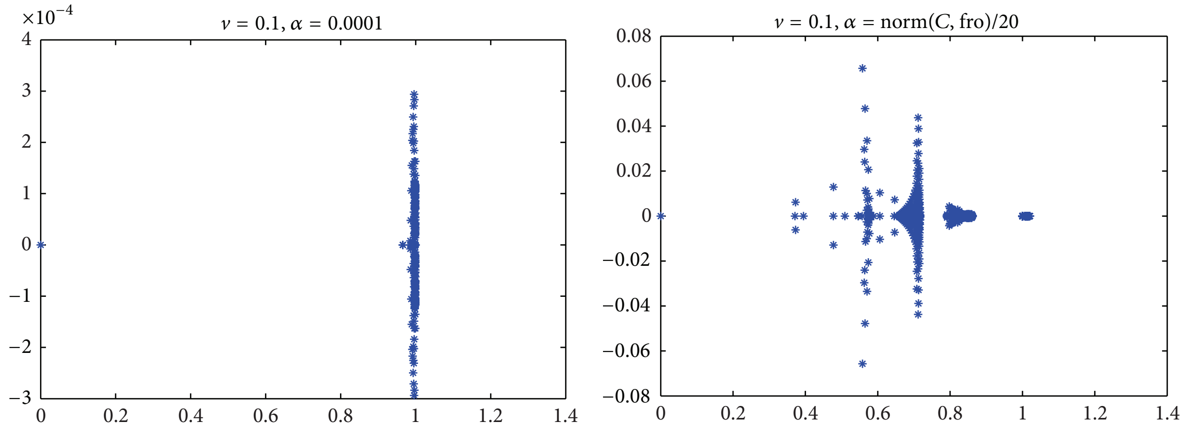
Remark 3. In Theorem 1, if the matrix $C = 0$, then for nonnil eigenvalues we have

$$\lim_{\alpha \to 0} \lambda = 1. \tag{13}$$

Figures 1, 2, and 3 are the eigenvalues plots of the preconditioned matrices obtained with our preconditioner. As we can see in the following numerical experiments, this good phenomenon is useful for accelerating convergence of iterative methods in Krylov subspace.

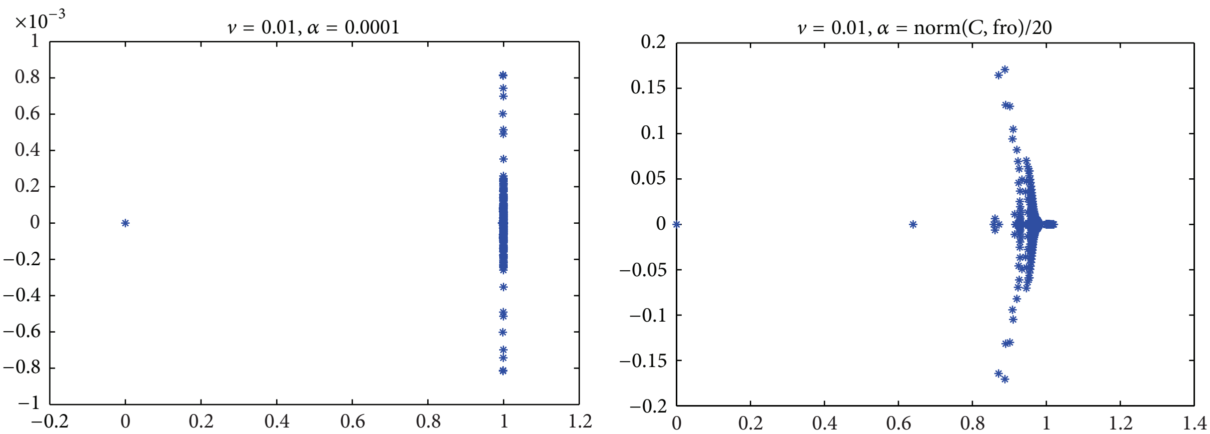
Additionally, for the purpose of practically executing our preconditioner

$$H = \begin{pmatrix} \left(A + \frac{B^TB}{\alpha}\right)^{-1} & \frac{-A^{-1}B^T}{\alpha}\left(I + \frac{BA^{-1}B^T}{\alpha}\right)^{-1} \\ \frac{B}{\alpha}\left(A + \frac{B^TB}{\alpha}\right)^{-1} & \frac{1}{\alpha}\left(I + \frac{BA^{-1}B^T}{\alpha}\right)^{-1} \end{pmatrix}, \tag{14}$$



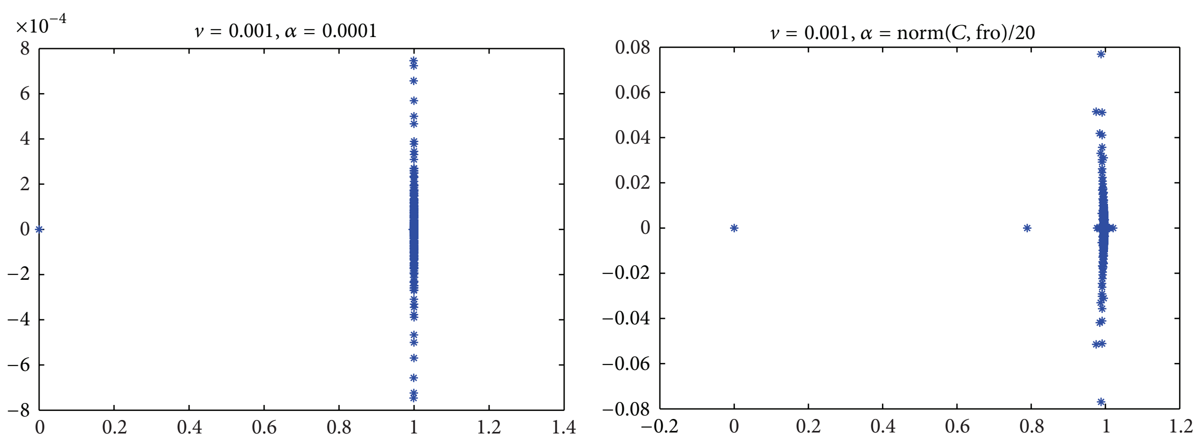
(a) (b)

Figure 1: Spectrum of the preconditioned steady Oseen matrix with viscosity coefficient $\nu = 0.1$, 32×32 grid. (a) Q2-Q1 FEM, $C = 0$; (b) Q1-P0 FEM, $C \neq 0$.



(a) (b)

Figure 2: Spectrum of the preconditioned steady Oseen matrix with viscosity coefficient $\nu = 0.01$, 32×32 grid. (a) Q2-Q1 FEM, $C = 0$; (b) Q1-P0 FEM, $C \neq 0$.



(a) (b)

Figure 3: Spectrum of the preconditioned steady Oseen matrix with viscosity coefficient $\nu = 0.001$, 32×32 grid. (a) Q2-Q1 FEM, $C = 0$; (b) Q1-P0 FEM, $C \neq 0$.

Table 1: Preconditioned GMRES(20) on Stokes problems with different grid sizes (uniform grids).

Grid	α	its	LU time	its time	Total time
16×16	0.0001	4	0.0457	0.0267	0.0724
32×32	0.0001	6	0.3912	0.0813	0.4725
64×64	0.0001	9	5.4472	0.5519	5.9991
128×128	0.0001	14	91.4698	5.7732	97.2430

Table 2: Preconditioned GMRES(20) on steady Oseen problems with different grid sizes (uniform grids), viscosity $\nu = 0.1$.

Grid	α	its	LU time	its time	Total time
16×16	0.0001	3	0.0479	0.0244	0.0723
32×32	0.0001	3	0.6312	0.0696	0.7009
64×64	0.0001	4	13.2959	0.3811	13.6770
128×128	0.0001	6	130.5463	3.7727	134.3190

Table 3: Preconditioned GMRES(20) on steady Oseen problems with different grid sizes (uniform grids), viscosity $\nu = 0.01$.

Grid	α	its	LU time	its time	Total time
16×16	0.0001	2	0.0442	0.0236	0.0678
32×32	0.0001	3	0.4160	0.0557	0.4717
64×64	0.0001	3	7.3645	0.2623	7.6268
128×128	0.0001	4	169.4009	3.1709	172.5718

Table 4: Preconditioned GMRES(20) on steady Oseen problems with different grid sizes (uniform grids), viscosity $\nu = 0.001$.

Grid	α	its	LU time	its time	Total time
16×16	0.0001	2	0.0481	0.0206	0.0687
32×32	0.0001	3	0.4661	0.0585	0.5246
64×64	0.0001	3	6.6240	0.2728	6.8969
128×128	0.0001	4	177.3130	2.9814	180.2944

Table 5: Preconditioned GMRES(20) on Stokes problems with different grid sizes (uniform grids).

Grid	α	its	LU time	its time	Total time
16×16	10000	5	0.0471	0.0272	0.0743
32×32	10000	7	0.3914	0.0906	0.4820
64×64	10000	9	5.6107	0.4145	6.0252
128×128	10000	14	92.7154	4.8908	97.6062

Table 6: Preconditioned GMRES(20) on steady Oseen problems with different grid sizes (uniform grids), viscosity $\nu = 0.1$.

Grid	α	its	LU time	its time	Total time
16×16	10000	4	0.0458	0.0223	0.0681
32×32	10000	4	0.5748	0.0670	0.6418
64×64	10000	5	12.2642	0.4179	12.6821
128×128	10000	7	128.2758	1.6275	129.9033

Table 7: Preconditioned GMRES(20) on steady Oseen problems with different grid sizes (uniform grids), viscosity $\nu = 0.01$.

Grid	α	its	LU time	its time	Total time
16×16	10000	3	0.0458	0.0224	0.0682
32×32	10000	3	0.4309	0.0451	0.4760
64×64	10000	4	7.6537	0.1712	7.8249
128×128	10000	5	175.1587	3.4554	178.6141

Table 8: Preconditioned GMRES(20) on steady Oseen problems with different grid sizes (uniform grids), viscosity $\nu = 0.001$.

Grid	α	its	LU time	its time	Total time
16×16	10000	3	0.0507	0.0212	0.0719
32×32	10000	3	0.4735	0.0449	0.5184
64×64	10000	4	6.6482	0.1645	6.8127
128×128	10000	4	172.0516	2.1216	174.1732

Table 9: Preconditioned GMRES(20) on Stokes problems with different grid sizes (uniform grids).

Grid	α	its	LU time	its time	Total time
16×16	norm(C, fro)	21	0.0240	0.0473	0.0713
32×32	norm(C, fro)	20	0.0890	0.1497	0.2387
64×64	norm(C, fro)	20	0.8387	0.6191	1.4578
128×128	norm(C, fro)	20	6.8917	3.0866	9.9783

Table 10: Preconditioned GMRES(20) on steady Oseen problems with different grid sizes (uniform grids), viscosity $\nu = 0.1$.

Grid	α	its	LU time	its time	Total time
16×16	norm(C, fro)/20	10	0.0250	0.0302	0.0552
32×32	norm(C, fro)/20	10	0.0816	0.0832	0.1648
64×64	norm(C, fro)/20	12	0.8466	0.3648	1.2114
128×128	norm(C, fro)/20	14	6.9019	2.0398	8.9417

Table 11: Preconditioned GMRES(20) on steady Oseen problems with different grid sizes (uniform grids), viscosity $\nu = 0.01$.

Grid	α	its	LU time	its time	Total time
16×16	norm(C, fro)/20	7	0.0238	0.0286	0.0524
32×32	norm(C, fro)/20	7	0.0850	0.0552	0.1402
64×64	norm(C, fro)/20	11	0.8400	0.3177	1.1577
128×128	norm(C, fro)/20	16	6.9537	2.2306	9.1844

Table 12: Preconditioned GMRES(20) on steady Oseen problems with different grid sizes (uniform grids), viscosity $\nu = 0.001$.

Grid	α	its	LU time	its time	Total time
16×16	norm(C, fro)/20	5	0.0245	0.0250	0.0495
32×32	norm(C, fro)/20	5	0.0905	0.0587	0.1492
64×64	norm(C, fro)/20	8	1.2916	0.3200	1.6116
128×128	norm(C, fro)/20	15	10.4399	2.7468	13.1867

Table 13: Size and number of non-nil elements of the coefficient matrix generalized by using Q2-Q1 FEM in steady Stokes problems.

Grid	dim(A)	nnz(A)	dim(B)	nnz(B)	nnz($A + (1/\alpha)B^T B$)
16×16	578×578	6178	81×578	2318	36904
32×32	2178×2178	28418	289×2178	10460	193220
64×64	8450×8450	122206	1089×8450	44314	875966
128×128	33282×33282	506376	4225×33282	184316	3741110

we should efficiently deal with the computation of $(I + (BA^{-1}B^T/\alpha))^{-1}$. This can been tackled by the well-known Sherman-Morrison-Woodbury formula:

$$\left(A_1 + X_1 A_2 X_2^T\right)^{-1} = A_1^{-1} - A_1^{-1} X_1 \left(A_2^{-1} + X_2^T A_1^{-1} X_1\right)^{-1} X_2^T A_1^{-1}, \tag{15}$$

where $A_1 \in \mathbb{R}^{n\times n}$, and $A_2 \in \mathbb{R}^{r_1\times r_1}$ are invertible matrices, $X_1 \in \mathbb{R}^{n\times r_1}$, and $X_2 \in \mathbb{R}^{n\times r_1}$ are any matrices, and n, r_1 are any positive integers.

From (15) we immediately get

$$\left(I + \frac{BA^{-1}B^T}{\alpha}\right)^{-1} = I - B\left(\alpha A + B^T B\right)^{-1} B^T. \tag{16}$$

In the following numerical examples we will always use (16) to compute $(I + (BA^{-1}B^T/\alpha))^{-1}$ in (14).

3. Numerical Examples

In this section, we give numerical experiments to illustrate the behavior of our preconditioner. The numerical experiments are done by using MATLAB 7.1. The linear systems are obtained by using finite element methods in the Stokes problems and steady Oseen problems, and they are respectively the cases of

(1) $C = 0$, which is caused by using Q2-Q1 FEM;

(2) $C \neq 0$, which is caused by using Q1-P0 FEM.

Furthermore, we compare our preconditioner with that of [9] in the case of general saddle point problems (namely, $C = 0$). For the general saddle point problem, [9] has presented the preconditioner

$$\widehat{H} = \begin{pmatrix} A + tB^T B & 0 \\ -2B & \dfrac{I}{t} \end{pmatrix}^{-1} \tag{17}$$

with t as a parameter and has proved that when A is symmetric positive definite, the preconditioned matrix has an eigenvalue 1 with multiplicity at n, and the remaining eigenvalues satisfy

$$\lambda = \frac{t\sigma_i^2}{1 + t\sigma_i^2}, \tag{18}$$

$$\lim_{t\to\infty} \lambda = 1, \tag{19}$$

where σ_i, $i = 1, 2, \ldots, m$, are m positive singular values of the matrix $BA^{-1/2}$.

All these systems can be generalized by using IFISS software package [11] (this is a free package that can be downloaded from the site http://www.maths.manchester.ac.uk/~djs/ifiss/). We use restarted GMRES(20) as the Krylov subspace method, and we always take a zero initial guess. The stopping criterion is

$$\frac{\|r_k\|_2}{\|r_0\|_2} \leq 10^{-6}, \tag{20}$$

where r_k is the residual vector at the kth iteration.

In the whole course of computation, we always replace $(I + (BA^{-1}B^T/\alpha))^{-1}$ in (14) with (16) and use the LU factorization of $A + (B^T B/\alpha)$ to tackle $(A + (B^T B/\alpha))^{-1}v$, where v is a corresponding vector in the iteration. Concretely, let $A + (B^T B/\alpha) = LU$; then we complete the matrix-vector product $(A + (B^T B/\alpha))^{-1}v$ by $U \setminus L \setminus v$ in MATLAB term. In the following tables, the denotation norm (C, fro) means the Frobenius form of the matrix C. The total time is the sum of LU time and iterative time, and the LU time is the time to compute LU factorization of $A + (B^T B/\alpha)$.

Case 1 (*for our preconditioner*). $C = 0$ (using Q2-Q1 FEM in Stokes problems and steady Oseen problems with different viscosity coefficients. The results are in Tables 1, 2, 3, 4).

Case 1' (*for preconditioner of [9]*). $C = 0$ (using Q2-Q1 FEM in Stokes problems and steady Oseen problems with different viscosity coefficients The results are in Tables 5, 6, 7, 8).

Case 2. $C \neq 0$ (using Q1-P0 FEM in Stokes problems and steady Oseen problems with different viscosity coefficients The results are in Tables 9, 10, 11, 12).

From Tables 1, 2, 3, 4, 5, 6, 7, and 8 we can see that these results are in agreement with the theoretical analyses (13) and (19), respectively. Additionally, comparing with the results in Tables 9, 10, 11, and 12, we find that, although the iterations used in Case 1 (either for the preconditioner of [9] or our preconditioner) are less than those in Case 2, the time spent by Case 1 is much more than that of Case 2. This is because the density of the coefficient matrix generalized by Q2-Q1 FEM is much larger than that generalized by Q1-P0 FEM. This can be partly illustrated by Tables 13 and 14, and the others can be illustrated similarly.

Table 14: Size and number of non-nil elements of the coefficient matrix generalized by using Q1-P0 FEM in steady Stokes problems.

Grid	$\dim(A)$	$\mathrm{nnz}(A)$	$\dim(B)$	$\mathrm{nnz}(B)$	$\mathrm{nnz}(A + (1/\alpha)B^T B)$
16×16	578×578	3826	256×578	1800	7076
32×32	2178×2178	16818	1024×2178	7688	31874
64×64	8450×8450	70450	4096×8450	31752	136582
128×128	33282×33282	288306	16384×33282	129032	567192

4. Conclusions

In this paper, we have introduced a splitting preconditioner for solving generalized saddle point systems. Theoretical analysis showed the modulus of eigenvalues of the preconditioned matrix would be located in interval $(0, 1)$ when the parameter is big enough. Particularly when the submatrix $C = 0$, the eigenvalues will tend to 1 as the parameter $\alpha \rightarrow 0$. These performances are tested by some examples, and the results are in agreement with the theoretical analysis.

There are still some future works to be done: how to properly choose a parameter α so that the preconditioned matrix has better properties? How to further precondition submatrix $(I + (BA^{-1}B^T/\alpha))^{-1}((BA^{-1}B^T/\alpha) + (C/\alpha))$ to improve our preconditioner?

Acknowledgments

The authors would express their great thankfulness to the referees and the editor Professor P. N. Shivakumar for their helpful suggestions for revising this paper. The authors would like to thank H. C. Elman, A. Ramage, and D. J. Silvester for their free IFISS software package. This research is supported by Chinese Universities Specialized Research Fund for the Doctoral Program (20110185110020), Sichuan Province Sci. & Tech. Research Project (2012GZX0080), and the Fundamental Research Funds for the Central Universities.

References

[1] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, vol. 15 of *Springer Series in Computational Mathematics*, Springer, New York, NY, USA, 1991.

[2] Z. Chen, Q. Du, and J. Zou, "Finite element methods with matching and nonmatching meshes for Maxwell equations with discontinuous coefficients," *SIAM Journal on Numerical Analysis*, vol. 37, no. 5, pp. 1542–1570, 2000.

[3] H. C. Elman, D. J. Silvester, and A. J. Wathen, *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, NY, USA, 2005.

[4] C. Cuvelier, A. Segal, and A. A. van Steenhoven, *Finite Element Methods and Navier-Stokes Equations*, vol. 22 of *Mathematics and its Applications*, D. Reidel, Dordrecht, The Netherlands, 1986.

[5] Y. Saad, *Iterative Methods for Sparse Linear Systems*, Society for Industrial and Applied Mathematics, Philadelphia, Pa, USA, 2nd edition, 2003.

[6] K. Chen, *Matrix Preconditioning Techniques and Applications*, vol. 19 of *Cambridge Monographs on Applied and Computational Mathematics*, Cambridge University Press, Cambridge, UK, 2005.

[7] M. Benzi and X.-P. Guo, "A dimensional split preconditioner for Stokes and linearized Navier-Stokes equations," *Applied Numerical Mathematics*, vol. 61, no. 1, pp. 66–76, 2011.

[8] M. Benzi, M. Ng, Q. Niu, and Z. Wang, "A relaxed dimensional factorization preconditioner for the incompressible Navier-Stokes equations," *Journal of Computational Physics*, vol. 230, no. 16, pp. 6185–6202, 2011.

[9] Y. Cao, M.-Q. Jiang, and Y.-L. Zheng, "A splitting preconditioner for saddle point problems," *Numerical Linear Algebra with Applications*, vol. 18, no. 5, pp. 875–895, 2011.

[10] V. Simoncini and M. Benzi, "Spectral properties of the Hermitian and skew-Hermitian splitting preconditioner for saddle point problems," *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 2, pp. 377–389, 2004.

[11] H. C. Elman, A. Ramage, and D. J. Silvester, "Algorithm 886: IFISS, a Matlab toolbox for modelling incompressible flow," *ACM Transactions on Mathematical Software*, vol. 33, no. 2, article 14, 2007.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 593549, 11 pages
http://dx.doi.org/10.1155/2013/593549

Research Article

Solving Optimization Problems on Hermitian Matrix Functions with Applications

Xiang Zhang and Shu-Wen Xiang

Department of Computer Science and Information, Guizhou University, Guiyang 550025, China

Correspondence should be addressed to Xiang Zhang; zxjnsc@163.com

Received 17 October 2012; Accepted 20 March 2013

Academic Editor: K. Sivakumar

We consider the extremal inertias and ranks of the matrix expressions $f(X,Y) = A_3 - B_3X - (B_3X)^* - C_3YD_3 - (C_3YD_3)^*$, where $A_3 = A_3^*$, B_3, C_3, and D_3 are known matrices and Y and X are the solutions to the matrix equations $A_1Y = C_1$, $YB_1 = D_1$, and $A_2X = C_2$, respectively. As applications, we present necessary and sufficient condition for the previous matrix function $f(X,Y)$ to be positive (negative), non-negative (positive) definite or nonsingular. We also characterize the relations between the Hermitian part of the solutions of the above-mentioned matrix equations. Furthermore, we establish necessary and sufficient conditions for the solvability of the system of matrix equations $A_1Y = C_1$, $YB_1 = D_1$, $A_2X = C_2$, and $B_3X + (B_3X)^* + C_3YD_3 + (C_3YD_3)^* = A_3$, and give an expression of the general solution to the above-mentioned system when it is solvable.

1. Introduction

Throughout, we denote the field of complex numbers by $\mathbb{C}$, the set of all $m\times n$ matrices over $\mathbb{C}$ by $\mathbb{C}^{m\times n}$, and the set of all $m\times m$ Hermitian matrices by $\mathbb{C}_h^{m\times m}$. The symbols A^* and $\mathscr{R}(A)$ stand for the conjugate transpose, the column space of a complex matrix A respectively. I_n denotes the $n\times n$ identity matrix. The Moore-Penrose inverse [1] $A^\dagger$ of A, is the unique solution X to the four matrix equations:

$$\begin{aligned} &\text{(i)}\ AXA = A,\\ &\text{(ii)}\ XAX = X,\\ &\text{(iii)}\ (AX)^* = AX,\\ &\text{(iv)}\ (XA)^* = XA. \end{aligned} \tag{1}$$

Moreover, L_A and R_A stand for the projectors $L_A = I - A^\dagger A$, $R_A = I - AA^\dagger$ induced by A. It is well known that the eigenvalues of a Hermitian matrix $A\in\mathbb{C}^{n\times n}$ are real, and the inertia of A is defined to be the triplet

$$\mathbb{I}_n(A) = \{i_+(A), i_-(A), i_0(A)\}, \tag{2}$$

where $i_+(A)$, $i_-(A)$, and $i_0(A)$ stand for the numbers of positive, negative, and zero eigenvalues of A, respectively. The symbols $i_+(A)$ and $i_-(A)$ are called the positive index and the negative index of inertia, respectively. For two Hermitian matrices A and B of the same sizes, we say $A \ge B$ ($A \le B$) in the Löwner partial ordering if $A - B$ is positive (negative) semidefinite. The Hermitian part of X is defined as $H(X) = X + X^*$. We will say that X is Re-nnd (Re-nonnegative semidefinite) if $H(X) \ge 0$, X is Re-pd (Re-positive definite) if $H(X) > 0$, and X is Re-ns if $H(X)$ is nonsingular.

It is well known that investigation on the solvability conditions and the general solution to linear matrix equations is very active (e.g., [2–9]). In 1999, Braden [10] gave the general solution to

$$BX + (BX)^* = A. \tag{3}$$

In 2007, Djordjević [11] considered the explicit solution to (3) for linear bounded operators on Hilbert spaces. Moreover, Cao [12] investigated the general explicit solution to

$$BXC + (BXC)^* = A. \tag{4}$$

Xu et al. [13] obtained the general expression of the solution of operator equation (4). In 2012, Wang and He [14] studied some necessary and sufficient conditions for the consistence of the matrix equation

$$A_1X + (A_1X)^* + B_1YC_1 + (B_1YC_1)^* = E_1 \tag{5}$$

and presented an expression of the general solution to (5).

Note that (5) is a special case of the following system:

$$\begin{gathered} A_1Y = C_1, \qquad YB_1 = D_1, \qquad A_2X = C_2, \\ B_3X + (B_3X)^* + C_3YD_3 + (C_3YD_3)^* = A_3. \end{gathered} \tag{6}$$

To our knowledge, there has been little information about (6). One goal of this paper is to give some necessary and sufficient conditions for the solvability of the system of matrix (6) and present an expression of the general solution to system (6) when it is solvable.

In order to find necessary and sufficient conditions for the solvability of the system of matrix equations (6), we need to consider the extremal ranks and inertias of (10) subject to (13) and (11).

There have been many papers to discuss the extremal ranks and inertias of the following Hermitian expressions:

$$p(X) = A_3 - B_3X - (B_3X)^*, \tag{7}$$

$$g(Y) = A - BYC - (BYC)^*, \tag{8}$$

$$h(X, Y) = A_1 - B_1XB_1^* - C_1YC_1^*, \tag{9}$$

$$f(X, Y) = A_3 - B_3X - (B_3X)^* - C_3YD_3 - (C_3YD_3)^*. \tag{10}$$

Tian has contributed much in this field. One of his works [15] considered the extremal ranks and inertias of (7). He and Wang [16] derived the extremal ranks and inertias of (7) subject to $A_1X = C_1$, $A_2XB_2 = C_2$. Liu and Tian [17] studied the extremal ranks and inertias of (8). Chu et al. [18] and Liu and Tian [19] derived the extremal ranks and inertias of (9). Zhang et al. [20] presented the extremal ranks and inertias of (9), where X and Y are Hermitian solutions of

$$A_2X = C_2, \tag{11}$$

$$YB_2 = D_2, \tag{12}$$

respectively. He and Wang [16] derived the extremal ranks and inertias of (10). We consider the extremal ranks and inertias of (10) subject to (11) and

$$A_1Y = C_1, \qquad YB_1 = D_1, \tag{13}$$

which is not only the generalization of the above matrix functions, but also can be used to investigate the solvability conditions for the existence of the general solution to the system (6). Moreover, it can be applied to characterize the relations between Hermitian part of the solutions of (11) and (13).

The remainder of this paper is organized as follows. In Section 2, we consider the extremal ranks and inertias of (10) subject to (11) and (13). In Section 3, we characterize the relations between the Hermitian part of the solution to (11) and (13). In Section 4, we establish the solvability conditions for the existence of a solution to (6) and obtain an expression of the general solution to (6).

2. Extremal Ranks and Inertias of Hermitian Matrix Function (10) with Some Restrictions

In this section, we consider formulas for the extremal ranks and inertias of (10) subject to (11) and (13). We begin with the following Lemmas.

Lemma 1 (see [21]). *(a) Let A_1, C_1, B_1, and D_1 be given. Then the following statements are equivalent:*

(1) *system* (13) *is consistent,*

(2)

$$R_{A_1}C_1 = 0, \qquad D_1L_{B_1} = 0, \qquad A_1D_1 = C_1B_1. \tag{14}$$

(3)

$$\begin{gathered} r\begin{bmatrix} A_1 & C_1 \end{bmatrix} = r(A_1), \\ \begin{bmatrix} D_1 \\ B_1 \end{bmatrix} = r(B_1), \\ A_1D_1 = C_1B_1. \end{gathered} \tag{15}$$

In this case, the general solution can be written as

$$Y = A_1^{\dagger}C_1 + L_{A_1}D_1B_1^{\dagger} + L_{A_1}VR_{B_1}, \tag{16}$$

where V is arbitrary.

(b) Let A_2 and C_2 be given. Then the following statements are equivalent:

(1) *equation* (11) *is consistent,*

(2)

$$R_{A_2}C_2 = 0, \tag{17}$$

(3)

$$r\begin{bmatrix} A_2 & C_2 \end{bmatrix} = r(A_2). \tag{18}$$

In this case, the general solution can be written as

$$X = A^{\dagger}C + L_AW, \tag{19}$$

where W is arbitrary.

Lemma 2 ([22, Lemma 1.5, Theorem 2.3]). *Let $A \in \mathbb{C}_h^{m\times m}$, $B \in \mathbb{C}^{m\times n}$, and $D \in \mathbb{C}_h^{n\times n}$, and denote that*

$$\begin{gathered} M = \begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix}, \\ N = \begin{bmatrix} 0 & Q \\ Q^* & 0 \end{bmatrix}, \\ L = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix}, \\ G = \begin{bmatrix} P & ML_N \\ L_NM^* & 0 \end{bmatrix}. \end{gathered} \tag{20}$$

Then one has the following

(a) *the following equalities hold*

$$i_{\pm}(M) = r(B) + i_{\pm}(R_B A R_B), \tag{21}$$

$$i_{\pm}(N) = r(Q), \tag{22}$$

(b) *if* $\mathscr{R}(B) \subseteq \mathscr{R}(A)$, *then* $i_{\pm}(L) = i_{\pm}(A) + i_{\pm}(D - B^* A^{\dagger} B)$). *Thus* $i_{\pm}(L) = i_{\pm}(A)$ *if and only if* $\mathscr{R}(B) \subseteq \mathscr{R}(A)$ *and* $i_{\pm}(D - B^* A^{\dagger} B) = 0$,

(c)

$$i_{\pm}(G) = \begin{bmatrix} P & M & 0 \\ M^* & 0 & N^* \\ 0 & N & 0 \end{bmatrix} - r(N). \tag{23}$$

Lemma 3 (see [23]). *Let* $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{m\times k}$, *and* $C \in \mathbb{C}^{l\times n}$. *Then they satisfy the following rank equalities:*

(a) $r[A \quad B] = r(A) + r(E_A B) = r(B) + r(E_B A)$,

(b) $r\begin{bmatrix} A \\ C \end{bmatrix} = r(A) + r(CF_A) = r(C) + r(AF_C)$,

(c) $r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r(B) + r(C) + r(E_B A F_C)$,

(d) $r[B \quad AF_C] = r\begin{bmatrix} B & A \\ 0 & C \end{bmatrix} - r(C)$,

(e) $r\begin{bmatrix} C \\ E_B A \end{bmatrix} = r\begin{bmatrix} C & 0 \\ A & B \end{bmatrix} - r(B)$,

(f) $r\begin{bmatrix} A & BF_D \\ E_E C & 0 \end{bmatrix} = r\begin{bmatrix} A & B & 0 \\ C & 0 & E \\ 0 & D & 0 \end{bmatrix} - r(D) - r(E)$,

Lemma 4 (see [15]). *Let* $A \in \mathbb{C}_h^{m\times m}$, $B \in \mathbb{C}^{m\times n}$, $C \in \mathbb{C}_h^{n\times n}$, $Q \in \mathbb{C}^{m\times n}$, *and* $P \in \mathbb{C}^{p\times n}$ *be given, and* $T \in \mathbb{C}^{m\times m}$ *be nonsingular. Then one has the following*

(1) $i_{\pm}(TAT^*) = i_{\pm}(A)$,

(2) $i_{\pm}\begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix} = i_{\pm}(A) + i_{\pm}(C)$,

(3) $i_{\pm}\begin{bmatrix} 0 & Q \\ Q^* & 0 \end{bmatrix} = r(Q)$,

(4) $i_{\pm}\begin{bmatrix} A & BL_P \\ L_P B^* & 0 \end{bmatrix} + r(P) = i_{\pm}\begin{bmatrix} A & B & 0 \\ B^* & 0 & P^* \\ 0 & P & 0 \end{bmatrix}$.

Lemma 5 (see [22, Lemma 1.4]). *Let* S *be a set consisting of (square) matrices over* $\mathbb{C}^{m\times m}$, *and let* H *be a set consisting of (square) matrices over* $\mathbb{C}_h^{m\times m}$. *Then Then one has the following*

(a) S *has a nonsingular matrix if and only if* $\max_{X\in S} r(X) = m$;

(b) *any* $X \in S$ *is nonsingular if and only if* $\min_{X\in S} r(X) = m$;

(c) $\{0\} \in S$ *if and only if* $\min_{X\in S} r(X) = 0$;

(d) $S = \{0\}$ *if and only if* $\max_{X\in S} r(X) = 0$;

(e) H *has a matrix* $X > 0\,(X < 0)$ *if and only if* $\max_{X\in H} i_+(X) = m(\max_{X\in H} i_-(X) = m)$;

(f) *any* $X \in H$ *satisfies* $X > 0\,(X < 0)$ *if and only if* $\min_{X\in H} i_+(X) = m\,(\min_{X\in H} i_-(X) = m)$;

(g) H *has a matrix* $X \geq 0\,(X \leq 0)$ *if and only if* $\min_{X\in H} i_-(X) = 0\,(\min_{X\in H} i_+(X) = 0)$;

(h) *any* $X \in H$ *satisfies* $X \geq 0\,(X \leq 0)$ *if and only if* $\max_{X\in H} i_-(X) = 0\,(\max_{X\in H} i_+(X) = 0)$.

Lemma 6 (see [16]). *Let* $p(X,Y) = A - BX - (BX)^* - CYD - (CYD)^*$, *where* A, B, C, *and* D *are given with appropriate sizes, and denote that*

$$M_1 = \begin{bmatrix} A & B & C \\ B^* & 0 & 0 \\ C^* & 0 & 0 \end{bmatrix}, \quad M_2 = \begin{bmatrix} A & B & D^* \\ B^* & 0 & 0 \\ D & 0 & 0 \end{bmatrix}, \quad M_3 = \begin{bmatrix} A & B & C & D^* \\ B^* & 0 & 0 & 0 \end{bmatrix}, \quad M_4 = \begin{bmatrix} A & B & C & D^* \\ B^* & 0 & 0 & 0 \\ C^* & 0 & 0 & 0 \end{bmatrix}, \quad M_5 = \begin{bmatrix} A & B & C & D^* \\ B^* & 0 & 0 & 0 \\ D & 0 & 0 & 0 \end{bmatrix}. \tag{24}$$

Then one has the following:

(1) *the maximal rank of* $p(X,Y)$ *is*

$$\max_{X\in\mathbb{C}^{n\times m},Y} r[p(X,Y)] = \min\{m, r(M_1), r(M_2), r(M_3)\}, \tag{25}$$

(2) *the minimal rank of* $p(X,Y)$ *is*

$$\begin{aligned}\min_{X\in\mathbb{C}^{n\times m},Y} r[p(X,Y)] &= 2r(M_3) - 2r(B) \\ &\quad + \max\{u_+ + u_-, v_+ + v_-, u_+ + v_-, u_- + v_+\},\end{aligned} \tag{26}$$

(3) *the maximal inertia of* $p(X,Y)$ *is*

$$\max_{X\in\mathbb{C}^{n\times m},Y} i_{\pm}[p(X,Y)] = \min\{i_{\pm}(M_1), i_{\pm}(M_2)\}, \tag{27}$$

(4) *the minimal inertias of* $p(X,Y)$ *is*

$$\begin{aligned}\min_{X\in\mathbb{C}^{n\times m},Y} i_{\pm}[p(X,Y)] &= r(M_3) - r(B) \\ &\quad + \max\{i_{\pm}(M_1) - r(M_4), \\ &\qquad i_{\pm}(M_2) - r(M_5)\},\end{aligned} \tag{28}$$

where

$$u_{\pm} = i_{\pm}(M_1) - r(M_4), \qquad v_{\pm} = i_{\pm}(M_2) - r(M_5). \tag{29}$$

Now we present the main theorem of this section.

Theorem 7. *Let $A_1 \in \mathbb{C}^{m\times n}$, $C_1 \in \mathbb{C}^{m\times k}$, $B_1 \in \mathbb{C}^{k\times l}$, $D_1 \in \mathbb{C}^{n\times l}$, $A_2 \in \mathbb{C}^{t\times q}$, $C_2 \in \mathbb{C}^{t\times p}$, $A_3 \in \mathbb{C}_h^{p\times p}$, $B_3 \in \mathbb{C}^{p\times q}$, $C_3 \in \mathbb{C}^{p\times n}$, and $D_3 \in \mathbb{C}^{p\times n}$ be given, and suppose that the system of matrix equations* (13) *and* (11) *is consistent, respectively. Denote the set of all solutions to* (13) *by S and* (11) *by G. Put*

$$E_1 = \begin{bmatrix} A_3 & C_3 & D_3^*C_1^* & B_3 & C_2^* \\ C_3^* & 0 & A_1^* & 0 & 0 \\ C_1D_3 & A_1 & 0 & 0 & 0 \\ B_3^* & 0 & 0 & 0 & A_2^* \\ C_2 & 0 & 0 & A_2 & 0 \end{bmatrix},$$

$$E_2 = r\begin{bmatrix} A_3 & D_3^* & C_3D_1 & B_3 & C_2^* \\ D_3 & 0 & B_1 & 0 & 0 \\ D_1^*C_3^* & B_1^* & 0 & 0 & 0 \\ B_3^* & 0 & 0 & 0 & A_2^* \\ C_2 & 0 & 0 & A_2 & 0 \end{bmatrix},$$

$$E_3 = \begin{bmatrix} A_3 & B_3 & C_3^* & D_3^* & C_2^* \\ B_3^* & 0 & 0 & 0 & A_2^* \\ D_1^*C_3^* & 0 & 0 & B_1^* & 0 \\ C_1D_3 & 0 & A_1 & 0 & 0 \\ C_2 & A_2 & 0 & 0 & 0 \end{bmatrix}, \quad (30)$$

$$E_4 = \begin{bmatrix} A_3 & B_3 & C_3 & D_3^* & C_2^* & D_3^*C_1^* \\ B_3^* & 0 & 0 & 0 & A_2^* & 0 \\ C_3^* & 0 & 0 & 0 & 0 & A_1^* \\ 0 & 0 & 0 & B_1^* & 0 & 0 \\ C_1D_3 & 0 & A_1 & 0 & 0 & 0 \\ C_2 & A_2 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$E_5 = \begin{bmatrix} A_3 & B_3 & C_3 & D_3^* & C_2^* & C_3D_1 \\ B_3^* & 0 & 0 & 0 & A_2^* & 0 \\ D_3 & 0 & 0 & 0 & 0 & B_1 \\ D_1^*C_3^* & 0 & 0 & A_1^* & 0 & 0 \\ 0 & 0 & A_1 & 0 & 0 & 0 \\ C_2 & A_2 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then one has the following:

(a) *the maximal rank of* (10) *subject to* (13) *and* (11) *is*

$$\begin{aligned} &\max_{X\in G, Y\in S} r[f(X,Y)] \\ &= \min\{p, r(E_1) - 2r(A_1) - 2r(A_2), \\ &\quad r(E_2) - 2r(B_1) - 2r(A_2), \\ &\quad r(E_3) - 2r(A_2) - r(A_1) - r(B_1)\}, \end{aligned} \quad (31)$$

(b) *the minimal rank of* (10) *subject to* (13) *and* (11) *is*

$$\begin{aligned} &\min_{X\in G, Y\in S} r[f(X,Y)] \\ &= 2r(E_3) - 2r\begin{bmatrix} B_3 \\ A_2 \end{bmatrix} \\ &\quad + \max\{r(E_1) - 2r(E_4), r(E_2) - 2r(E_5), \\ &\quad i_+(E_1) + i_-(E_2) - r(E_4) - r(E_5), \\ &\quad i_-(E_1) + i_+(E_2) - r(E_4) - r(E_5)\}, \end{aligned} \quad (32)$$

(c) *the maximal inertia of* (10) *subject to* (13) *and* (11) *is*

$$\begin{aligned} \max_{X\in G, Y\in Sx} i_\pm[f(X,Y)] &= \min\{i_\pm(E_1) - r(A_1) - r(A_2), \\ &\quad i_\pm(E_2) - r(B_1) - r(A_2)\}, \end{aligned} \quad (33)$$

(d) *the minimal inertia of* (10) *subject to* (13) *and* (11) *is*

$$\begin{aligned} \min_{X\in G, Y\in S} i_\pm[f(X,Y)] &= r(E_3) - r\begin{bmatrix} B_3 \\ A_2 \end{bmatrix} \\ &\quad + \max\{i_\pm(E_1) - r(E_4), \\ &\quad i_\pm(E_2) - r(E_5)\}. \end{aligned} \quad (34)$$

Proof. By Lemma 1, the general solutions to (13) and (11) can be written as

$$\begin{aligned} X &= A_2^\dagger C_2 + L_{A_2}W, \\ Y &= A_1^\dagger C_1 + L_{A_1}D_1B_1^\dagger + L_{A_1}ZR_{B_1}, \end{aligned} \quad (35)$$

where W and Z are arbitrary matrices with appropriate sizes. Put

$$\begin{aligned} Q &= B_3L_{A_2}, \qquad T = C_3L_{A_1}, \qquad J = R_{B_1}D_3, \\ P &= A_3 - B_3A_2^\dagger C_2 - \left(B_3A_2^\dagger C_2\right)^* \\ &\quad - C_3\left(A_1^\dagger C_1 + L_{A_1}D_1B_1^\dagger\right)D_3 \\ &\quad - \left(C_3\left(A_1^\dagger C_1 + L_{A_1}D_1B_1^\dagger\right)D_3\right)^*. \end{aligned} \quad (36)$$

Substituting (36) into (10) yields

$$f(X,Y) = P - QW - (QW)^* - TZJ - (TZJ)^*. \quad (37)$$

Clearly P is Hermitian. It follows from Lemma 6 that

$$\begin{aligned} &\max_{X\in G, Y\in S} r[f(X,Y)] \\ &= \max_{W,Z} r(P - QW - (QW)^* - TZJ - (TZJ)^*) \\ &= \min\{m, r(N_1), r(N_2), r(N_3)\}, \end{aligned} \quad (38)$$

$$\begin{aligned} &\min_{X\in G, Y\in S} r[f(X,Y)] \\ &= \max_{W,Z} r(P - QW - (QW)^* - TZJ - (TZJ)^*) \\ &= 2r(N_3) - 2r(Q) \\ &\quad + \max\{s_+ + s_-, t_+ + t_-, s_+ + t_-, s_- + t_+\}, \end{aligned} \quad (39)$$

$$\max_{X\in G,Y\in S} i_{\pm}\left[f(X,Y)\right] = \max_{W,Z} r\left(P-QW-(QW)^{*}-TZJ-(TZJ)^{*}\right) = \min\left\{i_{\pm}(N_1), i_{\pm}(N_2)\right\}, \tag{40}$$

$$\min_{X\in G,Y\in S} i_{\pm}\left[f(X,Y)\right] = \max_{W,Z} r\left(P-QW-(QW)^{*}-TZJ-(TZJ)^{*}\right) = r(N_3)-r(Q)+\max\left\{s_{\pm}, t_{\pm}\right\}, \tag{41}$$

where

$$N_1=\begin{bmatrix} P & Q & T\\ Q^{*} & 0 & 0\\ T^{*} & 0 & 0\end{bmatrix},\quad N_2=\begin{bmatrix} P & Q & J^{*}\\ Q^{*} & 0 & 0\\ J & 0 & 0\end{bmatrix},\quad N_3=\begin{bmatrix} P & Q & T & J^{*}\\ Q^{*} & 0 & 0 & 0\end{bmatrix},$$
$$N_4=\begin{bmatrix} P & Q & T & J^{*}\\ Q^{*} & 0 & 0 & 0\\ T^{*} & 0 & 0 & 0\end{bmatrix},\quad N_5=\begin{bmatrix} P & Q & T & J^{*}\\ Q^{*} & 0 & 0 & 0\\ J & 0 & 0 & 0\end{bmatrix},$$
$$s_{\pm}=i_{\pm}(N_1)-r(N_4),\quad t_{\pm}=i_{\pm}(N_2)-r(N_5). \tag{42}$$

Now, we simplify the ranks and inertias of block matrices in (38)–(41).

By Lemma 4, block Gaussian elimination, and noting that

$$L_S^{*}=\left(I-S^{\dagger}S\right)^{*}=I-S^{*}\left(S^{*}\right)^{\dagger}=R_{S^{*}}, \tag{43}$$

we have the following:

$$r(N_1)=r\begin{bmatrix} P & Q & T\\ Q^{*} & 0 & 0\\ T^{*} & 0 & 0\end{bmatrix} = r\begin{bmatrix} A_3 & C_3 & D_3^{*}C_1^{*} & B_3 & C_2^{*}\\ C_3^{*} & 0 & A_1^{*} & 0 & 0\\ C_1D_3 & A_1 & 0 & 0 & 0\\ B_3^{*} & 0 & 0 & 0 & A_2^{*}\\ C_2 & 0 & 0 & A_2 & 0\end{bmatrix} -2r(A_1)-2r(A_2). \tag{44}$$

By $C_1B_1=A_1D_1$, we obtain

$$r(N_2)=r\begin{bmatrix} P & Q & J^{*}\\ Q^{*} & 0 & 0\\ J & 0 & 0\end{bmatrix} = r\begin{bmatrix} A_3 & D_3^{*} & C_3D_1 & B_3 & C_2^{*}\\ D_3 & 0 & B_1 & 0 & 0\\ D_1^{*}C_3^{*} & B_1^{*} & 0 & 0 & 0\\ B_3^{*} & 0 & 0 & 0 & A_2^{*}\\ C_2 & 0 & 0 & A_2 & 0\end{bmatrix} -2r(B_1)-2r(A_2),$$

$$r(N_3)=r\begin{bmatrix} P & Q & T & J^{*}\\ Q^{*} & 0 & 0 & 0\end{bmatrix} = r\begin{bmatrix} A_3 & B_3 & C_3^{*} & D_3^{*} & C_2^{*}\\ B_3^{*} & 0 & 0 & 0 & A_2^{*}\\ D_1^{*}C_3^{*} & 0 & 0 & B_1^{*} & 0\\ C_1D_3 & 0 & A_1 & 0 & 0\\ C_2 & A_2 & 0 & 0 & 0\end{bmatrix} -r(B_1)-2r(A_2)-r(A_1),$$

$$r(N_4)=r\begin{bmatrix} P & Q & T & J^{*}\\ Q^{*} & 0 & 0 & 0\\ T^{*} & 0 & 0 & 0\end{bmatrix} = r\begin{bmatrix} A_3 & B_3 & C_3 & D_3^{*} & C_2^{*} & D_3^{*}C_1^{*}\\ B_3^{*} & 0 & 0 & 0 & A_2^{*} & 0\\ C_3^{*} & 0 & 0 & 0 & 0 & A_1^{*}\\ 0 & 0 & 0 & B_1^{*} & 0 & 0\\ C_1D_3 & 0 & A_1 & 0 & 0 & 0\\ C_2 & A_2 & 0 & 0 & 0 & 0\end{bmatrix} -r(B_1)-2r(A_2)-2r(A_1),$$

$$r(N_5)=r\begin{bmatrix} P & Q & T & J^{*}\\ Q^{*} & 0 & 0 & 0\\ T^{*} & 0 & 0 & 0\end{bmatrix} = r\begin{bmatrix} A_3 & B_3 & C_3 & D_3^{*} & C_2^{*} & C_3D_1\\ B_3^{*} & 0 & 0 & 0 & A_2^{*} & 0\\ D_3 & 0 & 0 & 0 & 0 & B_1\\ D_1^{*}C_3^{*} & 0 & 0 & A_1^{*} & 0 & 0\\ 0 & 0 & A_1 & 0 & 0 & 0\\ C_2 & A_2 & 0 & 0 & 0 & 0\end{bmatrix} -2r(B_1)-2r(A_2)-r(A_1). \tag{45}$$

By Lemma 2, we can get the following:

$$i_{\pm}(N_1)=i_{\pm}\begin{bmatrix} P & Q & T\\ Q^{*} & 0 & 0\\ T^{*} & 0 & 0\end{bmatrix}$$

$$= i_{\pm}\begin{bmatrix} A_3 & C_3 & D_3^*C_1^* & B_3 & C_2^* \\ C_3^* & 0 & A_1^* & 0 & 0 \\ C_1D_3 & A_1 & 0 & 0 & 0 \\ B_3^* & 0 & 0 & 0 & A_2^* \\ C_2 & 0 & 0 & A_2 & 0 \end{bmatrix} - r(A_1) - r(A_2), \tag{46}$$

$$i_{\pm}(N_2) = i_{\pm}\begin{bmatrix} P & Q & J^* \\ Q^* & 0 & 0 \\ J & 0 & 0 \end{bmatrix} = i_{\pm}\begin{bmatrix} A_3 & D_3^* & C_3D_1 & B_3 & C_2^* \\ D_3 & 0 & B_1 & 0 & 0 \\ D_1^*C_3^* & B_1^* & 0 & 0 & 0 \\ B_3^* & 0 & 0 & 0 & A_2^* \\ C_2 & 0 & 0 & A_2 & 0 \end{bmatrix} - r(B_1) - r(A_2). \tag{47}$$

Substituting (44)-(47) into (38) and (41) yields (31)–(34), respectively. □

Corollary 8. *Let A_1, C_1, B_1, D_1, A_2, C_2, A_3, B_3, C_3, D_3, and E_i, $(i = 1, 2, \ldots, 5)$ be as in Theorem 7, and suppose that the system of matrix equations* (13) *and* (11) *is consistent, respectively. Denote the set of all solutions to* (13) *by S and* (11) *by G. Then, one has the following:*

(a) *there exist $X \in G$ and $Y \in S$ such that $A_3 - B_3X - (B_3X)^* - C_3YD_3 - (C_3YD_3)^* > 0$ if and only if*

$$i_+(E_1) - r(A_1) - r(A_2) \geq p, \quad i_+(E_2) - r(B_1) - r(A_2) \geq p. \tag{48}$$

(b) *there exist $X \in G$ and $Y \in S$ such that $A_3 - B_3X - (B_3X)^* - C_3YD_3 - (C_3YD_3)^* < 0$ if and only if*

$$i_-(E_1) - r(A_1) - r(A_2) \geq p, \quad i_-(E_2) - r(B_1) - r(A_2) \geq p, \tag{49}$$

(c) *there exist $X \in G$ and $Y \in S$ such that $A_3 - B_3X - (B_3X)^* - C_3YD_3 - (C_3YD_3)^* \geq 0$ if and only if*

$$r(E_3) - r\begin{bmatrix} B_3 \\ A_2 \end{bmatrix} + i_-(E_1) - r(E_4) \leq 0, \quad r(E_3) - r\begin{bmatrix} B_3 \\ A_2 \end{bmatrix} + i_-(E_2) - r(E_5) \leq 0. \tag{50}$$

(d) *there exist $X \in G$ and $Y \in S$ such that $A_3 - B_3X - (B_3X)^* - C_3YD_3 - (C_3YD_3)^* \leq 0$ if and only if*

$$r(E_3) - r\begin{bmatrix} B_3 \\ A_2 \end{bmatrix} + i_+(E_1) - r(E_4) \leq 0, \quad r(E_3) - r\begin{bmatrix} B_3 \\ A_2 \end{bmatrix} + i_+(E_2) - r(E_5) \leq 0, \tag{51}$$

(e) *$A_3 - B_3X - (B_3X)^* - C_3YD_3 - (C_3YD_3)^* > 0$ for all $X \in G$ and $Y \in S$ if and only if*

$$r(E_3) - r\begin{bmatrix} B_3 \\ A_2 \end{bmatrix} + i_+(E_1) - r(E_4) = p \quad \text{or} \quad r(E_3) - r\begin{bmatrix} B_3 \\ A_2 \end{bmatrix} + i_+(E_2) - r(E_5) = p, \tag{52}$$

(f) *$A_3 - B_3X - (B_3X)^* - C_3YD_3 - (C_3YD_3)^* < 0$ for all $X \in G$ and $Y \in S$ if and only if*

$$r(E_3) - r\begin{bmatrix} B_3 \\ A_2 \end{bmatrix} + i_-(E_1) - r(E_4) = p \quad \text{or} \quad r(E_3) - r\begin{bmatrix} B_3 \\ A_2 \end{bmatrix} + i_-(E_2) - r(E_5) = p, \tag{53}$$

(g) *$A_3 - B_3X - (B_3X)^* - C_3YD_3 - (C_3YD_3)^* \geq 0$ for all $X \in G$ and $Y \in S$ if and only if*

$$i_-(E_1) - r(A_1) - r(A_2) \leq 0 \quad \text{or} \quad i_-(E_2) - r(B_1) - r(A_2) \leq 0, \tag{54}$$

(h) *$A_3 - B_3X - (B_3X)^* - C_3YD_3 - (C_3YD_3)^* \leq 0$ for all $X \in G$ and $Y \in S$ if and only if*

$$i_+(E_1) - r(A_1) - r(A_2) \leq 0 \quad \text{or} \quad i_+(E_2) - r(B_1) - r(A_2) \leq 0, \tag{55}$$

(i) *there exist $X \in G$ and $Y \in S$ such that $A_3 - B_3X - (B_3X)^* - C_3YD_3 - (C_3YD_3)^*$ is nonsingular if and only if*

$$r(E_1) - 2r(A_1) - 2r(A_2) \geq p, \quad r(E_2) - 2r(B_1) - 2r(A_2) \geq p, \quad r(E_3) - 2r(A_2) - r(A_1) - r(B_1) \geq p. \tag{56}$$

3. Relations between the Hermitian Part of the Solutions to (13) and (11)

Now we consider the extremal ranks and inertias of the difference between the Hermitian part of the solutions to (13) and (11).

Theorem 9. *Let $A_1 \in \mathbb{C}^{m\times p}$, $C_1 \in \mathbb{C}^{m\times p}$, $B_1 \in \mathbb{C}^{p\times l}$, $D_1 \in \mathbb{C}^{p\times l}$, $A_2 \in \mathbb{C}^{t\times p}$, and $C_2 \in \mathbb{C}^{t\times p}$, be given. Suppose that the system of matrix equations* (13) *and* (11) *is consistent,* *respectively. Denote the set of all solutions to* (13) *by S and* (11) *by G. Put*

$$H_1 = \begin{bmatrix} 0 & I & C_1^* & -I & C_2^* \\ I & 0 & A_1^* & 0 & 0 \\ C_1 & A_1 & 0 & 0 & 0 \\ -I & 0 & 0 & 0 & A_2^* \\ C_2 & 0 & 0 & A_2 & 0 \end{bmatrix},$$

$$H_2 = r\begin{bmatrix} 0 & I & D_1 & -I & C_2^* \\ I & 0 & B_1 & 0 & 0 \\ D_1^* & B_1^* & 0 & 0 & 0 \\ -I & 0 & 0 & 0 & A_2^* \\ C_2 & 0 & 0 & A_2 & 0 \end{bmatrix},$$

$$H_3 = \begin{bmatrix} 0 & -I & I & I & C_2^* \\ -I & 0 & 0 & 0 & A_2^* \\ D_1^* & 0 & 0 & B_1^* & 0 \\ C_1 & 0 & A_1 & 0 & 0 \\ C_2 & A_2 & 0 & 0 & 0 \end{bmatrix}, \quad (57)$$

$$H_4 = \begin{bmatrix} 0 & -I & I & I & C_2^* & C_1^* \\ -I & 0 & 0 & 0 & A_2^* & 0 \\ I & 0 & 0 & 0 & 0 & A_1^* \\ 0 & 0 & 0 & B_1^* & 0 & 0 \\ C_1 & 0 & A_1 & 0 & 0 & 0 \\ C_2 & A_2 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$H_5 = \begin{bmatrix} 0 & -I & I & I & C_2^* & D_1 \\ -I & 0 & 0 & 0 & A_2^* & 0 \\ I & 0 & 0 & 0 & 0 & B_1 \\ D_1^* & 0 & 0 & A_1^* & 0 & 0 \\ 0 & 0 & A_1 & 0 & 0 & 0 \\ C_2 & A_2 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then one has the following:

$$\begin{aligned} &\max_{X\in G, Y\in S} r\left[(X+X^*)-(Y+Y^*)\right] \\ &\quad = \min\{p, r(H_1)-2r(A_1)-2r(A_2), r(H_2)-2r(B_1) \\ &\qquad -2r(A_2), r(H_3)-2r(A_2)-r(A_1)-r(B_1)\}, \\ &\min_{X\in G, Y\in S} r\left[(X+X^*)-(Y+Y^*)\right] \\ &\quad = 2r(H_3)-2p \\ &\qquad + \max\{r(H_1)-2r(H_4), r(H_2)-2r(H_5), \\ &\qquad\quad i_+(H_1)+i_-(H_2)-r(H_4)-r(H_5), \\ &\qquad\quad i_-(H_1)+i_+(H_2)-r(H_4)-r(H_5)\}, \\ &\max_{X\in G, Y\in S} i_\pm\left[(X+X^*)-(Y+Y^*)\right] \\ &\quad = \min\{i_\pm(H_1)-r(A_1)-r(A_2), \\ &\qquad i_\pm(H_2)-r(B_1)-r(A_2)\}, \\ &\min_{X\in G, Y\in S} i_\pm\left[(X+X^*)-(Y+Y^*)\right] \\ &\quad = r(E_3)-p+\max\{i_\pm(H_1)-r(H_4), i_\pm(H_2)-r(H_5)\}. \end{aligned} \quad (58)$$

Proof. By letting $A_3 = 0$, $B_3 = -I$, $C_3 = I$, and $D_3 = I$ in Theorem 7, we can get the results. □

Corollary 10. *Let A_1, C_1, B_1, D_1, A_2, C_2, and H_i, ($i = 1, 2, \ldots, 5$) be as in Theorem 9, and suppose that the system of matrix equations* (13) *and* (11) *is consistent, respectively. Denote the set of all solutions to* (13) *by S and* (11) *by G. Then, one has the following:*

(a) *there exist $X \in G$ and $Y \in S$ such that $(X+X^*) > (Y+Y^*)$ if and only if*

$$\begin{aligned} i_+(H_1)-r(A_1)-r(A_2) &\ge p, \\ i_+(H_2)-r(B_1)-r(A_2) &\ge p. \end{aligned} \quad (59)$$

(b) *there exist $X \in G$ and $Y \in S$ such that $(X+X^*) < (Y+Y^*)$ if and only if*

$$\begin{aligned} i_-(H_1)-r(A_1)-r(A_2) &\ge p, \\ i_-(H_2)-r(B_1)-r(A_2) &\ge p, \end{aligned} \quad (60)$$

(c) *there exist $X \in G$ and $Y \in S$ such that $(X+X^*) \ge (Y+Y^*)$ if and only if*

$$\begin{aligned} r(H_3)-p+i_-(H_1)-r(H_4) &\le 0, \\ r(H_3)-p+i_-(H_1)-r(H_4) &\le 0, \end{aligned} \quad (61)$$

(d) *there exist $X \in G$ and $Y \in S$ such that $(X+X^*) \le (Y+Y^*)$ if and only if*

$$\begin{aligned} r(H_3)-p+i_+(H_1)-r(H_4) &\le 0, \\ r(H_3)-p+i_+(H_2)-r(H_5) &\le 0, \end{aligned} \quad (62)$$

(e) $(X+X^*) > (Y+Y^*)$ *for all $X \in G$ and $Y \in S$ if and only if*

$$\begin{aligned} &r(H_3)-p+i_+(H_1)-r(H_4) = p \\ &\text{or} \quad r(H_3)-p+i_+(H_2)-r(H_5) = p, \end{aligned} \quad (63)$$

(f) $(X+X^*) < (Y+Y^*)$ *for all $X \in G$ and $Y \in S$ if and only if*

$$\begin{aligned} &r(H_3)-p+i_-(H_1)-r(H_4) = p \\ &\text{or} \quad r(H_3)-p+i_-(H_2)-r(H_5) = p, \end{aligned} \quad (64)$$

(g) $(X+X^*) \ge (Y+Y^*)$ *for all $X \in G$ and $Y \in S$ if and only if*

$$\begin{aligned} &i_-(H_1)-r(A_1)-r(A_2) \le 0 \\ &\text{or} \quad i_-(H_2)-r(B_1)-r(A_2) \le 0, \end{aligned} \quad (65)$$

(h) $(X + X^*) \leq (Y + Y^*)$ *for all* $X \in G$ *and* $Y \in S$ *if and only if*

$$i_+(H_1) - r(A_1) - r(A_2) \leq 0 \quad \text{or} \quad i_+(H_2) - r(B_1) - r(A_2) \leq 0, \tag{66}$$

(i) *there exist* $X \in G$ *and* $Y \in S$ *such that* $(X + X^*) - (Y + Y^*)$ *is nonsingular if and only if*

$$\begin{gathered} r(H_1) - 2r(A_1) - 2r(A_2) \geq p, \\ r(H_2) - 2r(B_1) - 2r(A_2) \geq p, \\ r(H_3) - 2r(A_2) - r(A_1) - r(B_1) \geq p. \end{gathered} \tag{67}$$

4. The Solvability Conditions and the General Solution to System (6)

We now turn our attention to (6). We in this section use Theorem 9 to give some necessary and sufficient conditions for the existence of a solution to (6) and present an expression of the general solution to (6). We begin with a lemma which is used in the latter part of this section.

Lemma 11 (see [14]). *Let* $A_1 \in \mathbb{C}^{m\times n_1}$, $B_1 \in \mathbb{C}^{m\times n_2}$, $C_1 \in \mathbb{C}^{q\times m}$, *and* $E_1 \in \mathbb{C}_h^{m\times m}$ *be given. Let* $A = R_{A_1}B_1$, $B = C_1R_{A_1}$, $E = R_{A_1}E_1R_{A_1}$, $M = R_AB^*$, $N = A^*L_B$, *and* $S = B^*L_M$. *Then the following statements are equivalent:*

(1) *equation* (5) *is consistent,*

(2)

$$R_MR_AE = 0, \qquad R_AER_A = 0, \qquad L_BEL_B = 0, \tag{68}$$

(3)

$$\begin{gathered} r\begin{bmatrix} E_1 & B_1 & C_1^* & A_1 \\ A_1^* & 0 & 0 & 0 \end{bmatrix} = r\begin{bmatrix} B_1 & C_1^* & A_1 \end{bmatrix} + r(A_1), \\ r\begin{bmatrix} E_1 & B_1 & A_1 \\ A_1^* & 0 & 0 \\ B_1^* & 0 & 0 \end{bmatrix} = 2r\begin{bmatrix} B_1 & A_1 \end{bmatrix}, \\ r\begin{bmatrix} E_1 & C_1^* & A_1 \\ A_1^* & 0 & 0 \\ C_1 & 0 & 0 \end{bmatrix} = 2r\begin{bmatrix} C_1^* & A_1 \end{bmatrix}. \end{gathered} \tag{69}$$

In this case, the general solution of (5) *can be expressed as*

$$\begin{aligned} Y = {} & \frac{1}{2}\left[A^\dagger EB^\dagger - A^\dagger B^* M^\dagger EB^\dagger - A^\dagger S\left(B^\dagger\right)^* EN^\dagger A^* B^\dagger \right. \\ & \left. + A^\dagger E\left(M^\dagger\right)^* + \left(N^\dagger\right)^* EB^\dagger S^\dagger S\right] + L_AV_1 + V_2R_B \\ & + U_1L_SL_M + R_NU_2^*L_M - A^\dagger SU_2R_NA^*B^\dagger, \\ X = {} & A_1^\dagger\left[E_1 - B_1YC_1 - (B_1YC_1)^*\right] \\ & - \frac{1}{2}A_1^\dagger\left[E_1 - B_1YC_1 - (B_1YC_1)^*\right]A_1A_1^\dagger \\ & - A_1^\dagger W_1A_1^* + W_1^*A_1A_1^\dagger + L_{A_1}W_2, \end{aligned} \tag{70}$$

where U_1, U_2, V_1, V_2, W_1, *and* W_2 *are arbitrary matrices over* $\mathbb{C}$ *with appropriate sizes.*

Now we give the main theorem of this section.

Theorem 12. *Let* A_i, C_i, $(i = 1, 2, 3)$, B_j, *and* D_j, $(j = 1, 3)$ *be given. Set*

$$\begin{gathered} A = B_3L_{A_2}, \qquad B = C_3L_{A_1}, \\ C = R_{B_1}D_3, \qquad F = R_AB, \\ G = CR_A, \qquad M = R_FG^*, \\ N = F^*L_G, \qquad S = G^*L_M, \end{gathered} \tag{71}$$

$$\begin{aligned} D = {} & A_3 - B_3A_2^\dagger C_2 - \left(B_3A_2^\dagger C_2\right)^* \\ & - C_3\left(A_1^\dagger C_1 + L_{A_1}D_1B_1^\dagger\right)D_3 \\ & - D_3^*\left(A_1^\dagger C_1 + L_{A_1}D_1B_1^\dagger\right)^*C_3^*, \end{aligned} \tag{72}$$

$$E = R_ADR_A. \tag{73}$$

Then the following statements are equivalent:

(1) *system* (6) *is consistent,*

(2) *the equalities in* (14) *and* (17) *hold, and*

$$R_MR_FE = 0, \qquad R_FER_F = 0, \qquad L_GEL_G = 0, \tag{74}$$

(3) *the equalities in* (15) *and* (18) *hold, and*

$$\begin{gathered} r\begin{bmatrix} A_3 & C_3 & D_3^* & B_3 & C_2^* \\ B_3^* & 0 & 0 & 0 & A_2^* \\ C_1D_3 & A_1 & 0 & 0 & 0 \\ D_1^*C_3^* & 0 & B_1^* & 0 & 0 \\ C_2 & 0 & 0 & A_2 & 0 \end{bmatrix} = r\begin{bmatrix} C_3 & D_3^* & B_3 \\ A_1 & 0 & 0 \\ 0 & B_1^* & 0 \\ 0 & 0 & A_2 \end{bmatrix} + r\begin{bmatrix} A_2 \\ B_3 \end{bmatrix}, \\ r\begin{bmatrix} A_3 & C_3 & B_3 & D_3^*C_1^* & C_2^* \\ C_3^* & 0 & 0 & A_1^* & 0 \\ B_3^* & 0 & 0 & 0 & A_2^* \\ C_1D_3 & A_1 & 0 & 0 & 0 \\ C_2 & 0 & A_2 & 0 & 0 \end{bmatrix} = 2r\begin{bmatrix} C_3 & B_3 \\ A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \\ r\begin{bmatrix} A_3 & D_3^* & B_3 & C_3D_1 & C_2^* \\ D_3 & 0 & 0 & B_1 & 0 \\ B_3^* & 0 & 0 & 0 & A_2^* \\ D_1^*C_3^* & B_1^* & 0 & 0 & 0 \\ C_2 & 0 & A_2 & 0 & 0 \end{bmatrix} = 2r\begin{bmatrix} D_3^* & B_3 \\ B_1^* & 0 \\ 0 & A_2 \end{bmatrix}. \end{gathered} \tag{75}$$

In this case, the general solution of system (6) *can be expressed as*

$$\begin{gathered} X = A_2^\dagger C_2 + L_{A_2}U, \\ Y = A_1^\dagger C_1 + L_{A_1}D_1B_1^\dagger + L_{A_1}VR_{B_1}, \end{gathered} \tag{76}$$

where

$$V = \frac{1}{2}\left[F^{\dagger}EG^{\dagger} - F^{\dagger}G^{*}M^{\dagger}EG^{\dagger} - F^{\dagger}S\left(G^{\dagger}\right)^{*}EN^{\dagger}F^{*}G^{\dagger} + F^{\dagger}E\left(M^{\dagger}\right)^{*} + \left(N^{\dagger}\right)^{*}EG^{\dagger}S^{\dagger}S\right] + L_F V_1 + V_2 R_G + U_1 L_S L_M + R_N U_2^{*} L_M - F^{\dagger}SU_2R_NF^{*}G^{\dagger},$$
$$U = A^{\dagger}\left[D - BVC - (BVC)^{*}\right] - \frac{1}{2}A^{\dagger}\left[D - BVC - (BVC)^{*}\right]AA^{\dagger} - A^{\dagger}W_1A^{*} + W_1^{*}AA^{\dagger} + L_AW_2, \tag{77}$$

where U_1, U_2, V_1, V_2, W_1, and W_2 are arbitrary matrices over $\mathbb{C}$ with appropriate sizes.

Proof. (2) ⇔ (3): Applying Lemma 3 and Lemma 11 gives

$$\begin{aligned}
R_MR_FE = 0 &\Longleftrightarrow r\left(R_MR_FE\right) = 0 \Longleftrightarrow r\begin{bmatrix} D & B & C^{*} & A \\ A^{*} & 0 & 0 & 0\end{bmatrix} = r\begin{bmatrix} B & C^{*} & A\end{bmatrix} + r(A) \\
&\Longleftrightarrow r\begin{bmatrix} D & C_3 & D_3^{*} & B_3 & 0 \\ B_3^{*} & 0 & 0 & 0 & A_2^{*} \\ 0 & A_1 & 0 & 0 & 0 \\ 0 & 0 & B_1^{*} & 0 & 0 \\ 0 & 0 & 0 & A_2 & 0\end{bmatrix} = r\begin{bmatrix} C_3 & D_3^{*} & B_3 \\ A_1 & 0 & 0 \\ 0 & B_1^{*} & 0 \\ 0 & 0 & A_2\end{bmatrix} + r\begin{bmatrix} A_2 \\ B_3\end{bmatrix} \\
&\Longleftrightarrow r\begin{bmatrix} A_3 & C_3 & D_3^{*} & B_3 & C_2^{*} \\ B_3^{*} & 0 & 0 & 0 & A_2^{*} \\ C_1D_3 & A_1 & 0 & 0 & 0 \\ D_1^{*}C_3^{*} & 0 & B_1^{*} & 0 & 0 \\ C_2 & 0 & 0 & A_2 & 0\end{bmatrix} = r\begin{bmatrix} C_3 & D_3^{*} & B_3 \\ A_1 & 0 & 0 \\ 0 & B_1^{*} & 0 \\ 0 & 0 & A_2\end{bmatrix} + r\begin{bmatrix} A_2 \\ B_3\end{bmatrix}.
\end{aligned} \tag{78}$$

By a similar approach, we can obtain that

$$R_FER_F = 0 \Longleftrightarrow r\begin{bmatrix} A_3 & C_3 & B_3 & D_3^{*}C_1^{*} & C_2^{*} \\ C_3^{*} & 0 & 0 & A_1^{*} & 0 \\ B_3^{*} & 0 & 0 & 0 & A_2^{*} \\ C_1D_3 & A_1 & 0 & 0 & 0 \\ C_2 & 0 & A_2 & 0 & 0\end{bmatrix} = 2r\begin{bmatrix} C_3 & B_3 \\ A_1 & 0 \\ 0 & A_2\end{bmatrix},$$
$$L_GEL_G = 0 \Longleftrightarrow r\begin{bmatrix} A_3 & D_3^{*} & B_3 & C_3D_1 & C_2^{*} \\ D_3 & 0 & 0 & B_1 & 0 \\ B_3^{*} & 0 & 0 & 0 & A_2^{*} \\ D_1^{*}C_3^{*} & B_1^{*} & 0 & 0 & 0 \\ C_2 & 0 & A_2 & 0 & 0\end{bmatrix} = 2r\begin{bmatrix} D_3^{*} & B_3 \\ B_1^{*} & 0 \\ 0 & A_2\end{bmatrix}. \tag{79}$$

(1) ⇔ (2): We separate the four equations in system (6) into three groups:

$$A_1Y = C_1, \qquad YB_1 = D_1, \tag{80}$$

$$A_2X = C_2, \tag{81}$$

$$B_3X + \left(B_3X\right)^{*} + C_3YD_3 + \left(C_3YD_3\right)^{*} = A_3. \tag{82}$$

By Lemma 1, we obtain that system (80) is solvable if and only if (14), (81) is consistent if and only if (17). The general solutions to system (80) and (81) can be expressed as (16) and (19), respectively. Substituting (16) and (19) into (82) yields

$$AU + (AU)^{*} + BVC + (BVC)^{*} = D. \tag{83}$$

Hence, the system (5) is consistent if and only if (80), (81), and (83) are consistent, respectively. It follows from Lemma 11 that (83) is solvable if and only if

$$R_MR_FE = 0, \qquad R_FER_F = 0, \qquad L_GEL_G = 0. \tag{84}$$

We know by Lemma 11 that the general solution of (83) can be expressed as (77). □

In Theorem 12, let A_1 and D_1 vanish. Then we can obtain the general solution to the following system:

$$A_2X = C_2, \qquad YB_1 = D_1, \qquad B_3X + \left(B_3X\right)^{*} + C_3YD_3 + \left(C_3YD_3\right)^{*} = A_3. \tag{85}$$

Corollary 13. *Let A_2, C_2, B_1, D_1, B_3, C_3, D_3, and $A_3 = A_3^{*}$ be given. Set*

$$\begin{gathered}
A = B_3L_{A_2}, \qquad C = R_{B_1}D_3, \\
F = R_AC_3, \qquad G = CR_A, \\
M = R_FG^{*}, \qquad N = F^{*}L_G, \\
S = G^{*}L_M, \\
D = A_3 - B_3A_2^{\dagger}C_2 - \left(B_3A_2^{\dagger}C_2\right)^{*} - C_3D_1B_1^{\dagger}D_3 - \left(C_3D_1B_1^{\dagger}D_3\right)^{*}, \\
E = R_ADR_A.
\end{gathered} \tag{86}$$

Then the following statements are equivalent:

(1) *system* (85) *is consistent*

(2)

$$R_{A_2}C_2 = 0, \qquad D_1L_{B_1} = 0, \qquad R_MR_FE = 0, \tag{87}$$

$$R_FER_F = 0, \qquad L_GEL_G = 0, \tag{88}$$

(3)

$$r\left[A_2 \;\; C_2\right] = r\left(A_2\right), \qquad \begin{bmatrix} D_1 \\ B_1 \end{bmatrix} = r\left(B_1\right),$$

$$r\begin{bmatrix} A_3 & C_3 & D_3^* & B_3 & C_2^* \\ B_3^* & 0 & 0 & 0 & A_2^* \\ D_1^*C_3^* & 0 & B_1^* & 0 & 0 \\ C_2 & 0 & 0 & A_2 & 0 \end{bmatrix} = r\begin{bmatrix} C_3 & D_3^* & B_3 \\ 0 & B_1^* & 0 \\ 0 & 0 & A_2 \end{bmatrix} + r\begin{bmatrix} A_2 \\ B_3 \end{bmatrix},$$

$$r\begin{bmatrix} A_3 & C_3 & B_3 & C_2^* \\ C_3^* & 0 & 0 & 0 \\ B_3^* & 0 & 0 & A_2^* \\ C_2 & 0 & A_2 & 0 \end{bmatrix} = 2r\begin{bmatrix} C_3 & B_3 \\ 0 & A_2 \end{bmatrix},$$

$$r\begin{bmatrix} A_3 & D_3^* & B_3 & C_3D_1 & C_2^* \\ D_3 & 0 & 0 & B_1 & 0 \\ B_3^* & 0 & 0 & 0 & A_2^* \\ D_1^*C_3^* & B_1^* & 0 & 0 & 0 \\ C_2 & 0 & A_2 & 0 & 0 \end{bmatrix} = 2r\begin{bmatrix} D_3^* & B_3 \\ B_1^* & 0 \\ 0 & A_2 \end{bmatrix}. \tag{89}$$

In this case, the general solution of system (6) *can be expressed as*

$$X = A_2^{\dagger}C_2 + L_{A_2}U, \qquad Y = D_1B_1^{\dagger} + VR_{B_1}, \tag{90}$$

where

$$\begin{aligned} V = \frac{1}{2}\Big[& F^{\dagger}EG^{\dagger} - F^{\dagger}G^{*}M^{\dagger}EG^{\dagger} - F^{\dagger}S\left(G^{\dagger}\right)^{*}EN^{\dagger}F^{*}G^{\dagger} \\ & +F^{\dagger}E\left(M^{\dagger}\right)^{*} + \left(N^{\dagger}\right)^{*}EG^{\dagger}S^{\dagger}S\Big] + L_FV_1 \\ & + V_2R_G + U_1L_SL_M + R_NU_2^{*}L_M - F^{\dagger}SU_2R_NF^{*}G^{\dagger}, \\ U = A^{\dagger} & \left[D - C_3VC - \left(C_3VC\right)^{*}\right] \\ & - \frac{1}{2}A^{\dagger}\left[D - C_3VC - \left(C_3VC\right)^{*}\right]AA^{\dagger} \\ & - A^{\dagger}W_1A^{*} + W_1^{*}AA^{\dagger} + L_AW_2, \end{aligned} \tag{91}$$

where U_1, U_2, V_1, V_2, W_1, *and* W_2 *are arbitrary matrices over* $\mathbb{C}$ *with appropriate sizes.*

Acknowledgments

The authors would like to thank Dr. Mohamed, Dr. Sivakumar, and a referee very much for their valuable suggestions and comments, which resulted in a great improvement of the original paper. This research was supported by the Grants from the National Natural Science Foundation of China (NSFC (11161008)) and Doctoral Program Fund of Ministry of Education of P.R.China (20115201110002).

References

[1] H. Zhang, "Rank equalities for Moore-Penrose inverse and Drazin inverse over quaternion," *Annals of Functional Analysis*, vol. 3, no. 2, pp. 115–127, 2012.

[2] L. Bao, Y. Lin, and Y. Wei, "A new projection method for solving large Sylvester equations," *Applied Numerical Mathematics*, vol. 57, no. 5–7, pp. 521–532, 2007.

[3] M. Dehghan and M. Hajarian, "On the generalized reflexive and anti-reflexive solutions to a system of matrix equations," *Linear Algebra and Its Applications*, vol. 437, no. 11, pp. 2793–2812, 2012.

[4] F. O. Farid, M. S. Moslehian, Q.-W. Wang, and Z.-C. Wu, "On the Hermitian solutions to a system of adjointable operator equations," *Linear Algebra and Its Applications*, vol. 437, no. 7, pp. 1854–1891, 2012.

[5] Z. H. He and Q. W. Wang, "A real quaternion matrix equation with with applications," *Linear Multilinear Algebra*, vol. 61, pp. 725–740, 2013.

[6] Q.-W. Wang and Z.-C. Wu, "Common Hermitian solutions to some operator equations on Hilbert C^x-modules," *Linear Algebra and Its Applications*, vol. 432, no. 12, pp. 3159–3171, 2010.

[7] Q.-W. Wang and C.-Z. Dong, "Positive solutions to a system of adjointable operator equations over Hilbert C^x-modules," *Linear Algebra and Its Applications*, vol. 433, no. 7, pp. 1481–1489, 2010.

[8] Q.-W. Wang, J. W. van der Woude, and H.-X. Chang, "A system of real quaternion matrix equations with applications," *Linear Algebra and Its Applications*, vol. 431, no. 12, pp. 2291–2303, 2009.

[9] Q. W. Wang, H.-S. Zhang, and G.-J. Song, "A new solvable condition for a pair of generalized Sylvester equations," *Electronic Journal of Linear Algebra*, vol. 18, pp. 289–301, 2009.

[10] H. W. Braden, "The equations $A^TX \pm X^TA = B$," *SIAM Journal on Matrix Analysis and Applications*, vol. 20, no. 2, pp. 295–302, 1999.

[11] D. S. Djordjević, "Explicit solution of the operator equation $A^*X + X^{\pm}A = B$," *Journal of Computational and Applied Mathematics*, vol. 200, no. 2, pp. 701–704, 2007.

[12] W. Cao, "Solvability of a quaternion matrix equation," *Applied Mathematics*, vol. 17, no. 4, pp. 490–498, 2002.

[13] Q. Xu, L. Sheng, and Y. Gu, "The solutions to some operator equations," *Linear Algebra and Its Applications*, vol. 429, no. 8-9, pp. 1997–2024, 2008.

[14] Q.-W. Wang and Z.-H. He, "Some matrix equations with applications," *Linear and Multilinear Algebra*, vol. 60, no. 11-12, pp. 1327–1353, 2012.

[15] Y. Tian, "Maximization and minimization of the rank and inertia of the Hermitian matrix expression $A - BX - (BX)^x$," *Linear Algebra and Its Applications*, vol. 434, no. 10, pp. 2109–2139, 2011.

[16] Z.-H. He and Q.-W. Wang, "Solutions to optimization problems on ranks and inertias of a matrix function with applications," *Applied Mathematics and Computation*, vol. 219, no. 6, pp. 2989–3001, 2012.

[17] Y. Liu and Y. Tian, "Max-min problems on the ranks and inertias of the matrix expressions $A - BXC \pm (BXC)^*$," *Journal of*

Optimization Theory and Applications, vol. 148, no. 3, pp. 593–622, 2011.

[18] D. Chu, Y. S. Hung, and H. J. Woerdeman, "Inertia and rank characterizations of some matrix expressions," *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 3, pp. 1187–1226, 2009.

[19] Y. Liu and Y. Tian, "A simultaneous decomposition of a matrix triplet with applications," *Numerical Linear Algebra with Applications*, vol. 18, no. 1, pp. 69–85, 2011.

[20] X. Zhang, Q.-W. Wang, and X. Liu, "Inertias and ranks of some Hermitian matrix functions with applications," *Central European Journal of Mathematics*, vol. 10, no. 1, pp. 329–351, 2012.

[21] Q.-W. Wang, "The general solution to a system of real quaternion matrix equations," *Computers & Mathematics with Applications*, vol. 49, no. 5-6, pp. 665–675, 2005.

[22] Y. Tian, "Equalities and inequalities for inertias of Hermitian matrices with applications," *Linear Algebra and Its Applications*, vol. 433, no. 1, pp. 263–296, 2010.

[23] G. Marsaglia and G. P. H. Styan, "Equalities and inequalities for ranks of matrices," *Linear and Multilinear Algebra*, vol. 2, pp. 269–292, 1974.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 230408, 6 pages
http://dx.doi.org/10.1155/2013/230408

Research Article

Left and Right Inverse Eigenpairs Problem for κ-Hermitian Matrices

Fan-Liang Li,[1] Xi-Yan Hu,[2] and Lei Zhang[2]

[1] *Institute of Mathematics and Physics, School of Sciences, Central South University of Forestry and Technology, Changsha 410004, China*
[2] *College of Mathematics and Econometrics, Hunan University, Changsha 410082, China*

Correspondence should be addressed to Fan-Liang Li; lfl302@tom.com

Received 6 December 2012; Accepted 20 March 2013

Academic Editor: Panayiotis J. Psarrakos

Left and right inverse eigenpairs problem for κ-hermitian matrices and its optimal approximate problem are considered. Based on the special properties of κ-hermitian matrices, the equivalent problem is obtained. Combining a new inner product of matrices, the necessary and sufficient conditions for the solvability of the problem and its general solutions are derived. Furthermore, the optimal approximate solution and a calculation procedure to obtain the optimal approximate solution are provided.

1. Introduction

Throughout this paper we use some notations as follows. Let $C^{n\times m}$ be the set of all $n\times m$ complex matrices, $UC^{n\times n}$, $HC^{n\times n}$, $SHC^{n\times n}$ denote the set of all $n\times n$ unitary matrices, hermitian matrices, skew-hermitian matrices, respectively. Let $\overline{A}$, A^H, and A^+ be the conjugate, conjugate transpose, and the Moore-Penrose generalized inverse of A, respectively. For $A, B \in C^{n\times m}$, $\langle A, B\rangle = \operatorname{re}(\operatorname{tr}(B^H A))$, where $\operatorname{re}(\operatorname{tr}(B^H A))$ denotes the real part of $\operatorname{tr}(B^H A)$, the inner product of matrices A and B. The induced matrix norm is called Frobenius norm. That is, $\|A\| = \langle A, A\rangle^{1/2} = (\operatorname{tr}(A^H A))^{1/2}$.

Left and right inverse eigenpairs problem is a special inverse eigenvalue problem. That is, giving partial left and right eigenpairs (eigenvalue and corresponding eigenvector), (λ_i, x_i), $i = 1, \ldots, h$; (μ_j, y_j), $j = 1, \ldots, l$, a special matrix set S, finding a matrix $A \in S$ such that

$$\begin{aligned} Ax_i &= \lambda_i x_i, \quad i = 1, \ldots, h, \\ y_j^T A &= \mu_j y_j^T, \quad j = 1, \ldots, l. \end{aligned} \tag{1}$$

This problem, which usually arises in perturbation analysis of matrix eigenvalues and in recursive matters, has profound application background [1–6]. When the matrix set S is different, it is easy to obtain different left and right inverse eigenpairs problem. For example, we studied the left and right inverse eigenpairs problem of skew-centrosymmetric matrices and generalized centrosymmetric matrices, respectively [5, 6]. Based on the special properties of left and right eigenpairs of these matrices, we derived the solvability conditions of the problem and its general solutions. In this paper, combining the special properties of κ-hermitian matrices and a new inner product of matrices, we first obtain the equivalent problem, then derive the necessary and sufficient conditions for the solvability of the problem and its general solutions.

Hill and Waters [7] introduced the following matrices.

Definition 1. Let κ be a fixed product of disjoint transpositions, and let K be the associated permutation matrix, that is, $\overline{K} = K^H = K$, $K^2 = I_n$, a matrix $A \in C^{n\times n}$ is said to be κ-hermitian matrices (skew κ-hermitian matrices) if and only if $a_{ij} = \overline{a}_{k(j)k(i)}$ $(a_{ij} = -\overline{a}_{k(j)k(i)})$, $i, j = 1, \ldots, n$. We denote the set of κ-hermitian matrices (skew κ-hermitian matrices) by $KHC^{n\times n}$ $(SKHC^{n\times n})$.

From Definition 1, it is easy to see that hermitian matrices and perhermitian matrices are special cases of κ-hermitian matrices, with $k(i) = i$ and $k(i) = n - i + 1$, respectively. Hermitian matrices and perhermitian matrices, which are one of twelve symmetry patterns of matrices [8], are applied in engineering, statistics, and so on [9, 10].

From Definition 1, it is also easy to prove the following conclusions.

(1) $A \in KHC^{n\times n}$ if and only if $A = KA^HK$.

(2) $A \in SKHC^{n\times n}$ if and only if $A = -KA^HK$.

(3) If K is a fixed permutation matrix, then $KHC^{n\times n}$ and $SKHC^{n\times n}$ are the closed linear subspaces of $C^{n\times n}$ and satisfy

$$C^{n\times n} = KHC^{n\times n} \bigoplus SKHC^{n\times n}. \tag{2}$$

The notation $V_1 \oplus V_2$ stands for the orthogonal direct sum of linear subspace V_1 and V_2.

(4) $A \in KHC^{n\times n}$ if and only if there is a matrix $\widetilde{A} \in HC^{n\times n}$ such that $\widetilde{A} = KA$.

(5) $A \in SKHC^{n\times n}$ if and only if there is a matrix $\widetilde{A} \in SHC^{n\times n}$ such that $\widetilde{A} = KA$.

Proof. (1) From Definition 1, if $A = (a_{ij}) \in KHC^{n\times n}$, then $a_{ij} = \overline{a}_{k(j)k(i)}$, this implies $A = KA^HK$, for $KA^HK = (\overline{a}_{k(j)k(i)})$.

(2) With the same method, we can prove (2). So, the proof is omitted.

(3) (a) For any $A \in C^{n\times n}$, there exist $A_1 \in KHC^{n\times n}$, $A_2 \in SKHC^{n\times n}$ such that

$$A = A_1 + A_2, \tag{3}$$

where $A_1 = (1/2)(A + KA^HK)$, $A_2 = (1/2)(A - KA^HK)$.

(b) If there exist another $\overline{A}_1 \in KHC^{n\times n}$, $\overline{A}_2 \in SKHC^{n\times n}$ such that

$$A = \overline{A}_1 + \overline{A}_2, \tag{4}$$

(3)-(4) yields

$$A_1 - \overline{A}_1 = -\left(A_2 - \overline{A}_2\right). \tag{5}$$

Multiplying (5) on the left and on the right by K, respectively, and according to (1) and (2), we obtain

$$A_1 - \overline{A}_1 = A_2 - \overline{A}_2. \tag{6}$$

Combining (5) and (6) gives $A_1 = \overline{A}_1$, $A_2 = \overline{A}_2$.

(c) For any $A_1 \in KHC^{n\times n}$, $A_2 \in SKHC^{n\times n}$, we have

$$\begin{aligned}\langle A_1, A_2\rangle &= \mathrm{re}\left(\mathrm{tr}\left(A_2^HA_1\right)\right) = \mathrm{re}\left(\mathrm{tr}\left(KA_2^HKKA_1K\right)\right)\\ &= \mathrm{re}\left(\mathrm{tr}\left(-A_2^HA_1\right)\right) = -\langle A_1, A_2\rangle.\end{aligned} \tag{7}$$

This implies $\langle A_1, A_2\rangle = 0$. Combining (a), (b), and (c) gives (3).

(4) Let $\widetilde{A} = KA$, if $A \in KHC^{n\times n}$, then $\widetilde{A}^H = \widetilde{A} \in HC^{n\times n}$. If $\widetilde{A}^H = \widetilde{A} \in HC^{n\times n}$, then $A = K\widetilde{A}$ and $KA^HK = K\widetilde{A}^HKK = K\widetilde{A} = A \in KHC^{n\times n}$.

(5) With the same method, we can prove (5). So, the proof is omitted. □

In this paper, we suppose that K is a fixed permutation matrix and assume (λ_i, x_i), $i = 1,\ldots,h$, be right eigenpairs of A; (μ_j, y_j), $j = 1,\ldots,l$, be left eigenpairs of A. If we let $X = (x_1,\ldots,x_h) \in C^{n\times h}$, $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_h) \in C^{h\times h}$; $Y = (y_1,\ldots,y_l) \in C^{n\times l}$, $\Gamma = \mathrm{diag}(\mu_1,\ldots,\mu_l) \in C^{l\times l}$, then the problems studied in this paper can be described as follows.

Problem 2. Giving $X \in C^{n\times h}$, $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_h) \in C^{h\times h}$; $Y \in C^{n\times l}$, $\Gamma = \mathrm{diag}(\mu_1,\ldots,\mu_l) \in C^{l\times l}$, find $A \in KHC^{n\times n}$ such that

$$\begin{gathered}AX = X\Lambda,\\ Y^TA = \Gamma Y^T.\end{gathered} \tag{8}$$

Problem 3. Giving $B \in C^{n\times n}$, find $\widehat{A} \in S_E$ such that

$$\left\|B - \widehat{A}\right\| = \min_{\forall A\in S_E} \|B - A\|, \tag{9}$$

where S_E is the solution set of Problem 2.

This paper is organized as follows. In Section 2, we first obtain the equivalent problem with the properties of $KHC^{n\times n}$ and then derive the solvability conditions of Problem 2 and its general solution's expression. In Section 3, we first attest the existence and uniqueness theorem of Problem 3 then present the unique approximation solution. Finally, we provide a calculation procedure to compute the unique approximation solution and numerical experiment to illustrate the results obtained in this paper correction.

2. Solvability Conditions of Problem 2

We first discuss the properties of $KHC^{n\times n}$

Lemma 4. *Denoting $M = KEKGE$, and $E \in HC^{n\times n}$, one has the following conclusions.*

(1) *If $G \in KHC^{n\times n}$, then $M \in KHC^{n\times n}$.*

(2) *If $G \in SKHC^{n\times n}$, then $M \in SKHC^{n\times n}$.*

(3) *If $G = G_1 + G_2$, where $G_1 \in KHC^{n\times n}$, $G_2 \in SKHC^{n\times n}$, then $M \in KHC^{n\times n}$ if and only if $KEKG_2E = 0$. In addition, one has $M = KEKG_1E$.*

Proof. (1) $KM^HK = KEG^HKEKK = KE(KGK)KE = KEKGE = M$.

Hence, we have $M \in KHC^{n\times n}$.

(2) $KM^HK = KEG^HKEKK = KE(-KGK)KE = -KEKGE = -M$.

Hence, we have $M \in SKHC^{n\times n}$.

(3) $M = KEK(G_1 + G_2)E = KEKG_1E + KEKG_2E$, we have $KEKG_1E \in KHC^{n\times n}$, $KEKG_2E \in SKHC^{n\times n}$ from (1) and (2). If $M \in KHC^{n\times n}$, then $M - KEKG_1E \in KHC^{n\times n}$, while $M - KEKG_1E = KEKG_2E \in SKHC^{n\times n}$. Therefore from the conclusion (3) of Definition 1, we have $KEKG_2E = 0$, that is, $M = KEKG_1E$. On the contrary, if $KEKG_2E = 0$, it is clear that $M = KEKG_1E \in KHC^{n\times n}$. The proof is completed. □

Lemma 5. *Let $A \in KHC^{n\times n}$, if (λ, x) is a right eigenpair of A, then $(\overline{\lambda}, K\overline{x})$ is a left eigenpair of A.*

Proof. If (λ, x) is a right eigenpair of A, then we have

$$Ax = \lambda x. \tag{10}$$

From the conclusion (1) of Definition 1, it follows that

$$KA^H Kx = \lambda x. \tag{11}$$

This implies

$$(K\overline{x})^T A = \overline{\lambda}(K\overline{x})^T. \tag{12}$$

So $(\overline{\lambda}, K\overline{x})$ is a left eigenpair of A. □

From Lemma 5, without loss of the generality, we may assume that Problem 2 is as follows.

$$\begin{aligned} &X \in C^{n\times h}, \qquad \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_h) \in C^{h\times h}, \\ &Y = K\overline{X} \in C^{n\times h}, \qquad \Gamma = \overline{\Lambda} \in C^{h\times h}. \end{aligned} \tag{13}$$

Combining (13) and the conclusion (4) of Definition 1, it is easy to derive the following lemma.

Lemma 6. *If X, Λ, Y, Γ are given by (13), then Problem 2 is equivalent to the following problem. If X, Λ, Y, Γ are given by (13), find $KA \in HC^{n\times n}$ such that*

$$KAX = KX\Lambda. \tag{14}$$

Lemma 7 (see [11]). *If giving $X \in C^{n\times h}$, $B \in C^{n\times h}$, then matrix equation $\widetilde{A}X = B$ has solution $\widetilde{A} \in HC^{n\times n}$ if and only if*

$$B = BX^+X, \qquad B^H X = X^H B. \tag{15}$$

Moreover, the general solution $\widetilde{A}$ can be expressed as

$$\begin{aligned} \widetilde{A} = {} & BX^+ + (BX^+)^H (I_n - XX^+) \\ & + (I_n - XX^+)\widetilde{G}(I_n - XX^+), \quad \forall \widetilde{G} \in HC^{n\times n}. \end{aligned} \tag{16}$$

Theorem 8. *If X, Λ, Y, Γ are given by (13), then Problem 2 has a solution in $KHC^{n\times n}$ if and only if*

$$X^H KX\Lambda = \overline{\Lambda}X^H KX, \quad X\Lambda = X\Lambda X^+ X. \tag{17}$$

Moreover, the general solution can be expressed as

$$A = A_0 + KEKGE, \quad \forall G \in KHC^{n\times n}, \tag{18}$$

where

$$A_0 = X\Lambda X^+ + K(X\Lambda X^+)^H KE, \qquad E = I_n - XX^+. \tag{19}$$

Proof. Necessity: If there is a matrix $A \in KHC^{n\times n}$ such that $(AX = X\Lambda, Y^T A = \Gamma Y^T)$, then from Lemma 6, there exists a matrix $KA \in HC^{n\times n}$ such that $KAX = KX\Lambda$, and according to Lemma 7, we have

$$KX\Lambda = KX\Lambda X^+ X, \qquad (KX\Lambda)^H X = X^H (KX\Lambda). \tag{20}$$

It is easy to see that (20) is equivalent to (17).

Sufficiency: If (17) holds, then (20) holds. Hence, matrix equation $KAX = KX\Lambda$ has solution $KA \in HC^{n\times n}$. Moreover, the general solution can be expressed as follows:

$$\begin{aligned} KA = {} & KX\Lambda X^+ + (KX\Lambda X^+)^H (I_n - XX^+) \\ & + (I_n - XX^+)\widetilde{G}(I_n - XX^+), \quad \forall \widetilde{G} \in HC^{n\times n}. \end{aligned} \tag{21}$$

Let

$$A_0 = X\Lambda X^+ + K(X\Lambda X^+)^H KE, \qquad E = I_n - XX^+. \tag{22}$$

This implies $A = A_0 + KE\widetilde{G}E$. Combining the definition of K, E and the first equation of (17), we have

$$\begin{aligned} KA_0^H K &= K(X\Lambda X^+)^H K + X\Lambda X^+ - K(XX^+)^H K (X\Lambda X^+) \\ &= X\Lambda X^+ + K(X\Lambda X^+)^H K - K(X\Lambda X^+)^T KXX^+ \\ &= A_0. \end{aligned} \tag{23}$$

Hence, $A_0 \in KHC^{n\times n}$. Combining the definition of K, E, (13) and (17), we have

$$\begin{aligned} A_0 X &= X\Lambda X^+ X + K(X\Lambda X^+)^H K (I_n - XX^+) X = X\Lambda, \\ Y^T A_0 &= X^H KX\Lambda X^+ + X^H KK(X\Lambda X^+)^H KE \\ &= \overline{\Lambda}X^H KXX^+ + (KX\Lambda X^+ X)^H (I_n - XX^+) \\ &= \overline{\Lambda}X^H KXX^+ + \overline{\Lambda}X^H K (I_n - XX^+) \\ &= \overline{\Lambda}X^H K = \Gamma Y^T. \end{aligned} \tag{24}$$

Therefore, A_0 is a special solution of Problem 2. Combining the conclusion (4) of Definition 1, Lemma 4, and $E = I_n - XX^+ \in HC^{n\times n}$, it is easy to prove that $A = A_0 + KEKGE \in KHC^{n\times n}$ if and only if $G \in KHC^{n\times n}$. Hence, the solution set of Problem 2 can be expressed as (18). □

3. An Expression of the Solution of Problem 3

From (18), it is easy to prove that the solution set S_E of Problem 2 is a nonempty closed convex set if Problem 2 has a solution in $KHC^{n\times n}$. We claim that for any given $B \in R^{n\times n}$, there exists a unique optimal approximation for Problem 3.

Theorem 9. *Giving $B \in C^{n\times n}$, if the conditions of X, Y, Λ, Γ are the same as those in Theorem 8, then Problem 3 has a unique solution $\widehat{A} \in S_E$. Moreover, $\widehat{A}$ can be expressed as*

$$\widehat{A} = A_0 + KEKB_1E, \tag{25}$$

where A_0, E are given by (19) and $B_1 = (1/2)(B + KB^H K)$.

Proof. Denoting $E_1 = I_n - E$, it is easy to prove that matrices E and E_1 are orthogonal projection matrices satisfying $EE_1 = 0$. It is clear that matrices KEK and KE_1K are also orthogonal projection matrices satisfying $(KEK)(KE_1K) = 0$. According to the conclusion (3) of Definition 1, for any $B \in C^{n\times n}$, there exists unique

$$B_1 \in KHC^{n\times n}, \qquad B_2 \in SKHC^{n\times n} \tag{26}$$

such that

$$B = B_1 + B_2, \quad \langle B_1, B_2 \rangle = 0, \tag{27}$$

where

$$B_1 = \frac{1}{2}\left(B + KB^H K\right), \qquad B_2 = \frac{1}{2}\left(B - KB^H K\right). \tag{28}$$

Combining Theorem 8, for any $A \in S_E$, we have

$$\begin{aligned}
&\|B - A\|^2 \\
&\quad = \|B - A_0 - KEKGE\|^2 \\
&\quad = \|B_1 + B_2 - A_0 - KEKGE\|^2 \\
&\quad = \|B_1 - A_0 - KEKGE\|^2 + \|B_2\|^2 \\
&\quad = \|(B_1 - A_0)(E + E_1) - KEKGE\|^2 + \|B_2\|^2 \\
&\quad = \|(B_1 - A_0)E - KEKGE\|^2 \\
&\qquad + \|(B_1 - A_0)E_1\|^2 + \|B_2\|^2 \\
&\quad = \|K(E + E_1)K(B_1 - A_0)E - KEKGE\|^2 \\
&\qquad + \|(B_1 - A_0)E_1\|^2 + \|B_2\|^2 \\
&\quad = \|KEK(B_1 - A_0)E - KEKGE\|^2 \\
&\qquad + \|KE_1K(B_1 - A_0)E\|^2 \\
&\qquad + \|(B_1 - A_0)E_1\|^2 + \|B_2\|^2.
\end{aligned} \tag{29}$$

It is easy to prove that $KEKA_0E = 0$ according to the definitions of A_0, E. So we have

$$\begin{aligned}
\|B - A\|^2 = {} & \|KEKB_1E - KEKGE\|^2 + \|KE_1K(B_1 - A_0)E\|^2 \\
& + \|(B_1 - A_0)E_1\|^2 + \|B_2\|^2.
\end{aligned} \tag{30}$$

Obviously, $\min_{A\in S_E}\|B - A\|$ is equivalent to

$$\min_{G\in KHC^{n\times n}} \|KEKB_1E - KEKGE\|. \tag{31}$$

Since $EE_1 = 0, (KEK)(KE_1K) = 0$, it is clear that $G = B_1 + KE_1K\widehat{G}E_1$, for any $\widehat{G} \in KHC^{n\times n}$, is a solution of (31). Substituting this result to (18), we can obtain (25). □

Algorithm 10. (1) Input X, Λ, Y, Γ according to (13). (2) Compute $X^HKX\Lambda$, $\overline{\Lambda}X^HKX$, $X\Lambda X^+X$, $X\Lambda$, if (17) holds, then continue; otherwise stop. (3) Compute A_0 according to (19), and compute B_1 according to (28). (4) According to (25) calculate $\widehat{A}$.

Example 11 ($n = 8$, $h = l = 4$).

$$X = \begin{pmatrix}
0.5661 & -0.2014 - 0.1422i & 0.1446 + 0.2138i & 0.524 \\
-0.2627 + 0.1875i & 0.5336 & -0.2110 - 0.4370i & -0.0897 + 0.3467i \\
-0.4132 + 0.2409i & 0.0226 - 0.0271i & -0.1095 + 0.2115i & -0.3531 - 0.0642i \\
-0.0306 + 0.2109i & -0.3887 - 0.0425i & 0.2531 + 0.2542i & 0.0094 + 0.2991i \\
0.0842 - 0.1778i & -0.0004 - 0.3733i & 0.3228 - 0.1113i & 0.1669 + 0.1952i \\
0.0139 - 0.3757i & -0.2363 + 0.3856i & 0.2583 + 0.0721i & 0.1841 - 0.2202i \\
0.0460 + 0.3276i & -0.1114 + 0.0654i & -0.0521 - 0.2556i & -0.2351 + 0.3002i \\
0.0085 - 0.1079i & 0.0974 + 0.3610i & 0.5060 & -0.2901 - 0.0268i
\end{pmatrix},$$

$$\Lambda = \begin{pmatrix}
-0.3967 - 0.4050i & 0 & 0 & 0 \\
0 & -0.3967 + 0.4050i & 0 & 0 \\
0 & 0 & 0.0001 & 0 \\
0 & 0 & 0 & -0.0001i
\end{pmatrix},$$

$$K = \begin{pmatrix}
0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&1&0\\
0&0&0&0&0&1&0&0\\
0&0&0&1&0&0&0&0\\
0&0&0&0&1&0&0&0\\
0&0&1&0&0&0&0&0\\
0&1&0&0&0&0&0&0\\
1&0&0&0&0&0&0&0
\end{pmatrix},$$

$$Y = K\overline{X}, \qquad \Gamma = \overline{\Lambda} \tag{32}$$

$B =$

From the first column to the fourth column

$$\begin{pmatrix} -0.5218+0.0406i & 0.2267-0.0560i & -0.1202+0.0820i & -0.0072-0.3362i \\ 0.3909-0.3288i & 0.2823-0.2064i & -0.0438-0.0403i & 0.2707+0.0547i \\ 0.2162-0.1144i & -0.4307+0.2474i & -0.0010-0.0412i & 0.2164-0.1314i \\ -0.1872-0.0599i & -0.0061+0.4698i & 0.3605-0.0247i & 0.4251+0.1869i \\ -0.1227-0.0194i & 0.2477-0.0606i & 0.3918+0.6340i & 0.1226+0.0636i \\ -0.0893+0.4335i & 0.0662+0.0199i & -0.0177-0.1412i & 0.4047+0.2288i \\ 0.1040-0.2015i & 0.1840+0.2276i & 0.2681-0.3526i & -0.5252+0.1022i \\ 0.1808+0.2669i & 0.2264+0.3860i & -0.1791+0.1976i & -0.0961-0.0117i \end{pmatrix} \quad (33)$$

From the fifth column to the eighth column

$$\begin{pmatrix} -0.2638-0.4952i & -0.0863-0.1664i & 0.2687+0.1958i & -0.2544-0.1099i \\ -0.2741-0.1656i & -0.0227+0.2684i & 0.1846+0.2456i & -0.0298+0.5163i \\ -0.1495-0.3205i & 0.1391+0.2434i & 0.1942-0.5211i & -0.3052-0.1468i \\ -0.2554+0.2690i & -0.4222-0.1080i & 0.2232+0.0774i & 0.0965-0.0421i \\ 0.3856-0.0619i & 0.1217-0.0270i & 0.1106-0.3090i & -0.1122+0.2379i \\ -0.1130+0.0766i & 0.7102-0.0901i & 0.1017+0.1397i & -0.0445+0.0038i \\ 0.1216+0.0076i & 0.2343-0.1772i & 0.5242-0.0089i & -0.0613+0.0258i \\ -0.0750-0.3581i & 0.0125+0.0964i & 0.0779-0.1074i & 0.6735-0.0266i \end{pmatrix}. \quad (34)$$

It is easy to see that matrices X, Λ, Y, Γ satisfy (17). Hence, there exists the unique solution for Problem 3. Using the software "MATLAB", we obtain the unique solution $\widehat{A}$ of Problem 3.

From the first column to the fourth column

$$\begin{pmatrix} -0.1983+0.0491i & 0.1648+0.0032i & 0.0002+0.1065i & 0.1308+0.2690i \\ 0.1071-0.2992i & 0.2106+0.2381i & -0.0533-0.3856i & -0.0946+0.0488i \\ 0.1935-0.1724i & -0.0855-0.0370i & 0.0200-0.0665i & -0.2155-0.0636i \\ 0.0085-0.2373i & -0.0843-0.1920i & 0.0136+0.0382i & -0.0328-0.0000i \\ -0.0529+0.1703i & 0.1948-0.0719i & 0.1266+0.1752i & 0.0232-0.2351i \\ 0.0855+0.1065i & 0.0325-0.2068i & 0.2624+0.0000i & 0.0136-0.0382i \\ 0.1283-0.1463i & -0.0467+0.0000i & 0.0325+0.2067i & -0.0843+0.1920i \\ 0.2498+0.0000i & 0.1283+0.1463i & 0.0855-0.1065i & 0.0086+0.2373i \end{pmatrix} \quad (35)$$

From the fifth column to the eighth column

$$\begin{pmatrix} 0.2399-0.1019i & 0.1928-0.1488i & -0.3480-0.2574i & 0.1017-0.0000i \\ -0.1955+0.0644i & 0.2925+0.1872i & 0.3869+0.0000i & -0.3481+0.2574i \\ 0.0074+0.0339i & -0.3132-0.0000i & 0.2926-0.1872i & 0.1928+0.1488i \\ 0.0232+0.2351i & -0.2154+0.0636i & -0.0946-0.0489i & 0.1309-0.2691i \\ -0.0545-0.0000i & 0.0074-0.0339i & -0.1955-0.0643i & 0.2399+0.1019i \\ 0.1266-0.1752i & 0.0200+0.0665i & -0.0533+0.3857i & 0.0002-0.1065i \\ 0.1949+0.0719i & -0.0855+0.0370i & 0.2106-0.2381i & 0.1648-0.0032i \\ -0.0529-0.1703i & 0.1935+0.1724i & 0.1071+0.2992i & -0.1983-0.0491i \end{pmatrix}. \quad (36)$$

Conflict of Interests

There is no conflict of interests between the authors.

Acknowledgments

This research was supported by National Natural Science Foundation of China (31170532). The authors are very grateful to the referees for their valuable comments and also thank the editor for his helpful suggestions.

References

[1] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, UK, 1965.

[2] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, vol. 9 of *Classics in Applied Mathematics*, SIAM, Philadelphia, Pa, USA, 1994.

[3] M. Arav, D. Hershkowitz, V. Mehrmann, and H. Schneider, "The recursive inverse eigenvalue problem," *SIAM Journal on Matrix Analysis and Applications*, vol. 22, no. 2, pp. 392–412, 2000.

[4] R. Loewy and V. Mehrmann, "A note on the symmetric recursive inverse eigenvalue problem," *SIAM Journal on Matrix Analysis and Applications*, vol. 25, no. 1, pp. 180–187, 2003.

[5] F. L. Li, X. Y. Hu, and L. Zhang, "Left and right inverse eigenpairs problem of skew-centrosymmetric matrices," *Applied Mathematics and Computation*, vol. 177, no. 1, pp. 105–110, 2006.

[6] F. L. Li, X. Y. Hu, and L. Zhang, "Left and right inverse eigenpairs problem of generalized centrosymmetric matrices and its optimal approximation problem," *Applied Mathematics and Computation*, vol. 212, no. 2, pp. 481–487, 2009.

[7] R. D. Hill and S. R. Waters, "On κ-real and κ-Hermitian matrices," *Linear Algebra and its Applications*, vol. 169, pp. 17–29, 1992.

[8] S. P. Irwin, "Matrices with multiple symmetry properties: applications of centro-Hermitian and per-Hermitian matrices," *Linear Algebra and its Applications*, vol. 284, no. 1–3, pp. 239–258, 1998.

[9] L. C. Biedenharn and J. D. Louck, *Angular Momentum in Quantum Physics*, vol. 8 of *Encyclopedia of Mathematics and its Applications*, Addison-Wesley, Reading, Mass, USA, 1981.

[10] W. M. Gibson and B. R. Pollard, *Symmetry Principles in Elementary Particle Physics*, Cambridge University Press, Cambridge, UK, 1976.

[11] W. F. Trench, "Hermitian, Hermitian R-symmetric, and Hermitian R-skew symmetric Procrustes problems," *Linear Algebra and its Applications*, vol. 387, pp. 83–98, 2004.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 271978, 5 pages
http://dx.doi.org/10.1155/2013/271978

Research Article

Completing a 2×2 Block Matrix of Real Quaternions with a Partial Specified Inverse

Yong Lin[1,2] and Qing-Wen Wang[1]

[1] *Department of Mathematics, Shanghai University, Shanghai 200444, China*
[2] *School of Mathematics and Statistics, Suzhou University, Suzhou 234000, China*

Correspondence should be addressed to Qing-Wen Wang; wqw858@yahoo.com.cn

Received 4 December 2012; Revised 23 February 2013; Accepted 20 March 2013

Academic Editor: K. Sivakumar

This paper considers a completion problem of a nonsingular 2×2 block matrix over the real quaternion algebra $\mathbb{H}$: Let m_1, m_2, n_1, n_2 be nonnegative integers, $m_1 + m_2 = n_1 + n_2 = n > 0$, and $A_{12} \in \mathbb{H}^{m_1 \times n_2}$, $A_{21} \in \mathbb{H}^{m_2 \times n_1}$, $A_{22} \in \mathbb{H}^{m_2 \times n_2}$, $B_{11} \in \mathbb{H}^{n_1 \times m_1}$ be given. We determine necessary and sufficient conditions so that there exists a variant block entry matrix $A_{11} \in \mathbb{H}^{m_1 \times n_1}$ such that $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \in \mathbb{H}^{n \times n}$ is nonsingular, and B_{11} is the upper left block of a partitioning of A^{-1}. The general expression for A_{11} is also obtained. Finally, a numerical example is presented to verify the theoretical findings.

1. Introduction

The problem of completing a block-partitioned matrix of a specified type with some of its blocks given has been studied by many authors. Fiedler and Markham [1] considered the following completion problem over the real number field $\mathbb{R}$. Suppose m_1, m_2, n_1, n_2 are nonnegative integers, $m_1 + m_2 = n_1 + n_2 = n > 0$, $A_{11} \in \mathbb{R}^{m_1 \times n_1}$, $A_{12} \in \mathbb{R}^{m_1 \times n_2}$, $A_{21} \in \mathbb{R}^{m_2 \times n_1}$, and $B_{22} \in \mathbb{R}^{n_2 \times m_2}$. Determine a matrix $A_{22} \in \mathbb{R}^{m_2 \times n_2}$ such that

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \tag{1}$$

is nonsingular and B_{22} is the lower right block of a partitioning of A^{-1}. This problem has the form of

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & ? \end{pmatrix}^{-1} = \begin{pmatrix} ? & ? \\ ? & B_{22} \end{pmatrix}, \tag{2}$$

and the solution and the expression for A_{22} were obtained in [1]. Dai [2] considered this form of completion problems with symmetric and symmetric positive definite matrices over $\mathbb{R}$.

Some other particular forms for 2×2 block matrices over $\mathbb{R}$ have also been examined (see, e.g., [3]), such as

$$\begin{gathered} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & ? \end{pmatrix}^{-1} = \begin{pmatrix} B_{11} & ? \\ ? & ? \end{pmatrix}, \\ \begin{pmatrix} A_{11} & ? \\ ? & ? \end{pmatrix}^{-1} = \begin{pmatrix} ? & ? \\ ? & B_{22} \end{pmatrix}, \\ \begin{pmatrix} A_{11} & ? \\ ? & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} ? & B_{12} \\ B_{21} & ? \end{pmatrix}. \end{gathered} \tag{3}$$

The real quaternion matrices play a role in computer science, quantum physics, and so on (e.g., [4–6]). Quaternion matrices are receiving much attention as witnessed recently (e.g., [7–9]). Motivated by the work of [1, 10] and keeping such applications of quaternion matrices in view, in this paper we consider the following completion problem over the real quaternion algebra:

$$\begin{aligned} \mathbb{H} = \{ & a_0 + a_1 i + a_2 j + a_3 k \mid \\ & i^2 = j^2 = k^2 = ijk = -1 \text{ and } a_0, a_1, a_2, a_3 \in \mathbb{R} \}. \end{aligned} \tag{4}$$

Problem 1. Suppose m_1, m_2, n_1, n_2 are nonnegative integers, $m_1 + m_2 = n_1 + n_2 = n > 0$, and $A_{12} \in \mathbb{H}^{m_1 \times n_2}$,

$A_{21} \in \mathbb{H}^{m_2 \times n_1}$, $A_{22} \in R^{m_2 \times n_2}$, $B_{11} \in \mathbb{H}^{n_1 \times m_1}$. Find a matrix $A_{11} \in \mathbb{H}^{m_1 \times n_1}$ such that

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \in \mathbb{H}^{n \times n} \tag{5}$$

is nonsingular, and B_{11} is the upper left block of a partitioning of A^{-1}. That is

$$\begin{pmatrix} ? & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} B_{11} & ? \\ ? & ? \end{pmatrix}, \tag{6}$$

where $\mathbb{H}^{m \times n}$ denotes the set of all $m \times n$ matrices over $\mathbb{H}$ and A^{-1} denotes the inverse matrix of A.

Throughout, over the real quaternion algebra $\mathbb{H}$, we denote the identity matrix with the appropriate size by I, the transpose of A by A^T, the rank of A by $r(A)$, the conjugate transpose of A by $A^* = (\overline{A})^T$, a reflexive inverse of a matrix A over $\mathbb{H}$ by A^+ which satisfies simultaneously $AA^+A = A$ and $A^+AA^+ = A^+$. Moreover, $L_A = I - A^+A$, $R_A = I - AA^+$, where A^+ is an arbitrary but fixed reflexive inverse of A. Clearly, L_A and R_A are idempotent, and each is a reflexive inverse of itself. $\mathscr{R}(A)$ denotes the right column space of the matrix A.

The rest of this paper is organized as follows. In Section 2, we establish some necessary and sufficient conditions to solve Problem 1 over $\mathbb{H}$, and the general expression for A_{11} is also obtained. In Section 3, we present a numerical example to illustrate the developed theory.

2. Main Results

In this section, we begin with the following lemmas.

Lemma 1 (singular-value decomposition [9]). *Let $A \in \mathbb{H}^{m \times n}$ be of rank r. Then there exist unitary quaternion matrices $U \in \mathbb{H}^{m \times m}$ and $V \in \mathbb{H}^{n \times n}$ such that*

$$UAV = \begin{pmatrix} D_r & 0 \\ 0 & 0 \end{pmatrix}, \tag{7}$$

where $D_r = \operatorname{diag}(d_1, \ldots, d_r)$ and the d_j's are the positive singular values of A.

Let $\mathbb{H}_c^n$ denote the collection of column vectors with n components of quaternions and A be an $m \times n$ quaternion matrix. Then the solutions of $Ax = 0$ form a subspace of $\mathbb{H}_c^n$ of dimension $n(A)$. We have the following lemma.

Lemma 2. *Let*

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \tag{8}$$

be a partitioning of a nonsingular matrix $A \in \mathbb{H}^{n \times n}$, and let

$$\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \tag{9}$$

be the corresponding (i.e., transpose) partitioning of A^{-1}. Then $n(A_{11}) = n(B_{22})$.

Proof. It is readily seen that

$$\begin{pmatrix} B_{22} & B_{21} \\ B_{12} & B_{11} \end{pmatrix}, \qquad \begin{pmatrix} A_{22} & A_{21} \\ A_{12} & A_{11} \end{pmatrix} \tag{10}$$

are inverse to each other, so we may suppose that $n(A_{11}) < n(B_{22})$.

If $n(B_{22}) = 0$, necessarily $n(A_{11}) = 0$ and we are finished. Let $n(B_{22}) = c > 0$, then there exists a matrix F with c right linearly independent columns, such that $B_{22}F = 0$. Then, using

$$A_{11}B_{12} + A_{12}B_{22} = 0, \tag{11}$$

we have

$$A_{11}B_{12}F = 0. \tag{12}$$

From

$$A_{21}B_{12} + A_{22}B_{22} = I, \tag{13}$$

we have

$$A_{21}B_{12}F = F. \tag{14}$$

It follows that the rank $r(B_{12}F) \geq c$. In view of (12), this implies

$$n(A_{11}) \geq r(B_{12}F) \geq c = n(B_{22}). \tag{15}$$

Thus

$$n(A_{11}) = n(B_{22}). \tag{16}$$

□

Lemma 3 (see [10]). *Let $A \in \mathbb{H}^{m \times n}$, $B \in \mathbb{H}^{p \times q}$, $D \in \mathbb{H}^{m \times q}$ be known and $X \in \mathbb{H}^{n \times p}$ unknown. Then the matrix equation*

$$AXB = D \tag{17}$$

is consistent if and only if

$$AA^+DB^+B = D. \tag{18}$$

In that case, the general solution is

$$X = A^+DB^+ + L_AY_1 + Y_2R_B, \tag{19}$$

where Y_1, Y_2 are any matrices with compatible dimensions over $\mathbb{H}$.

By Lemma 1, let the singular value decomposition of the matrix A_{22} and B_{11} in Problem 1 be

$$A_{22} = Q\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}R^*, \tag{20}$$

$$B_{11} = U\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}V^*, \tag{21}$$

where $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_s)$ is a positive diagonal matrix, $\lambda_i \neq 0$ $(i = 1, \ldots, s)$ are the singular values of A_{22}, $s = r(A_{22})$, $\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$ is a positive diagonal matrix, $\sigma_i \neq 0$ $(i = 1, \ldots, r)$ are the singular values of B_{11} and $r = r(B_{11})$.

$Q = (Q_1\ \ Q_2) \in \mathbb{H}^{m_2 \times m_2}$, $R = (R_1\ \ R_2) \in \mathbb{H}^{n_2 \times n_2}$, $U = (U_1\ \ U_2) \in \mathbb{H}^{n_1 \times n_1}$, $V = (V_1\ \ V_2) \in \mathbb{H}^{m_1 \times m_1}$ are unitary quaternion matrices, where $Q_1 \in \mathbb{H}^{m_2 \times s}$, $R_1 \in \mathbb{H}^{n_2 \times s}$, $U_1 \in \mathbb{H}^{n_1 \times r}$, and $V_1 \in \mathbb{H}^{m_1 \times r}$.

Theorem 4. *Problem 1 has a solution if and only if the following conditions are satisfied:*

(a) $r\begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix} = n_2$,

(b) $n_2 - r(A_{22}) = m_1 - r(B_{11})$, *that is* $n_2 - s = m_1 - r$,

(c) $\mathscr{R}(A_{21}B_{11}) \subset \mathscr{R}(A_{22})$,

(d) $\mathscr{R}(A_{12}^* B_{11}^*) \subset \mathscr{R}(A_{22}^*)$.

In that case, the general solution has the form of

$$A_{11} = B_{11}^{+} + A_{12}R\begin{pmatrix} \Lambda^{-1}Q_1^* A_{21}U_1\Sigma & 0 \\ H & -(V_2^* A_{12}R_2)^{-1} \end{pmatrix} \times V^* B_{11}^{+} + Y - YB_{11}B_{11}^{+}, \tag{22}$$

where H is an arbitrary matrix in $\mathbb{H}^{(n_2-s)\times r}$ and Y is an arbitrary matrix in $\mathbb{H}^{m_1 \times n_1}$.

Proof. If there exists an $m_1 \times n_1$ matrix A_{11} such that A is nonsingular and B_{11} is the corresponding block of A^{-1}, then (a) is satisfied. From $AB = BA = I$, we have that

$$\begin{aligned} A_{21}B_{11} + A_{22}B_{21} &= 0, \\ B_{11}A_{12} + B_{12}A_{22} &= 0, \end{aligned} \tag{23}$$

so that (c) and (d) are satisfied.

By (11), we have

$$r(A_{22}) + n(A_{22}) = n_2, \qquad r(B_{11}) + n(B_{11}) = m_1. \tag{24}$$

From Lemma 2, Notice that $\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ is the corresponding partitioning of B^{-1}, we have

$$n(B_{11}) = n(A_{22}), \tag{25}$$

implying that (b) is satisfied.

Conversely, from (c), we know that there exists a matrix $K \in \mathbb{H}^{n_2 \times m_1}$ such that

$$A_{21}B_{11} = A_{22}K. \tag{26}$$

Let

$$B_{21} = -K. \tag{27}$$

From (20), (21), and (26), we have

$$A_{21}U\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}V^* = Q\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}R^* K. \tag{28}$$

It follows that

$$Q^* A_{21}U\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}V^* V = Q^* Q\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}R^* KV. \tag{29}$$

This implies that

$$\begin{aligned} &\begin{pmatrix} Q_1^* A_{21}U_1 & Q_1^* A_{21}U_2 \\ Q_2^* A_{21}U_1 & Q_2^* A_{21}U_2 \end{pmatrix}\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \\ &\quad = \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} R_1^* KV_1 & R_1^* KV_2 \\ R_2^* KV_1 & R_2^* KV_2 \end{pmatrix}. \end{aligned} \tag{30}$$

Comparing corresponding blocks in (30), we obtain

$$Q_2^* A_{21}U_1 = 0. \tag{31}$$

Let $R^* KV = \widehat{K}$. From (29), (30), we have

$$\begin{gathered} \widehat{K} = \begin{pmatrix} \Lambda^{-1}Q_1^* A_{21}U_1\Sigma & 0 \\ H & K_{22} \end{pmatrix}, \\ H \in \mathbb{H}^{(n_2-s)\times r}, \quad K_{22} \in \mathbb{H}^{(n_2-s)\times(m_1-r)}. \end{gathered} \tag{32}$$

In the same way, from (d), we can obtain

$$V_1^* A_{12}R_2 = 0. \tag{33}$$

Notice that $\begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix}$ in (a) is a full column rank matrix. By (20), (21), and (33), we have

$$\begin{pmatrix} 0 & Q^* \\ V^* & 0 \end{pmatrix}\begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix}R = \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \\ V_1^* A_{12}R_1 & V_1^* A_{12}R_2 \\ V_2^* A_{12}R_1 & V_2^* A_{12}R_2 \end{pmatrix}, \tag{34}$$

so that

$$\begin{aligned} n_2 &= r\begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix} = r\left(\begin{pmatrix} 0 & Q^* \\ V^* & 0 \end{pmatrix}\begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix}R\right) \\ &= r\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \\ V_1^* A_{12}R_1 & V_1^* A_{12}R_2 \\ V_2^* A_{12}R_1 & V_2^* A_{12}R_2 \end{pmatrix} \\ &= r(\Lambda) + r(V_2^* A_{12}R_2) \\ &= s + r(V_2^* A_{12}R_2). \end{aligned} \tag{35}$$

It follows from (b) and (35) that $V_2^T A_{12}R_2$ is a full column rank matrix, so it is nonsingular.

From $AB = I$, we have the following matrix equation:

$$A_{11}B_{11} + A_{12}B_{21} = I, \tag{36}$$

that is

$$A_{11}B_{11} = I - A_{12}B_{21}, \quad I \in \mathbb{H}^{m_1 \times m_1}, \tag{37}$$

where B_{11}, A_{12} were given, $B_{21} = -K$ (from (27)). By Lemma 3, the matrix equation (37) has a solution if and only if

$$(I - A_{12}B_{21}) B_{11}^{+} B_{11} = I - A_{12}B_{21}. \tag{38}$$

By (21), (27), (32), and (33), we have that (38) is equivalent to:

$$(I + A_{12}K) V \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^{*} U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^{*} = I + A_{12}K. \tag{39}$$

We simplify the equation above. The left hand side reduces to $(I + A_{12}K)V_1V_1^{*}$ and so we have

$$A_{12}KV_1V_1^{*} - A_{12}K = I - V_1V_1^{*}. \tag{40}$$

So,

$$A_{12}R\widehat{K}V^{*}V_1V_1^{*} - A_{12}R\widehat{K}V^{*} = \begin{pmatrix} V_1 & V_2 \end{pmatrix} \begin{pmatrix} V_1^{*} \\ V_2^{*} \end{pmatrix} - V_1V_1^{*}. \tag{41}$$

This implies that

$$A_{12}R\widehat{K} \begin{pmatrix} V_1^{*}V_1 \\ V_2^{*}V_1 \end{pmatrix} V_1^{*} - A_{12}R\widehat{K} \begin{pmatrix} V_1^{*} \\ V_2^{*} \end{pmatrix} = V_2V_2^{*}, \tag{42}$$

so that

$$A_{12}R\widehat{K} \begin{pmatrix} I \\ 0 \end{pmatrix} V_1^{*} - A_{12}R\widehat{K} \begin{pmatrix} V_1^{*} \\ V_2^{*} \end{pmatrix} = V_2V_2^{*}. \tag{43}$$

So,

$$-A_{12}R\widehat{K} \begin{pmatrix} 0 \\ V_2^{*} \end{pmatrix} = V_2V_2^{*}, \tag{44}$$

and hence,

$$-\begin{pmatrix} A_{12}R_1 & A_{12}R_2 \end{pmatrix} \begin{pmatrix} \Lambda^{-1}Q_1^{*}A_{21}U_1\Sigma & 0 \\ H & K_{22} \end{pmatrix} \begin{pmatrix} 0 \\ V_2^{*} \end{pmatrix} = V_2V_2^{*}. \tag{45}$$

Finally, we obtain

$$A_{12}R_2K_{22}V_2^{*} = -V_2V_2^{*}. \tag{46}$$

Multiplying both sides of (46) by V^{*} from the left, considering (33) and the fact that $V_2^{*}A_{12}R_2$ is nonsingular, we have

$$K_{22} = -(V_2^{*}A_{12}R_2)^{-1}. \tag{47}$$

From Lemma 3, (38), (47), Problem 1 has a solution and the general solution is

$$A_{11} = B_{11}^{+} + A_{12}R \begin{pmatrix} \Lambda^{-1}Q_1^{*}A_{21}U_1\Sigma & 0 \\ H & -(V_2^{*}A_{12}R_2)^{-1} \end{pmatrix} \times V^{*}B_{11}^{+} + Y - YB_{11}B_{11}^{+}, \tag{48}$$

where H is an arbitrary matrix in $\mathbb{H}^{(n_2-s)\times r}$ and Y is an arbitrary matrix in $\mathbb{H}^{m_1\times n_1}$. □

3. An Example

In this section, we give a numerical example to illustrate the theoretical results.

Example 5. Consider Problem 1 with the parameter matrices as follows:

$$\begin{gathered} A_{12} = \begin{pmatrix} 2+j & \frac{1}{2}k \\ -k & 1+\frac{1}{2}j \end{pmatrix}, \\ A_{21} = \begin{pmatrix} \frac{3}{2}+\frac{1}{2}i & -\frac{1}{2}j-\frac{1}{2}k \\ \frac{1}{2}j+\frac{1}{2}k & \frac{3}{2}+\frac{1}{2}i \end{pmatrix}, \\ A_{22} = \begin{pmatrix} 2 & i \\ 2j & k \end{pmatrix}, \qquad B_{11} = \begin{pmatrix} 1 & i \\ j & k \end{pmatrix}. \end{gathered} \tag{49}$$

It is easy to show that (c), (d) are satisfied, and that

$$\begin{gathered} n_2 = r\begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix} = 2, \\ n_2 - r(A_{22}) = m_1 - r(B_{11}) = 0, \end{gathered} \tag{50}$$

so (a), (b) are satisfied too. Therefore, we have

$$\begin{gathered} B_{11}^{+} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2}j \\ -\frac{1}{2}i & -\frac{1}{2}k \end{pmatrix}, \\ A_{22} = Q\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}R^{*}, \qquad B_{11} = U\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}V^{*}, \end{gathered} \tag{51}$$

where

$$\begin{gathered} Q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ j & k \end{pmatrix}, \qquad \Lambda = \begin{pmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}, \\ R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ j & k \end{pmatrix}, \\ \Sigma = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}, \qquad V = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{gathered} \tag{52}$$

We also have

$$\begin{gathered} Q_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ j & k \end{pmatrix}, \qquad R_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\ U_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ j & k \end{pmatrix}, \qquad V_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{gathered} \tag{53}$$

By Theorem 4, for an arbitrary matrices $Y \in \mathbb{H}^{2\times 2}$, we have

$$\begin{aligned} A_{11} &= B_{11}^{+} + A_{12}R\left(\Lambda^{-1}Q_1^{*}A_{21}U_1\Sigma\right)V^{*}B_{11}^{+} + Y - YB_{11}B_{11}^{+} \\ &= \begin{pmatrix} \frac{3}{2}+\frac{1}{4}j+\frac{1}{4}k & \frac{3}{4}+\frac{1}{4}i-\frac{3}{2}j \\ \frac{1}{2}-i+\frac{1}{4}j-\frac{1}{4}k & \frac{1}{4}-\frac{3}{4}i-\frac{1}{2}j-k \end{pmatrix}, \end{aligned} \tag{54}$$

it follows that

$$A = \begin{pmatrix} \frac{3}{2}+\frac{1}{4}j+\frac{1}{4}k & \frac{3}{4}+\frac{1}{4}i-\frac{3}{2}j & 2+j & \frac{1}{2}k \\ \frac{1}{2}-i+\frac{1}{4}j-\frac{1}{4}k & \frac{1}{4}-\frac{3}{4}i-\frac{1}{2}j-k & -k & 1+\frac{1}{2}j \\ \frac{3}{2}+\frac{1}{2}i & -\frac{1}{2}j-\frac{1}{2}k & 2 & i \\ \frac{1}{2}j+\frac{1}{2}k & \frac{3}{2}+\frac{1}{2}i & 2j & k \end{pmatrix},$$

$$A^{-1} = \begin{pmatrix} 1 & i & -1 & -1 \\ j & k & 0 & -1 \\ -1 & 0 & \frac{3}{4} & \frac{1}{2}-\frac{3}{4}j \\ -1 & -1 & \frac{1}{2}-i & \frac{1}{2}-\frac{1}{2}i-\frac{1}{2}j-k \end{pmatrix}. \tag{55}$$

The results verify the theoretical findings of Theorem 4.

Acknowledgments

The authors would like to give many thanks to the referees and Professor K. C. Sivakumar for their valuable suggestions and comments, which resulted in a great improvement of the paper. This research was supported by Grants from the Key Project of Scientific Research Innovation Foundation of Shanghai Municipal Education Commission (13ZZ080), the National Natural Science Foundation of China (11171205), the Natural Science Foundation of Shanghai (11ZR1412500), the Discipline Project at the corresponding level of Shanghai (A. 13-0101-12-005), and Shanghai Leading Academic Discipline Project (J50101).

References

[1] M. Fiedler and T. L. Markham, "Completing a matrix when certain entries of its inverse are specified," *Linear Algebra and Its Applications*, vol. 74, pp. 225–237, 1986.

[2] H. Dai, "Completing a symmetric 2 × 2 block matrix and its inverse," *Linear Algebra and Its Applications*, vol. 235, pp. 235–245, 1996.

[3] W. W. Barrett, M. E. Lundquist, C. R. Johnson, and H. J. Woerdeman, "Completing a block diagonal matrix with a partially prescribed inverse," *Linear Algebra and Its Applications*, vol. 223/224, pp. 73–87, 1995.

[4] S. L. Adler, *Quaternionic Quantum Mechanics and Quantum Fields*, Oxford University Press, New York, NY, USA, 1994.

[5] A. Razon and L. P. Horwitz, "Uniqueness of the scalar product in the tensor product of quaternion Hilbert modules," *Journal of Mathematical Physics*, vol. 33, no. 9, pp. 3098–3104, 1992.

[6] J. W. Shuai, "The quaternion NN model: the identification of colour images," *Acta Comput. Sinica*, vol. 18, no. 5, pp. 372–379, 1995 (Chinese).

[7] F. Zhang, "On numerical range of normal matrices of quaternions," *Journal of Mathematical and Physical Sciences*, vol. 29, no. 6, pp. 235–251, 1995.

[8] F. Z. Zhang, *Permanent Inequalities and Quaternion Matrices [Ph.D. thesis]*, University of California, Santa Barbara, Calif, USA, 1993.

[9] F. Zhang, "Quaternions and matrices of quaternions," *Linear Algebra and Its Applications*, vol. 251, pp. 21–57, 1997.

[10] Q.-W. Wang, "The general solution to a system of real quaternion matrix equations," *Computers & Mathematics with Applications*, vol. 49, no. 5-6, pp. 665–675, 2005.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 178209, 7 pages
http://dx.doi.org/10.1155/2013/178209

Research Article

Fuzzy Approximate Solution of Positive Fully Fuzzy Linear Matrix Equations

Xiaobin Guo[1] and Dequan Shang[2]

[1] *College of Mathematics and Statistics, Northwest Normal University, Lanzhou 730070, China*
[2] *Department of Public Courses, Gansu College of Traditional Chinese Medicine, Lanzhou 730000, China*

Correspondence should be addressed to Dequan Shang; gxbglz@163.com

Received 22 November 2012; Accepted 19 January 2013

Academic Editor: Panayiotis J. Psarrakos

The fuzzy matrix equations $\widetilde{A}\otimes\widetilde{X}\otimes\widetilde{B}=\widetilde{C}$ in which $\widetilde{A}$, $\widetilde{B}$, and $\widetilde{C}$ are $m\times m$, $n\times n$, and $m\times n$ nonnegative LR fuzzy numbers matrices, respectively, are investigated. The fuzzy matrix systems is extended into three crisp systems of linear matrix equations according to arithmetic operations of LR fuzzy numbers. Based on pseudoinverse of matrix, the fuzzy approximate solution of original fuzzy systems is obtained by solving the crisp linear matrix systems. In addition, the existence condition of nonnegative fuzzy solution is discussed. Two examples are calculated to illustrate the proposed method.

1. Introduction

Since many real-world engineering systems are too complex to be defined in precise terms, imprecision is often involved in any engineering design process. Fuzzy systems have an essential role in this fuzzy modeling, which can formulate uncertainty in actual environment. In many matrix equations, some or all of the system parameters are vague or imprecise, and fuzzy mathematics is a better tool than crisp mathematics for modeling these problems, and hence solving a fuzzy matrix equation is becoming more important. The concept of fuzzy numbers and arithmetic operations with these numbers were first introduced and investigated by Zadeh [1, 2], Dubois and Prade [3], and Nahmias [4]. A different approach to fuzzy numbers and the structure of fuzzy number spaces was given by Puri and Ralescu [5], Goetschel, Jr. and Voxman [6], and Wu and Ma [7, 8].

In the past decades, many researchers have studied the fuzzy linear equations such as fuzzy linear systems (FLS), dual fuzzy linear systems (DFLS), general fuzzy linear systems (GFLS), fully fuzzy linear systems (FFLS), dual fully fuzzy linear systems (DFFLS), and general dual fuzzy linear systems (GDFLS). These works were performed mainly by Friedman et al. [9, 10], Allahviranloo et al. [11–17], Abbasbandy et al. [18–21], Wang and Zheng [22, 23], and Dehghan et al. [24, 25]. The general method they applied is the fuzzy linear equations were converted to a crisp function system of linear equations with high order according to the embedding principles and algebraic operations of fuzzy numbers. Then the fuzzy solution of the original fuzzy linear systems was derived from solving the crisp function linear systems. However, for a fuzzy matrix equation which always has a wide use in control theory and control engineering, few works have been done in the past. In 2009, Allahviranloo et al. [26] discussed the fuzzy linear matrix equations (FLME) of the form $A\widetilde{X}B=\widetilde{C}$ in which the matrices A and B are known $m\times m$ and $n\times n$ real matrices, respectively; $\widetilde{C}$ is a given $m\times n$ fuzzy matrix. By using the parametric form of fuzzy number, they derived necessary and sufficient conditions for the existence condition of fuzzy solutions and designed a numerical procedure for calculating the solutions of the fuzzy matrix equations. In 2011, Guo et al. [27–29] investigated a class of fuzzy matrix equations $A\widetilde{x}=\widetilde{B}$ by means of the block Gaussian elimination method and studied the least squares solutions of the inconsistent fuzzy matrix equation $A\widetilde{x}=\widetilde{B}$ by using generalized inverses of the matrix, and discussed fuzzy symmetric solutions of fuzzy matrix equations $A\widetilde{X}=\widetilde{B}$. What is more, there are two shortcomings in the above fuzzy systems. The first is that in these fuzzy linear systems and fuzzy matrix systems the fuzzy elements were denoted by triangular fuzzy numbers, so the extended model equations always contain parameter r, $0 \le r \le 1$ which makes their computation especially numerical implementation inconvenient in some sense. The other one is that the weak fuzzy solution of fuzzy linear systems $A\tilde{x} = \tilde{b}$ does not exist sometimes see; [30].

To make the multiplication of fuzzy numbers easy and handle the fully fuzzy systems, Dubois and Prade [3] introduced the LR fuzzy number in 1978. We know that triangular fuzzy numbers and trapezoidal fuzzy numbers [31] are just specious cases of LR fuzzy numbers. In 2006, Dehghan et al. [25] discussed firstly computational methods for fully fuzzy linear systems $\widetilde{A}\tilde{x} = \tilde{b}$ whose coefficient matrix and the right-hand side vector are LR fuzzy numbers. In 2012, Allahviranloo et al. studied LR fuzzy linear systems [32] by the linear programming with equality constraints. Otadi and Mosleh considered the nonnegative fuzzy solution [33] of fully fuzzy matrix equations $\widetilde{A}\widetilde{X} = \widetilde{B}$ by employing linear programming with equality constraints at the same year. Recently, Guo and Shang [34] investigated the fuzzy approximate solution of LR fuzzy Sylvester matrix equations $A\widetilde{X} + \widetilde{X}B = \widetilde{C}$. In this paper we propose a general model for solving the fuzzy linear matrix equation $\widetilde{A} \otimes \widetilde{X} \otimes \widetilde{B} = \widetilde{C}$, where $\widetilde{A}$, $\widetilde{B}$, and $\widetilde{C}$ are $m \times m$, $n \times n$, and $m \times n$ nonnegative LR fuzzy numbers matrices, respectively. The model is proposed in this way; that is, we extend the fuzzy linear matrix system into a system of linear matrix equations according to arithmetic operations of LR fuzzy numbers. The LR fuzzy solution of the original matrix equation is derived from solving crisp systems of linear matrix equations. The structure of this paper is organized as follows.

In Section 2, we recall the LR fuzzy numbers and present the concept of fully fuzzy linear matrix equation. The computing model to the positive fully fuzzy linear matrix equation is proposed in detail and the fuzzy approximate solution of the fuzzy linear matrix equation is obtained by using pseudo-inverse in Section 3. Some examples are given to illustrate our method in Section 4 and the conclusion is drawn in Section 5.

2. Preliminaries

2.1. The LR Fuzzy Number

Definition 1 (see [1]). A fuzzy number is a fuzzy set like $u : R \to I = [0, 1]$ which satisfies the following.

(1) u is upper semicontinuous,

(2) u is fuzzy convex, that is, $u(\lambda x + (1 - \lambda)y) \ge \min\{u(x), u(y)\}$ for all $x, y \in R$, $\lambda \in [0, 1]$,

(3) u is normal, that is, there exists $x_0 \in R$ such that $u(x_0) = 1$,

(4) $\operatorname{supp} u = \{x \in R \mid u(x) > 0\}$ is the support of the u, and its closure $\operatorname{cl}(\operatorname{supp} u)$ is compact.

Let E^1 be the set of all fuzzy numbers on R.

Definition 2 (see [3]). A fuzzy number $\widetilde{M}$ is said to be an LR fuzzy number if

$$\mu_{\widetilde{M}}(x) = \begin{cases} L\left(\dfrac{m - x}{\alpha}\right), & x \le m, \ \alpha > 0, \\ R\left(\dfrac{x - m}{\beta}\right), & x \ge m, \ \beta > 0, \end{cases} \tag{1}$$

where m, α, and β are called the mean value, left, and right spreads of $\widetilde{M}$, respectively. The function $L(\cdot)$, which is called left shape function satisfies

(1) $L(x) = L(-x)$,

(2) $L(0) = 1$ and $L(1) = 0$,

(3) $L(x)$ is nonincreasing on $[0, \infty)$.

The definition of a right shape function $R(\cdot)$ is similar to that of $L(\cdot)$.

Clearly, $\widetilde{M} = (m, \alpha, \beta)_{\mathrm{LR}}$ is positive (negative) if and only if $m - \alpha > 0$ $(m + \beta < 0)$.

Also, two LR fuzzy numbers $\widetilde{M} = (m, \alpha, \beta)_{\mathrm{LR}}$ and $\widetilde{N} = (n, \gamma, \delta)_{\mathrm{LR}}$ are said to be equal, if and only if $m = n$, $\alpha = \gamma$, and $\beta = \delta$.

Definition 3 (see [5]). For arbitrary LR fuzzy numbers $\widetilde{M} = (m, \alpha, \beta)_{\mathrm{LR}}$ and $\widetilde{N} = (n, \gamma, \delta)_{\mathrm{LR}}$, we have the following.

(1) Addition:

$$\widetilde{M} \oplus \widetilde{N} = (m, \alpha, \beta)_{\mathrm{LR}} \oplus (n, \gamma, \delta)_{\mathrm{LR}} = (m + n, \alpha + \gamma, \beta + \delta)_{\mathrm{LR}}. \tag{2}$$

(2) Multiplication:

(i) If $\widetilde{M} > 0$ and $\widetilde{N} > 0$, then

$$\begin{aligned} \widetilde{M} \otimes \widetilde{N} &= (m, \alpha, \beta)_{\mathrm{LR}} \otimes (n, \gamma, \delta)_{\mathrm{LR}} \\ &\cong (mn, m\gamma + n\alpha, m\delta + n\beta)_{\mathrm{LR}}. \end{aligned} \tag{3}$$

(ii) if $\widetilde{M} < 0$ and $\widetilde{N} > 0$, then

$$\begin{aligned} \widetilde{M} \otimes \widetilde{N} &= (m, \alpha, \beta)_{\mathrm{RL}} \otimes (n, \gamma, \delta)_{\mathrm{LR}} \\ &\cong (mn, n\alpha - m\delta, n\beta - m\gamma)_{\mathrm{RL}}, \end{aligned} \tag{4}$$

(iii) if $\widetilde{M} < 0$ and $\widetilde{N} < 0$, then

$$\begin{aligned} \widetilde{M} \otimes \widetilde{N} &= (m, \alpha, \beta)_{\mathrm{RL}} \otimes (n, \gamma, \delta)_{\mathrm{LR}} \\ &\cong (mn, -m\delta - n\beta, -m\gamma - n\alpha)_{\mathrm{RL}}, \end{aligned} \tag{5}$$

(3) Scalar multiplication:

$$\begin{aligned} \lambda \times \widetilde{M} &= \lambda \times (m, \alpha, \beta)_{\mathrm{LR}} \\ &= \begin{cases} (\lambda m, \lambda\alpha, \lambda\beta)_{\mathrm{LR}}, & \lambda \ge 0, \\ (\lambda m, -\lambda\beta, -\lambda\alpha)_{\mathrm{RL}}, & \lambda < 0. \end{cases} \end{aligned} \tag{6}$$

2.2. The LR Fuzzy Matrix

Definition 4. A matrix $\widetilde{A} = (\widetilde{a}_{ij})$ is called an LR fuzzy matrix, if each element $\widetilde{a}_{ij}$ of $\widetilde{A}$ is an LR fuzzy number.

$\widetilde{A}$ will be positive (negative) and denoted by $\widetilde{A} > 0$ ($\widetilde{A} < 0$) if each element $\widetilde{a}_{ij}$ of $\widetilde{A}$ is positive (negative), where $\widetilde{a}_{ij} = (a_{ij}, a^l_{ij}, a^r_{ij})_{\mathrm{LR}}$. Up to the rest of this paper, we use positive LR fuzzy numbers and formulas given in Definition 3. For example, we represent $m \times n$ LR fuzzy matrix $\widetilde{A} = (\widetilde{a}_{ij})$, that $\widetilde{a}_{ij} = (a_{ij}, \alpha_{ij}, \beta_{ij})_{\mathrm{LR}}$ with new notation $\widetilde{A} = (A, M, N)$, where $A = (a_{ij})$, $M = (\alpha_{ij})$ and $N = (\beta_{ij})$ are three $m \times n$ crisp matrices. In particular, an n dimensions LR fuzzy numbers vector $\widetilde{x}$ can be denoted by (x, x^l, x^r), where $x = (x_i)$, $x^l = (x^l_i)$ and $x^r = (x^r_i)$ are three n dimensions crisp vectors.

Definition 5. Let $\widetilde{A} = (\widetilde{a}_{ij})$ and $\widetilde{B} = (\widetilde{b}_{ij})$ be two $m \times n$ and $n \times p$ fuzzy matrices; we define $\widetilde{A} \otimes \widetilde{B} = \widetilde{C} = (\widetilde{c}_{ij})$ which is an $m \times p$ fuzzy matrix, where

$$\widetilde{c}_{ij} = \sum_{k=1,2,\ldots,n}^{\oplus} \widetilde{a}_{ik} \otimes \widetilde{b}_{kj}. \tag{7}$$

2.3. The Fully Fuzzy Linear Matrix Equation

Definition 6. The matrix system:

$$\begin{pmatrix} \widetilde{a}_{11} & \widetilde{a}_{12} & \cdots & \widetilde{a}_{1m} \\ \widetilde{a}_{21} & \widetilde{a}_{22} & \cdots & \widetilde{a}_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \widetilde{a}_{m1} & \widetilde{a}_{m2} & \cdots & \widetilde{a}_{mm} \end{pmatrix} \otimes \begin{pmatrix} \widetilde{x}_{11} & \widetilde{x}_{12} & \cdots & \widetilde{x}_{1n} \\ \widetilde{x}_{21} & \widetilde{x}_{22} & \cdots & \widetilde{x}_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \widetilde{x}_{m1} & \widetilde{x}_{m2} & \cdots & \widetilde{x}_{mn} \end{pmatrix}$$
$$\otimes \begin{pmatrix} \widetilde{b}_{11} & \widetilde{b}_{12} & \cdots & \widetilde{b}_{1n} \\ \widetilde{b}_{21} & \widetilde{b}_{22} & \cdots & \widetilde{b}_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \widetilde{b}_{n1} & \widetilde{b}_{n2} & \cdots & \widetilde{b}_{nn} \end{pmatrix} \tag{8}$$
$$= \begin{pmatrix} \widetilde{c}_{11} & \widetilde{c}_{12} & \cdots & \widetilde{c}_{1n} \\ \widetilde{c}_{21} & \widetilde{c}_{22} & \cdots & \widetilde{c}_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \widetilde{c}_{m1} & \widetilde{c}_{m2} & \cdots & \widetilde{c}_{mn} \end{pmatrix},$$

where $\widetilde{a}_{ij}$, $1 \le i, j \le m$, $\widetilde{b}_{ij}$, $1 \le i, j \le n$, and $\widetilde{c}_{ij}$, $1 \le i \le m$, $1 \le j \le n$ are LR fuzzy numbers, is called an LR fully fuzzy linear matrix equation (FFLME).

Using matrix notation, we have

$$\widetilde{A} \otimes \widetilde{X} \otimes \widetilde{B} = \widetilde{C}. \tag{9}$$

A fuzzy numbers matrix

$$\widetilde{X} = (x_{ij}, y_{ij}, z_{ij})_{\mathrm{LR}}, \quad 1 \le i \le m,\ 1 \le j \le n \tag{10}$$

is called an LR fuzzy approximate solution of fully fuzzy linear matrix equation (8) if $\widetilde{X}$ satisfies (9).

Up to the rest of this paper, we will discuss the nonnegative solution $\widetilde{X} = (X, Y, Z) \ge 0$ of FFLME $\widetilde{A} \otimes \widetilde{X} \otimes \widetilde{B} = \widetilde{C}$, where $\widetilde{A} = (A, M, N) \ge 0$, $\widetilde{B} = (B, E, F) \ge 0$, and $\widetilde{C} = (C, G, H) \ge 0$.

3. Method for Solving FFLME

First, we extend the fuzzy linear matrix system (9) into three systems of linear matrix equations according to the LR fuzzy number and its arithmetic operations.

Theorem 7. *The fuzzy linear matrix system* (9) *can be extended into the following model:*

$$\begin{gathered} AXB = C, \\ AXE + AYB + MXB = G, \\ AXF + AZB + NXB = H. \end{gathered} \tag{11}$$

Proof. We denote $\widetilde{A} = (A, M, N)$, $\widetilde{B} = (B, E, F)$ and $\widetilde{C} = (C, G, H)$, and assume $\widetilde{X} = (X, Y, Z) \ge 0$, then

$$\begin{aligned} \widetilde{A}\widetilde{X}\widetilde{B} &= (A, M, N) \otimes (X, Y, Z) \otimes (B, E, F) \\ &= (AX, AY + MX, AZ + NX) \otimes (B, E, F) \\ &= (AXB, AXE + AYB + MXB, AXF + AZB + NXB) \\ &= (C, G, H), \end{aligned} \tag{12}$$

according to multiplication of nonnegative LR fuzzy numbers of Definition 2. Thus we obtain a model for solving FFLME (9) as follows:

$$\begin{gathered} AXB = C, \\ AXE + AYB + MXB = G, \\ AXF + AZB + NXB = H. \end{gathered} \tag{13}$$

□

Secondly, in order to solve the fuzzy linear matrix equation (9), we need to consider the crisp systems of linear matrix equation (11). Supposing A and B are nonsingular crisp matrices, we have

$$\begin{gathered} X = A^{-1}CB^{-1}, \\ Y = A^{-1}G - XEB^{-1} - A^{-1}MX, \\ Z = A^{-1}H - XFB^{-1} - A^{-1}NX, \end{gathered} \tag{14}$$

that is,

$$\begin{gathered} X = A^{-1}CB^{-1}, \\ Y = A^{-1}G - A^{-1}CB^{-1}EB^{-1} - A^{-1}MA^{-1}CB^{-1}, \\ Z = A^{-1}H - A^{-1}CB^{-1}FB^{-1} - A^{-1}NA^{-1}CB^{-1}. \end{gathered} \tag{15}$$

Definition 8. Let $\widetilde{X} = (X, Y, Z)$ be an LR fuzzy matrix. If (X, Y, Z) is an exact solution of (11) such that $X \ge 0$, $Y \ge 0$, $Z \ge 0$, and $X - Y \ge 0$; we call $\widetilde{X} = (X, Y, Z)$ a nonnegative LR fuzzy approximate solution of (9).

Theorem 9. *Let $\widetilde{A} = (A, M, N)$, $\widetilde{B} = (B, E, F)$, and $\widetilde{C} = (C, G, H)$ be three nonnegative fuzzy matrices, respectively, and let A and B be the product of a permutation matrix by a diagonal matrix with positive diagonal entries. Moreover, let $GB \geq CB^{-1}E + MA^{-1}C$, $HB \geq CB^{-1}F + NA^{-1}C$, and $C + CB^{-1}E + MA^{-1}C \geq GB$. Then the systems $\widetilde{A} \otimes \widetilde{X} \otimes \widetilde{B} = \widetilde{C}$ has nonnegative fuzzy solutions.*

Proof. Our hypotheses on A and B imply that A^{-1} and B^{-1} exist and they are nonnegative matrices. Thus $X = A^{-1}CB^{-1} \geq 0$.

On the other hand, because $GB \geq CB^{-1}E + MA^{-1}C$ and $HB \geq CB^{-1}F + NA^{-1}C$, so with $Y = A^{-1}(GB - CB^{-1}E - MA^{-1}C)B^{-1}$ and $Z = A^{-1}(HB - CB^{-1}F - NA^{-1}C)B^{-1}$, we have $Y \geq 0$ and $Z \geq 0$. Thus $\widetilde{X} = (X, Y, Z)$ is a fuzzy matrix which satisfies $\widetilde{A} \otimes \widetilde{X} \otimes \widetilde{B} = \widetilde{C}$. Since $X - Y = A^{-1}(C - GB + CB^{-1}E + MA^{-1}C)B^{-1}$, the positivity property of $\widetilde{X}$ can be obtained from the condition $C + CB^{-1}E + MA^{-1}C \geq GB$. □

When A or B is a singular crisp matrix, the following result is obvious.

Theorem 10 (see [35]). *For linear matrix equations $AXB = C$, where $A \in R^{m\times m}$, $B \in R^{n\times n}$, and $C \in R^{m\times n}$. Then*

$$X = A^{\dagger}CB^{\dagger} \tag{16}$$

is its minimal norm least squares solution.

By the pseudoinverse of matrices, we solve model (11) *and obtain its minimal norm least squares solution as follows:*

$$\begin{aligned} X &= A^{\dagger}CB^{\dagger}, \\ Y &= A^{\dagger}G - XEB^{\dagger} - A^{\dagger}MX, \\ Z &= A^{\dagger}H - XFB^{\dagger} - A^{\dagger}NX, \end{aligned} \tag{17}$$

that is,

$$\begin{aligned} X &= A^{\dagger}CB^{\dagger}, \\ Y &= A^{\dagger}G - A^{\dagger}CB^{\dagger}EB^{\dagger} - A^{\dagger}MA^{\dagger}CB^{\dagger}, \\ Z &= A^{\dagger}H - A^{\dagger}CB^{\dagger}FB^{\dagger} - A^{\dagger}NA^{\dagger}CB^{\dagger}. \end{aligned} \tag{18}$$

Definition 11. Let $\widetilde{X} = (X, Y, Z)$ be an LR fuzzy matrix. If (X, Y, Z) is a minimal norm least squares solution of (11) such that $X \geq 0$, $Y \geq 0$, $Z \geq 0$, and $X - Y \geq 0$, we call $\widetilde{X} = (X, Y, Z)$ a nonnegative LR fuzzy minimal norm least squares solution of (9).

At last, we give a sufficient condition for nonnegative fuzzy minimal norm least squares solution of FFLME (9) in the same way.

Theorem 12. *Let $A^{\dagger}$ and $B^{\dagger}$ be nonnegative matrices. Moreover, let $GB \geq CB^{\dagger}E + MA^{\dagger}C$, $HB \geq CB^{\dagger}F + NA^{\dagger}C$, and $C + CB^{\dagger}E + MA^{\dagger}C \geq GB$. Then the systems $\widetilde{A} \otimes \widetilde{X} \otimes \widetilde{B} = \widetilde{C}$ has nonnegative fuzzy minimal norm least squares solutions.*

Proof. Since $A^{\dagger}$ and $B^{\dagger}$ are nonnegative matrices, we have $X = A^{\dagger}CB^{\dagger} \geq 0$.

Now that $GB \geq CB^{\dagger}E + MA^{\dagger}C$ and $HB \geq CB^{\dagger}F + NA^{\dagger}C$, therefore, with $Y = A^{\dagger}(GB - CB^{\dagger}E - MA^{\dagger}C)B^{\dagger}$ and $Z = A^{\dagger}(HB - CB^{\dagger}F - NA^{\dagger}C)B^{\dagger}$, we have $Y \geq 0$ and $Z \geq 0$. Thus $\widetilde{X} = (X, Y, Z)$ is a fuzzy matrix which satisfies $\widetilde{A} \otimes \widetilde{X} \otimes \widetilde{B} = \widetilde{C}$. Since $X - Y = A^{\dagger}(C - GB + CB^{\dagger}E + MA^{\dagger}C)B^{\dagger}$, the nonnegativity property of $\widetilde{X}$ can be obtained from the condition $C + CB^{\dagger}E + MA^{\dagger}C \geq GB$. □

The following Theorems give some results for such S^{-1} and $S^{\dagger}$ to be nonnegative. As usual, $(\cdot)^{\top}$ denotes the transpose of a matrix $(\cdot)$.

Theorem 13 (see [36]). *The inverse of a nonnegative matrix A is nonnegative if and only if A is a generalized permutation matrix.*

Theorem 14 (see [37]). *Let A be an $m \times m$ nonnegative matrix with rank r. Then the following assertions are equivalent:*

(a) $A^{\dagger} \geq 0$.

(b) *There exists a permutation matrix P, such that PA has the form*

$$PA = \begin{pmatrix} S_1 \\ S_2 \\ \vdots \\ S_r \\ O \end{pmatrix}, \tag{19}$$

where each S_i has rank 1 and the rows of S_i are orthogonal to the rows of S_j, whenever $i \neq j$, the zero matrix may be absent.

(c) $A^{\dagger} = \begin{pmatrix} KP^{\top} & KQ^{\top} \\ KP^{\top} & KP^{\top} \end{pmatrix}$ *for some positive diagonal matrix K. In this case,*

$$(P + Q)^{\dagger} = K(P + Q)^{\top}, \qquad (P - Q)^{\dagger} = K(P - Q)^{\top}. \tag{20}$$

4. Numerical Examples

Example 15. Consider the fully fuzzy linear matrix system

$$\begin{pmatrix} (2,1,1)_{\mathrm{LR}} & (1,0,1)_{\mathrm{LR}} \\ (1,0,0)_{\mathrm{LR}} & (2,1,0)_{\mathrm{LR}} \end{pmatrix} \otimes \begin{pmatrix} \widetilde{x}_{11} & \widetilde{x}_{12} & \widetilde{x}_{13} \\ \widetilde{x}_{21} & \widetilde{x}_{22} & \widetilde{x}_{23} \end{pmatrix} \otimes \begin{pmatrix} (1,0,1)_{\mathrm{LR}} & (2,1,1)_{\mathrm{LR}} & (1,1,0)_{\mathrm{LR}} \\ (2,1,0)_{\mathrm{LR}} & (1,0,0)_{\mathrm{LR}} & (2,1,1)_{\mathrm{LR}} \\ (1,0,0)_{\mathrm{LR}} & (2,1,1)_{\mathrm{LR}} & (1,0,1)_{\mathrm{LR}} \end{pmatrix} = \begin{pmatrix} (17,12,23)_{\mathrm{LR}} & (22,22,29)_{\mathrm{LR}} & (17,16,28)_{\mathrm{LR}} \\ (19,15,13)_{\mathrm{LR}} & (23,23,16)_{\mathrm{LR}} & (19,16,17)_{\mathrm{LR}} \end{pmatrix}. \tag{21}$$

By Theorem 7, the model of the above fuzzy linear matrix system is made of the following three crisp systems of linear matrix equations

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 17 & 22 & 17 \\ 19 & 23 & 19 \end{pmatrix},$$

$$\begin{aligned} &\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{pmatrix}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ &+ \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix} \\ &+ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 12 & 22 & 16 \\ 15 & 23 & 20 \end{pmatrix}, \end{aligned}$$

$$\begin{aligned} &\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \\ &+ \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} z_{11} & z_{12} & z_{13} \\ z_{21} & z_{22} & z_{23} \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix} \\ &+ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 23 & 29 & 28 \\ 13 & 16 & 17 \end{pmatrix}. \end{aligned} \tag{22}$$

Now that the matrix B is singular, according to formula (18), the solutions of the above three systems of linear matrix equations are as follows:

$$\begin{aligned} X &= A^{\dagger}CB^{\dagger} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{\dagger}\begin{pmatrix} 17 & 22 & 17 \\ 19 & 23 & 19 \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix}^{\dagger} \\ &= \begin{pmatrix} 1.5000 & 1.0000 & 1.5000 \\ 1.5000 & 2.0000 & 1.5000 \end{pmatrix}, \\ Y &= A^{\dagger}G - XEB^{\dagger} - A^{\dagger}MX = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{\dagger}\begin{pmatrix} 12 & 22 & 16 \\ 19 & 23 & 19 \end{pmatrix} \\ &\quad - X\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix}^{\dagger} - \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{\dagger}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}X \\ &= \begin{pmatrix} 1.0333 & 0.6667 & 0.8725 \\ 1.1250 & 1.6553 & 1.2578 \end{pmatrix}, \\ Z &= A^{\dagger}H - XFB^{\dagger} - A^{\dagger}NX = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{\dagger}\begin{pmatrix} 23 & 29 & 28 \\ 13 & 16 & 17 \end{pmatrix} \\ &\quad - X\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix}^{\dagger} - \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{\dagger}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}X \\ &= \begin{pmatrix} 8.3333 & 11.6667 & 9.3333 \\ 1.4167 & 1.3333 & 2.4167 \end{pmatrix}. \end{aligned} \tag{23}$$

By Definition 11, we know that the original fuzzy linear matrix equations have a nonnegative LR fuzzy solution

$$\widetilde{X} = \begin{pmatrix} \widetilde{x}_{11} & \widetilde{x}_{12} & \widetilde{x}_{13} \\ \widetilde{x}_{21} & \widetilde{x}_{22} & \widetilde{x}_{23} \end{pmatrix} = \begin{pmatrix} (1.5000, 1.0333, 8.3333)_{LR} & (1.0000, 0.6667, 11.6667)_{LR} & (1.5000, 0.8725, 9.3333)_{LR} \\ (1.5000, 1.1250, 1.4167)_{LR} & (2.0000, 1.6553, 1.3333)_{LR} & (1.5000, 1.2578, 2.4167)_{LR} \end{pmatrix}, \tag{24}$$

since $X \geq 0$, $Y \geq 0$, $Z \geq 0$, and $X - Y \geq 0$.

Example 16. Consider the following fuzzy matrix system:

$$\begin{aligned} &\begin{pmatrix} (1,1,0)_{LR} & (2,0,1)_{LR} \\ (2,0,1)_{LR} & (1,0,0)_{LR} \end{pmatrix}\begin{pmatrix} \widetilde{x}_{11} & \widetilde{x}_{12} \\ \widetilde{x}_{21} & \widetilde{x}_{22} \end{pmatrix}\begin{pmatrix} (1,0,1)_{LR} & (2,1,0)_{LR} \\ (2,1,1)_{LR} & (3,1,2)_{LR} \end{pmatrix} \\ &= \begin{pmatrix} (15,14,18)_{LR} & (25,24,24)_{LR} \\ (12,10,12)_{LR} & (20,17,15)_{LR} \end{pmatrix}. \end{aligned} \tag{25}$$

By Theorem 7, the model of the above fuzzy linear matrix system is made of following three crisp systems of linear matrix equations:

$$\begin{aligned} &\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 15 & 25 \\ 10 & 20 \end{pmatrix}, \\ &\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \\ &\quad + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 14 & 24 \\ 10 & 17 \end{pmatrix}, \\ &\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \\ &\quad + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 18 & 24 \\ 12 & 15 \end{pmatrix}. \end{aligned} \tag{26}$$

By the same way, we obtain that the solutions of the above three systems of linear matrix equations are as follows:

$$X = A^{-1}CB^{-1} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix},$$
$$Y = A^{-1}G - XEB^{-1} - A^{-1}MX = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad (27)$$
$$Z = A^{-1}H - XFB^{-1} - A^{-1}NX = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Since $X \geq 0$, $Y \geq 0$, $Z \geq 0$, and $X - Y \geq 0$, we know that the original fuzzy linear matrix equations have a nonnegative LR fuzzy solution given by

$$\widetilde{X} = \begin{pmatrix} \widetilde{x}_{11} & \widetilde{x}_{12} \\ \widetilde{x}_{21} & \widetilde{x}_{22} \end{pmatrix} = \begin{pmatrix} (1.000, 1.000, 0.000)_{\mathrm{LR}} & (1.000, 1.000, 0.000)_{\mathrm{LR}} \\ (2.000, 0.000, 1.000)_{\mathrm{LR}} & (2.000, 1.000, 0.000)_{\mathrm{LR}} \end{pmatrix}. \quad (28)$$

5. Conclusion

In this work we presented a model for solving fuzzy linear matrix equations $\widetilde{A} \otimes \widetilde{X} \otimes \widetilde{B} = \widetilde{C}$ in which $\widetilde{A}$ and $\widetilde{B}$ are $m \times m$ and $n \times n$ fuzzy matrices, respectively, and $\widetilde{C}$ is an $m \times n$ arbitrary LR fuzzy numbers matrix. The model was made of three crisp systems of linear equations which determined the mean value and the left and right spreads of the solution. The LR fuzzy approximate solution of the fuzzy linear matrix equation was derived from solving the crisp systems of linear matrix equations. In addition, the existence condition of strong LR fuzzy solution was studied. Numerical examples showed that our method is feasible to solve this type of fuzzy matrix equations. Based on LR fuzzy numbers and their operations, we can investigate all kinds of fully fuzzy matrix equations in future.

Acknowledgments

The work is supported by the Natural Scientific Funds of China (no. 71061013) and the Youth Research Ability Project of Northwest Normal University (NWNU-LKQN-1120).

References

[1] L. A. Zadeh, "Fuzzy sets," *Information and Computation*, vol. 8, pp. 338–353, 1965.

[2] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning. I," *Information Science*, vol. 8, pp. 199–249, 1975.

[3] D. Dubois and H. Prade, "Operations on fuzzy numbers," *International Journal of Systems Science*, vol. 9, no. 6, pp. 613–626, 1978.

[4] S. Nahmias, "Fuzzy variables," *Fuzzy Sets and Systems. An International Journal in Information Science and Engineering*, vol. 1, no. 2, pp. 97–110, 1978.

[5] M. L. Puri and D. A. Ralescu, "Differentials of fuzzy functions," *Journal of Mathematical Analysis and Applications*, vol. 91, no. 2, pp. 552–558, 1983.

[6] R. Goetschel, Jr. and W. Voxman, "Elementary fuzzy calculus," *Fuzzy Sets and Systems*, vol. 18, no. 1, pp. 31–43, 1986.

[7] C. X. Wu and M. Ma, "Embedding problem of fuzzy number space. I," *Fuzzy Sets and Systems*, vol. 44, no. 1, pp. 33–38, 1991.

[8] C. X. Wu and M. Ma, "Embedding problem of fuzzy number space. III," *Fuzzy Sets and Systems*, vol. 46, no. 2, pp. 281–286, 1992.

[9] M. Friedman, M. Ming, and A. Kandel, "Fuzzy linear systems," *Fuzzy Sets and Systems*, vol. 96, no. 2, pp. 201–209, 1998.

[10] M. Ma, M. Friedman, and A. Kandel, "Duality in fuzzy linear systems," *Fuzzy Sets and Systems*, vol. 109, no. 1, pp. 55–58, 2000.

[11] T. Allahviranloo and M. Ghanbari, "On the algebraic solution of fuzzy linear systems based on interval theory," *Applied Mathematical Modelling*, vol. 36, no. 11, pp. 5360–5379, 2012.

[12] T. Allahviranloo and S. Salahshour, "Fuzzy symmetric solutions of fuzzy linear systems," *Journal of Computational and Applied Mathematics*, vol. 235, no. 16, pp. 4545–4553, 2011.

[13] T. Allahviranloo and M. Ghanbari, "A new approach to obtain algebraic solution of interval linear systems," *Soft Computing*, vol. 16, no. 1, pp. 121–133, 2012.

[14] T. Allahviranloo, "Numerical methods for fuzzy system of linear equations," *Applied Mathematics and Computation*, vol. 155, no. 2, pp. 493–502, 2004.

[15] T. Allahviranloo, "Successive over relaxation iterative method for fuzzy system of linear equations," *Applied Mathematics and Computation*, vol. 162, no. 1, pp. 189–196, 2005.

[16] T. Allahviranloo, "The Adomian decomposition method for fuzzy system of linear equations," *Applied Mathematics and Computation*, vol. 163, no. 2, pp. 553–563, 2005.

[17] R. Nuraei, T. Allahviranloo, and M. Ghanbari, "Finding an inner estimation of the so- lution set of a fuzzy linear system," *Applied Mathematical Modelling*, vol. 37, no. 7, pp. 5148–5161, 2013.

[18] S. Abbasbandy, R. Ezzati, and A. Jafarian, "LU decomposition method for solving fuzzy system of linear equations," *Applied Mathematics and Computation*, vol. 172, no. 1, pp. 633–643, 2006.

[19] S. Abbasbandy, A. Jafarian, and R. Ezzati, "Conjugate gradient method for fuzzy symmetric positive definite system of linear equations," *Applied Mathematics and Computation*, vol. 171, no. 2, pp. 1184–1191, 2005.

[20] S. Abbasbandy, M. Otadi, and M. Mosleh, "Minimal solution of general dual fuzzy linear systems," *Chaos, Solitons & Fractals*, vol. 37, no. 4, pp. 1113–1124, 2008.

[21] B. Asady, S. Abbasbandy, and M. Alavi, "Fuzzy general linear systems," *Applied Mathematics and Computation*, vol. 169, no. 1, pp. 34–40, 2005.

[22] K. Wang and B. Zheng, "Inconsistent fuzzy linear systems," *Applied Mathematics and Computation*, vol. 181, no. 2, pp. 973–981, 2006.

[23] B. Zheng and K. Wang, "General fuzzy linear systems," *Applied Mathematics and Computation*, vol. 181, no. 2, pp. 1276–1286, 2006.

[24] M. Dehghan and B. Hashemi, "Iterative solution of fuzzy linear systems," *Applied Mathematics and Computation*, vol. 175, no. 1, pp. 645–674, 2006.

[25] M. Dehghan, B. Hashemi, and M. Ghatee, "Solution of the fully fuzzy linear systems using iterative techniques," *Chaos, Solitons & Fractals*, vol. 34, no. 2, pp. 316–336, 2007.

[26] T. Allahviranloo, N. Mikaeilvand, and M. Barkhordary, "Fuzzy linear matrix equation," *Fuzzy Optimization and Decision Making*, vol. 8, no. 2, pp. 165–177, 2009.

[27] X. Guo and Z. Gong, "Block Gaussian elimination methods for fuzzy matrix equations," *International Journal of Pure and Applied Mathematics*, vol. 58, no. 2, pp. 157–168, 2010.

[28] Z. Gong and X. Guo, "Inconsistent fuzzy matrix equations and its fuzzy least squares solutions," *Applied Mathematical Modelling*, vol. 35, no. 3, pp. 1456–1469, 2011.

[29] X. Guo and D. Shang, "Fuzzy symmetric solutions of fuzzy matrix equations," *Advances in Fuzzy Systems*, vol. 2012, Article ID 318069, 9 pages, 2012.

[30] T. Allahviranloo, M. Ghanbari, A. A. Hosseinzadeh, E. Haghi, and R. Nuraei, "A note on 'Fuzzy linear systems'," *Fuzzy Sets and Systems*, vol. 177, pp. 87–92, 2011.

[31] C.-T. Yeh, "Weighted semi-trapezoidal approximations of fuzzy numbers," *Fuzzy Sets and Systems*, vol. 165, pp. 61–80, 2011.

[32] T. Allahviranloo, F. H. Lotfi, M. K. Kiasari, and M. Khezerloo, "On the fuzzy solution of LR fuzzy linear systems," *Applied Mathematical Modelling*, vol. 37, no. 3, pp. 1170–1176, 2013.

[33] M. Otadi and M. Mosleh, "Solving fully fuzzy matrix equations," *Applied Mathematical Modelling*, vol. 36, no. 12, pp. 6114–6121, 2012.

[34] X. B. Guo and D. Q. Shang, "Approximate solutions of LR fuzzy Sylvester matrix equations," *Journal of Applied Mathematics*. In press.

[35] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, Springer-Verlag, New York, NY, USA, 2nd edition, 2003.

[36] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, NY, USA, 1979.

[37] R. J. Plemmons, "Regular nonnegative matrices," *Proceedings of the American Mathematical Society*, vol. 39, pp. 26–32, 1973.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 514984, 9 pages
http://dx.doi.org/10.1155/2013/514984

Research Article

Ranks of a Constrained Hermitian Matrix Expression with Applications

Shao-Wen Yu

Department of Mathematics, East China University of Science and Technology, Shanghai 200237, China

Correspondence should be addressed to Shao-Wen Yu; yushaowen@ecust.edu.cn

Received 12 November 2012; Accepted 4 January 2013

Academic Editor: Yang Zhang

We establish the formulas of the maximal and minimal ranks of the quaternion Hermitian matrix expression $C_4 - A_4XA_4^*$ where X is a Hermitian solution to quaternion matrix equations $A_1X = C_1$, $XB_1 = C_2$, and $A_3XA_3^* = C_3$. As applications, we give a new necessary and sufficient condition for the existence of Hermitian solution to the system of matrix equations $A_1X = C_1$, $XB_1 = C_2$, $A_3XA_3^* = C_3$, and $A_4XA_4^* = C_4$, which was investigated by Wang and Wu, 2010, by rank equalities. In addition, extremal ranks of the generalized Hermitian Schur complement $C_4 - A_4A_3^{\sim}A_4^*$ with respect to a Hermitian g-inverse $A_3^{\sim}$ of A_3, which is a common solution to quaternion matrix equations $A_1X = C_1$ and $XB_1 = C_2$, are also considered.

1. Introduction

Throughout this paper, we denote the real number field by $\mathbb{R}$, the complex number field by $\mathbb{C}$, the set of all $m \times n$ matrices over the quaternion algebra

$$\mathbb{H} = \left\{a_0 + a_1 i + a_2 j + a_3 k \mid i^2 = j^2 = k^2 = ijk = -1,\ a_0, a_1, a_2, a_3 \in \mathbb{R}\right\} \tag{1}$$

by $\mathbb{H}^{m\times n}$, the identity matrix with the appropriate size by I, the column right space, the row left space of a matrix A over $\mathbb{H}$ by $\mathscr{R}(A)$, $\mathscr{N}(A)$, respectively, the dimension of $\mathscr{R}(A)$ by $\dim \mathscr{R}(A)$, a Hermitian g-inverse of a matrix A by $X = A^{\sim}$ which satisfies $AA^{\sim}A = A$ and $X = X^*$, and the Moore-Penrose inverse of matrix A over $\mathbb{H}$ by $A^{\dagger}$ which satisfies four Penrose equations $AA^{\dagger}A = A$, $A^{\dagger}AA^{\dagger} = A^{\dagger}$, $(AA^{\dagger})^* = AA^{\dagger}$, and $(A^{\dagger}A)^* = A^{\dagger}A$. In this case $A^{\dagger}$ is unique and $(A^{\dagger})^* = (A^*)^{\dagger}$. Moreover, R_A and L_A stand for the two projectors $L_A = I - A^{\dagger}A$, $R_A = I - AA^{\dagger}$ induced by A. Clearly, R_A and L_A are idempotent, Hermitian and $R_A = L_{A^*}$. By [1], for a quaternion matrix A, $\dim \mathscr{R}(A) = \dim \mathscr{N}(A)$. $\dim \mathscr{R}(A)$ is called the rank of a quaternion matrix A and denoted by $r(A)$.

Mitra [2] investigated the system of matrix equations

$$A_1X = C_1, \qquad XB_1 = C_2. \tag{2}$$

Khatri and Mitra [3] gave necessary and sufficient conditions for the existence of the common Hermitian solution to (2) and presented an explicit expression for the general Hermitian solution to (2) by generalized inverses. Using the singular value decomposition (SVD), Yuan [4] investigated the general symmetric solution of (2) over the real number field $\mathbb{R}$. By the SVD, Dai and Lancaster [5] considered the symmetric solution of equation

$$AXA^* = C \tag{3}$$

over $\mathbb{R}$, which was motivated and illustrated with an inverse problem of vibration theory. Groß [6], Tian and Liu [7] gave the solvability conditions for Hermitian solution and its expressions of (3) over $\mathbb{C}$ in terms of generalized inverses, respectively. Liu, Tian and Takane [8] investigated ranks of Hermitian and skew-Hermitian solutions to the matrix equation (3). By using the generalized SVD, Chang and Wang [9] examined the symmetric solution to the matrix equations

$$A_3XA_3^* = C_3, \qquad A_4XA_4^* = C_4 \tag{4}$$

over $\mathbb{R}$. Note that all the matrix equations mentioned above are special cases of

$$\begin{aligned} A_1X = C_1, \qquad XB_1 = C_2, \\ A_3XA_3^* = C_3, \qquad A_4XA_4^* = C_4. \end{aligned} \tag{5}$$

Wang and Wu [10] gave some necessary and sufficient conditions for the existence of the common Hermitian solution to (5) for operators between Hilbert C^*-modules by generalized inverses and range inclusion of matrices. In view of the complicated computations of the generalized inverses of matrices, we naturally hope to establish a more practical, necessary, and sufficient condition for system (5) over quaternion algebra to have Hermitian solution by rank equalities.

As is known to us, solutions to matrix equations and ranks of solutions to matrix equations have been considered previously by many authors [10–34], and extremal ranks of matrix expressions can be used to characterize their rank invariance, nonsingularity, range inclusion, and solvability conditions of matrix equations. Tian and Cheng [35] investigated the maximal and minimal ranks of $A - BXC$ with respect to X with applications; Tian [36] gave the maximal and minimal ranks of $A_1 - B_1XC_1$ subject to a consistent matrix equation $B_2XC_2 = A_2$. Tian and Liu [7] established the solvability conditions for (4) to have a Hermitian solution over $\mathbb{C}$ by the ranks of coefficient matrices. Wang and Jiang [20] derived extreme ranks of (skew)Hermitian solutions to a quaternion matrix equation $AXA^* + BYB^* = C$. Wang, Yu and Lin [31] derived the extremal ranks of $C_4 - A_4XB_4$ subject to a consistent system of matrix equations

$$A_1X = C_1, \qquad XB_1 = C_2, \qquad A_3XB_3 = C_3 \tag{6}$$

over $\mathbb{H}$ and gave a new solvability condition to system

$$\begin{aligned} A_1X = C_1, \qquad XB_1 = C_2, \\ A_3XB_3 = C_3, \qquad A_4XB_4 = C_4. \end{aligned} \tag{7}$$

In matrix theory and its applications, there are many matrix expressions that have symmetric patterns or involve Hermitian (skew-Hermitian) matrices. For example,

$$\begin{aligned} &A - BXB^*, \qquad A - BX \pm X^*B^*, \\ &A - BXB^* - CYC^*, \qquad A - BXC \pm (BXC)^*, \end{aligned} \tag{8}$$

where $A = \pm A^*$, B, and C are given and X and Y are variable matrices. In recent papers [7, 8, 37, 38], Liu and Tian considered some maximization and minimization problems on the ranks of Hermitian matrix expressions (8).

Define a Hermitian matrix expression

$$f(X) = C_4 - A_4XA_4^*, \tag{9}$$

where $C_4 = C_4^*$; we have an observation that by investigating extremal ranks of (9), where X is a Hermitian solution to a system of matrix equations

$$A_1X = C_1, \qquad XB_1 = C_2, \qquad A_3XA_3^* = C_3. \tag{10}$$

A new necessary and sufficient condition for system (5) to have Hermitian solution can be given by rank equalities, which is more practical than one given by generalized inverses and range inclusion of matrices.

It is well known that Schur complement is one of the most important matrix expressions in matrix theory; there have been many results in the literature on Schur complements and their applications [39–41]. Tian [36, 42] has investigated the maximal and minimal ranks of Schur complements with applications.

Motivated by the work mentioned above, we in this paper investigate the extremal ranks of the quaternion Hermitian matrix expression (9) subject to the consistent system of quaternion matrix equations (10) and its applications. In Section 2, we derive the formulas of extremal ranks of (9) with respect to Hermitian solution of (10). As applications, in Section 3, we give a new, necessary, and sufficient condition for the existence of Hermitian solution to system (5) by rank equalities. In Section 4, we derive extremal ranks of generalized Hermitian Schur complement subject to (2). We also consider the rank invariance problem in Section 5.

2. Extremal Ranks of (9) Subject to System (10)

Corollary 8 in [10] over Hilbert C^*-modules can be changed into the following lemma over $\mathbb{H}$.

Lemma 1. *Let $A_1, C_1 \in \mathbb{H}^{m\times n}$, $B_1, C_2 \in \mathbb{H}^{n\times s}$, $A_3 \in \mathbb{H}^{r\times n}$, $C_3 \in \mathbb{H}^{r\times r}$ be given, and $F = B_1^*L_{A_1}$, $M = SL_F$, $S = A_3L_{A_1}$, $D = C_2^* - B_1^*A_1^\dagger C_1$, $J = A_1^\dagger C_1 + F^\dagger D$, $G = C_3 - A_3(J + L_{A_1}L_F^*J^*)A_3^*$; then the following statements are equivalent:*

(1) *the system* (10) *have a Hermitian solution,*

(2) $C_3 = C_3^*$,

$$A_1C_2 = C_1B_1, \qquad A_1C_1^* = C_1A_1^*, \qquad B_1^*C_2 = C_2^*B_1, \tag{11}$$

$$R_{A_1}C_1 = 0, \qquad R_FD = 0, \qquad R_MG = 0, \tag{12}$$

(3) $C_3 = C_3^*$; *the equalities in* (11) *hold and*

$$\begin{gathered} r\begin{bmatrix} A_1 & C_1 \end{bmatrix} = r(A_1), \qquad r\begin{bmatrix} A_1 & C_1 \\ B_1^* & C_2^* \end{bmatrix} = r\begin{bmatrix} A_1 \\ B_1^* \end{bmatrix}, \\ r\begin{bmatrix} A_1 & C_1A_3^* \\ B_1^* & C_2^*A_3^* \\ A_3 & C_3 \end{bmatrix} = r\begin{bmatrix} A_1 \\ B_1^* \\ A_3 \end{bmatrix}. \end{gathered} \tag{13}$$

In that case, the general Hermitian solution of (10) *can be expressed as*

$$\begin{aligned} X = {} & J + L_{A_1}L_FJ^* + L_{A_1}L_FM^\dagger G\left(M^\dagger\right)^*L_FL_{A_1} \\ & + L_{A_1}L_FL_MVL_FL_{A_1} + L_{A_1}L_FV^*L_ML_FL_{A_1}, \end{aligned} \tag{14}$$

where V is Hermitian matrix over $\mathbb{H}$ with compatible size.

Lemma 2 (see Lemma 2.4 in [24]). *Let $A \in \mathbb{H}^{m\times n}$, $B \in \mathbb{H}^{m\times k}$, $C \in \mathbb{H}^{l\times n}$, $D \in \mathbb{H}^{j\times k}$, and $E \in \mathbb{H}^{l\times i}$. Then the following rank equalities hold:*

(a) $r(CL_A) = r\begin{bmatrix} A \\ C \end{bmatrix} - r(A)$,

(b) $r\begin{bmatrix} B & AL_C \end{bmatrix} = r\begin{bmatrix} B & A \\ 0 & C \end{bmatrix} - r(C)$,

(c) $r\begin{bmatrix} C \\ R_BA \end{bmatrix} = r\begin{bmatrix} C & 0 \\ A & B \end{bmatrix} - r(B)$,

(d) $r\begin{bmatrix} A & BL_D \\ R_EC & 0 \end{bmatrix} = r\begin{bmatrix} A & B & 0 \\ C & 0 & E \\ 0 & D & 0 \end{bmatrix} - r(D) - r(E)$.

Lemma 2 plays an important role in simplifying ranks of various block matrices.

Liu and Tian [38] has given the following lemma over a field. The result can be generalized to $\mathbb{H}$.

Lemma 3. *Let $A = \pm A^* \in \mathbb{H}^{m\times m}$, $B \in \mathbb{H}^{m\times n}$, and $C \in \mathbb{H}^{p\times m}$ be given; then*

$$\begin{aligned} &\max_{X\in\mathbb{H}^{n\times p}} r\left[A - BXC \mp (BXC)^*\right] \\ &\quad = \min\left\{r\begin{bmatrix} A & B & C^* \end{bmatrix}, r\begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix}, r\begin{bmatrix} A & C^* \\ C & 0 \end{bmatrix}\right\}, \\ &\min_{X\in\mathbb{H}^{n\times p}} r\left[A - BXC \mp (BXC)^*\right] \\ &\quad = 2r\begin{bmatrix} A & B & C^* \end{bmatrix} + \max\left\{s_1, s_2\right\}, \end{aligned} \tag{15}$$

where

$$\begin{aligned} s_1 &= r\begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix} - 2r\begin{bmatrix} A & B & C^* \\ B^* & 0 & 0 \end{bmatrix}, \\ s_2 &= r\begin{bmatrix} A & C^* \\ C & 0 \end{bmatrix} - 2r\begin{bmatrix} A & B & C^* \\ C & 0 & 0 \end{bmatrix}. \end{aligned} \tag{16}$$

If $\mathscr{R}(B) \subseteq \mathscr{R}(C^*)$,

$$\begin{aligned} \max_X r\left[A - BXC - (BXC)^*\right] &= \min\left\{r\begin{bmatrix} A & C^* \end{bmatrix}, r\begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix}\right\}, \\ \max_X r\left[A - BXC - (BXC)^*\right] &= \min\left\{r\begin{bmatrix} A & C^* \end{bmatrix}, r\begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix}\right\}. \end{aligned} \tag{17}$$

Now we consider the extremal ranks of the matrix expression (9) subject to the consistent system (10).

Theorem 4. *Let A_1, C_1, B_1, C_2, A_3, and C_3 be defined as Lemma 1, $C_4 \in \mathbb{H}^{t\times t}$, and $A_4 \in \mathbb{H}^{t\times n}$. Then the extremal ranks of the quaternion matrix expression $f(X)$ defined as (9) subject to system (10) are the following:*

$$\max r\left[f(X)\right] = \min\{a, b\}, \tag{18}$$

where

$$\begin{aligned} a &= r\begin{bmatrix} C_4 & A_4 \\ C_2^*A_4^* & B_1^* \\ C_1A_4^* & A_1 \end{bmatrix} - r\begin{bmatrix} B_1^* \\ A_1 \end{bmatrix}, \\ b &= r\begin{bmatrix} 0 & A_4^* & A_3^* & B_1 & A_1^* \\ A_4 & C_4 & 0 & 0 & 0 \\ A_3 & 0 & -C_3 & -A_3C_2 & -A_3C_1^* \\ B_1^* & 0 & -C_2^*A_3^* & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1A_3^* & -C_1B_1 & -C_1A_1^* \end{bmatrix} - 2r\begin{bmatrix} A_3 \\ B_1^* \\ A_1 \end{bmatrix}, \end{aligned} \tag{19}$$

$$\begin{aligned} \min r\left[f(X)\right] &= 2r\begin{bmatrix} C_4 & A_4 \\ C_2^*A_4^* & B_1^* \\ C_1A_4^* & A_1 \end{bmatrix} \\ &\quad + r\begin{bmatrix} 0 & A_4^* & A_3^* & B_1 & A_1^* \\ A_4 & C_4 & 0 & 0 & 0 \\ A_3 & 0 & -C_3 & -A_3C_2 & -A_3C_1^* \\ B_1^* & 0 & -C_2^*A_3^* & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1A_3^* & -C_1B_1 & -C_1A_1^* \end{bmatrix} \\ &\quad - 2r\begin{bmatrix} 0 & A_4^* & B_1 & A_1^* \\ A_4 & C_4 & 0 & 0 \\ A_3 & 0 & -A_3C_2 & -A_3C_1^* \\ B_1^* & 0 & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1B_1 & -C_1A_1^* \end{bmatrix}. \end{aligned} \tag{20}$$

Proof. By Lemma 1, the general Hermitian solution of the system (10) can be expressed as

$$\begin{aligned} X = {} & J + L_{A_1}L_FJ^* + L_{A_1}L_FM^\dagger G\left(M^\dagger\right)^* L_FL_{A_1} \\ & + L_{A_1}L_FL_MVL_FL_{A_1} + L_{A_1}L_FV^*L_ML_FL_{A_1}, \end{aligned} \tag{21}$$

where V is Hermitian matrix over $\mathbb{H}$ with appropriate size. Substituting (21) into (9) yields

$$\begin{aligned} f(X) = {} & C_4 - A_4\Big(J + L_{A_1}L_FJ^* \\ & + L_{A_1}L_FM^\dagger G\left(M^\dagger\right)^* L_FL_{A_1}\Big)A_4^* \\ & - A_4L_{A_1}L_FL_MVL_FL_{A_1}A_4^* \\ & - A_4L_{A_1}L_FV^*L_ML_FL_{A_1}A_4^*. \end{aligned} \tag{22}$$

Put

$$C_4 - A_4\left(J + L_{A_1}L_FJ^* + L_{A_1}L_FM^{\dagger}G\left(M^{\dagger}\right)^*L_FL_{A_1}\right)A_4^* = A,$$
$$J + L_{A_1}L_FJ^* + L_{A_1}L_FM^{\dagger}G\left(M^{\dagger}\right)^*L_FL_{A_1} = J',$$
$$A_4L_{A_1}L_FL_M = N,$$
$$L_FL_{A_1}A_4^* = P; \tag{23}$$

then

$$f(X) = A - NVP - (NVP)^*. \tag{24}$$

Note that $A = A^*$ and $\mathscr{R}(N) \subseteq \mathscr{R}(P^*)$. Thus, applying (17) to (24), we get the following:

$$\begin{aligned}\max r\left[f(X)\right] &= \max_V r\left(A - NVP - (NVP)^*\right)\\ &= \min\left\{r\begin{bmatrix}A & P^*\end{bmatrix}, r\begin{bmatrix}A & N\\ N^* & 0\end{bmatrix}\right\},\\ \min r\left[f(X)\right] &= \min_V r\left(A - NVP - (NVP)^*\right)\\ &= 2r\begin{bmatrix}A & P^*\end{bmatrix} + r\begin{bmatrix}A & N\\ N^* & 0\end{bmatrix} - 2r\begin{bmatrix}A & N\\ P & 0\end{bmatrix}.\end{aligned} \tag{25}$$

Now we simplify the ranks of block matrices in (25).

In view of Lemma 2, block Gaussian elimination, (11), (12), and (23), we have the following:

$$r(F) = r\left(B_1^*L_{A_1}\right) = r\begin{bmatrix}B_1^*\\ A_1\end{bmatrix} - r(A_1),$$

$$\begin{aligned}r(M) = r\left(SL_F\right) &= r\begin{bmatrix}S\\ F\end{bmatrix} - r(F)\\ &= r\begin{bmatrix}A_3L_{A_1}\\ B_1^*L_{A_1}\end{bmatrix} - r(F)\\ &= r\begin{bmatrix}A_3\\ B_1^*\\ A_1\end{bmatrix} - r(A_1) - r(F),\end{aligned}$$

$$\begin{aligned}r\begin{bmatrix}A & P^*\end{bmatrix} &= r\begin{bmatrix}C_4 - A_4JA_4^* & P^*\end{bmatrix}\\ &= r\begin{bmatrix}C_4 - A_4JA_4^* & A_4L_{A_1}\\ 0 & F\end{bmatrix} - r(F)\\ &= r\begin{bmatrix}C_4 - A_4JA_4^* & A_4\\ 0 & B_1^*\\ 0 & A_1\end{bmatrix} - r(F) - r(A_1)\\ &= r\begin{bmatrix}C_4 & A_4\\ C_2^*A_4^* & B_1^*\\ C_1A_4^* & A_1\end{bmatrix} - r\begin{bmatrix}B_1^*\\ A_1\end{bmatrix},\end{aligned}$$

$$\begin{aligned}r\begin{bmatrix}A & N\\ N^* & 0\end{bmatrix} &= r\begin{bmatrix}C_4 - A_4J'A_4^* & A_4L_{A_1}L_FL_M\\ R_{M^*}R_{F^*}R_{A_1^*}A_4^* & 0\end{bmatrix}\\ &= r\begin{bmatrix}C_4 - A_4J'A_4^* & A_4 & 0 & 0 & 0\\ A_4^* & 0 & A_3^* & B_1 & A_1^*\\ 0 & A_3 & 0 & 0 & 0\\ 0 & B_1^* & 0 & 0 & 0\\ 0 & A_1 & 0 & 0 & 0\end{bmatrix}\\ &\quad - 2r(M) - 2r(F) - 2r(A_1)\\ &= r\begin{bmatrix}C_4 & A_4 & 0 & 0 & 0\\ A_4^* & 0 & A_3^* & B_1 & A_1^*\\ A_3J'A_4^* & A_3 & 0 & 0 & 0\\ B_1^*J'A_4^* & B_1^* & 0 & 0 & 0\\ A_1J'A_4^* & A_1 & 0 & 0 & 0\end{bmatrix} - 2r\begin{bmatrix}A_3\\ B_1^*\\ A_1\end{bmatrix}\\ &= r\begin{bmatrix}C_4 & A_4 & 0 & 0 & 0\\ A_4^* & 0 & A_3^* & B_1 & A_1^*\\ 0 & A_3 & -C_3 & -A_3C_2 & -A_3C_1^*\\ 0 & B_1^* & -C_2^*A_3^* & -C_2^*B_1 & -C_2^*A_1^*\\ 0 & A_1 & -C_1A_3^* & -C_1B_1 & -C_1A_1^*\end{bmatrix}\\ &\quad - 2r\begin{bmatrix}A_3\\ B_1^*\\ A_1\end{bmatrix}\\ &= r\begin{bmatrix}0 & A_4^* & A_3^* & B_1 & A_1^*\\ A_4 & C_4 & 0 & 0 & 0\\ A_3 & 0 & -C_3 & -A_3C_2 & -A_3C_1^*\\ B_1^* & 0 & -C_2^*A_3^* & -C_2^*B_1 & -C_2^*A_1^*\\ A_1 & 0 & -C_1A_3^* & -C_1B_1 & -C_1A_1^*\end{bmatrix}\\ &\quad - 2r\begin{bmatrix}A_3\\ B_1^*\\ A_1\end{bmatrix},\end{aligned}$$

$$\begin{aligned}r\begin{bmatrix}A & N\\ P & 0\end{bmatrix} &= r\begin{bmatrix}C_4 - A_4J'A_4^* & A_4L_{A_1}L_FL_M\\ R_{F^*}R_{A_1^*}A_4^* & 0\end{bmatrix}\\ &= r\begin{bmatrix}C_4 & A_4 & 0 & 0\\ A_4^* & 0 & B_1 & A_1^*\\ 0 & A_3 & -A_3C_2 & -A_3C_1^*\\ 0 & B_1^* & -C_2^*B_1 & -C_2^*A_1^*\\ 0 & A_1 & -C_1B_1 & -C_1A_1^*\end{bmatrix}\\ &\quad - r\begin{bmatrix}A_3\\ B_1^*\\ A_1\end{bmatrix} - r\begin{bmatrix}B_1^*\\ A_1\end{bmatrix}\end{aligned}$$

$$= r\begin{bmatrix} 0 & A_4^* & B_1 & A_1^* \\ A_4 & C_4 & 0 & 0 \\ A_3 & 0 & -A_3C_2 & -A_3C_1^* \\ B_1^* & 0 & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1B_1 & -C_1A_1^* \end{bmatrix}$$

$$- r\begin{bmatrix} A_3 \\ B_1^* \\ A_1 \end{bmatrix} - r\begin{bmatrix} B_1^* \\ A_1 \end{bmatrix}. \tag{26}$$

Substituting (26) into (25) yields (18) and (20). □

In Theorem 4, letting C_4 vanish and A_4 be I with appropriate size, respectively, we have the following.

Corollary 5. *Assume that* A_1, $C_1 \in \mathbb{H}^{m\times n}$, $B_1, C_2 \in \mathbb{H}^{n\times s}$, $A_3 \in \mathbb{H}^{r\times n}$, *and* $C_3 \in \mathbb{H}^{r\times r}$ *are given; then the maximal and minimal ranks of the Hermitian solution* X *to the system* (10) *can be expressed as*

$$\max r(X) = \min\{a, b\}, \tag{27}$$

where

$$a = n + r\begin{bmatrix} C_2^* \\ C_1 \end{bmatrix} - r\begin{bmatrix} B_1^* \\ A_1 \end{bmatrix},$$

$$b = 2n + r\begin{bmatrix} C_3 & A_3C_2 & A_3C_1^* \\ C_2^*A_3^* & C_2^*B_1 & C_2^*A_1^* \\ C_1A_3^* & C_1B_1 & C_1A_1^* \end{bmatrix} - 2r\begin{bmatrix} A_3 \\ B_1^* \\ A_1 \end{bmatrix},$$

$$\min r(X) = 2r\begin{bmatrix} C_2^* \\ C_1 \end{bmatrix}$$

$$+ r\begin{bmatrix} C_3 & A_3C_2 & A_3C_1^* \\ C_2^*A_3^* & C_2^*B_1 & C_2^*A_1^* \\ C_1A_3^* & C_1B_1 & C_1A_1^* \end{bmatrix}$$

$$- 2r\begin{bmatrix} A_3C_2 & A_3C_1^* \\ C_2^*B_1 & C_2^*A_1^* \\ C_1B_1 & C_1A_1^* \end{bmatrix}. \tag{28}$$

In Theorem 4, assuming that A_1, B_1, C_1, and C_2 vanish, we have the following.

Corollary 6. *Suppose that the matrix equation* $A_3XA_3^* = C_3$ *is consistent; then the extremal ranks of the quaternion matrix expression* $f(X)$ *defined as* (9) *subject to* $A_3XA_3^* = C_3$ *are the following:*

$$\max r[f(X)] = \min\left\{ r[C_4 \ \ A_4], r\begin{bmatrix} 0 & A_4^* & A_3^* \\ A_4 & C_4 & 0 \\ A_3 & 0 & -C_3 \end{bmatrix} - 2r(A_3) \right\},$$

$$\min r[f(X)] = 2r[C_4 \ \ A_4] + r\begin{bmatrix} 0 & A_4^* & A_3^* \\ A_4 & C_4 & 0 \\ A_3 & 0 & -C_3 \end{bmatrix} - 2r\begin{bmatrix} 0 & A_4^* \\ A_4 & C_4 \\ A_3 & 0 \end{bmatrix}. \tag{29}$$

3. A Practical Solvability Condition for Hermitian Solution to System (5)

In this section, we use Theorem 4 to give a necessary and sufficient condition for the existence of Hermitian solution to system (5) by rank equalities.

Theorem 7. *Let* A_1, $C_1 \in \mathbb{H}^{m\times n}$, $B_1, C_2 \in \mathbb{H}^{n\times s}$, $A_3 \in \mathbb{H}^{r\times n}$, $C_3 \in \mathbb{H}^{r\times r}$, $A_4 \in \mathbb{H}^{t\times n}$, *and* $C_4 \in \mathbb{H}^{t\times t}$ *be given; then the system* (5) *have Hermitian solution if and only if* $C_3 = C_3^*$, (11), (13) *hold, and the following equalities are all satisfied:*

$$r[A_4 \ \ C_4] = r(A_4), \tag{30}$$

$$r\begin{bmatrix} C_4 & A_4 \\ C_2^*A_4^* & B_1^* \\ C_1A_4^* & A_1 \end{bmatrix} = r\begin{bmatrix} A_4 \\ B_1^* \\ A_1 \end{bmatrix}, \tag{31}$$

$$r\begin{bmatrix} 0 & A_4^* & A_3^* & B_1 & A_1^* \\ A_4 & C_4 & 0 & 0 & 0 \\ A_3 & 0 & -C_3 & -A_3C_2 & -A_3C_1^* \\ B_1^* & 0 & -C_2^*A_3^* & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1A_3^* & -C_1B_1 & -C_1A_1^* \end{bmatrix} = 2r\begin{bmatrix} A_4 \\ A_3 \\ B_1^* \\ A_1 \end{bmatrix}. \tag{32}$$

Proof. It is obvious that the system (5) have Hermitian solution if and only if the system (10) have Hermitian solution and

$$\min r[f(X)] = 0, \tag{33}$$

where $f(X)$ is defined as (9) subject to system (10). Let X_0 be a Hermitian solution to the system (5); then X_0 is a Hermitian solution to system (10) and X_0 satisfies $A_4X_0A_4^* = C_4$. Hence, Lemma 1 yields $C_3 = C_3^*$, (11), (13), and (30). It follows from

$$\begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ A_3X_0 & 0 & I & 0 & 0 \\ B_1^*X_0 & 0 & 0 & I & 0 \\ A_1X_0 & 0 & 0 & 0 & I \end{bmatrix} \times \begin{bmatrix} 0 & A_4^* & A_3^* & B_1 & A_1^* \\ A_4 & C_4 & 0 & 0 & 0 \\ A_3 & 0 & -C_3 & -A_3C_2 & -A_3C_1^* \\ B_1^* & 0 & -C_2^*A_3^* & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1A_3^* & -C_1B_1 & -C_1A_1^* \end{bmatrix} \times \begin{bmatrix} I & -X_0A_4^* & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix} = \begin{bmatrix} 0 & A_4^* & A_3^* & B_1 & A_1^* \\ A_4 & 0 & 0 & 0 & 0 \\ A_3 & 0 & 0 & 0 & 0 \\ B_1^* & 0 & 0 & 0 & 0 \\ A_1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (34)$$

that (32) holds. Similarly, we can obtain (31).

Conversely, assume that $C_3 = C_3^*$, (11), (13) hold; then by Lemma 1, system (10) have Hermitian solution. By (20), (31)-(32), and

$$r\begin{bmatrix} 0 & A_4^* & B_1 & A_1^* \\ A_4 & C_4 & 0 & 0 \\ A_3 & 0 & -A_3C_2 & -A_3C_1^* \\ B_1^* & 0 & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1B_1 & -C_1A_1^* \end{bmatrix} \geq r\begin{bmatrix} A_4 \\ A_3 \\ B_1^* \\ A_1 \end{bmatrix} + r\begin{bmatrix} A_4 \\ B_1^* \\ A_1 \end{bmatrix} \quad (35)$$

we can get

$$\min r\left[f(X)\right] \leq 0. \quad (36)$$

However,

$$\min r\left[f(X)\right] \geq 0. \quad (37)$$

Hence (33) holds, implying that the system (5) have Hermitian solution. □

By Theorem 7, we can also get the following.

Corollary 8. *Suppose that A_3, C_3, A_4, and C_4 are those in Theorem 7; then the quaternion matrix equations $A_3XA_3^* = C_3$ and $A_4XA_4^* = C_4$ have common Hermitian solution if and only if* (30) *hold and the following equalities are satisfied:*

$$r\begin{bmatrix} A_3 & C_3 \end{bmatrix} = r(A_3), \qquad r\begin{bmatrix} 0 & A_4^* & A_3^* \\ A_4 & C_4 & 0 \\ A_3 & 0 & -C_3 \end{bmatrix} = 2r\begin{bmatrix} A_3 \\ A_4 \end{bmatrix}. \quad (38)$$

Corollary 9. *Suppose that $A_1, C_1 \in \mathbb{H}^{m\times n}$, $B_1, C_2 \in \mathbb{H}^{n\times s}$, and A, $B \in \mathbb{H}^{n\times n}$ are Hermitian. Then A and B have a common Hermitian g-inverse which is a solution to the system* (2) *if and only if* (11) *holds and the following equalities are all satisfied:*

$$r\begin{bmatrix} A_1 & C_1A \\ B_1^* & C_2^*A \\ A & A \end{bmatrix} = r\begin{bmatrix} A_1 \\ B_1^* \\ A \end{bmatrix}, \qquad r\begin{bmatrix} A_1 & C_1B \\ B_1^* & C_2^*B \\ B & B \end{bmatrix} = r\begin{bmatrix} A_1 \\ B_1^* \\ B \end{bmatrix}, \quad (39)$$

$$r\begin{bmatrix} 0 & B & A & B_1 & A_1^* \\ B & B & 0 & 0 & 0 \\ A & 0 & -A & -AC_2 & -AC_1^* \\ B_1^* & 0 & -C_2^*A & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1A & -C_1B_1 & -C_1A_1^* \end{bmatrix} = 2r\begin{bmatrix} B \\ A \\ B_1^* \\ A_1 \end{bmatrix}. \quad (40)$$

4. Extremal Ranks of Schur Complement Subject to (2)

As is well known, for a given block matrix

$$M = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix}, \quad (41)$$

where A and D are Hermitian quaternion matrices with appropriate sizes, then the Hermitian Schur complement of A in M is defined as

$$S_A = D - B^*A^{\sim}B, \quad (42)$$

where $A^{\sim}$ is a Hermitian g-inverse of A, that is, $A^{\sim} \in \{X \mid AXA = A, X = X^*\}$.

Now we use Theorem 4 to establish the extremal ranks of S_A given by (42) with respect to $A^{\sim}$ which is a solution to system (2).

Theorem 10. *Suppose $A_1, C_1 \in \mathbb{H}^{m\times n}$, $B_1, C_2 \in \mathbb{H}^{n\times s}$, $D \in \mathbb{H}^{t\times t}$, $B \in \mathbb{H}^{n\times t}$, and $A \in \mathbb{H}^{n\times n}$ are given and system* (2) *is consistent; then the extreme ranks of S_A given by* (42) *with respect to $A^{\sim}$ which is a solution of* (2) *are the following:*

$$\max_{\substack{A_1A^{\sim}=C_1 \\ A^{\sim}B_1=C_2}} r(S_A) = \min\{a, b\}, \quad (43)$$

where

$$a=r\begin{bmatrix} D & B^* \\ C_2^*B & B_1^* \\ C_1B & A_1 \end{bmatrix}-r\begin{bmatrix} B_1^* \\ A_1 \end{bmatrix},$$

$$b=r\begin{bmatrix} 0 & B & A & B_1 & A_1^* \\ B^* & D & 0 & 0 & 0 \\ A & 0 & -A & -AC_2 & -AC_1^* \\ B_1^* & 0 & -C_2^*A & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1A & -C_1B_1 & -C_1A_1^* \end{bmatrix}-2r\begin{bmatrix} A \\ B_1^* \\ A_1 \end{bmatrix},$$

$$\min_{\substack{A_1A^{\sim}=C_1\\ A^{\sim}B_1=C_2}} r(S_A)=2r\begin{bmatrix} D & B^* \\ C_2^*B & B_1^* \\ C_1B & A_1 \end{bmatrix}$$
$$+r\begin{bmatrix} 0 & B & A & B_1 & A_1^* \\ B^* & D & 0 & 0 & 0 \\ A & 0 & -A & -AC_2 & -AC_1^* \\ B_1^* & 0 & -C_2^*A & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1A & -C_1B_1 & -C_1A_1^* \end{bmatrix}$$
$$-2r\begin{bmatrix} 0 & B & B_1 & A_1^* \\ B^* & D & 0 & 0 \\ A & 0 & -AC_2 & -AC_1^* \\ B_1^* & 0 & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1B_1 & -C_1A_1^* \end{bmatrix}. \tag{44}$$

Proof. It is obvious that

$$\max_{A_1A^{\sim}=C_1, A^{\sim}B_1=C_2} r(D-B^*A^{\sim}B) = \max_{A_1X=C_1, XB_1=C_2, AXA=A} r(D-B^*XB),$$
$$\min_{A_1A^{\sim}=C_1, A^{\sim}B_1=C_2} r(D-B^*A^{\sim}B) = \min_{A_1X=C_1, XB_1=C_2, AXA=A} r(D-B^*XB). \tag{45}$$

Thus in Theorem 4 and its proof, letting $A_3=A_3^*=C_3=A$, $A_4=B^*$, and $C_4=D$, we can easily get the proof. □

In Theorem 10, let A_1, C_1, B_1, and C_2 vanish. Then we can easily get the following.

Corollary 11. *The extreme ranks of S_A given by (42) with respect to $A^{\sim}$ are the following:*

$$\max_{A^{\sim}} r(S_A)=\min\left\{r\begin{bmatrix} D & B^* \end{bmatrix}, r\begin{bmatrix} 0 & B & A \\ B^* & D & 0 \\ A & 0 & -A \end{bmatrix}-2r(A)\right\},$$
$$\min_{A^{\sim}} r(S_A)=2r\begin{bmatrix} D & B^* \end{bmatrix}+r\begin{bmatrix} 0 & B & A \\ B^* & D & 0 \\ A & 0 & -A \end{bmatrix}-2r\begin{bmatrix} 0 & B \\ B^* & D \\ A & 0 \end{bmatrix}. \tag{46}$$

5. The Rank Invariance of (9)

As another application of Theorem 4, we in this section consider the rank invariance of the matrix expression (9) with respect to the Hermitian solution of system (10).

Theorem 12. *Suppose that (10) have Hermitian solution; then the rank of $f(X)$ defined by (9) with respect to the Hermitian solution of (10) is invariant if and only if*

$$r\begin{bmatrix} 0 & A_4^* & B_1 & A_1^* \\ A_4 & C_4 & 0 & 0 \\ A_3 & 0 & -A_3C_2 & -A_3C_1^* \\ B_1^* & 0 & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1B_1 & -C_1A_1^* \end{bmatrix} = r\begin{bmatrix} C_4 & A_4 \\ C_2^*A_4^* & B_1^* \\ C_1A_4^* & A_1 \end{bmatrix}+r\begin{bmatrix} A_3 \\ B_1^* \\ A_1 \end{bmatrix},$$
$$r\begin{bmatrix} 0 & A_4^* & A_3^* & B_1 & A_1^* \\ A_4 & C_4 & 0 & 0 & 0 \\ A_3 & 0 & -C_3 & -A_3C_2 & -A_3C_1^* \\ B_1^* & 0 & -C_2^*A_3^* & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1A_3^* & -C_1B_1 & -C_1A_1^* \end{bmatrix}+r\begin{bmatrix} B_1^* \\ A_1 \end{bmatrix} = r\begin{bmatrix} 0 & A_4^* & B_1 & A_1^* \\ A_4 & C_4 & 0 & 0 \\ A_3 & 0 & -A_3C_2 & -A_3C_1^* \\ B_1^* & 0 & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1B_1 & -C_1A_1^* \end{bmatrix}+r\begin{bmatrix} A_3 \\ B_1^* \\ A_1 \end{bmatrix}, \tag{47}$$

or

$$r\begin{bmatrix} 0 & A_4^* & B_1 & A_1^* \\ A_4 & C_4 & 0 & 0 \\ A_3 & 0 & -A_3C_2 & -A_3C_1^* \\ B_1^* & 0 & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1B_1 & -C_1A_1^* \end{bmatrix} = r\begin{bmatrix} C_4 & A_4 \\ C_2^*A_4^* & B_1^* \\ C_1A_4^* & A_1 \end{bmatrix}+r\begin{bmatrix} A_3 \\ B_1^* \\ A_1 \end{bmatrix}. \tag{48}$$

Proof. It is obvious that the rank of $f(X)$ with respect to Hermitian solution of system (10) is invariant if and only if

$$\max r[f(X)]-\min r[f(X)]=0. \tag{49}$$

By (49), Theorem 4, and simplifications, we can get (47) and (48). □

Corollary 13. *The rank of* S_A *defined by* (42) *with respect to* $A^{\sim}$ *which is a solution to system* (2) *is invariant if and only if*

$$r\begin{bmatrix} 0 & B & B_1 & A_1^* \\ B^* & D & 0 & 0 \\ A & 0 & -AC_2 & -AC_1^* \\ B_1^* & 0 & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1B_1 & -C_1A_1^* \end{bmatrix} = r\begin{bmatrix} D & B^* \\ C_2^*B & B_1^* \\ C_1B & A_1 \end{bmatrix} + r\begin{bmatrix} A \\ B_1^* \\ A_1 \end{bmatrix},$$

$$r\begin{bmatrix} 0 & B & A & B_1 & A_1^* \\ B^* & D & 0 & 0 & 0 \\ A & 0 & -A & -AC_2 & -AC_1^* \\ B_1^* & 0 & -C_2^*A & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1A & -C_1B_1 & -C_1A_1^* \end{bmatrix} + r\begin{bmatrix} B_1^* \\ A_1 \end{bmatrix}$$
$$= r\begin{bmatrix} 0 & B & B_1 & A_1^* \\ B^* & D & 0 & 0 \\ A & 0 & -AC_2 & -AC_1^* \\ B_1^* & 0 & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1B_1 & -C_1A_1^* \end{bmatrix} + r\begin{bmatrix} A \\ B_1^* \\ A_1 \end{bmatrix}, \tag{50}$$

or

$$r\begin{bmatrix} 0 & B & B_1 & A_1^* \\ B^* & D & 0 & 0 \\ A & 0 & -AC_2 & -AC_1^* \\ B_1^* & 0 & -C_2^*B_1 & -C_2^*A_1^* \\ A_1 & 0 & -C_1B_1 & -C_1A_1^* \end{bmatrix} \tag{51}$$
$$= r\begin{bmatrix} D & B^* \\ C_2^*B & B_1^* \\ C_1B & A_1 \end{bmatrix} + r\begin{bmatrix} A \\ B_1^* \\ A_1 \end{bmatrix}.$$

Acknowledgments

This research was supported by the National Natural Science Foundation of China, Tian Yuan Foundation (11226067), the Fundamental Research Funds for the Central Universities (WM1214063), and China Postdoctoral Science Foundation (2012M511014).

References

[1] T. W. Hungerford, *Algebra*, Springer, New York, NY, USA, 1980.

[2] S. K. Mitra, "The matrix equations $AX = C$, $XB = D$," *Linear Algebra and its Applications*, vol. 59, pp. 171–181, 1984.

[3] C. G. Khatri and S. K. Mitra, "Hermitian and nonnegative definite solutions of linear matrix equations," *SIAM Journal on Applied Mathematics*, vol. 31, no. 4, pp. 579–585, 1976.

[4] Y. X. Yuan, "On the symmetric solutions of matrix equation $(AX, XC) = (B, D)$," *Journal of East China Shipbuilding Institute*, vol. 15, no. 4, pp. 82–85, 2001 (Chinese).

[5] H. Dai and P. Lancaster, "Linear matrix equations from an inverse problem of vibration theory," *Linear Algebra and its Applications*, vol. 246, pp. 31–47, 1996.

[6] J. Groß, "A note on the general Hermitian solution to $AXA^* = B$," *Malaysian Mathematical Society*, vol. 21, no. 2, pp. 57–62, 1998.

[7] Y. G. Tian and Y. H. Liu, "Extremal ranks of some symmetric matrix expressions with applications," *SIAM Journal on Matrix Analysis and Applications*, vol. 28, no. 3, pp. 890–905, 2006.

[8] Y. H. Liu, Y. G. Tian, and Y. Takane, "Ranks of Hermitian and skew-Hermitian solutions to the matrix equation $AXA^* = B$," *Linear Algebra and its Applications*, vol. 431, no. 12, pp. 2359–2372, 2009.

[9] X. W. Chang and J. S. Wang, "The symmetric solution of the matrix equations $AX + YA = C$, $AXA^T + BYB^T = C$ and $(A^TXA, B^TXB) = (C, D)$," *Linear Algebra and its Applications*, vol. 179, pp. 171–189, 1993.

[10] Q.-W. Wang and Z.-C. Wu, "Common Hermitian solutions to some operator equations on Hilbert C^*-modules," *Linear Algebra and its Applications*, vol. 432, no. 12, pp. 3159–3171, 2010.

[11] F. O. Farid, M. S. Moslehian, Q.-W. Wang, and Z.-C. Wu, "On the Hermitian solutions to a system of adjointable operator equations," *Linear Algebra and its Applications*, vol. 437, no. 7, pp. 1854–1891, 2012.

[12] Z.-H. He and Q.-W. Wang, "Solutions to optimization problems on ranks and inertias of a matrix function with applications," *Applied Mathematics and Computation*, vol. 219, no. 6, pp. 2989–3001, 2012.

[13] Q.-W. Wang, "The general solution to a system of real quaternion matrix equations," *Computers & Mathematics with Applications*, vol. 49, no. 5-6, pp. 665–675, 2005.

[14] Q.-W. Wang, "Bisymmetric and centrosymmetric solutions to systems of real quaternion matrix equations," *Computers & Mathematics with Applications*, vol. 49, no. 5-6, pp. 641–650, 2005.

[15] Q. W. Wang, "A system of four matrix equations over von Neumann regular rings and its applications," *Acta Mathematica Sinica*, vol. 21, no. 2, pp. 323–334, 2005.

[16] Q.-W. Wang, "A system of matrix equations and a linear matrix equation over arbitrary regular rings with identity," *Linear Algebra and its Applications*, vol. 384, pp. 43–54, 2004.

[17] Q.-W. Wang, H.-X. Chang, and Q. Ning, "The common solution to six quaternion matrix equations with applications," *Applied Mathematics and Computation*, vol. 198, no. 1, pp. 209–226, 2008.

[18] Q.-W. Wang, H.-X. Chang, and C.-Y. Lin, "*P*-(skew)symmetric common solutions to a pair of quaternion matrix equations," *Applied Mathematics and Computation*, vol. 195, no. 2, pp. 721–732, 2008.

[19] Q. W. Wang and Z. H. He, "Some matrix equations with applications," *Linear and Multilinear Algebra*, vol. 60, no. 11-12, pp. 1327–1353, 2012.

[20] Q. W. Wang and J. Jiang, "Extreme ranks of (skew-)Hermitian solutions to a quaternion matrix equation," *Electronic Journal of Linear Algebra*, vol. 20, pp. 552–573, 2010.

[21] Q.-W. Wang and C.-K. Li, "Ranks and the least-norm of the general solution to a system of quaternion matrix equations," *Linear Algebra and its Applications*, vol. 430, no. 5-6, pp. 1626–1640, 2009.

[22] Q.-W. Wang, X. Liu, and S.-W. Yu, "The common bisymmetric nonnegative definite solutions with extreme ranks and inertias to a pair of matrix equations," *Applied Mathematics and Computation*, vol. 218, no. 6, pp. 2761–2771, 2011.

[23] Q.-W. Wang, F. Qin, and C.-Y. Lin, "The common solution to matrix equations over a regular ring with applications," *Indian Journal of Pure and Applied Mathematics*, vol. 36, no. 12, pp. 655–672, 2005.

[24] Q.-W. Wang, G.-J. Song, and C.-Y. Lin, "Extreme ranks of the solution to a consistent system of linear quaternion matrix equations with an application," *Applied Mathematics and Computation*, vol. 189, no. 2, pp. 1517–1532, 2007.

[25] Q.-W. Wang, G.-J. Song, and C.-Y. Lin, "Rank equalities related to the generalized inverse $A^{(2)}_{T,S}$ with applications," *Applied Mathematics and Computation*, vol. 205, no. 1, pp. 370–382, 2008.

[26] Q. W. Wang, G. J. Song, and X. Liu, "Maximal and minimal ranks of the common solution of some linear matrix equations over an arbitrary division ring with applications," *Algebra Colloquium*, vol. 16, no. 2, pp. 293–308, 2009.

[27] Q.-W. Wang, Z.-C. Wu, and C.-Y. Lin, "Extremal ranks of a quaternion matrix expression subject to consistent systems of quaternion matrix equations with applications," *Applied Mathematics and Computation*, vol. 182, no. 2, pp. 1755–1764, 2006.

[28] Q. W. Wang and S. W. Yu, "Ranks of the common solution to some quaternion matrix equations with applications," *Bulletin of Iranian Mathematical Society*, vol. 38, no. 1, pp. 131–157, 2012.

[29] Q. W. Wang, S. W. Yu, and W. Xie, "Extreme ranks of real matrices in solution of the quaternion matrix equation $AXB = C$ with applications," *Algebra Colloquium*, vol. 17, no. 2, pp. 345–360, 2010.

[30] Q.-W. Wang, S.-W. Yu, and Q. Zhang, "The real solutions to a system of quaternion matrix equations with applications," *Communications in Algebra*, vol. 37, no. 6, pp. 2060–2079, 2009.

[31] Q.-W. Wang, S.-W. Yu, and C.-Y. Lin, "Extreme ranks of a linear quaternion matrix expression subject to triple quaternion matrix equations with applications," *Applied Mathematics and Computation*, vol. 195, no. 2, pp. 733–744, 2008.

[32] Q.-W. Wang and F. Zhang, "The reflexive re-nonnegative definite solution to a quaternion matrix equation," *Electronic Journal of Linear Algebra*, vol. 17, pp. 88–101, 2008.

[33] Q. W. Wang, X. Zhang, and Z. H. He, "On the Hermitian structures of the solution to a pair of matrix equations," *Linear and Multilinear Algebra*, vol. 61, no. 1, pp. 73–90, 2013.

[34] X. Zhang, Q.-W. Wang, and X. Liu, "Inertias and ranks of some Hermitian matrix functions with applications," *Central European Journal of Mathematics*, vol. 10, no. 1, pp. 329–351, 2012.

[35] Y. G. Tian and S. Z. Cheng, "The maximal and minimal ranks of $A - BXC$ with applications," *New York Journal of Mathematics*, vol. 9, pp. 345–362, 2003.

[36] Y. G. Tian, "Upper and lower bounds for ranks of matrix expressions using generalized inverses," *Linear Algebra and its Applications*, vol. 355, pp. 187–214, 2002.

[37] Y. H. Liu and Y. G. Tian, "More on extremal ranks of the matrix expressions $A - BX \pm X^* B^*$ with statistical applications," *Numerical Linear Algebra with Applications*, vol. 15, no. 4, pp. 307–325, 2008.

[38] Y. H. Liu and Y. G. Tian, "Max-min problems on the ranks and inertias of the matrix expressions $A - BXC \pm (BXC)^*$ with applications," *Journal of Optimization Theory and Applications*, vol. 148, no. 3, pp. 593–622, 2011.

[39] T. Ando, "Generalized Schur complements," *Linear Algebra and its Applications*, vol. 27, pp. 173–186, 1979.

[40] D. Carlson, E. Haynsworth, and T. Markham, "A generalization of the Schur complement by means of the Moore-Penrose inverse," *SIAM Journal on Applied Mathematics*, vol. 26, pp. 169–175, 1974.

[41] M. Fiedler, "Remarks on the Schur complement," *Linear Algebra and its Applications*, vol. 39, pp. 189–195, 1981.

[42] Y. G. Tian, "More on maximal and minimal ranks of Schur complements with applications," *Applied Mathematics and Computation*, vol. 152, no. 3, pp. 675–692, 2004.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 217540, 14 pages
http://dx.doi.org/10.1155/2013/217540

Research Article

An Efficient Algorithm for the Reflexive Solution of the Quaternion Matrix Equation $AXB + CX^H D = F$

Ning Li,[1,2] Qing-Wen Wang,[1] and Jing Jiang[3]

[1] *Department of Mathematics, Shanghai University, Shanghai 200444, China*
[2] *School of Mathematics and Quantitative Economics, Shandong University of Finance and Economics, Jinan 250002, China*
[3] *Department of Mathematics, Qilu Normal University, Jinan 250013, China*

Correspondence should be addressed to Qing-Wen Wang; wqw858@yahoo.com.cn

Received 3 October 2012; Accepted 4 December 2012

Academic Editor: P. N. Shivakumar

We propose an iterative algorithm for solving the reflexive solution of the quaternion matrix equation $AXB + CX^H D = F$. When the matrix equation is consistent over reflexive matrix X, a reflexive solution can be obtained within finite iteration steps in the absence of roundoff errors. By the proposed iterative algorithm, the least Frobenius norm reflexive solution of the matrix equation can be derived when an appropriate initial iterative matrix is chosen. Furthermore, the optimal approximate reflexive solution to a given reflexive matrix X_0 can be derived by finding the least Frobenius norm reflexive solution of a new corresponding quaternion matrix equation. Finally, two numerical examples are given to illustrate the efficiency of the proposed methods.

1. Introduction

Throughout the paper, the notations $\mathbb{R}^{m\times n}$ and $\mathbb{H}^{m\times n}$ represent the set of all $m \times n$ real matrices and the set of all $m \times n$ matrices over the quaternion algebra $\mathbb{H} = \{a_1 + a_2 i + a_3 j + a_4 k \mid i^2 = j^2 = k^2 = ijk = -1,\ a_1, a_2, a_3, a_4 \in \mathbb{R}\}$. We denote the identity matrix with the appropriate size by I. We denote the conjugate transpose, the transpose, the conjugate, the trace, the column space, the real part, the $mn \times 1$ vector formed by the vertical concatenation of the respective columns of a matrix A by $A^H, A^T, \overline{A}, \mathrm{tr}(A), R(A), \mathrm{Re}(A)$, and $\mathrm{vec}(A)$, respectively. The Frobenius norm of A is denoted by $\|A\|$, that is, $\|A\| = \sqrt{\mathrm{tr}(A^H A)}$. Moreover, $A \otimes B$ and $A \odot B$ stand for the Kronecker matrix product and Hadmard matrix product of the matrices A and B.

Let $Q \in \mathbb{H}^{n\times n}$ be a generalized reflection matrix, that is, $Q^2 = I$ and $Q^H = Q$. A matrix A is called reflexive with respect to the generalized reflection matrix Q, if $A = QAQ$. It is obvious that any matrix is reflexive with respect to I. Let $\mathbb{RH}^{n\times n}(Q)$ denote the set of order n reflexive matrices with respect to Q. The reflexive matrices with respect to a generalized reflection matrix Q have been widely used in engineering and scientific computations [1, 2].

In the field of matrix algebra, quaternion matrix equations have received much attention. Wang et al. [3] gave necessary and sufficient conditions for the existence and the representations of P-symmetric and P-skew-symmetric solutions of quaternion matrix equations $A_a X = C_a$ and $A_b X B_b = C_b$. Yuan and Wang [4] derived the expressions of the least squares η-Hermitian solution with the least norm and the expressions of the least squares anti-η-Hermitian solution with the least norm for the quaternion matrix equation $AXB + CXD = E$. Jiang and Wei [5] derived the explicit solution of the quaternion matrix equation $X - A\widetilde{X}B = C$. Li and Wu [6] gave the expressions of symmetric and skew-antisymmetric solutions of the quaternion matrix equations $A_1 X = C_1$ and $XB_3 = C_3$. Feng and Cheng [7] gave a clear description of the solution set of the quaternion matrix equation $AX - \overline{X}B = 0$.

The iterative method is a very important method to solve matrix equations. Peng [8] constructed a finite iteration method to solve the least squares symmetric solutions of linear matrix equation $AXB = C$. Also Peng [9–11] presented several efficient iteration methods to solve the constrained least squares solutions of linear matrix equations $AXB = C$ and $AXB + CYD = E$, by using Paige's algorithm [12] as the frame method. Duan et al. [13–17] proposed iterative algorithms for the (Hermitian) positive definite solutions of some nonlinear matrix equations. Ding et al. proposed the hierarchical gradient-based iterative algorithms [18] and hierarchical least squares iterative algorithms [19] for solving general (coupled) matrix equations, based on the hierarchical identification principle [20]. Wang et al. [21] proposed an iterative method for the least squares minimum-norm symmetric solution of $AXB = E$. Dehghan and Hajarian constructed finite iterative algorithms to solve several linear matrix equations over (anti)reflexive [22–24], generalized centrosymmetric [25, 26], and generalized bisymmetric [27, 28] matrices. Recently, Wu et al. [29–31] proposed iterative algorithms for solving various complex matrix equations.

However, to the best of our knowledge, there has been little information on iterative methods for finding a solution of a quaternion matrix equation. Due to the noncommutative multiplication of quaternions, the study of quaternion matrix equations is more complex than that of real and complex equations. Motivated by the work mentioned above and keeping the interests and wide applications of quaternion matrices in view (e.g., [32–45]), we, in this paper, consider an iterative algorithm for the following two problems.

Problem 1. For given matrices $A, C \in \mathbb{H}^{m\times n}$, $B, D \in \mathbb{H}^{n\times p}$, $F \in \mathbb{H}^{m\times p}$ and the generalized reflection matrix Q, find $X \in \mathbb{RH}^{n\times n}(Q)$, such that

$$AXB + CX^H D = F. \tag{1}$$

Problem 2. When Problem 1 is consistent, let its solution set be denoted by S_H. For a given reflexive matrix $X_0 \in \mathbb{RH}^{n\times n}(Q)$, find $\check{X} \in \mathbb{RH}^{n\times n}(Q)$, such that

$$\left\|\check{X} - X_0\right\| = \min_{X\in S_H} \left\|X - X_0\right\|. \tag{2}$$

The remainder of this paper is organized as follows. In Section 2, we give some preliminaries. In Section 3, we introduce an iterative algorithm for solving Problem 1. Then we prove that the given algorithm can be used to obtain a reflexive solution for any initial matrix within finite steps in the absence of roundoff errors. Also we prove that the least Frobenius norm reflexive solution can be obtained by choosing a special kind of initial matrix. In addition, the optimal reflexive solution of Problem 2 by finding the least Frobenius norm reflexive solution of a new matrix equation is given. In Section 4, we give two numerical examples to illustrate our results. In Section 5, we give some conclusions to end this paper.

2. Preliminary

In this section, we provide some results which will play important roles in this paper. First, we give a real inner product for the space $\mathbb{H}^{m\times n}$ over the real field $\mathbb{R}$.

Theorem 3. *In the space $\mathbb{H}^{m\times n}$ over the field $\mathbb{R}$, a real inner product can be defined as*

$$\langle A, B\rangle = \operatorname{Re}\left[\operatorname{tr}\left(B^H A\right)\right] \tag{3}$$

for $A, B \in \mathbb{H}^{m\times n}$. This real inner product space is denoted as $(\mathbb{H}^{m\times n}, \mathbb{R}, \langle\cdot,\cdot\rangle)$.

Proof. (1) For $A \in \mathbb{H}^{m\times n}$, let $A = A_1 + A_2 i + A_3 j + A_4 k$, then

$$\begin{aligned}\langle A, A\rangle &= \operatorname{Re}\left[\operatorname{tr}\left(A^H A\right)\right]\\ &= \operatorname{tr}\left(A_1^T A_1 + A_2^T A_2 + A_3^T A_3 + A_4^T A_4\right).\end{aligned} \tag{4}$$

It is obvious that $\langle A, A\rangle > 0$ and $\langle A, A\rangle = 0 \Leftrightarrow A = 0$.

(2) For $A, B \in \mathbb{H}^{m\times n}$, let $A = A_1 + A_2 i + A_3 j + A_4 k$ and $B = B_1 + B_2 i + B_3 j + B_4 k$, then we have

$$\begin{aligned}\langle A, B\rangle &= \operatorname{Re}\left[\operatorname{tr}\left(B^H A\right)\right]\\ &= \operatorname{tr}\left(B_1^T A_1 + B_2^T A_2 + B_3^T A_3 + B_4^T A_4\right)\\ &= \operatorname{tr}\left(A_1^T B_1 + A_2^T B_2 + A_3^T B_3 + A_4^T B_4\right)\\ &= \operatorname{Re}\left[\operatorname{tr}\left(A^H B\right)\right] = \langle B, A\rangle.\end{aligned} \tag{5}$$

(3) For $A, B, C \in \mathbb{H}^{m\times n}$

$$\begin{aligned}\langle A + B, C\rangle &= \operatorname{Re}\left\{\operatorname{tr}\left[C^H (A + B)\right]\right\}\\ &= \operatorname{Re}\left[\operatorname{tr}\left(C^H A + C^H B\right)\right]\\ &= \operatorname{Re}\left[\operatorname{tr}\left(C^H A\right)\right] + \operatorname{Re}\left[\operatorname{tr}\left(C^H B\right)\right]\\ &= \langle A, C\rangle + \langle B, C\rangle.\end{aligned} \tag{6}$$

(4) For $A, B \in \mathbb{H}^{m\times n}$ and $a \in \mathbb{R}$,

$$\begin{aligned}\langle aA, B\rangle &= \operatorname{Re}\left\{\operatorname{tr}\left[B^H (aA)\right]\right\} = \operatorname{Re}\left[\operatorname{tr}\left(aB^H A\right)\right]\\ &= a\operatorname{Re}\left[\operatorname{tr}\left(B^H A\right)\right] = a\langle A, B\rangle.\end{aligned} \tag{7}$$

All the above arguments reveal that the space $\mathbb{H}^{m\times n}$ over field $\mathbb{R}$ with the inner product defined in (3) is an inner product space. □

Let $\|\cdot\|_\Delta$ represent the matrix norm induced by the inner product $\langle\cdot,\cdot\rangle$. For an arbitrary quaternion matrix $A \in \mathbb{H}^{m\times n}$, it is obvious that the following equalities hold:

$$\|A\|_\Delta = \sqrt{\langle A, A\rangle} = \sqrt{\operatorname{Re}\left[\operatorname{tr}\left(A^H A\right)\right]} = \sqrt{\operatorname{tr}\left(A^H A\right)} = \|A\|, \tag{8}$$

which reveals that the induced matrix norm is exactly the Frobenius norm. For convenience, we still use $\|\cdot\|$ to denote the induced matrix norm.

Let E_{ij} denote the $m \times n$ quaternion matrix whose (i, j) entry is 1, and the other elements are zeros. In inner product space $(\mathbb{H}^{m\times n}, \mathbb{R}, \langle\cdot,\cdot\rangle)$, it is easy to verify that $E_{ij}, E_{ij}i, E_{ij}j, E_{ij}k$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, is an orthonormal basis, which reveals that the dimension of the inner product space $(\mathbb{H}^{m\times n}, \mathbb{R}, \langle\cdot,\cdot\rangle)$ is $4mn$.

Next, we introduce a real representation of a quaternion matrix.

For an arbitrary quaternion matrix $M = M_1 + M_2 i + M_3 j + M_4 k$, a map $\phi(\cdot)$, from $\mathbb{H}^{m\times n}$ to $\mathbb{R}^{4m\times 4n}$, can be defined as

$$\phi(M) = \begin{bmatrix} M_1 & -M_2 & -M_3 & -M_4 \\ M_2 & M_1 & -M_4 & M_3 \\ M_3 & M_4 & M_1 & -M_2 \\ M_4 & -M_3 & M_2 & M_1 \end{bmatrix}. \tag{9}$$

Lemma 4 (see [41]). *Let M and N be two arbitrary quaternion matrices with appropriate size. The map $\phi(\cdot)$ defined by* (9) *satisfies the following properties.*

(a) $M = N \Leftrightarrow \phi(M) = \phi(N)$.

(b) $\phi(M + N) = \phi(M) + \phi(N)$, $\phi(MN) = \phi(M)\phi(N)$, $\phi(kM) = k\phi(M)$, $k \in \mathbb{R}$.

(c) $\phi(M^H) = \phi^T(M)$.

(d) $\phi(M) = T_m^{-1}\phi(M)T_n = R_m^{-1}\phi(M)R_n = S_m^{-1}\phi(M)S_n$, *where*

$$T_t = \begin{bmatrix} 0 & -I_t & 0 & 0 \\ I_t & 0 & 0 & 0 \\ 0 & 0 & 0 & I_t \\ 0 & 0 & -I_t & 0 \end{bmatrix}, \qquad R_t = \begin{bmatrix} 0 & -I_{2t} \\ I_{2t} & 0 \end{bmatrix},$$
$$S_t = \begin{bmatrix} 0 & 0 & 0 & -I_t \\ 0 & 0 & I_t & 0 \\ 0 & -I_t & 0 & 0 \\ I_t & 0 & 0 & 0 \end{bmatrix}, \quad t = m, n. \tag{10}$$

By (9), it is easy to verify that

$$\|\phi(M)\| = 2\|M\|. \tag{11}$$

Finally, we introduce the commutation matrix.

A commutation matrix $P(m,n)$ is a $mn \times mn$ matrix which has the following explicit form:

$$P(m,n) = \sum_{i=1}^{m}\sum_{j=1}^{n} E_{ij} \otimes E_{ij}^T = \left[E_{ij}^T\right]_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}. \tag{12}$$

Moreover, $P(m,n)$ is a permutation matrix and $P(m,n) = P^T(n,m) = P^{-1}(n,m)$. We have the following lemmas on the commutation matrix.

Lemma 5 (see [46]). *Let A be a $m \times n$ matrix. There is a commutation matrix $P(m,n)$ such that*

$$\operatorname{vec}\left(A^T\right) = P(m,n)\operatorname{vec}(A). \tag{13}$$

Lemma 6 (see [46]). *Let A be a $m \times n$ matrix and B a $p \times q$ matrix. There exist two commutation matrices $P(m,p)$ and $P(n,q)$ such that*

$$B \otimes A = P^T(m,p)(A \otimes B)P(n,q). \tag{14}$$

3. Main Results

3.1. The Solution of Problem 1. In this subsection, we will construct an algorithm for solving Problem 1. Then some lemmas will be given to analyse the properties of the proposed algorithm. Using these lemmas, we prove that the proposed algorithm is convergent.

Algorithm 7 (Iterative algorithm for Problem 1).

(1) Choose an initial matrix $X(1) \in \mathbb{RH}^{n\times n}(Q)$.

(2) Calculate

$$\begin{aligned} R(1) &= F - AX(1)B - CX^H(1)D; \\ P(1) &= \frac{1}{2}\Big(A^H R(1)B^H + DR^H(1)C \\ &\quad + QA^H R(1)B^H Q + QDR^H(1)CQ\Big); \\ k &:= 1. \end{aligned} \tag{15}$$

(3) If $R(k) = 0$, then stop and $X(k)$ is the solution of Problem 1; else if $R(k) \neq 0$ and $P(k) = 0$, then stop and Problem 1 is not consistent; else $k := k + 1$.

(4) Calculate

$$\begin{aligned} X(k) &= X(k-1) + \frac{\|R(k-1)\|^2}{\|P(k-1)\|^2}P(k-1); \\ R(k) &= R(k-1) - \frac{\|R(k-1)\|^2}{\|P(k-1)\|^2} \\ &\quad \times \Big(AP(k-1)B + CP^H(k-1)D\Big); \\ P(k) &= \frac{1}{2}\Big(A^H R(k)B^H + DR^H(k)C \\ &\quad + QA^H R(k)B^H Q + QDR^H(k)CQ\Big) \\ &\quad + \frac{\|R(k)\|^2}{\|R(k-1)\|^2}P(k-1). \end{aligned} \tag{16}$$

(5) Go to Step (3).

Lemma 8. *Assume that the sequences $\{R(i)\}$ and $\{P(i)\}$ are generated by Algorithm 7, then $\langle R(i), R(j)\rangle = 0$ and $\langle P(i), P(j)\rangle = 0$ for $i, j = 1, 2, \ldots$, $i \neq j$.*

Proof. Since $\langle R(i), R(j)\rangle = \langle R(j), R(i)\rangle$ and $\langle P(i), P(j)\rangle = \langle P(j), P(i)\rangle$ for $i, j = 1, 2, \ldots$, we only need to prove that $\langle R(i), R(j)\rangle = 0$ and $\langle P(i), P(j)\rangle = 0$ for $1 \leq i < j$.

Now we prove this conclusion by induction.

Step 1. We show that

$$\begin{aligned} &\langle R(i), R(i+1)\rangle = 0, \\ &\langle P(i), P(i+1)\rangle = 0 \quad \text{for } i = 1, 2, \ldots. \end{aligned} \tag{17}$$

We also prove (17) by induction. When $i = 1$, we have

$$\begin{aligned}
\langle R(1), R(2)\rangle
&= \operatorname{Re}\left\{\operatorname{tr}\left[R^H(2)R(1)\right]\right\}\\
&= \operatorname{Re}\left\{\operatorname{tr}\left[\left(R(1) - \frac{\|R(1)\|^2}{\|P(1)\|^2}\left(AP(1)B + CP^H(1)D\right)\right)^H \times R(1)\right]\right\}\\
&= \|R(1)\|^2 - \frac{\|R(1)\|^2}{\|P(1)\|^2} \times \operatorname{Re}\left\{\operatorname{tr}\left[P^H(1)\left(A^H R(1)B^H + DR^H(1)C\right)\right]\right\}\\
&= \|R(1)\|^2 - \frac{\|R(1)\|^2}{\|P(1)\|^2} \times \operatorname{Re}\left\{\operatorname{tr}\left[P^H(1)\left(\left(A^H R(1)B^H + DR^H(1)C + QA^H R(1) \times B^H Q + QDR^H(1)CQ\right) \times (2)^{-1}\right)\right]\right\}\\
&= \|R(1)\|^2 - \frac{\|R(1)\|^2}{\|P(1)\|^2}\|P(1)\|^2\\
&= 0.
\end{aligned} \tag{18}$$

Also we can write

$$\begin{aligned}
\langle P(1), P(2)\rangle
&= \operatorname{Re}\left\{\operatorname{tr}\left[P^H(2)P(1)\right]\right\}\\
&= \operatorname{Re}\left\{\operatorname{tr}\left[\left(\frac{1}{2}\left(A^H R(2)B^H + DR^H(2)C + QA^H R(2)B^H Q + QDR^H(2)CQ\right) + \frac{\|R(2)\|^2}{\|R(1)\|^2}P(1)\right)^H P(1)\right]\right\}\\
&= \frac{\|R(2)\|^2}{\|R(1)\|^2}\|P(1)\|^2 + \operatorname{Re}\left\{\operatorname{tr}\left[P^H(1) \times \left(\left(A^H R(2)B^H + DR^H(2)C + QA^H R(2)B^H Q + QDR^H(2)CQ\right) \times (2)^{-1}\right)\right]\right\}\\
&= \frac{\|R(2)\|^2}{\|R(1)\|^2}\|P(1)\|^2 + \operatorname{Re}\left\{\operatorname{tr}\left[P^H(1)\left(A^H R(2)B^H + DR^H(2)C\right)\right]\right\}\\
&= \frac{\|R(2)\|^2}{\|R(1)\|^2}\|P(1)\|^2 + \operatorname{Re}\left\{\operatorname{tr}\left[R^H(2)\left(AP(1)B + CP^H(1)D\right)\right]\right\}\\
&= \frac{\|R(2)\|^2}{\|R(1)\|^2}\|P(1)\|^2 + \frac{\|P(1)\|^2}{\|R(1)\|^2}\operatorname{Re}\left\{\operatorname{tr}\left[R^H(2)(R(1) - R(2))\right]\right\}\\
&= \frac{\|R(2)\|^2}{\|R(1)\|^2}\|P(1)\|^2 - \frac{\|P(1)\|^2}{\|R(1)\|^2}\|R(2)\|^2\\
&= 0.
\end{aligned} \tag{19}$$

Now assume that conclusion (17) holds for $1 \le i \le s-1$, then

$$\begin{aligned}
\langle R(s), R(s+1)\rangle
&= \operatorname{Re}\left\{\operatorname{tr}\left[R^H(s+1)R(s)\right]\right\}\\
&= \operatorname{Re}\left\{\operatorname{tr}\left[\left(R(s) - \frac{\|R(s)\|^2}{\|P(s)\|^2}\left(AP(s)B + CP^H(s)D\right)\right)^H \times R(s)\right]\right\}\\
&= \|R(s)\|^2 - \frac{\|R(s)\|^2}{\|P(s)\|^2} \times \operatorname{Re}\left\{\operatorname{tr}\left[P^H(s)\left(A^H R(s)B^H + DR^H(s)C\right)\right]\right\}\\
&= \|R(s)\|^2 - \frac{\|R(s)\|^2}{\|P(s)\|^2} \times \operatorname{Re}\left\{\operatorname{tr}\left[P^H(s) \times \left(\left(A^H R(s)B^H + DR^H(s)C + QA^H R(s)B^H Q + QDR^H(s)CQ\right) \times (2)^{-1}\right)\right]\right\}\\
&= \|R(s)\|^2 - \frac{\|R(s)\|^2}{\|P(s)\|^2} \times \operatorname{Re}\left\{\operatorname{tr}\left[P^H(s)\left(P(s) - \frac{\|R(s)\|^2}{\|R(s-1)\|^2}P(s-1)\right)\right]\right\}\\
&= 0.
\end{aligned} \tag{20}$$

And it can also be obtained that

$$\begin{aligned}
&\langle P(s), P(s+1)\rangle \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[P^H(s+1)P(s)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[\left(\left(\left(A^H R(s+1)B^H + DR^H(s+1)C\right.\right.\right.\right.\right. \\
&\qquad\qquad + QA^H R(s+1)B^H Q \\
&\qquad\qquad \left.+ QDR^H(s+1)CQ\right)\times(2)^{-1}\Big) \\
&\qquad\qquad \left.\left.\left.+\frac{\|R(s+1)\|^2}{\|R(s)\|^2}P(s)\right)^H P(s)\right]\right\} \\
&\quad = \frac{\|R(s+1)\|^2}{\|R(s)\|^2}\|P(s)\|^2 \\
&\qquad + \operatorname{Re}\left\{\operatorname{tr}\left[P^H(s)\left(A^H R(s+1)B^H + DR^H(s+1)C\right)\right]\right\} \\
&\quad = \frac{\|R(s+1)\|^2}{\|R(s)\|^2}\|P(s)\|^2 \\
&\qquad + \operatorname{Re}\left\{\operatorname{tr}\left[R^H(s+1)\left(AP(s)B + CP^H(s)D\right)\right]\right\} \\
&\quad = \frac{\|R(s+1)\|^2}{\|R(s)\|^2}\|P(s)\|^2 \\
&\qquad + \frac{\|P(s)\|^2}{\|R(s)\|^2}\operatorname{Re}\left\{\operatorname{tr}\left[R^H(s+1)(R(s)-R(s+1))\right]\right\} \\
&\quad = 0.
\end{aligned} \tag{21}$$

Therefore, the conclusion (17) holds for $i = 1, 2, \ldots$.

Step 2. Assume that $\langle R(i), R(i+r)\rangle = 0$ and $\langle P(i), P(i+r)\rangle = 0$ for $i \geq 1$ and $r \geq 1$. We will show that

$$\langle R(i), R(i+r+1)\rangle = 0, \qquad \langle P(i), P(i+r+1)\rangle = 0. \tag{22}$$

We prove the conclusion (22) in two substeps.

Substep 2.1. In this substep, we show that

$$\langle R(1), R(r+2)\rangle = 0, \qquad \langle P(1), P(r+2)\rangle = 0. \tag{23}$$

It follows from Algorithm 7 that

$$\begin{aligned}
&\langle R(1), R(r+2)\rangle \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[R^H(r+2)R(1)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[\left(R(r+1) - \frac{\|R(r+1)\|^2}{\|P(r+1)\|^2}\right.\right.\right. \\
&\qquad\qquad \left.\left.\left.\times\left(AP(r+1)B + CP^H(r+1)D\right)\right)^H R(1)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[R^H(r+1)R(1)\right]\right\} - \frac{\|R(r+1)\|^2}{\|P(r+1)\|^2} \\
&\qquad \times \operatorname{Re}\left\{\operatorname{tr}\left[P^H(r+1)\left(A^H R(1)B^H + DR^H(1)C\right)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[R^H(r+1)R(1)\right]\right\} - \frac{\|R(r+1)\|^2}{\|P(r+1)\|^2} \\
&\qquad \times \operatorname{Re}\left\{\operatorname{tr}\left[P^H(r+1)\left(\left(A^H R(1)B^H + DR^H(1)C\right.\right.\right.\right. \\
&\qquad\qquad + QA^H R(1)B^H Q \\
&\qquad\qquad \left.\left.\left.\left.+ QDR^H(1)CQ\right)\times(2)^{-1}\right)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[R^H(r+1)R(1)\right]\right\} \\
&\qquad - \frac{\|R(r+1)\|^2}{\|P(r+1)\|^2}\operatorname{Re}\left\{\operatorname{tr}\left[P^H(r+1)P(1)\right]\right\} \\
&\quad = 0.
\end{aligned} \tag{24}$$

Also we can write

$$\begin{aligned}
&\langle P(1), P(r+2)\rangle \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[P^H(r+2)P(1)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[\left(\left(\left(A^H R(r+2)B^H + DR^H(r+2)C\right.\right.\right.\right.\right. \\
&\qquad\qquad \left.+ QA^H R(r+2)B^H Q + QDR^H(r+2)CQ\right) \\
&\qquad\qquad \left.\left.\left.\times(2)^{-1}\Big) + \frac{\|R(r+2)\|^2}{\|R(r+1)\|^2}P(r+1)\right)^H P(1)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[\left(A^H R(r+2)B^H + DR^H(r+2)C\right)^H P(1)\right]\right\} \\
&\qquad + \frac{\|R(r+2)\|^2}{\|R(r+1)\|^2}\operatorname{Re}\left\{\operatorname{tr}\left[P^H(r+1)P(1)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[R^H(r+2)\left(AP(1)B + CP^H(1)D\right)\right]\right\} \\
&\qquad + \frac{\|R(r+2)\|^2}{\|R(r+1)\|^2}\operatorname{Re}\left\{\operatorname{tr}\left[P^H(r+1)P(1)\right]\right\}
\end{aligned}$$

$$= \frac{\|P(1)\|^2}{\|R(1)\|^2} \operatorname{Re}\left\{\operatorname{tr}\left[R^H(r+2)(R(1)-R(2))\right]\right\}$$
$$+ \frac{\|R(r+2)\|^2}{\|R(r+1)\|^2} \operatorname{Re}\left\{\operatorname{tr}\left[P^H(r+1)P(1)\right]\right\}$$
$$= 0. \tag{25}$$

Substep 2.2. In this substep, we prove the conclusion (22) in Step 2. It follows from Algorithm 7 that

$$\begin{aligned}
&\langle R(i), R(i+r+1)\rangle \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[R^H(i+r+1)R(i)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[\left(R(i+r) - \frac{\|R(i+r)\|^2}{\|P(i+r)\|^2} \times \left(AP(i+r)B + CP^H(i+r)D\right)\right)^H R(i)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[R^H(i+r)R(i)\right]\right\} - \frac{\|R(i+r)\|^2}{\|P(i+r)\|^2} \times \operatorname{Re}\left\{\operatorname{tr}\left[P^H(i+r)\left(A^H R(i)B^H + DR^H(i)C\right)\right]\right\} \\
&\quad = -\frac{\|R(i+r)\|^2}{\|P(i+r)\|^2} \times \operatorname{Re}\left\{\operatorname{tr}\left[P^H(i+r)\left(\left(A^H R(i)B^H + DR^H(i)C + QA^H R(i)B^H Q + QDR^H(i)CQ\right)\times(2)^{-1}\right)\right]\right\} \\
&\quad = -\frac{\|R(i+r)\|^2}{\|P(i+r)\|^2} \times \operatorname{Re}\left\{\operatorname{tr}\left[P^H(i+r)\left(P(i) - \frac{\|R(i)\|^2}{\|R(i-1)\|^2}P(i-1)\right)\right]\right\} \\
&\quad = \frac{\|R(i+r)\|^2\|R(i)\|^2}{\|P(i+r)\|^2\|R(i-1)\|^2} \operatorname{Re}\left\{\operatorname{tr}\left[P^H(i+r)P(i-1)\right]\right\}, \\
&\langle P(i), P(i+r+1)\rangle \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[P^H(i+r+1)P(i)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[\left(\frac{1}{2}\left(A^H R(i+r+1)B^H + DR^H(i+r+1)C + QA^H R(i+r+1)B^H Q + QDR^H(i+r+1)CQ\right) + \frac{\|R(i+r+1)\|^2}{\|R(i+r)\|^2}P(i+r)\right)^H P(i)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[\left(A^H R(i+r+1)B^H + DR^H(i+r+1)C\right)^H P(i)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[R^H(i+r+1)\left(AP(i)B + CP^H(i)D\right)\right]\right\} \\
&\quad = \frac{\|P(i)\|^2}{\|R(i)\|^2} \operatorname{Re}\left\{\operatorname{tr}\left[R^H(i+r+1)(R(i)-R(i+1))\right]\right\} \\
&\quad = \frac{\|P(i)\|^2}{\|R(i)\|^2} \operatorname{Re}\left\{\operatorname{tr}\left[R^H(i+r+1)R(i)\right]\right\} - \frac{\|P(i)\|^2}{\|R(i)\|^2} \operatorname{Re}\left\{\operatorname{tr}\left[R^H(i+r+1)R(i+1)\right]\right\} \\
&\quad = \frac{\|P(i)\|^2\|R(i+r)\|^2\|R(i)\|^2}{\|R(i)\|^2\|P(i+r)\|^2\|R(i-1)\|^2} \times \operatorname{Re}\left\{\operatorname{tr}\left[P^H(i+r)P(i-1)\right]\right\}.
\end{aligned} \tag{26}$$

Repeating the above process (26), we can obtain

$$\begin{aligned}
\langle R(i), R(i+r+1)\rangle &= \cdots = \alpha \operatorname{Re}\left\{\operatorname{tr}\left[P^H(r+2)P(1)\right]\right\}; \\
\langle P(i), P(i+r+1)\rangle &= \cdots = \beta \operatorname{Re}\left\{\operatorname{tr}\left[P^H(r+2)P(1)\right]\right\}.
\end{aligned} \tag{27}$$

Combining these two relations with (24) and (25), it implies that (22) holds. So, by the principle of induction, we know that Lemma 8 holds. □

Lemma 9. *Assume that Problem 1 is consistent, and let $\widetilde{X} \in \mathbb{RH}^{n\times n}(Q)$ be its solution. Then, for any initial matrix $X(1) \in \mathbb{RH}^{n\times n}(Q)$, the sequences $\{R(i)\}$, $\{P(i)\}$, and $\{X(i)\}$ generated by Algorithm 7 satisfy*

$$\left\langle P(i), \widetilde{X} - X(i)\right\rangle = \|R(i)\|^2, \quad i = 1, 2, \ldots. \tag{28}$$

Proof. We also prove this conclusion by induction.

When $i = 1$, it follows from Algorithm 7 that

$$\begin{aligned}
&\left\langle P(1), \widetilde{X} - X(1)\right\rangle \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[\left(\widetilde{X} - X(1)\right)^H P(1)\right]\right\} \\
&\quad = \operatorname{Re}\left\{\operatorname{tr}\left[\left(\widetilde{X} - X(1)\right)^H \times \left(\left(A^H R(1)B^H + DR^H(1)C + QA^H R(1)B^H Q + QDR^H(1)CQ\right)\times(2)^{-1}\right)\right]\right\}
\end{aligned}$$

$$\begin{aligned}
&= \mathrm{Re}\left\{\mathrm{tr}\left[\left(\widetilde{X}-X(1)\right)^H\left(A^HR(1)B^H+DR^H(1)C\right)\right]\right\}\\
&= \mathrm{Re}\left\{\mathrm{tr}\left[R^H(1)\left(A\left(\widetilde{X}-X(1)\right)B+C\left(\widetilde{X}-X(1)\right)^HD\right)\right]\right\}\\
&= \mathrm{Re}\left\{\mathrm{tr}\left[R^H(1)R(1)\right]\right\}\\
&= \|R(1)\|^2.
\end{aligned} \tag{29}$$

This implies that (28) holds for $i = 1$.

Now it is assumed that (28) holds for $i = s$, that is

$$\left\langle P(s), \widetilde{X}-X(s)\right\rangle = \|R(s)\|^2. \tag{30}$$

Then, when $i = s+1$

$$\begin{aligned}
&\left\langle P(s+1), \widetilde{X}-X(s+1)\right\rangle\\
&\quad= \mathrm{Re}\left\{\mathrm{tr}\left[\left(\widetilde{X}-X(s+1)\right)^HP(s+1)\right]\right\}\\
&\quad= \mathrm{Re}\left\{\mathrm{tr}\left[\left(\widetilde{X}-X(s+1)\right)^H\right.\right.\\
&\qquad\times\left(\frac{1}{2}\left(A^HR(s+1)B^H+DR^H(s+1)C\right.\right.\\
&\qquad\qquad+QA^HR(s+1)B^HQ\\
&\qquad\qquad\left.+QDR^H(s+1)CQ\right)\\
&\qquad\left.\left.\left.+\frac{\|R(s+1)\|^2}{\|R(s)\|^2}P(s)\right)\right]\right\}\\
&\quad= \mathrm{Re}\left\{\mathrm{tr}\left[\left(\widetilde{X}-X(s+1)\right)^H\right.\right.\\
&\qquad\left.\left.\times\left(A^HR(s+1)B^H+DR^H(s+1)C\right)\right]\right\}\\
&\qquad+\frac{\|R(s+1)\|^2}{\|R(s)\|^2}\mathrm{Re}\left\{\mathrm{tr}\left[\left(\widetilde{X}-X(s+1)\right)^HP(s)\right]\right\}\\
&\quad= \mathrm{Re}\left\{\mathrm{tr}\left[R^H(s+1)\right.\right.\\
&\qquad\times\left(A\left(\widetilde{X}-X(s+1)\right)B\right.\\
&\qquad\left.\left.\left.+C\left(\widetilde{X}-X(s+1)\right)^HD\right)\right]\right\}\\
&\qquad+\frac{\|R(s+1)\|^2}{\|R(s)\|^2}\left\{\mathrm{Re}\left\{\mathrm{tr}\left[\left(\widetilde{X}-X(s)\right)^HP(s)\right]\right\}\right.\\
&\qquad\qquad\left.-\mathrm{Re}\left\{\mathrm{tr}\left[\left(X(s+1)-X(s)\right)^HP(s)\right]\right\}\right\}\\
&\quad= \|R(s+1)\|^2+\frac{\|R(s+1)\|^2}{\|R(s)\|^2}\\
&\qquad\times\left\{\|R(s)\|^2-\frac{\|R(s)\|^2}{\|P(s)\|^2}\mathrm{Re}\left\{\mathrm{tr}\left[P^H(s)P(s)\right]\right\}\right\}\\
&\quad= \|R(s+1)\|^2.
\end{aligned} \tag{31}$$

Therefore, Lemma 9 holds by the principle of induction. □

From the above two lemmas, we have the following conclusions.

Remark 10. If there exists a positive number i such that $R(i) \neq 0$ and $P(i) = 0$, then we can get from Lemma 9 that Problem 1 is not consistent. Hence, the solvability of Problem 1 can be determined by Algorithm 7 automatically in the absence of roundoff errors.

Theorem 11. *Suppose that Problem 1 is consistent. Then for any initial matrix $X(1) \in \mathbb{RH}^{n\times n}(Q)$, a solution of Problem 1 can be obtained within finite iteration steps in the absence of roundoff errors.*

Proof. In Section 2, it is known that the inner product space $(\mathbb{H}^{m\times p}, R, \langle\cdot,\cdot\rangle)$ is $4mp$-dimensional. According to Lemma 9, if $R(i) \neq 0$, $i = 1, 2, \ldots, 4mp$, then we have $P(i) \neq 0$, $i = 1, 2, \ldots, 4mp$. Hence $R(4mp+1)$ and $P(4mp+1)$ can be computed. From Lemma 8, it is not difficult to get

$$\langle R(i), R(j)\rangle = 0 \quad \text{for } i, j = 1, 2, \ldots, 4mp,\ i \neq j. \tag{32}$$

Then $R(1), R(2), \ldots, R(4mp)$ is an orthogonal basis of the inner product space $(\mathbb{H}^{m\times p}, R, \langle\cdot,\cdot\rangle)$. In addition, we can get from Lemma 8 that

$$\langle R(i), R(4mp+1)\rangle = 0 \quad \text{for } i = 1, 2, \ldots, 4mp. \tag{33}$$

It follows that $R(4mp+1) = 0$, which implies that $X(4mp+1)$ is a solution of Problem 1. □

3.2. The Solution of Problem 2. In this subsection, firstly we introduce some lemmas. Then, we will prove that the least Frobenius norm reflexive solution of (1) can be derived by choosing a suitable initial iterative matrix. Finally, we solve Problem 2 by finding the least Frobenius norm reflexive solution of a new-constructed quaternion matrix equation.

Lemma 12 (see [47]). *Assume that the consistent system of linear equations $My = b$ has a solution $y_0 \in R(M^T)$ then y_0 is the unique least Frobenius norm solution of the system of linear equations.*

Lemma 13. *Problem 1 is consistent if and only if the system of quaternion matrix equations*

$$\begin{aligned}
AXB + CX^HD &= F,\\
AQXQB + CQX^HQD &= F
\end{aligned} \tag{34}$$

is consistent. Furthermore, if the solution sets of Problem 1 and (34) *are denoted by* S_H *and* S_H^1*, respectively, then, we have* $S_H \subseteq S_H^1$.

Proof. First, we assume that Problem 1 has a solution X. By $AXB + CX^HD = F$ and $QXQ = X$, we can obtain $AXB + CX^HD = F$ and $AQXQB + CQX^HQD = F$, which implies that X is a solution of quaternion matrix equations (34), and $S_H \subseteq S_H^1$.

Conversely, suppose (34) is consistent. Let X be a solution of (34). Set $X_a = (X + QXQ)/2$. It is obvious that $X_a \in \mathbb{RH}^{n\times n}(Q)$. Now we can write

$$\begin{aligned} AX_aB + CX_a^HD &= \frac{1}{2}\left[A(X+QXQ)B + C(X+QXQ)^HD\right] \\ &= \frac{1}{2}\left[AXB + CX^HD + AQXQB + CQX^HQD\right] \\ &= \frac{1}{2}[F+F] \\ &= F. \end{aligned} \tag{35}$$

Hence X_a is a solution of Problem 1. The proof is completed. □

Lemma 14. *The system of quaternion matrix equations* (34) *is consistent if and only if the system of real matrix equations*

$$\begin{aligned} &\phi(A)\left[X_{ij}\right]_{4\times 4}\phi(B) + \phi(C)\left[X_{ij}\right]_{4\times 4}^T\phi(D) = \phi(F), \\ &\phi(A)\phi(Q)\left[X_{ij}\right]_{4\times 4}\phi(Q)\phi(B) \\ &\quad + \phi(C)\phi(Q)\left[X_{ij}\right]_{4\times 4}^T\phi(Q)\phi(D) = \phi(F) \end{aligned} \tag{36}$$

is consistent, where $X_{ij} \in \mathbb{H}^{n\times n}$, $i, j = 1, 2, 3, 4$, *are submatrices of the unknown matrix. Furthermore, if the solution sets of* (34) *and* (36) *are denoted by* S_H^1 *and* S_R^2*, respectively, then, we have* $\phi(S_H^1) \subseteq S_R^2$.

Proof. Suppose that (34) has a solution

$$X = X_1 + X_2 i + X_3 j + X_4 k. \tag{37}$$

Applying (a), (b), and (c) in Lemma 4 to (34) yields

$$\begin{aligned} &\phi(A)\phi(X)\phi(B) + \phi(C)\phi^T(X)\phi(D) = \phi(F), \\ &\phi(A)\phi(Q)\phi(X)\phi(Q)\phi(B) \\ &\quad + \phi(C)\phi(Q)\phi^T(X)\phi(Q)\phi(D) = \phi(F), \end{aligned} \tag{38}$$

which implies that $\phi(X)$ is a solution of (36) and $\phi(S_H^1) \subseteq S_R^2$. Conversely, suppose that (36) has a solution

$$\check{X} = \left[X_{ij}\right]_{4\times 4}. \tag{39}$$

By (d) in Lemma 4, we have that

$$\begin{aligned} &T_m^{-1}\phi(A)T_n\check{X}T_n^{-1}\phi(B)T_p \\ &\quad + T_m^{-1}\phi(C)T_n\check{X}^TT_n^{-1}\phi(D)T_p = T_m^{-1}\phi(F)T_p, \\ &R_m^{-1}\phi(A)R_n\check{X}R_n^{-1}\phi(B)R_p \\ &\quad + R_m^{-1}\phi(C)R_n\check{X}^TR_n^{-1}\phi(D)R_p = R_m^{-1}\phi(F)R_p, \\ &S_m^{-1}\phi(A)S_n\check{X}S_n^{-1}\phi(B)S_p \\ &\quad + S_m^{-1}\phi(C)S_n\check{X}^TS_n^{-1}\phi(D)S_p = S_m^{-1}\phi(F)S_p, \\ &T_m^{-1}\phi(A)\phi(Q)T_n\check{X}T_n^{-1}\phi(Q)\phi(B)T_p \\ &\quad + T_m^{-1}\phi(C)\phi(Q)T_n\check{X}^TT_n^{-1}\phi(Q)\phi(D)T_p = T_m^{-1}\phi(F)T_p, \\ &R_m^{-1}\phi(A)\phi(Q)R_n\check{X}R_n^{-1}\phi(Q)\phi(B)R_p \\ &\quad + R_m^{-1}\phi(C)\phi(Q)R_n\check{X}^TR_n^{-1}\phi(Q)\phi(D)R_p = R_m^{-1}\phi(F)R_p, \\ &S_m^{-1}\phi(A)\phi(Q)S_n\check{X}S_n^{-1}\phi(Q)\phi(B)S_p \\ &\quad + S_m^{-1}\phi(C)\phi(Q)S_n\check{X}^TS_n^{-1}\phi(Q)\phi(D)S_p = S_m^{-1}\phi(F)S_p. \end{aligned} \tag{40}$$

Hence

$$\begin{aligned} &\phi(A)T_n\check{X}T_n^{-1}\phi(B) + \phi(C)\left(T_n\check{X}T_n^{-1}\right)^T\phi(D) = \phi(F), \\ &\phi(A)R_n\check{X}R_n^{-1}\phi(B) + \phi(C)\left(R_n\check{X}R_n^{-1}\right)^T\phi(D) = \phi(F), \\ &\phi(A)S_n\check{X}S_n^{-1}\phi(B) + \phi(C)\left(S_n\check{X}S_n^{-1}\right)^T\phi(D) = \phi(F), \\ &\phi(A)\phi(Q)T_n\check{X}T_n^{-1}\phi(Q)\phi(B) \\ &\quad + \phi(C)\phi(Q)\left(T_n\check{X}T_n^{-1}\right)^T\phi(Q)\phi(D) = \phi(F), \\ &\phi(A)\phi(Q)R_n\check{X}R_n^{-1}\phi(Q)\phi(B) \\ &\quad + \phi(C)\phi(Q)\left(R_n\check{X}R_n^{-1}\right)^T\phi(Q)\phi(D) = \phi(F), \\ &\phi(A)\phi(Q)S_n\check{X}S_n^{-1}\phi(Q)\phi(B) \\ &\quad + \phi(C)\phi(Q)\left(S_n\check{X}S_n^{-1}\right)^T\phi(Q)\phi(D) = \phi(F), \end{aligned} \tag{41}$$

which implies that $T_n\check{X}T_n^{-1}, R_n\check{X}R_n^{-1}$, and $S_n\check{X}S_n^{-1}$ are also solutions of (36). Thus,

$$\frac{1}{4}\left(\check{X} + T_n\check{X}T_n^{-1} + R_n\check{X}R_n^{-1} + S_n\check{X}S_n^{-1}\right) \tag{42}$$

is also a solution of (36), where

$$\begin{aligned}
&\check{X} + T_n\check{X}T_n^{-1} + R_n\check{X}R_n^{-1} + S_n\check{X}S_n^{-1} \\
&\quad = \left[\widetilde{X_{ij}}\right]_{4\times 4}, \quad i,j = 1,2,3,4, \\
&\widetilde{X_{11}} = \ X_{11} + X_{22} + X_{33} + X_{44}, \\
&\widetilde{X_{12}} = \ X_{12} - X_{21} + X_{34} - X_{43}, \\
&\widetilde{X_{13}} = \ X_{13} - X_{24} - X_{31} + X_{42}, \\
&\widetilde{X_{14}} = \ X_{14} + X_{23} - X_{32} - X_{41}, \\
&\widetilde{X_{21}} = -X_{12} + X_{21} - X_{34} + X_{43}, \\
&\widetilde{X_{22}} = \ X_{11} + X_{22} + X_{33} + X_{44}, \\
&\widetilde{X_{23}} = \ X_{14} + X_{23} - X_{32} - X_{41}, \\
&\widetilde{X_{24}} = -X_{13} + X_{24} + X_{31} - X_{42}, \\
&\widetilde{X_{31}} = -X_{13} + X_{24} + X_{31} - X_{42}, \\
&\widetilde{X_{32}} = -X_{14} - X_{23} + X_{32} + X_{41}, \\
&\widetilde{X_{33}} = \ X_{11} + X_{22} + X_{33} + X_{44}, \\
&\widetilde{X_{34}} = X_{12} - X_{21} + X_{34} - X_{43}, \\
&\widetilde{X_{41}} = -X_{14} - X_{23} + X_{32} + X_{41}, \\
&\widetilde{X_{42}} = \ X_{13} - X_{24} - X_{31} + X_{42}, \\
&\widetilde{X_{43}} = -X_{12} + X_{21} - X_{34} + X_{43}, \\
&\widetilde{X_{44}} = X_{11} + X_{22} + X_{33} + X_{44}.
\end{aligned} \tag{43}$$

Let

$$\begin{aligned}
\widehat{X} = {} & \frac{1}{4}\left(X_{11} + X_{22} + X_{33} + X_{44}\right) \\
& + \frac{1}{4}\left(-X_{12} + X_{21} - X_{34} + X_{43}\right) i \\
& + \frac{1}{4}\left(-X_{13} + X_{24} + X_{31} - X_{42}\right) j \\
& + \frac{1}{4}\left(-X_{14} - X_{23} + X_{32} + X_{41}\right) k.
\end{aligned} \tag{44}$$

Then it is not difficult to verify that

$$\phi\left(\widehat{X}\right) = \frac{1}{4}\left(\check{X} + T_n\check{X}T_n^{-1} + R_n\check{X}R_n^{-1} + S_n\check{X}S_n^{-1}\right). \tag{45}$$

We have that $\widehat{X}$ is a solution of (34) by (a) in Lemma 4. The proof is completed. □

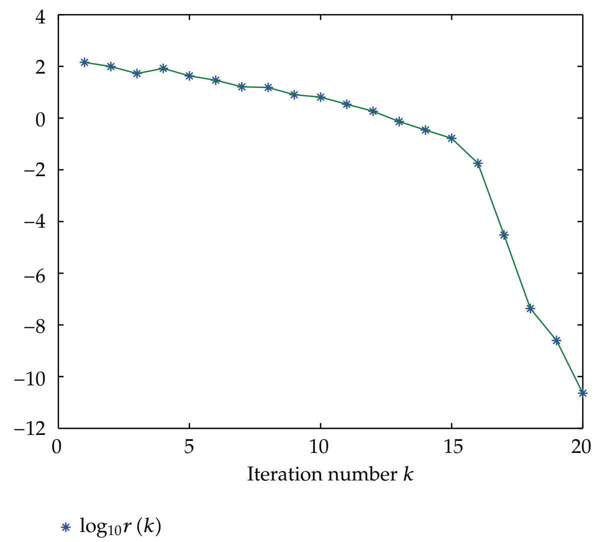


Figure 1: The convergence curve for the Frobenius norm of the residuals from Example 18.

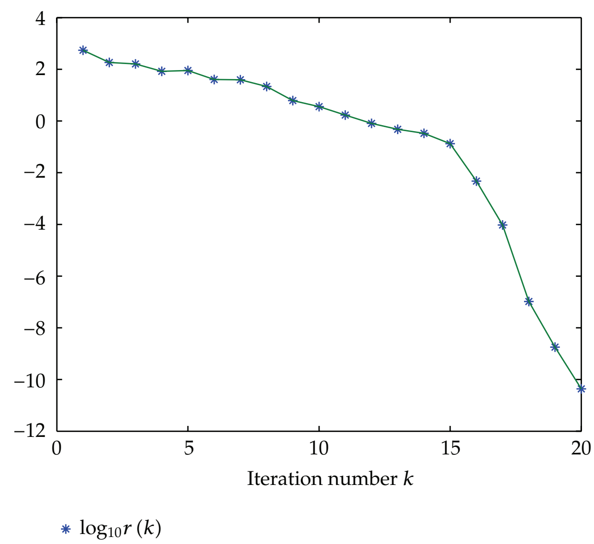


Figure 2: The convergence curve for the Frobenius norm of the residuals from Example 19.

Lemma 15. *There exists a permutation matrix $P(4n,4n)$ such that* (36) *is equivalent to*

$$\begin{gathered}
\begin{bmatrix} \phi^T(B)\otimes\phi(A) + \left(\phi^T(D)\otimes\phi(C)\right)P(4n,4n) \\ \phi^T(B)\phi(Q)\otimes\phi(A)\phi(Q) + \left(\phi^T(D)\phi(Q)\otimes\phi(C)\phi(Q)\right)P(4n,4n) \end{bmatrix} \\
\times \operatorname{vec}\left(\left[X_{ij}\right]_{4\times 4}\right) = \begin{bmatrix} \operatorname{vec}\left(\phi(F)\right) \\ \operatorname{vec}\left(\phi(F)\right) \end{bmatrix}.
\end{gathered} \tag{46}$$

Lemma 15 is easily proven through Lemma 5. So we omit it here.

Theorem 16. *When Problem 1 is consistent, let its solution set be denoted by S_H. If $\overset{\circ}{X} \in S_H$, and $\overset{\circ}{X}$ can be expressed as*

$$\overset{\circ}{X} = A^H G B^H + D G^H C + Q A^H G B^H Q + Q D G^H C Q, \quad G \in \mathbb{H}^{m\times p}, \tag{47}$$

then, $\overset{\circ}{X}$ is the least Frobenius norm solution of Problem 1.

Proof. By (a), (b), and (c) in Lemmas 4, 5, and 6, we have that

$$\begin{aligned}
&\operatorname{vec}\left(\phi\left(\overset{\circ}{X}\right)\right)\\
&= \operatorname{vec}\left(\phi^T(A)\phi(G)\phi^T(B) + \phi(D)\phi^T(G)\phi(C)\right.\\
&\qquad + \phi(Q)\phi^T(A)\phi(G)\phi^T(B)\phi(Q)\\
&\qquad \left.+ \phi(Q)\phi(D)\phi^T(G)\phi(C)\phi(Q)\right)\\
&= \left[\phi(B)\otimes\phi^T(A) + \left(\phi^T(C)\otimes\phi(D)\right)P(4m,4p),\right.\\
&\qquad \phi(Q)\phi(B)\otimes\phi(Q)\phi^T(A)\\
&\qquad \left.+\left(\phi(Q)\phi^T(C)\otimes\phi(Q)\phi(D)\right)P(4m,4p)\right]\\
&\qquad \times\begin{bmatrix}\operatorname{vec}(\phi(G))\\ \operatorname{vec}(\phi(G))\end{bmatrix}\\
&= \begin{bmatrix}\phi^T(B)\otimes\phi(A) + (\phi^T(D)\otimes\phi(C))P(4n,4n)\\ \phi^T(B)\phi(Q)\otimes\phi(A)\phi(Q) + (\phi^T(D)\phi(Q)\otimes\phi(C)\phi(Q))P(4n,4n)\end{bmatrix}^T\\
&\qquad \times\begin{bmatrix}\operatorname{vec}(\phi(G))\\ \operatorname{vec}(\phi(G))\end{bmatrix}\\
&\in R\begin{bmatrix}\phi^T(B)\otimes\phi(A) + \left(\phi^T(D)\otimes\phi(C)\right)P(4n,4n)\\ \phi^T(B)\phi(Q)\otimes\phi(A)\phi(Q) + \left(\phi^T(D)\phi(Q)\otimes\phi(C)\phi(Q)\right)P(4n,4n)\end{bmatrix}^T.
\end{aligned} \tag{48}$$

By Lemma 12, $\phi(\overset{\circ}{X})$ is the least Frobenius norm solution of matrix equations (46).

Noting (11), we derive from Lemmas 13, 14, and 15 that $\overset{\circ}{X}$ is the least Frobenius norm solution of Problem 1. □

From Algorithm 7, it is obvious that, if we consider

$$X(1) = A^H G B^H + D G^H C + Q A^H G B^H Q + Q D G^H C Q, \quad G \in \mathbb{H}^{m\times p}, \tag{49}$$

then all $X(k)$ generated by Algorithm 7 can be expressed as

$$X(k) = A^H G_k B^H + D G_k^H C + Q A^H G_k B^H Q + Q D G_k^H C Q, \quad G_k \in \mathbb{H}^{m\times p}. \tag{50}$$

Using the above conclusion and considering Theorem 16, we propose the following theorem.

Theorem 17. *Suppose that Problem 1 is consistent. Let the initial iteration matrix be*

$$X(1) = A^H G B^H + D G^H C + Q A^H G B^H Q + Q D G^H C Q, \tag{51}$$

where G is an arbitrary quaternion matrix, or especially, $X(1) = 0$, then the solution X^*, generated by Algorithm 7, is the least Frobenius norm solution of Problem 1.

Now we study Problem 2. When Problem 1 is consistent, the solution set of Problem 1 denoted by S_H is not empty. Then, For a given reflexive matrix $X_0 \in \mathbb{RH}^{n\times n}(Q)$,

$$\begin{aligned}AXB + CX^H D = F &\Longleftrightarrow A(X - X_0)B + C(X - X_0)^H D\\ &= F - AX_0B - CX_0^H D.\end{aligned} \tag{52}$$

Let $\dot{X} = X - X_0$ and $\dot{F} = F - AX_0B - CX_0^H D$, then Problem 2 is equivalent to finding the least Frobenius norm reflexive solution of the quaternion matrix equation

$$A\dot{X}B + C\dot{X}^H D = \dot{F}. \tag{53}$$

By using Algorithm 7, let the initial iteration matrix $\dot{X}(1) = A^H G B^H + D G^H C + Q A^H G B^H Q + Q D G^H C Q$, where G is an arbitrary quaternion matrix in $\mathbb{H}^{m\times p}$, or especially, $\dot{X}(1) = 0$, we can obtain the least Frobenius norm reflexive solution $\dot{X}^*$ of (53). Then we can obtain the solution of Problem 2, which is

$$\check{X} = \dot{X}^* + X_0. \tag{54}$$

4. Examples

In this section, we give two examples to illustrate the efficiency of the theoretical results.

Example 18. Consider the quaternion matrix equation

$$AXB + CX^H D = F, \tag{55}$$

with

$$A = \begin{bmatrix} 1+i+j+k & 2+2j+2k & i+3j+3k & 2+i+4j+4k \\ 3+2i+2j+k & 3+3i+j & 1+7i+3j+k & 6-7i+2j-4k \end{bmatrix},$$
$$B = \begin{bmatrix} -2+i+3j+k & 3+3i+4j+k \\ 5-4i-6j-5k & -1-5i+4j+3k \\ -1+2i+j+k & 2+2j+6k \\ 3-2i+j+k & 4+i-2j-4k \end{bmatrix},$$
$$C = \begin{bmatrix} 1+i+j+k & 2+2i+2j & k & 2+2i+2j+k \\ -1+i+3j+3k & -2+3i+4j+2k & 6-2i+6j+4k & 3+i+j+5k \end{bmatrix}, \tag{56}$$
$$D = \begin{bmatrix} -1+4j+2k & 1+2i+3j+3k \\ 7+6i+5j+6k & 3+7i-j+9k \\ 4+i+6j+k & 7+2i+9j+k \\ 1+i+3j-3k & 1+3i+2j+2k \end{bmatrix},$$
$$F = \begin{bmatrix} 1+3i+2j+k & 2-2i+3j+3k \\ -1+4i+2j+k & -2+2i-1j \end{bmatrix}.$$

We apply Algorithm 7 to find the reflexive solution with respect to the generalized reflection matrix.

$$Q = \begin{bmatrix} 0.28 & 0 & 0.96k & 0 \\ 0 & -1 & 0.0 & 0 \\ -0.96k & 0 & -0.28 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{57}$$

For the initial matrix $X(1) = Q$, we obtain a solution, that is

$$\begin{aligned} X^* &= X(20) \\ &= \left[\begin{matrix} 0.27257 - 0.23500i + 0.034789j + 0.054677k & 0.085841 - 0.064349i + 0.10387j - 0.26871k \\ 0.028151 - 0.013246i - 0.091097j + 0.073137k & -0.46775 + 0.048363i - 0.015746j + 0.21642k \\ 0.043349 + 0.079178i + 0.085167j - 0.84221k & 0.35828 - 0.13849i - 0.085799j + 0.11446k \\ 0.13454 - 0.028417i - 0.063010j - 0.10574k & 0.10491 - 0.039417i + 0.18066j - 0.066963k \end{matrix} \right. \\ &\quad \left. \begin{matrix} -0.043349 + 0.079178i + 0.085167j + 0.84221k & -0.014337 - 0.14868i - 0.011146j - 0.00036397k \\ 0.097516 + 0.12146i - 0.017662j - 0.037534k & -0.064483 + 0.070885i - 0.089331j + 0.074969k \\ -0.21872 + 0.18532i + 0.011399j + 0.029390k & 0.00048529 + 0.014861i - 0.19824j - 0.019117k \\ -0.14099 + 0.084013i - 0.037890j - 0.17938k & -0.63928 - 0.067488i + 0.042030j + 0.10106k \end{matrix} \right], \end{aligned} \tag{58}$$

with corresponding residual $\|R(20)\| = 2.2775 \times 10^{-11}$. The convergence curve for the Frobenius norm of the residuals $R(k)$ is given in Figure 1, where $r(k) = \|R(k)\|$.

Example 19. In this example, we choose the matrices A, B, C, D, F, and Q as same as in Example 18. Let

$$X_0 = \begin{bmatrix} 1 & -0.36i - 0.75k & 0 & -0.5625i - 0.48k \\ 0.36i & 1.28 & 0.48j & -0.96j \\ 0 & 1-0.48j & 1 & 0.64-0.75j \\ 0.48k & 0.96j & 0.64 & 0.72 \end{bmatrix} \in \mathbb{RH}^{n\times n}(Q). \tag{59}$$

In order to find the optimal approximation reflexive solution to the given matrix X_0, let $\dot{X} = X - X_0$ and $\dot{F} = F - AX_0B - CX_0^H D$. Now we can obtain the least Frobenius norm reflexive solution $\dot{X}^*$ of the quaternion matrix equation $A\dot{X}B + C\dot{X}^H D = \dot{F}$, by choosing the initial matrix $\dot{X}(1) = 0$, that is

$$\dot{X}^{*} = \dot{X}(20) = \begin{bmatrix} -0.47933 + 0.021111i + 0.11703j - 0.066934k & -0.52020 - 0.19377i + 0.17420j - 0.59403k & 0.24134 - 0.025941i - 0.31930j + 0.26961k & 0.089010 - 0.67960i + 0.43044j - 0.045772k \\ -0.31132 - 0.11705i - 0.33980j + 0.11991k & -0.86390 - 0.21715i - 0.037089j + 0.40609k & -0.089933 - 0.25485i + 0.087788j - 0.23349k & -0.17004 - 0.37655i + 0.80481j + 0.37636k \\ -0.24134 - 0.025941i - 0.31930j - 0.26961k & -0.44552 + 0.13065i + 0.14533j + 0.39015k & -0.63660 + 0.16515i - 0.13216j + 0.073845k & -0.034329 + 0.32283i + 0.50970j - 0.066757k \\ -0.11693 - 0.27043i + 0.22718j - 0.0000012004k & -0.44790 - 0.31260i - 1.0271j + 0.27275k & 0.00000090029 + 0.17038i + 0.20282j - 0.087697k & -0.44735 + 0.21846i - 0.21469j - 0.15347k \end{bmatrix}, \quad (60)$$

with corresponding residual $\|\dot{R}(20)\| = 4.3455 \times 10^{-11}$. The convergence curve for the Frobenius norm of the residuals $\dot{R}(k)$ is given in Figure 2, where $r(k) = \|\dot{R}(k)\|$.

Therefore, the optimal approximation reflexive solution to the given matrix X_0 is

$$\breve{X} = \dot{X}^{*} + X_0 = \begin{bmatrix} 0.52067 + 0.021111i + 0.11703j - 0.066934k & -0.52020 - 0.55377i + 0.17420j - 1.3440k & 0.24134 - 0.025941i - 0.31930j + 0.26961k & 0.089010 - 1.2421i + 0.43044j - 0.52577k \\ -0.31132 + 0.24295i - 0.33980j + 0.11991k & 0.41610 - 0.21715i - 0.037089j + 0.40609k & -0.089933 - 0.25485i + 0.56779j - 0.23349k & -0.17004 - 0.37655i - 0.15519j + 0.37636k \\ -0.24134 - 0.025941i - 0.31930j - 0.26961k & 0.55448 + 0.13065i - 0.33467j + 0.39015k & 0.36340 + 0.16515i - 0.13216j + 0.073845k & 0.60567 + 0.32283i - 0.24030j - 0.066757k \\ -0.11693 - 0.27043i + 0.22718j + 0.48000k & -0.44790 - 0.31260i - 0.067117j + 0.27275k & 0.64000 + 0.17038i + 0.20282j - 0.087697k & 0.27265 + 0.21846i - 0.21469j - 0.15347k \end{bmatrix}. \quad (61)$$

The results show that Algorithm 7 is quite efficient.

5. Conclusions

In this paper, an algorithm has been presented for solving the reflexive solution of the quaternion matrix equation $AXB + CX^HD = F$. By this algorithm, the solvability of the problem can be determined automatically. Also, when the quaternion matrix equation $AXB + CX^HD = F$ is consistent over reflexive matrix X, for any reflexive initial iterative matrix, a reflexive solution can be obtained within finite iteration steps in the absence of roundoff errors. It has been proven that by choosing a suitable initial iterative matrix, we can derive the least Frobenius norm reflexive solution of the quaternion matrix equation $AXB + CX^HD = F$ through Algorithm 7. Furthermore, by using Algorithm 7, we solved Problem 2. Finally, two numerical examples were given to show the efficiency of the presented algorithm.

Acknowledgments

This research was supported by Grants from Innovation Program of Shanghai Municipal Education Commission (13ZZ080), the National Natural Science Foundation of China (11171205), the Natural Science Foundation of Shanghai (11ZR1412500), the Discipline Project at the corresponding level of Shanghai (A. 13-0101-12-005), and Shanghai Leading Academic Discipline Project (J50101).

References

[1] H. C. Chen and A. H. Sameh, "A matrix decomposition method for orthotropic elasticity problems," *SIAM Journal on Matrix Analysis and Applications*, vol. 10, no. 1, pp. 39–64, 1989.

[2] H. C. Chen, "Generalized reflexive matrices: special properties and applications," *SIAM Journal on Matrix Analysis and Applications*, vol. 19, no. 1, pp. 140–153, 1998.

[3] Q. W. Wang, H. X. Chang, and C. Y. Lin, "P-(skew)symmetric common solutions to a pair of quaternion matrix equations," *Applied Mathematics and Computation*, vol. 195, no. 2, pp. 721–732, 2008.

[4] S. F. Yuan and Q. W. Wang, "Two special kinds of least squares solutions for the quaternion matrix equation $AXB+CXD = E$," *Electronic Journal of Linear Algebra*, vol. 23, pp. 257–274, 2012.

[5] T. S. Jiang and M. S. Wei, "On a solution of the quaternion matrix equation $X - A\widetilde{X}B = C$ and its application," *Acta Mathematica Sinica (English Series)*, vol. 21, no. 3, pp. 483–490, 2005.

[6] Y. T. Li and W. J. Wu, "Symmetric and skew-antisymmetric solutions to systems of real quaternion matrix equations," *Computers & Mathematics with Applications*, vol. 55, no. 6, pp. 1142–1147, 2008.

[7] L. G. Feng and W. Cheng, "The solution set to the quaternion matrix equation $AX - \overline{X}B = 0$," *Algebra Colloquium*, vol. 19, no. 1, pp. 175–180, 2012.

[8] Z. Y. Peng, "An iterative method for the least squares symmetric solution of the linear matrix equation $AXB = C$," *Applied Mathematics and Computation*, vol. 170, no. 1, pp. 711–723, 2005.

[9] Z. Y. Peng, "A matrix LSQR iterative method to solve matrix equation $AXB = C$," *International Journal of Computer Mathematics*, vol. 87, no. 8, pp. 1820–1830, 2010.

[10] Z. Y. Peng, "New matrix iterative methods for constraint solutions of the matrix equation $AXB = C$," *Journal of Computational and Applied Mathematics*, vol. 235, no. 3, pp. 726–735, 2010.

[11] Z. Y. Peng, "Solutions of symmetry-constrained least-squares problems," *Numerical Linear Algebra with Applications*, vol. 15, no. 4, pp. 373–389, 2008.

[12] C. C. Paige, "Bidiagonalization of matrices and solutions of the linear equations," *SIAM Journal on Numerical Analysis*, vol. 11, pp. 197–209, 1974.

[13] X. F. Duan, A. P. Liao, and B. Tang, "On the nonlinear matrix equation $X - \sum_{i=1}^{m} A^{x}2a;_{i}X^{\delta_i}A_i = Q$," *Linear Algebra and Its Applications*, vol. 429, no. 1, pp. 110–121, 2008.

[14] X. F. Duan and A. P. Liao, "On the existence of Hermitian positive definite solutions of the matrix equation $X^s + A^* X^{-t} A = Q$," *Linear Algebra and Its Applications*, vol. 429, no. 4, pp. 673–687, 2008.

[15] X. F. Duan and A. P. Liao, "On the nonlinear matrix equation $X + A^* X^{-q} A = Q (q \geq 1)$," *Mathematical and Computer Modelling*, vol. 49, no. 5-6, pp. 936–945, 2009.

[16] X. F. Duan and A. P. Liao, "On Hermitian positive definite solution of the matrix equation $X - \sum_{i=1}^{m} A_i^* X^r A_i = Q$," *Journal of Computational and Applied Mathematics*, vol. 229, no. 1, pp. 27–36, 2009.

[17] X. F. Duan, C. M. Li, and A. P. Liao, "Solutions and perturbation analysis for the nonlinear matrix equation $X + \sum_{i=1}^{m} A_i^* X^{-1} A_i = I$," *Applied Mathematics and Computation*, vol. 218, no. 8, pp. 4458–4466, 2011.

[18] F. Ding, P. X. Liu, and J. Ding, "Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle," *Applied Mathematics and Computation*, vol. 197, no. 1, pp. 41–50, 2008.

[19] F. Ding and T. Chen, "Iterative least-squares solutions of coupled Sylvester matrix equations," *Systems & Control Letters*, vol. 54, no. 2, pp. 95–107, 2005.

[20] F. Ding and T. Chen, "Hierarchical gradient-based identification of multivariable discrete-time systems," *Automatica*, vol. 41, no. 2, pp. 315–325, 2005.

[21] M. H. Wang, M. S. Wei, and S. H. Hu, "An iterative method for the least-squares minimum-norm symmetric solution," *CMES: Computer Modeling in Engineering & Sciences*, vol. 77, no. 3-4, pp. 173–182, 2011.

[22] M. Dehghan and M. Hajarian, "An iterative algorithm for the reflexive solutions of the generalized coupled Sylvester matrix equations and its optimal approximation," *Applied Mathematics and Computation*, vol. 202, no. 2, pp. 571–588, 2008.

[23] M. Dehghan and M. Hajarian, "Finite iterative algorithms for the reflexive and anti-reflexive solutions of the matrix equation $A_1X_1B_1 + A_2X_2B_2 = C$," *Mathematical and Computer Modelling*, vol. 49, no. 9-10, pp. 1937–1959, 2009.

[24] M. Hajarian and M. Dehghan, "Solving the generalized Sylvester matrix equation $\sum_{i=1}^{p} A_iXB_i + \sum_{j=1}^{q} C_jYD_j = E$ over reflexive and anti-reflexive matrices," *International Journal of Control and Automation*, vol. 9, no. 1, pp. 118–124, 2011.

[25] M. Dehghan and M. Hajarian, "An iterative algorithm for solving a pair of matrix equations $AYB = E$, $CYD = F$ over generalized centro-symmetric matrices," *Computers & Mathematics with Applications*, vol. 56, no. 12, pp. 3246–3260, 2008.

[26] M. Dehghan and M. Hajarian, "Analysis of an iterative algorithm to solve the generalized coupled Sylvester matrix equations," *Applied Mathematical Modelling*, vol. 35, no. 7, pp. 3285–3300, 2011.

[27] M. Dehghan and M. Hajarian, "The general coupled matrix equations over generalized bisymmetric matrices," *Linear Algebra and Its Applications*, vol. 432, no. 6, pp. 1531–1552, 2010.

[28] M. Dehghan and M. Hajarian, "An iterative method for solving the generalized coupled Sylvester matrix equations over generalized bisymmetric matrices," *Applied Mathematical Modelling*, vol. 34, no. 3, pp. 639–654, 2010.

[29] A. G. Wu, B. Li, Y. Zhang, and G. R. Duan, "Finite iterative solutions to coupled Sylvester-conjugate matrix equations," *Applied Mathematical Modelling*, vol. 35, no. 3, pp. 1065–1080, 2011.

[30] A. G. Wu, L. L. Lv, and G. R. Duan, "Iterative algorithms for solving a class of complex conjugate and transpose matrix equations," *Applied Mathematics and Computation*, vol. 217, no. 21, pp. 8343–8353, 2011.

[31] A. G. Wu, L. L. Lv, and M. Z. Hou, "Finite iterative algorithms for extended Sylvester-conjugate matrix equations," *Mathematical and Computer Modelling*, vol. 54, no. 9-10, pp. 2363–2384, 2011.

[32] N. L. Bihan and J. Mars, "Singular value decomposition of matrices of Quaternions matrices: a new tool for vector-sensor signal processing," *Signal Process*, vol. 84, no. 7, pp. 1177–1199, 2004.

[33] N. L. Bihan and S. J. Sangwine, "Jacobi method for quaternion matrix singular value decomposition," *Applied Mathematics and Computation*, vol. 187, no. 2, pp. 1265–1271, 2007.

[34] F. O. Farid, Q. W. Wang, and F. Zhang, "On the eigenvalues of quaternion matrices," *Linear and Multilinear Algebra*, vol. 59, no. 4, pp. 451–473, 2011.

[35] S. de Leo and G. Scolarici, "Right eigenvalue equation in quaternionic quantum mechanics," *Journal of Physics A*, vol. 33, no. 15, pp. 2971–2995, 2000.

[36] S. J. Sangwine and N. L. Bihan, "Quaternion singular value decomposition based on bidiagonalization to a real or complex matrix using quaternion Householder transformations," *Applied Mathematics and Computation*, vol. 182, no. 1, pp. 727–738, 2006.

[37] C. C. Took, D. P. Mandic, and F. Zhang, "On the unitary diagonalisation of a special class of quaternion matrices," *Applied Mathematics Letters*, vol. 24, no. 11, pp. 1806–1809, 2011.

[38] Q. W. Wang, H. X. Chang, and Q. Ning, "The common solution to six quaternion matrix equations with applications," *Applied Mathematics and Computation*, vol. 198, no. 1, pp. 209–226, 2008.

[39] Q. W. Wang, J. W. van der Woude, and H. X. Chang, "A system of real quaternion matrix equations with applications," *Linear Algebra and Its Applications*, vol. 431, no. 12, pp. 2291–2303, 2009.

[40] Q. W. Wang, S. W. Yu, and W. Xie, "Extreme ranks of real matrices in solution of the quaternion matrix equation $AXB = C$ with applications," *Algebra Colloquium*, vol. 17, no. 2, pp. 345–360, 2010.

[41] Q. W. Wang and J. Jiang, "Extreme ranks of (skew-)Hermitian solutions to a quaternion matrix equation," *Electronic Journal of Linear Algebra*, vol. 20, pp. 552–573, 2010.

[42] Q. W. Wang, X. Liu, and S. W. Yu, "The common bisymmetric nonnegative definite solutions with extreme ranks and inertias to a pair of matrix equations," *Applied Mathematics and Computation*, vol. 218, no. 6, pp. 2761–2771, 2011.

[43] Q. W. Wang, Y. Zhou, and Q. Zhang, "Ranks of the common solution to six quaternion matrix equations," *Acta Mathematicae Applicatae Sinica (English Series)*, vol. 27, no. 3, pp. 443–462, 2011.

[44] F. Zhang, "Geršgorin type theorems for quaternionic matrices," *Linear Algebra and Its Applications*, vol. 424, no. 1, pp. 139–153, 2007.

[45] F. Zhang, "Quaternions and matrices of quaternions," *Linear Algebra and Its Applications*, vol. 251, pp. 21–57, 1997.

[46] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.

[47] Y. X. Peng, X. Y. Hu, and L. Zhang, "An iteration method for the symmetric solutions and the optimal approximation solution of the matrix equation $AXB = C$," *Applied Mathematics and Computation*, vol. 160, no. 3, pp. 763–777, 2005.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 961568, 6 pages
http://dx.doi.org/10.1155/2013/961568

Research Article

The Optimization on Ranks and Inertias of a Quadratic Hermitian Matrix Function and Its Applications

Yirong Yao

Department of Mathematics, Shanghai University, Shanghai 200444, China

Correspondence should be addressed to Yirong Yao; yryao@staff.shu.edu.cn

Received 3 December 2012; Accepted 9 January 2013

Academic Editor: Yang Zhang

We solve optimization problems on the ranks and inertias of the quadratic Hermitian matrix function $Q - XPX^*$ subject to a consistent system of matrix equations $AX = C$ and $XB = D$. As applications, we derive necessary and sufficient conditions for the solvability to the systems of matrix equations and matrix inequalities $AX = C, XB = D$, and $XPX^* = (>, <, \geq, \leq)Q$ in the Löwner partial ordering to be feasible, respectively. The findings of this paper widely extend the known results in the literature.

1. Introduction

Throughout this paper, we denote the complex number field by $\mathbb{C}$. The notations $\mathbb{C}^{m\times n}$ and $\mathbb{C}_h^{m\times m}$ stand for the sets of all $m \times n$ complex matrices and all $m \times m$ complex Hermitian matrices, respectively. The identity matrix with an appropriate size is denoted by I. For a complex matrix A, the symbols A^* and $r(A)$ stand for the conjugate transpose and the rank of A, respectively. The Moore-Penrose inverse of $A \in \mathbb{C}^{m\times n}$, denoted by $A^\dagger$, is defined to be the unique solution X to the following four matrix equations

$$\begin{aligned} &(1)\ AXA = A, \qquad (2)\ XAX = X, \\ &(3)\ (AX)^* = AX, \qquad (4)\ (XA)^* = XA. \end{aligned} \tag{1}$$

Furthermore, L_A and R_A stand for the two projectors $L_A = I - A^\dagger A$ and $R_A = I - AA^\dagger$ induced by A, respectively. It is known that $L_A = L_A^*$ and $R_A = R_A^*$. For $A \in \mathbb{C}_h^{m\times m}$, its inertia

$$\mathbb{I}_n(A) = (i_+(A), i_-(A), i_0(A)) \tag{2}$$

is the triple consisting of the numbers of the positive, negative, and zero eigenvalues of A, counted with multiplicities, respectively. It is easy to see that $i_+(A) + i_-(A) = r(A)$. For two Hermitian matrices A and B of the same sizes, we say $A > B$ ($A \geq B$) in the Löwner partial ordering if $A - B$ is positive (nonnegative) definite.

The investigation on maximal and minimal ranks and inertias of linear and quadratic matrix function is active in recent years (see, e.g., [1–24]). Tian [21] considered the maximal and minimal ranks and inertias of the Hermitian quadratic matrix function

$$h(X) = AXBX^*A^* + AXC + C^*X^*A^* + D, \tag{3}$$

where B and D are Hermitian matrices. Moreover, Tian [22] investigated the maximal and minimal ranks and inertias of the quadratic Hermitian matrix function

$$f(X) = Q - XPX^* \tag{4}$$

such that $AX = C$.

The goal of this paper is to give the maximal and minimal ranks and inertias of the matrix function (4) subject to the consistent system of matrix equations

$$AX = C, \qquad XB = D, \tag{5}$$

where $Q \in \mathbb{C}_h^{n\times n}$, $P \in \mathbb{C}_h^{p\times p}$ are given complex matrices. As applications, we consider the necessary and sufficient conditions for the solvability to the systems of matrix equations and inequality

$$\begin{aligned} &AX = C, \quad XB = D, \quad XPX^* = Q, \\ &AX = C, \quad XB = D, \quad XPX^* > Q, \\ &AX = C, \quad XB = D, \quad XPX^* < Q, \\ &AX = C, \quad XB = D, \quad XPX^* \geq Q, \\ &AX = C, \quad XB = D, \quad XPX^* \leq Q, \end{aligned} \tag{6}$$

in the Löwner partial ordering to be feasible, respectively.

2. The Optimization on Ranks and Inertias of (4) Subject to (5)

In this section, we consider the maximal and minimal ranks and inertias of the quadratic Hermitian matrix function (4) subject to (5). We begin with the following lemmas.

Lemma 1 (see [3]). *Let $A \in \mathbb{C}_h^{m\times m}$, $B \in \mathbb{C}^{m\times p}$, and $C \in \mathbb{C}^{q\times m}$ be given and denote*

$$\begin{gathered} P_1 = \begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix}, \qquad P_2 = \begin{bmatrix} A & C^* \\ C & 0 \end{bmatrix}, \\ P_3 = \begin{bmatrix} A & B & C^* \\ B^* & 0 & 0 \end{bmatrix}, \qquad P_4 = \begin{bmatrix} A & B & C^* \\ C & 0 & 0 \end{bmatrix}. \end{gathered} \tag{7}$$

Then

$$\begin{aligned} &\max_{Y\in\mathbb{C}^{p\times q}} r\left[A - BYC - (BYC)^*\right] \\ &\quad = \min\left\{r\left[A \ \ B \ \ C^*\right], r(P_1), r(P_2)\right\}, \\ &\min_{Y\in\mathbb{C}^{p\times q}} r\left[A - BYC - (BYC)^*\right] \\ &\quad = 2r\left[A \ \ B \ \ C^*\right] \\ &\qquad + \max\left\{w_+ + w_-, g_+ + g_-, w_+ + g_-, w_- + g_+\right\}, \\ &\max_{Y\in\mathbb{C}^{p\times q}} i_\pm\left[A - BYC - (BYC)^*\right] = \min\left\{i_\pm(P_1), i_\pm(P_2)\right\}, \\ &\min_{Y\in\mathbb{C}^{p\times q}} i_\pm\left[A - BYC - (BYC)^*\right] \\ &\quad = r\left[A \ \ B \ \ C^*\right] + \max\left\{i_\pm(P_1) - r(P_3), i_\pm(P_2) - r(P_4)\right\}, \end{aligned} \tag{8}$$

where

$$w_\pm = i_\pm(P_1) - r(P_3), \qquad g_\pm = i_\pm(P_2) - r(P_4). \tag{9}$$

Lemma 2 (see [4]). *Let $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{m\times k}$, $C \in \mathbb{C}^{l\times n}$, $D \in \mathbb{C}^{m\times p}$, $E \in \mathbb{C}^{q\times n}$, $Q \in \mathbb{C}^{m_1\times k}$, and $P \in \mathbb{C}^{l\times n_1}$ be given. Then*

$$\begin{aligned} &(1)\ \ r(A) + r(R_A B) = r(B) + r(R_B A) = r\left[A \ \ B\right], \\ &(2)\ \ r(A) + r(CL_A) = r(C) + r(AL_C) = r\begin{bmatrix} A \\ C \end{bmatrix}, \\ &(3)\ \ r(B) + r(C) + r(R_B A L_C) = r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}, \\ &(4)\ \ r(P) + r(Q) + r\begin{bmatrix} A & BL_Q \\ R_P C & 0 \end{bmatrix} = r\begin{bmatrix} A & B & 0 \\ C & 0 & P \\ 0 & Q & 0 \end{bmatrix}, \\ &(5)\ \ r\begin{bmatrix} R_B A L_C & R_B D \\ E L_C & 0 \end{bmatrix} + r(B) + r(C) = r\begin{bmatrix} A & D & B \\ E & 0 & 0 \\ C & 0 & 0 \end{bmatrix}. \end{aligned} \tag{10}$$

Lemma 3 (see [23]). *Let $A \in \mathbb{C}_h^{m\times m}$, $B \in \mathbb{C}^{m\times n}$, $C \in \mathbb{C}_h^{n\times n}$, $Q \in \mathbb{C}^{m\times n}$, and $P \in \mathbb{C}^{p\times n}$ be given, and, $T \in \mathbb{C}^{m\times m}$ be nonsingular. Then*

$$\begin{aligned} &(1)\ \ i_\pm(TAT^*) = i_\pm(A), \\ &(2)\ \ i_\pm\begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix} = i_\pm(A) + i_\pm(C), \\ &(3)\ \ i_\pm\begin{bmatrix} 0 & Q \\ Q^* & 0 \end{bmatrix} = r(Q), \\ &(4)\ \ i_\pm\begin{bmatrix} A & BL_P \\ L_P B^* & 0 \end{bmatrix} + r(P) = i_\pm\begin{bmatrix} A & B & 0 \\ B^* & 0 & P^* \\ 0 & P & 0 \end{bmatrix}. \end{aligned} \tag{11}$$

Lemma 4. *Let A, C, B, and D be given. Then the following statements are equivalent.*

(1) *System* (5) *is consistent.*

(2) *Let*

$$r\left[A \ \ C\right] = r(A), \qquad \begin{bmatrix} D \\ B \end{bmatrix} = r(B), \qquad AD = CB. \tag{12}$$

In this case, the general solution can be written as

$$X = A^\dagger C + L_A D B^\dagger + L_A V R_B, \tag{13}$$

where V is an arbitrary matrix over $\mathbb{C}$ with appropriate size.

Now we give the fundamental theorem of this paper.

Theorem 5. *Let $f(X)$ be as given in* (4) *and assume that $AX = C$ and $XB = D$ in* (5) *is consistent. Then*

$$\max_{AX=C,XB=D} r(Q - XPX^*) = \min\left\{ n + r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix} - r(A) - r(B) - r(P), 2n + r(AQA^* - CPC^*) - 2r(A), r\begin{bmatrix} Q & 0 & D \\ 0 & -P & B \\ D^* & B^* & 0 \end{bmatrix} - 2r(B) \right\},$$

$$\min_{AX=C,XB=D} r(Q - XPX^*) = 2n + 2r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix} - 2r(A) - 2r(B) - r(P) + \max\{s_+ + s_-, t_+ + t_-, s_+ + t_-, s_- + t_+\},$$

$$\max_{AX=C,XB=D} i_\pm(Q - XPX^*) = \min\left\{ n + i_\pm(AQA^* - CPC^*) - r(A), i_\pm\begin{bmatrix} Q & 0 & D \\ 0 & -P & B \\ D^* & B^* & 0 \end{bmatrix} - r(B) \right\}, \quad (14)$$

$$\min_{AX=C,XB=D} i_\pm(Q - XPX^*) = n + r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix} - r(A) - r(B) - i_\pm(P) + \max\{s_\pm, t_\pm\}, \quad (15)$$

where

$$s_\pm = -n + r(A) - i_\mp(P) + i_\pm(AQA^* - CPC^*) - r\begin{bmatrix} CP & AQA^* \\ B^* & D^*A^* \end{bmatrix},$$

$$t_\pm = -n + r(A) - i_\mp(P) + i_\pm\begin{bmatrix} Q & 0 & D \\ 0 & -P & B \\ D^* & B^* & 0 \end{bmatrix} - \begin{bmatrix} 0 & P & B \\ AQ & CP & 0 \\ D^* & B^* & 0 \end{bmatrix}. \quad (16)$$

Proof. It follows from Lemma 4 that the general solution of (4) can be expressed as

$$X = X_0 + L_AVR_B, \quad (17)$$

where V is an arbitrary matrix over $\mathbb{C}$ and X_0 is a special solution of (5). Then

$$Q - XPX^* = Q - (X_0 + L_AVR_B)P(X_0 + L_AVR_B)^*. \quad (18)$$

Note that

$$\begin{aligned} r&\left[Q - (X_0 + L_AVR_B)P(X_0 + L_AVR_B)^*\right] \\ &= r\begin{bmatrix} Q & (X_0 + L_AVR_B)P \\ P(X_0 + L_AVR_B)^* & P \end{bmatrix} - r(P) \\ &= r\left[\begin{bmatrix} Q & X_0P \\ PX_0^* & P \end{bmatrix} + \begin{bmatrix} L_A \\ 0 \end{bmatrix} V \begin{bmatrix} 0 & R_BP \end{bmatrix} + \left(\begin{bmatrix} L_A \\ 0 \end{bmatrix} V \begin{bmatrix} 0 & R_BP \end{bmatrix}\right)^*\right] - r(P), \end{aligned} \quad (19)$$

$$\begin{aligned} i_\pm&\left[Q - (X_0 + L_AVR_B)P(X_0 + L_AVR_B)^*\right] \\ &= i_\pm\begin{bmatrix} Q & (X_0 + L_AVR_B)P \\ P(X_0 + L_AVR_B)^* & P \end{bmatrix} - i_\pm(P) \\ &= i_\pm\left[\begin{bmatrix} Q & X_0P \\ PX_0^* & P \end{bmatrix} + \begin{bmatrix} L_A \\ 0 \end{bmatrix} V \begin{bmatrix} 0 & R_BP \end{bmatrix} + \left(\begin{bmatrix} L_A \\ 0 \end{bmatrix} V \begin{bmatrix} 0 & R_BP \end{bmatrix}\right)^*\right] - i_\pm(P). \end{aligned} \quad (20)$$

Let

$$q(V) = \begin{bmatrix} Q & X_0P \\ PX_0^* & P \end{bmatrix} + \begin{bmatrix} L_A \\ 0 \end{bmatrix} V \begin{bmatrix} 0 & R_BP \end{bmatrix} + \left(\begin{bmatrix} L_A \\ 0 \end{bmatrix} V \begin{bmatrix} 0 & R_BP \end{bmatrix}\right)^*. \quad (21)$$

Applying Lemma 1 to (19) and (20) yields

$$\begin{aligned} &\max_V r[q(V)] = \min\{r(M), r(M_1), r(M_2)\}, \\ &\min_V r[q(V)] = 2r(M) + \max\{s_+ + s_-, t_+ + t_-, s_+ + t_-, s_- + t_+\}, \\ &\max_V i_\pm[q(V)] = \min\{i_\pm(M_1), i_\pm(M_2)\}, \\ &\min_V i_\pm[q(V)] = r(M) + \max\{s_\pm, t_\pm\}, \end{aligned} \quad (22)$$

where

$$M = \begin{bmatrix} Q & X_0P & L_A & 0 \\ PX_0^* & P & 0 & PR_B \end{bmatrix}, \quad M_1 = \begin{bmatrix} Q & X_0P & L_A \\ PX_0^* & P & 0 \\ L_A & 0 & 0 \end{bmatrix},$$

$$M_2 = \begin{bmatrix} Q & X_0P & 0 \\ PX_0^* & P & PR_B \\ 0 & R_BP & 0 \end{bmatrix}, \quad M_3 = \begin{bmatrix} Q & X_0P & L_A & 0 \\ PX_0^* & P & 0 & PR_B \\ L_A & 0 & 0 & 0 \end{bmatrix},$$

$$M_4 = \begin{bmatrix} Q & X_0P & L_A & 0 \\ PX_0^* & P & 0 & PR_B \\ 0 & R_BP & 0 & 0 \end{bmatrix},$$

$$s_\pm = i_\pm(M_1) - r(M_3), \qquad t_\pm = i_\pm(M_2) - r(M_4). \quad (23)$$

Applying Lemmas 2 and 3, elementary matrix operations and congruence matrix operations, we obtain

$$\begin{gathered} r(M) = n + r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix} - r(A) - r(B), \\ r(M_1) = 2n + r(AQA^* - CPC^*) - 2r(A) + r(P), \\ i_\pm(M_1) = n + i_\pm(AQA^* - CPC^*) - r(A) + i_\pm(P), \\ r(M_2) = r\begin{bmatrix} Q & 0 & D \\ 0 & -P & B \\ D^* & B^* & 0 \end{bmatrix} - 2r(B) + r(P), \\ i_\pm(M_2) = i_\pm\begin{bmatrix} Q & 0 & D \\ 0 & -P & B \\ D^* & B^* & 0 \end{bmatrix} - r(B) + i_\pm(P), \\ r(M_3) = 2n + r(P) - 2r(A) - r(B) + r\begin{bmatrix} CP & AQA^* \\ B^* & D^*A^* \end{bmatrix}, \\ r(M_4) = n + r(P) + r\begin{bmatrix} 0 & P & B \\ AQ & CP & 0 \\ D^* & B^* & 0 \end{bmatrix} - 2r(B) - r(A). \end{gathered} \tag{24}$$

Substituting (24) into (22), we obtain the results. □

Using immediately Theorem 5, we can easily get the following.

Theorem 6. *Let $f(X)$ be as given in (4), $s_\pm$ and let $t_\pm$ be as given in Theorem 5 and assume that $AX = C$ and $XB = D$ in (5) are consistent. Then we have the following.*

(a) *$AX = C$ and $XB = D$ have a common solution such that $Q - XPX^* \geq 0$ if and only if*

$$\begin{gathered} n + r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix} - r(A) - r(B) - i_-(P) + s_- \leq 0, \\ n + r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix} - r(A) - r(B) - i_-(P) + t_- \leq 0. \end{gathered} \tag{25}$$

(b) *$AX = C$ and $XB = D$ have a common solution such that $Q - XPX^* \leq 0$ if and only if*

$$\begin{gathered} n + r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix} - r(A) - r(B) - i_+(P) + s_+ \leq 0, \\ n + r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix} - r(A) - r(B) - i_+(P) + t_+ \leq 0. \end{gathered} \tag{26}$$

(c) *$AX = C$ and $XB = D$ have a common solution such that $Q - XPX^* > 0$ if and only if*

$$\begin{gathered} i_+(AQA^* - CPC^*) - r(A) \geq 0, \\ i_+\begin{bmatrix} Q & 0 & D \\ 0 & -P & B \\ D^* & B^* & 0 \end{bmatrix} - r(B) \geq n. \end{gathered} \tag{27}$$

(d) *$AX = C$ and $XB = D$ have a common solution such that $Q - XPX^* < 0$ if and only if*

$$\begin{gathered} i_-(AQA^* - CPC^*) - r(A) \geq 0, \\ i_-\begin{bmatrix} Q & 0 & D \\ 0 & -P & B \\ D^* & B^* & 0 \end{bmatrix} - r(B) \geq n. \end{gathered} \tag{28}$$

(e) *All common solutions of $AX = C$ and $XB = D$ satisfy $Q - XPX^* \geq 0$ if and only if*

$$\begin{gathered} n + i_-(AQA^* - CPC^*) - r(A) = 0, \\ \text{or, } i_-\begin{bmatrix} Q & 0 & D \\ 0 & -P & B \\ D^* & B^* & 0 \end{bmatrix} - r(B) = 0. \end{gathered} \tag{29}$$

(f) *All common solutions of $AX = C$ and $XB = D$ satisfy $Q - XPX^* \leq 0$ if and only if*

$$\begin{gathered} n + i_+(AQA^* - CPC^*) - r(A) = 0, \\ \text{or, } i_+\begin{bmatrix} Q & 0 & D \\ 0 & -P & B \\ D^* & B^* & 0 \end{bmatrix} - r(B) = 0. \end{gathered} \tag{30}$$

(g) *All common solutions of $AX = C$ and $XB = D$ satisfy $Q - XPX^* > 0$ if and only if*

$$n + r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix} - r(A) - r(B) - i_+(P) + s_+ = n, \tag{31}$$

or

$$n + r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix} - r(A) - r(B) - i_+(P) + t_+ = n. \tag{32}$$

(h) *All common solutions of $AX = C$ and $XB = D$ satisfy $Q - XPX^* < 0$ if and only if*

$$n + r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix} - r(A) - r(B) - i_-(P) + s_- = n, \tag{33}$$

or

$$n + r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix} - r(A) - r(B) - i_-(P) + t_- = n. \tag{34}$$

(i) $AX = C$, $XB = D$, *and* $Q = XPX^*$ *have a common solution if and only if*

$$2n+2r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix}-2r(A)-2r(B)-r(P)+s_++s_-\leq 0,$$
$$2n+2r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix}-2r(A)-2r(B)-r(P)+t_++t_-\leq 0,$$
$$2n+2r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix}-2r(A)-2r(B)-r(P)+s_++t_-\leq 0,$$
$$2n+2r\begin{bmatrix} 0 & P & P \\ AQ & CP & 0 \\ -D^* & 0 & B^* \end{bmatrix}-2r(A)-2r(B)-r(P)+s_-+t_+\leq 0. \tag{35}$$

Let $P = I$ in Theorem 5, we get the following corollary.

Corollary 7. *Let* $Q \in \mathbb{C}^{n\times n}$, A, B, C, *and* D *be given. Assume that* (5) *is consistent. Denote*

$$T_1 = \begin{bmatrix} C & AQ \\ B^* & D^* \end{bmatrix}, \qquad T_2 = AQA^* - CC^*,$$
$$T_3 = \begin{bmatrix} Q & D \\ D^* & B^*B \end{bmatrix}, \qquad T_4 = \begin{bmatrix} C & AQA^* \\ B^* & D^*A^* \end{bmatrix}, \tag{36}$$
$$T_5 = \begin{bmatrix} CB & AQ \\ B^*B & D^* \end{bmatrix}.$$

Then,

$$\begin{aligned}
&\max_{AX=C,XB=D} r(Q-XX^*) \\
&\quad = \min\{n + r(T_1) - r(A) - r(B), 2n + r(T_2) \\
&\qquad -2r(A), n + r(T_3) - 2r(B)\}, \\
&\min_{AX=C,XB=D} r(Q-XX^*) \\
&\quad = 2r(T_1) + \max\{r(T_2) - 2r(T_4), -n + r(T_3) \\
&\qquad - 2r(T_5), i_+(T_2) + i_-(T_3) \\
&\qquad - r(T_4) - r(T_5), -n + i_-(T_2) \\
&\qquad + i_+(T_3) - r(T_4) - r(T_5)\} \\
&\max_{AX=C,XB=D} i_+(Q-XX^*) \\
&\quad = \min\{n + i_+(T_2) - r(A), i_+(T_3) - r(B)\}, \\
&\max_{AX=C,XB=D} i_-(Q-XX^*) \\
&\quad = \min\{n + i_-(T_2) - r(A), n + i_-(T_3) - r(B)\}, \\
&\min_{AX=C,XB=D} i_+(Q-XX^*) \\
&\quad = r(T_1) + \max\{i_+(T_2) - r(T_4), i_+(T_3) - n - r(T_5)\}, \\
&\min_{AX=C,XB=D} i_-(Q-XX^*) \\
&\quad = r(T_1) + \max\{i_-(T_2) - r(T_4), i_-(T_3) - r(T_5)\}.
\end{aligned} \tag{37}$$

Remark 8. Corollary 7 is one of the results in [24].

Let B and D vanish in Theorem 5, then we can obtain the maximal and minimal ranks and inertias of (4) subject to $AX = C$.

Corollary 9. *Let* $f(X)$ *be as given in* (4) *and assume that* $AX = C$ *is consistent. Then*

$$\begin{aligned}
&\max_{AX=C} r(Q-XPX^*) \\
&\quad = \min\{n + r[AQ \;\; CP] - r(A) - r(B), \\
&\qquad 2n + r(AQA^* - CPC^*) - 2r(A), r(Q) + r(P)\} \\
&\min_{AX=C} r(Q-XPX^*) \\
&\quad = 2n + 2r[AQ \;\; CP] - 2r(A) \\
&\qquad + \max\{s_+ + s_-, t_+ + t_-, s_+ + t_-, s_- + t_+\}, \\
&\max_{AX=C} i_\pm(Q-XPX^*) \\
&\quad = \min\{n + i_\pm(AQA^* - CPC^*) - r(A), i_\pm(Q) + i_\mp(P)\}, \\
&\min_{AX=C} i_\pm(Q-XPX^*) \\
&\quad = n + r[AQ \;\; CP] - r(A) + i_\mp(P) + \max\{s_\pm, t_\pm\},
\end{aligned} \tag{38}$$

where

$$\begin{aligned}
s_\pm &= -n + r(A) - i_\mp(P) \\
&\quad + i_\pm(AQA^* - CPC^*) - r[CP \;\; AQA^*],
\end{aligned} \tag{39}$$
$$t_\pm = -n + r(A) + i_\pm(Q) - r(P) - [AQ \;\; CP].$$

Remark 10. Corollary 9 is one of the results in [22].

References

[1] D. L. Chu, Y. S. Hung, and H. J. Woerdeman, "Inertia and rank characterizations of some matrix expressions," *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 3, pp. 1187–1226, 2009.

[2] Z. H. He and Q. W. Wang, "Solutions to optimization problems on ranks and inertias of a matrix function with applications,"

Applied Mathematics and Computation, vol. 219, no. 6, pp. 2989–3001, 2012.

[3] Y. Liu and Y. Tian, "Max-min problems on the ranks and inertias of the matrix expressions $A - BXC \pm (BXC)^*$ with applications," *Journal of Optimization Theory and Applications*, vol. 148, no. 3, pp. 593–622, 2011.

[4] G. Marsaglia and G. P. H. Styan, "Equalities and inequalities for ranks of matrices," *Linear and Multilinear Algebra*, vol. 2, pp. 269–292, 1974/75.

[5] X. Zhang, Q. W. Wang, and X. Liu, "Inertias and ranks of some Hermitian matrix functions with applications," *Central European Journal of Mathematics*, vol. 10, no. 1, pp. 329–351, 2012.

[6] Q. W. Wang, Y. Zhou, and Q. Zhang, "Ranks of the common solution to six quaternion matrix equations," *Acta Mathematicae Applicatae Sinica*, vol. 27, no. 3, pp. 443–462, 2011.

[7] Q. Zhang and Q. W. Wang, "The (P, Q)-(skew)symmetric extremal rank solutions to a system of quaternion matrix equations," *Applied Mathematics and Computation*, vol. 217, no. 22, pp. 9286–9296, 2011.

[8] D. Z. Lian, Q. W. Wang, and Y. Tang, "Extreme ranks of a partial banded block quaternion matrix expression subject to some matrix equations with applications," *Algebra Colloquium*, vol. 18, no. 2, pp. 333–346, 2011.

[9] H. S. Zhang and Q. W. Wang, "Ranks of submatrices in a general solution to a quaternion system with applications," *Bulletin of the Korean Mathematical Society*, vol. 48, no. 5, pp. 969–990, 2011.

[10] Q. W. Wang and J. Jiang, "Extreme ranks of (skew-)Hermitian solutions to a quaternion matrix equation," *Electronic Journal of Linear Algebra*, vol. 20, pp. 552–573, 2010.

[11] I. A. Khan, Q. W. Wang, and G. J. Song, "Minimal ranks of some quaternion matrix expressions with applications," *Applied Mathematics and Computation*, vol. 217, no. 5, pp. 2031–2040, 2010.

[12] Q. Wang, S. Yu, and W. Xie, "Extreme ranks of real matrices in solution of the quaternion matrix equation $AXB = C$ with applications," *Algebra Colloquium*, vol. 17, no. 2, pp. 345–360, 2010.

[13] Q. W. Wang, H. S. Zhang, and G. J. Song, "A new solvable condition for a pair of generalized Sylvester equations," *Electronic Journal of Linear Algebra*, vol. 18, pp. 289–301, 2009.

[14] Q. W. Wang, S. W. Yu, and Q. Zhang, "The real solutions to a system of quaternion matrix equations with applications," *Communications in Algebra*, vol. 37, no. 6, pp. 2060–2079, 2009.

[15] Q. W. Wang and C. K. Li, "Ranks and the least-norm of the general solution to a system of quaternion matrix equations," *Linear Algebra and its Applications*, vol. 430, no. 5-6, pp. 1626–1640, 2009.

[16] Q. W. Wang, H. S. Zhang, and S. W. Yu, "On solutions to the quaternion matrix equation $AXB+CYD = E$," *Electronic Journal of Linear Algebra*, vol. 17, pp. 343–358, 2008.

[17] Q. W. Wang, G. J. Song, and X. Liu, "Maximal and minimal ranks of the common solution of some linear matrix equations over an arbitrary division ring with applications," *Algebra Colloquium*, vol. 16, no. 2, pp. 293–308, 2009.

[18] Q. W. Wang, G. J. Song, and C. Y. Lin, "Rank equalities related to the generalized inverse $A^{(2)}_{T,S}$ with applications," *Applied Mathematics and Computation*, vol. 205, no. 1, pp. 370–382, 2008.

[19] Q. W. Wang, S. W. Yu, and C. Y. Lin, "Extreme ranks of a linear quaternion matrix expression subject to triple quaternion matrix equations with applications," *Applied Mathematics and Computation*, vol. 195, no. 2, pp. 733–744, 2008.

[20] Q. W. Wang, G. J. Song, and C. Y. Lin, "Extreme ranks of the solution to a consistent system of linear quaternion matrix equations with an application," *Applied Mathematics and Computation*, vol. 189, no. 2, pp. 1517–1532, 2007.

[21] Y. Tian, "Formulas for calculating the extremum ranks and inertias of a four-term quadratic matrix-valued function and their applications," *Linear Algebra and its Applications*, vol. 437, no. 3, pp. 835–859, 2012.

[22] Y. Tian, "Solving optimization problems on ranks and inertias of some constrained nonlinear matrix functions via an algebraic linearization method," *Nonlinear Analysis. Theory, Methods & Applications*, vol. 75, no. 2, pp. 717–734, 2012.

[23] Y. Tian, "Maximization and minimization of the rank and inertia of the Hermitian matrix expression $A - BX - (BX)^*$ with applications," *Linear Algebra and its Applications*, vol. 434, no. 10, pp. 2109–2139, 2011.

[24] Q. W. Wang, X. Zhang, and Z. H. He, "On the Hermitian structures of the solution to a pair of matrix equations," *Linear and Multilinear Algebra*, vol. 61, no. 1, pp. 73–90, 2013.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2012, Article ID 471573, 19 pages
doi:10.1155/2012/471573

Research Article

Constrained Solutions of a System of Matrix Equations

Qing-Wen Wang[1] and Juan Yu[1, 2]

[1] *Department of Mathematics, Shanghai University, 99 Shangda Road, Shanghai 200444, China*
[2] *Department of Basic Mathematics, China University of Petroleum, Qingdao 266580, China*

Correspondence should be addressed to Qing-Wen Wang, wqw858@yahoo.com.cn

Received 26 September 2012; Accepted 7 December 2012

Academic Editor: Panayiotis J. Psarrakos

We derive the necessary and sufficient conditions of and the expressions for the orthogonal solutions, the symmetric orthogonal solutions, and the skew-symmetric orthogonal solutions of the system of matrix equations $AX = B$ and $XC = D$, respectively. When the matrix equations are not consistent, the least squares symmetric orthogonal solutions and the least squares skew-symmetric orthogonal solutions are respectively given. As an auxiliary, an algorithm is provided to compute the least squares symmetric orthogonal solutions, and meanwhile an example is presented to show that it is reasonable.

1. Introduction

Throughout this paper, the following notations will be used. $\mathbb{R}^{m\times n}$, $O\mathbb{R}^{n\times n}$, $S\mathbb{R}^{n\times n}$, and $AS\mathbb{R}^{n\times n}$ denote the set of all $m \times n$ real matrices, the set of all $n \times n$ orthogonal matrices, the set of all $n \times n$ symmetric matrices, and the set of all $n \times n$ skew-symmetric matrices, respectively. I_n is the identity matrix of order n. $(\cdot)^T$ and $\mathrm{tr}(\cdot)$ represent the transpose and the trace of the real matrix, respectively. $\|\cdot\|$ stands for the Frobenius norm induced by the inner product. The following two definitions will also be used.

Definition 1.1 (see [1]). A real matrix $X \in \mathbb{R}^{n\times n}$ is said to be a symmetric orthogonal matrix if $X^T = X$ and $X^T X = I_n$.

Definition 1.2 (see [2]). A real matrix $X \in \mathbb{R}^{2m\times 2m}$ is called a skew-symmetric orthogonal matrix if $X^T = -X$ and $X^T X = I_n$.

The set of all $n \times n$ symmetric orthogonal matrices and the set of all $2m \times 2m$ skew-symmetric orthogonal matrices are, respectively, denoted by $SO\mathbb{R}^{n\times n}$ and $SSO\mathbb{R}^{2m\times 2m}$. Since the linear matrix equation(s) and its optimal approximation problem have great applications in structural design, biology, control theory, and linear optimal control, and so forth, see, for example, [3–5], there has been much attention paid to the linear matrix equation(s). The well-known system of matrix equations

$$AX = B, \qquad XC = D, \tag{1.1}$$

as one kind of linear matrix equations, has been investigated by many authors, and a series of important and useful results have been obtained. For instance, the system (1.1) with unknown matrix X being bisymmetric, centrosymmetric, bisymmetric nonnegative definite, Hermitian and nonnegative definite, and (P, Q)-(skew) symmetric has been, respectively, investigated by Wang et al. [6, 7], Khatri and Mitra [8], and Zhang and Wang [9]. Of course, if the solvability conditions of system (1.1) are not satisfied, we may consider its least squares solution. For example, Li et al. [10] presented the least squares mirrorsymmetric solution. Yuan [11] got the least-squares solution. Some results concerning the system (1.1) can also be found in [12–18].

Symmetric orthogonal matrices and skew-symmetric orthogonal matrices play important roles in numerical analysis and numerical solutions of partial differential equations. Papers [1, 2], respectively, derived the symmetric orthogonal solution X of the matrix equation $XC = D$ and the skew-symmetric orthogonal solution X of the matrix equation $AX = B$. Motivated by the work mentioned above, we in this paper will, respectively, study the orthogonal solutions, symmetric orthogonal solutions, and skew-symmetric orthogonal solutions of the system (1.1). Furthermore, if the solvability conditions are not satisfied, the least squares skew-symmetric orthogonal solutions and the least squares symmetric orthogonal solutions of the system (1.1) will be also given.

The remainder of this paper is arranged as follows. In Section 2, some lemmas are provided to give the main results of this paper. In Sections 3, 4, and 5, the necessary and sufficient conditions of and the expression for the orthogonal, the symmetric orthogonal, and the skew-symmetric orthogonal solutions of the system (1.1) are, respectively, obtained. In Section 6, the least squares skew-symmetric orthogonal solutions and the least squares symmetric orthogonal solutions of the system (1.1) are presented, respectively. In addition, an algorithm is provided to compute the least squares symmetric orthogonal solutions, and meanwhile an example is presented to show that it is reasonable. Finally, in Section 7, some concluding remarks are given.

2. Preliminaries

In this section, we will recall some lemmas and the special C-S decomposition which will be used to get the main results of this paper.

Lemma 2.1 (see [1, Lemmas 1 and 2]). *Given $C \in \mathbb{R}^{2m\times n}$, $D \in \mathbb{R}^{2m\times n}$. The matrix equation $YC = D$ has a solution $Y \in O\mathbb{R}^{2m\times 2m}$ if and only if $D^T D = C^T C$. Let the singular value decompositions of C and D be, respectively,*

$$C = U\begin{pmatrix} \Pi & 0 \\ 0 & 0 \end{pmatrix}V^T, \qquad D = W\begin{pmatrix} \Pi & 0 \\ 0 & 0 \end{pmatrix}V^T, \tag{2.1}$$

where

$$\Pi = \operatorname{diag}(\sigma_1, \ldots, \sigma_k) > 0, \quad k = \operatorname{rank}(C) = \operatorname{rank}(D),$$
$$U = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \in O\mathbb{R}^{2m\times 2m}, \quad U_1 \in \mathbb{R}^{2m\times k}, \tag{2.2}$$
$$W = \begin{pmatrix} W_1 & W_2 \end{pmatrix} \in O\mathbb{R}^{2m\times 2m}, \quad W_1 \in \mathbb{R}^{2m\times k}, \qquad V = \begin{pmatrix} V_1 & V_2 \end{pmatrix} \in O\mathbb{R}^{n\times n}, \quad V_1 \in \mathbb{R}^{n\times k}.$$

Then the orthogonal solutions of $YC = D$ can be described as

$$Y = W\begin{pmatrix} I_k & 0 \\ 0 & P \end{pmatrix} U^T, \tag{2.3}$$

where $P \in O\mathbb{R}^{(2m-k)\times(2m-k)}$ is arbitrary.

Lemma 2.2 (see [2, Lemmas 1 and 2]). *Given $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{n\times m}$. The matrix equation $AX = B$ has a solution $X \in O\mathbb{R}^{m\times m}$ if and only if $AA^T = BB^T$. Let the singular value decompositions of A and B be, respectively,*

$$A = U\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^T, \qquad B = U\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} Q^T, \tag{2.4}$$

where

$$\Sigma = \operatorname{diag}(\delta_1, \ldots, \delta_l) > 0, \quad l = \operatorname{rank}(A) = \operatorname{rank}(B), \qquad U = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \in O\mathbb{R}^{n\times n}, \quad U_1 \in \mathbb{R}^{n\times l},$$
$$V = \begin{pmatrix} V_1 & V_2 \end{pmatrix} \in O\mathbb{R}^{m\times m}, \quad V_1 \in \mathbb{R}^{m\times l}, \qquad Q = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \in O\mathbb{R}^{m\times m}, \quad Q_1 \in \mathbb{R}^{m\times l}. \tag{2.5}$$

Then the orthogonal solutions of $AX = B$ can be described as

$$X = V\begin{pmatrix} I_l & 0 \\ 0 & W \end{pmatrix} Q^T, \tag{2.6}$$

where $W \in O\mathbb{R}^{(m-l)\times(m-l)}$ is arbitrary.

Lemma 2.3 (see [2, Theorem 1]). *If*

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} \in O\mathbb{R}^{2m\times 2m}, \quad X_{11} \in AS\mathbb{R}^{k\times k}, \tag{2.7}$$

then the C-S decomposition of X can be expressed as

$$\begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}^T \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} \begin{pmatrix} D_1 & 0 \\ 0 & R_2 \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \tag{2.8}$$

where $D_1 \in O\mathbb{R}^{k\times k}$, $D_2, R_2 \in O\mathbb{R}^{(2m-k)\times(2m-k)}$,

$$\Sigma_{11} = \begin{pmatrix} \widetilde{I} & 0 & 0 \\ 0 & \widetilde{C} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \Sigma_{12} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & I \end{pmatrix},$$

$$\Sigma_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & I \end{pmatrix}, \qquad \Sigma_{22} = \begin{pmatrix} I & 0 & 0 \\ 0 & -\widetilde{C}^T & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\widetilde{I} = \operatorname{diag}\left(\widetilde{I}_1, \ldots, \widetilde{I}_{r_1}\right), \quad \widetilde{I}_1 = \cdots = \widetilde{I}_{r_1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \widetilde{C} = \operatorname{diag}(C_1, \ldots, C_{l_1}), \tag{2.9}$$

$$C_i = \begin{pmatrix} 0 & c_i \\ -c_i & 0 \end{pmatrix}, \quad i = 1, \ldots, l_1; \qquad S = \operatorname{diag}(S_1, \ldots, S_{l_1}),$$

$$S_i = \begin{pmatrix} s_i & 0 \\ 0 & s_i \end{pmatrix}, \quad S_i^T S_i + C_i^T C_i = I_2, \quad i = 1, \ldots, l_1;$$

$$2r_1 + 2l_1 = \operatorname{rank}(X_{11}), \qquad \Sigma_{21} = \Sigma_{12}^T, \qquad k - 2r_1 = \operatorname{rank}(X_{21}).$$

Lemma 2.4 (see [1, Theorem 1]). *If*

$$K = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} \in O\mathbb{R}^{m\times m}, \quad K_{11} \in S\mathbb{R}^{l\times l}, \tag{2.10}$$

then the C-S decomposition of K can be described as

$$\begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}^T \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} \begin{pmatrix} D_1 & 0 \\ 0 & R_2 \end{pmatrix} = \begin{pmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{pmatrix}, \tag{2.11}$$

where $D_1 \in O\mathbb{R}^{l\times l}$, D_2, $R_2 \in O\mathbb{R}^{(m-l)\times(m-l)}$,

$$\Pi_{11} = \begin{pmatrix} \widetilde{I} & 0 & 0 \\ 0 & \widetilde{C} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \Pi_{12} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & I \end{pmatrix},$$

$$\Pi_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & I \end{pmatrix}, \qquad \Pi_{22} = \begin{pmatrix} I & 0 & 0 \\ 0 & -\widetilde{C} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\widetilde{I} = \operatorname{diag}(i_1, \ldots, i_r), \qquad i_j = \pm 1, \quad j = 1, \ldots, r; \tag{2.12}$$

$$S = \operatorname{diag}(s_{1'}, \ldots, s_{l'}), \qquad \widetilde{C} = \operatorname{diag}(c_{1'}, \ldots, c_{l'}),$$

$$s_{i'}^2 = \sqrt{1 - c_{i'}^2}, \quad i = 1', \ldots, l'; \qquad r + l' = \operatorname{rank}(K_{11}), \qquad \Pi_{21} = \Pi_{12}^T.$$

Remarks 2.5. In order to know the C-S decomposition of an orthogonal matrix with a $k \times k$ leading (skew-) symmetric submatrix for details, one can deeply study the proof of Theorem 1 in [1] and [2].

Lemma 2.6. *Given $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{n\times m}$. Then the matrix equation $AX = B$ has a solution $X \in SO\mathbb{R}^{m\times m}$ if and only if $AA^T = BB^T$ and $AB^T = BA^T$. When these conditions are satisfied, the general symmetric orthogonal solutions can be expressed as*

$$X = \widetilde{V}\begin{pmatrix} I_{2l-r} & 0 \\ 0 & G \end{pmatrix}\widetilde{Q}^T, \tag{2.13}$$

where

$$\widetilde{Q} = JQ\,\mathrm{diag}(I, D_2) \in O\mathbb{R}^{m\times m}, \qquad \widetilde{V} = JV\,\mathrm{diag}(I, R_2) \in O\mathbb{R}^{m\times m},$$
$$J = \begin{pmatrix} I & 0 & 0 \\ 0 & 0 & I \\ 0 & I & 0 \end{pmatrix}, \tag{2.14}$$

and $G \in SO\mathbb{R}^{(m-2l+r)\times(m-2l+r)}$ is arbitrary.

Proof. The Necessity. Assume $X \in SO\mathbb{R}^{m\times m}$ is a solution of the matrix equation $AX = B$, then we have

$$BB^T = AXX^TA^T = AA^T,$$
$$BA^T = AXA^T = AX^TA^T = AB^T. \tag{2.15}$$

The Sufficiency. Since the equality $AA^T = BB^T$ holds, then by Lemma 2.2, the singular value decompositions of A and B can be, respectively, expressed as (2.4). Moreover, the condition $AB^T = BA^T$ means

$$U\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}V^TQ\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}U^T = U\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}Q^TV\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}U^T, \tag{2.16}$$

which can be written as

$$V_1^TQ_1 = Q_1^TV_1. \tag{2.17}$$

From Lemma 2.2, the orthogonal solutions of the matrix equation $AX = B$ can be described as (2.6). Now we aim to find that X in (2.6) is also symmetric. Suppose that X is symmetric, then we have

$$\begin{pmatrix} I_l & 0 \\ 0 & W^T \end{pmatrix}V^TQ = Q^TV\begin{pmatrix} I_l & 0 \\ 0 & W \end{pmatrix}, \tag{2.18}$$

together with the partitions of the matrices Q and V in Lemma 2.2, we get

$$V_1^T Q_1 = Q_1^T V_1, \tag{2.19}$$

$$Q_1^T V_2 W = V_1^T Q_2, \tag{2.20}$$

$$Q_2^T V_2 W = W^T V_2^T Q_2. \tag{2.21}$$

By (2.17), we can get (2.19). Now we aim to find the orthogonal solutions of the system of matrix equations (2.20) and (2.21). Firstly, we obtain from (2.20) that $Q_1^T V_2 (Q_1^T V_2)^T = V_1^T Q_2 (V_1^T Q_2)^T$, then by Lemma 2.2, (2.20) has an orthogonal solution W. By (2.17), the $l \times l$ leading principal submatrix of the orthogonal matrix $V^T Q$ is symmetric. Then we have, from Lemma 2.4,

$$V_1^T Q_2 = D_1 \Pi_{12} R_2^T, \tag{2.22}$$

$$Q_1^T V_2 = D_1 \Pi_{12} D_2^T, \tag{2.23}$$

$$V_2^T Q_2 = D_2 \Pi_{22} R_2^T. \tag{2.24}$$

From (2.20), (2.22), and (2.23), the orthogonal solution W of (2.20) is

$$W = D_2 \begin{pmatrix} G & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix} R_2^T, \tag{2.25}$$

where $G \in O\mathbb{R}^{(m-2l+r)\times(m-2l+r)}$ is arbitrary. Combining (2.21), (2.24), and (2.25) yields $G^T = G$, that is, G is a symmetric orthogonal matrix. Denote

$$\widehat{V} = V \operatorname{diag}(I, D_2), \qquad \widehat{Q} = Q \operatorname{diag}(I, R_2), \tag{2.26}$$

then the symmetric orthogonal solutions of the matrix equation $AX = B$ can be expressed as

$$X = \widehat{V} \begin{pmatrix} I & 0 & 0 \\ 0 & G & 0 \\ 0 & 0 & I \end{pmatrix} \widehat{Q}^T. \tag{2.27}$$

Let the partition matrix J be

$$J = \begin{pmatrix} I & 0 & 0 \\ 0 & 0 & I \\ 0 & I & 0 \end{pmatrix}, \tag{2.28}$$

compatible with the block matrix

$$\begin{pmatrix} I & 0 & 0 \\ 0 & G & 0 \\ 0 & 0 & I \end{pmatrix}. \tag{2.29}$$

Put

$$\widetilde{V} = J\widehat{V}, \qquad \widetilde{Q} = J\widehat{Q}, \tag{2.30}$$

then the symmetric orthogonal solutions of the matrix equation $AX = B$ can be described as (2.13). □

Setting $A = C^T$, $B = D^T$, and $X = Y^T$ in [2, Theorem 2], and then by Lemmas 2.1 and 2.3, we can have the following result.

Lemma 2.7. *Given $C \in \mathbb{R}^{2m\times n}$, $D \in \mathbb{R}^{2m\times n}$. Then the equation has a solution $Y \in SSO\mathbb{R}^{2m\times 2m}$ if and only if $D^TD = C^TC$ and $D^TC = -C^TD$. When these conditions are satisfied, the skew-symmetric orthogonal solutions of the matrix equation $YC = D$ can be described as*

$$Y = \widetilde{W}\begin{pmatrix} I & 0 \\ 0 & H \end{pmatrix}\widetilde{U}^T, \tag{2.31}$$

where

$$\widetilde{W} = J'W \operatorname{diag}(I, -D_2) \in O\mathbb{R}^{2m\times 2m}, \qquad \widetilde{U} = J'U \operatorname{diag}(I, R_2) \in O\mathbb{R}^{2m\times 2m},$$
$$J' = \begin{pmatrix} I & 0 & 0 \\ 0 & 0 & I \\ 0 & I & 0 \end{pmatrix}, \tag{2.32}$$

and $H \in SSO\mathbb{R}^{2r\times 2r}$ is arbitrary.

3. The Orthogonal Solutions of the System (1.1)

The following theorems give the orthogonal solutions of the system (1.1).

Theorem 3.1. *Given $A, B \in \mathbb{R}^{n\times m}$ and $C, D \in \mathbb{R}^{m\times n}$, suppose the singular value decompositions of A and B are, respectively, as* (2.4). *Denote*

$$Q^TC = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}, \qquad V^TD = \begin{pmatrix} D_1 \\ D_2 \end{pmatrix}, \tag{3.1}$$

where $C_1, D_1 \in \mathbb{R}^{l\times n}$, and $C_2, D_2 \in \mathbb{R}^{(m-l)\times n}$. Let the singular value decompositions of C_2 and D_2 be, respectively,

$$C_2 = \widetilde{U}\begin{pmatrix}\Pi & 0\\ 0 & 0\end{pmatrix}\widetilde{V}^T, \qquad D_2 = \widetilde{W}\begin{pmatrix}\Pi & 0\\ 0 & 0\end{pmatrix}\widetilde{V}^T, \tag{3.2}$$

where $\widetilde{U}, \widetilde{W} \in O\mathbb{R}^{(m-l)\times(m-l)}$, $\widetilde{V} \in O\mathbb{R}^{n\times n}$, $\Pi \in \mathbb{R}^{k\times k}$ is diagonal, whose diagonal elements are nonzero singular values of C_2 or D_2. Then the system (1.1) *has orthogonal solutions if and only if*

$$AA^T = BB^T, \qquad C_1 = D_1, \qquad D_2^T D_2 = C_2^T C_2. \tag{3.3}$$

In which case, the orthogonal solutions can be expressed as

$$X = \widehat{V}\begin{pmatrix}I_{k+l} & 0\\ 0 & G'\end{pmatrix}\widehat{Q}^T, \tag{3.4}$$

where

$$\widehat{V} = V\begin{pmatrix}I_l & 0\\ 0 & \widetilde{W}\end{pmatrix} \in O\mathbb{R}^{m\times m}, \qquad \widehat{Q} = Q\begin{pmatrix}I_l & 0\\ 0 & \widetilde{U}\end{pmatrix} \in O\mathbb{R}^{m\times m}, \tag{3.5}$$

and $G' \in O\mathbb{R}^{(m-k-l)\times(m-k-l)}$ is arbitrary.

Proof. Let the singular value decompositions of A and B be, respectively, as (2.4). Since the matrix equation $AX = B$ has orthogonal solutions if and only if

$$AA^T = BB^T, \tag{3.6}$$

then by Lemma 2.2, its orthogonal solutions can be expressed as (2.6). Substituting (2.6) and (3.1) into the matrix equation $XC = D$, we have $C_1 = D_1$ and $WC_2 = D_2$. By Lemma 2.1, the matrix equation $WC_2 = D_2$ has orthogonal solution W if and only if

$$D_2^T D_2 = C_2^T C_2. \tag{3.7}$$

Let the singular value decompositions of C_2 and D_2 be, respectively,

$$C_2 = \widetilde{U}\begin{pmatrix}\Pi & 0\\ 0 & 0\end{pmatrix}\widetilde{V}^T, \qquad D_2 = \widetilde{W}\begin{pmatrix}\Pi & 0\\ 0 & 0\end{pmatrix}\widetilde{V}^T, \tag{3.8}$$

where $\widetilde{U}, \widetilde{W} \in O\mathbb{R}^{(m-l)\times(m-l)}$, $\widetilde{V} \in O\mathbb{R}^{n\times n}$, $\Pi \in \mathbb{R}^{k\times k}$ is diagonal, whose diagonal elements are nonzero singular values of C_2 or D_2. Then the orthogonal solutions can be described as

$$W = \widetilde{W}\begin{pmatrix}I_k & 0\\ 0 & G'\end{pmatrix}\widetilde{U}^T, \tag{3.9}$$

where $G' \in O\mathbb{R}^{(m-k-l)\times(m-k-l)}$ is arbitrary. Therefore, the common orthogonal solutions of the system (1.1) can be expressed as

$$X = V\begin{pmatrix} I_l & 0 \\ 0 & W \end{pmatrix}Q^T = V\begin{pmatrix} I_l & 0 \\ 0 & \widetilde{W} \end{pmatrix}\begin{pmatrix} I_l & 0 & 0 \\ 0 & I_k & 0 \\ 0 & 0 & G' \end{pmatrix}\begin{pmatrix} I_l & 0 \\ 0 & \widetilde{U}^T \end{pmatrix}Q^T = \widehat{V}\begin{pmatrix} I_{k+l} & 0 \\ 0 & G' \end{pmatrix}\widehat{Q}^T, \tag{3.10}$$

where

$$\widehat{V} = V\begin{pmatrix} I_l & 0 \\ 0 & \widetilde{W} \end{pmatrix} \in O\mathbb{R}^{m\times m}, \qquad \widehat{Q} = Q\begin{pmatrix} I_l & 0 \\ 0 & \widetilde{U} \end{pmatrix} \in O\mathbb{R}^{m\times m}, \tag{3.11}$$

and $G' \in O\mathbb{R}^{(m-k-l)\times(m-k-l)}$ is arbitrary. □

The following theorem can be shown similarly.

Theorem 3.2. *Given A, $B \in \mathbb{R}^{n\times m}$ and C, $D \in \mathbb{R}^{m\times n}$, let the singular value decompositions of C and D be, respectively, as* (2.1). *Partition*

$$AW = \begin{pmatrix} A_1 & A_2 \end{pmatrix}, \qquad BU = \begin{pmatrix} B_1 & B_2 \end{pmatrix}, \tag{3.12}$$

where $A_1, B_1 \in \mathbb{R}^{n\times k}$, $A_2, B_2 \in \mathbb{R}^{n\times(m-k)}$. Assume the singular value decompositions of A_2 and B_2 are, respectively,

$$A_2 = \widetilde{U}\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}\widetilde{V}^T, \qquad B_2 = \widetilde{U}\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}\widetilde{Q}^T, \tag{3.13}$$

where $\widetilde{V}$, $\widetilde{Q} \in O\mathbb{R}^{(m-k)\times(m-k)}$, $\widetilde{U} \in O\mathbb{R}^{n\times n}$, $\Sigma \in \mathbb{R}^{l'\times l'}$ is diagonal, whose diagonal elements are nonzero singular values of A_2 or B_2. Then the system (1.1) *has orthogonal solutions if and only if*

$$D^TD = C^TC, \qquad A_1 = B_1, \qquad A_2A_2^T = B_2B_2^T. \tag{3.14}$$

In which case, the orthogonal solutions can be expressed as

$$X = \widehat{W}\begin{pmatrix} I_{k+l'} & 0 \\ 0 & H' \end{pmatrix}\widehat{U}^T, \tag{3.15}$$

where

$$\widehat{W} = W\begin{pmatrix} I_k & 0 \\ 0 & \widetilde{W} \end{pmatrix} \in O\mathbb{R}^{m\times m}, \qquad \widehat{U} = U\begin{pmatrix} I_k & 0 \\ 0 & \widetilde{U} \end{pmatrix} \in O\mathbb{R}^{m\times m}, \tag{3.16}$$

and $H' \in O\mathbb{R}^{(m-k-l')\times(m-k-l')}$ is arbitrary.

4. The Symmetric Orthogonal Solutions of the System (1.1)

We now present the symmetric orthogonal solutions of the system (1.1).

Theorem 4.1. *Given* $A, B \in \mathbb{R}^{n\times m}$, $C, D \in \mathbb{R}^{m\times n}$. *Let the symmetric orthogonal solutions of the matrix equation* $AX = B$ *be described as in Lemma 2.6. Partition*

$$\widetilde{Q}^T C = \begin{pmatrix} C_1' \\ C_2' \end{pmatrix}, \qquad \widetilde{V}^T D = \begin{pmatrix} D_1' \\ D_2' \end{pmatrix}, \tag{4.1}$$

where $C_1', D_1' \in \mathbb{R}^{(2l-r)\times n}$, $C_2', D_2' \in \mathbb{R}^{(m-2l+r)\times n}$. *Then the system* (1.1) *has symmetric orthogonal solutions if and only if*

$$AA^T = BB^T, \qquad AB^T = BA^T, \qquad C_1' = D_1', \qquad D_2'^T D_2' = C_2'^T C_2', \qquad D_2'^T C_2' = C_2'^T D_2'. \tag{4.2}$$

In which case, the solutions can be expressed as

$$X = \widehat{V}\begin{pmatrix} I & 0 \\ 0 & G'' \end{pmatrix}\widehat{Q}^T, \tag{4.3}$$

where

$$\widehat{V} = V\begin{pmatrix} I_{2l-r} & 0 \\ 0 & \widetilde{W} \end{pmatrix} \in O\mathbb{R}^{m\times m}, \qquad \widehat{Q} = Q\begin{pmatrix} I_{2l-r} & 0 \\ 0 & \widetilde{U} \end{pmatrix} \in O\mathbb{R}^{m\times m}, \tag{4.4}$$

and $G'' \in SO\mathbb{R}^{(m-2l+r-2l'+r')\times(m-2l+r-2l'+r')}$ *is arbitrary.*

Proof. From Lemma 2.6, we obtain that the matrix equation $AX = B$ has symmetric orthogonal solutions if and only if $AA^T = BB^T$ and $AB^T = BA^T$. When these conditions are satisfied, the general symmetric orthogonal solutions can be expressed as

$$X = \widetilde{V}\begin{pmatrix} I_{2l-r} & 0 \\ 0 & G \end{pmatrix}\widetilde{Q}^T, \tag{4.5}$$

where $G \in SO\mathbb{R}^{(m-2l+r)\times(m-2l+r)}$ is arbitrary, $\widetilde{Q} \in O\mathbb{R}^{m\times m}$, $\widetilde{V} \in O\mathbb{R}^{m\times m}$. Inserting (4.1) and (4.5) into the matrix equation $XC = D$, we get $C_1' = D_1'$ and $GC_2' = D_2'$. By [1, Theorem 2], the matrix equation $GC_2' = D_2'$ has a symmetric orthogonal solution if and only if

$$D_2'^T D_2' = C_2'^T C_2', \qquad D_2'^T C_2' = C_2'^T D_2'. \tag{4.6}$$

In which case, the solutions can be described as

$$G = \widetilde{W}\begin{pmatrix} I_{2l'-r'} & 0 \\ 0 & G'' \end{pmatrix}\widetilde{U}^T, \tag{4.7}$$

where $G'' \in SO\mathbb{R}^{(m-2l+r-2l'+r')\times(m-2l+r-2l'+r')}$ is arbitrary, $\widetilde{W} \in O\mathbb{R}^{(m-2l+r)\times(m-2l+r)}$, and $\tilde{U} \in O\mathbb{R}^{(m-2l+r)\times(m-2l+r)}$. Hence the system (1.1) has symmetric orthogonal solutions if and only if all equalities in (4.2) hold. In which case, the solutions can be expressed as

$$X = \tilde{V}\begin{pmatrix} I_{2l-r} & 0 \\ 0 & G \end{pmatrix}\tilde{Q}^T = \tilde{V}\begin{pmatrix} I_{2l-r} & 0 \\ 0 & \widetilde{W} \end{pmatrix}\begin{pmatrix} I_{2l-r} & 0 \\ 0 & \begin{pmatrix} I & 0 \\ 0 & G'' \end{pmatrix} \end{pmatrix}\begin{pmatrix} I_{2l-r} & 0 \\ 0 & \tilde{U}^T \end{pmatrix}\tilde{Q}^T, \tag{4.8}$$

that is, the expression in (4.3). □

The following theorem can also be obtained by the method used in the proof of Theorem 4.1.

Theorem 4.2. *Given $A, B \in \mathbb{R}^{n\times m}$, $C, D \in \mathbb{R}^{m\times n}$. Let the symmetric orthogonal solutions of the matrix equation $XC = D$ be described as*

$$X = \widetilde{M}\begin{pmatrix} I_{2k-r} & 0 \\ 0 & G \end{pmatrix}\widetilde{N}^T, \tag{4.9}$$

where $\widetilde{M}, \widetilde{N} \in O\mathbb{R}^{m\times m}$, $G \in SO\mathbb{R}^{(m-2k+r)\times(m-2k+r)}$. Partition

$$A\widetilde{M} = \begin{pmatrix} M_1 & M_2 \end{pmatrix}, \quad B\widetilde{N} = \begin{pmatrix} N_1 & N_2 \end{pmatrix}, \quad M_1, N_1 \in \mathbb{R}^{n\times(2k-r)}, M_2, N_2 \in \mathbb{R}^{n\times(m-2k+r)}. \tag{4.10}$$

Then the system (1.1) *has symmetric orthogonal solutions if and only if*

$$D^TD = C^TC, \qquad D^TC = C^TD, \qquad M_1 = N_1, \qquad M_2M_2^T = N_2N_2^T, \qquad M_2N_2^T = N_2M_2^T. \tag{4.11}$$

In which case, the solutions can be expressed as

$$X = \widehat{M}\begin{pmatrix} I & 0 \\ 0 & H'' \end{pmatrix}\widehat{N}^T, \tag{4.12}$$

where

$$\widehat{M} = \widetilde{M}\begin{pmatrix} I_{2k-r} & 0 \\ 0 & \widetilde{W}_1 \end{pmatrix} \in O\mathbb{R}^{m\times m}, \qquad \widehat{N} = \widetilde{N}\begin{pmatrix} I_{2k-r} & 0 \\ 0 & \tilde{U}_1 \end{pmatrix} \in O\mathbb{R}^{m\times m}, \tag{4.13}$$

and $H'' \in SO\mathbb{R}^{(m-2k+r-2k'+r')\times(m-2k+r-2k'+r')}$ is arbitrary.

5. The Skew-Symmetric Orthogonal Solutions of the System (1.1)

In this section, we show the skew-symmetric orthogonal solutions of the system (1.1).

Theorem 5.1. *Given $A, B \in \mathbb{R}^{n\times 2m}$, $C, D \in \mathbb{R}^{2m\times n}$. Suppose the matrix equation $AX = B$ has skew-symmetric orthogonal solutions with the form*

$$X = \widetilde{V}_1 \begin{pmatrix} I & 0 \\ 0 & H \end{pmatrix} \widetilde{Q}_1^T, \tag{5.1}$$

where $H \in SSO\mathbb{R}^{2r\times 2r}$ is arbitrary, $\widetilde{V}_1$, $\widetilde{Q}_1 \in O\mathbb{R}^{2m\times 2m}$. Partition

$$\widetilde{Q}_1^T C = \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}, \qquad \widetilde{V}_1^T D = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}, \tag{5.2}$$

where $Q_1, V_1 \in \mathbb{R}^{(2m-2r)\times n}$, $Q_2, V_2 \in \mathbb{R}^{2r\times n}$. Then the system (1.1) has skew-symmetric orthogonal solutions if and only if

$$AA^T = BB^T, \qquad AB^T = -BA^T, \qquad Q_1 = V_1, \qquad Q_2^T Q_2 = V_2^T V_2, \qquad Q_2^T V_2 = -V_2^T Q_2. \tag{5.3}$$

In which case, the solutions can be expressed as

$$X = \widehat{V}_1 \begin{pmatrix} I & 0 \\ 0 & J' \end{pmatrix} \widehat{Q}_1^T, \tag{5.4}$$

where

$$\widehat{V} = \widetilde{V}_1 \begin{pmatrix} I_{2m-2r} & 0 \\ 0 & \widetilde{W} \end{pmatrix} \in O\mathbb{R}^{2m\times 2m}, \qquad \widehat{Q} = \widetilde{Q}_1 \begin{pmatrix} I_{2m-2r} & 0 \\ 0 & \widetilde{U} \end{pmatrix} \in O\mathbb{R}^{2m\times 2m}, \tag{5.5}$$

and $J' \in SSO\mathbb{R}^{2k'\times 2k'}$ is arbitrary.

Proof. By [2, Theorem 2], the matrix equation $AX = B$ has the skew-symmetric orthogonal solutions if and only if $AA^T = BB^T$ and $AB^T = -BA^T$. When these conditions are satisfied, the general skew-symmetric orthogonal solutions can be expressed as (5.1). Substituting (5.1) and (5.2) into the matrix equation $XC = D$, we get $Q_1 = V_1$ and $HQ_2 = V_2$. From Lemma 2.7, equation $HQ_2 = V_2$ has a skew-symmetric orthogonal solution H if and only if

$$Q_2^T Q_2 = V_2^T V_2, \qquad Q_2^T V_2 = -V_2^T Q_2. \tag{5.6}$$

When these conditions are satisfied, the solution can be described as

$$H = \widetilde{W} \begin{pmatrix} I & 0 \\ 0 & J' \end{pmatrix} \widetilde{U}^T, \tag{5.7}$$

where $J' \in SSO\mathbb{R}^{2k' \times 2k'}$ is arbitrary, $\widetilde{W}$, $\tilde{U} \in O\mathbb{R}^{2r \times 2r}$. Inserting (5.7) into (5.1) yields that the system (1.1) has skew-symmetric orthogonal solutions if and only if all equalities in (5.3) hold. In which case, the solutions can be expressed as (5.4). □

Similarly, the following theorem holds.

Theorem 5.2. *Given $A, B \in \mathbb{R}^{n \times 2m}$, $C, D \in \mathbb{R}^{2m \times n}$. Suppose the matrix equation $XC = D$ has skew-symmetric orthogonal solutions with the form*

$$X = \widetilde{W} \begin{pmatrix} I & 0 \\ 0 & K \end{pmatrix} \tilde{U}^T, \tag{5.8}$$

where $K \in SSO\mathbb{R}^{2p \times 2p}$ is arbitrary, $\widetilde{W}$, $\tilde{U} \in O\mathbb{R}^{2m \times 2m}$. Partition

$$A\widetilde{W} = (W_1 \ \ W_2), \qquad B\tilde{U} = (U_1 \ \ U_2), \tag{5.9}$$

where $W_1, U_1 \in \mathbb{R}^{n \times (2m-2p)}$, $W_2, U_2 \in \mathbb{R}^{n \times 2p}$. Then the system (1.1) *has skew-symmetric orthogonal solutions if and only if*

$$D^T D = C^T C, \qquad D^T C = -C^T D, \qquad W_1 = U_1, \qquad W_2 W_2^T = U_2 U_2^T, \qquad U_2 W_2^T = -W_2 U_2^T. \tag{5.10}$$

In which case, the solutions can be expressed as

$$X = \widehat{W}_1 \begin{pmatrix} I & 0 \\ 0 & J'' \end{pmatrix} \hat{U}_1^T, \tag{5.11}$$

where

$$\widehat{W}_1 = \widetilde{W} \begin{pmatrix} I_{2m-2p} & 0 \\ 0 & \widetilde{W}_1 \end{pmatrix} \in O\mathbb{R}^{2m \times 2m}, \qquad \hat{U}_1 = \begin{pmatrix} I_{2m-2p} & 0 \\ 0 & \tilde{U}_1^T \end{pmatrix} \tilde{U} \in O\mathbb{R}^{2m \times 2m}, \tag{5.12}$$

and $J'' \in SSO\mathbb{R}^{2q \times 2q}$ is arbitrary.

6. The Least Squares (Skew-) Symmetric Orthogonal Solutions of the System (1.1)

If the solvability conditions of a system of matrix equations are not satisfied, it is natural to consider its least squares solution. In this section, we get the least squares (skew-) symmetric orthogonal solutions of the system (1.1), that is, seek $X \in SSO\mathbb{R}^{2m \times 2m}(SO\mathbb{R}^{n \times n})$ such that

$$\min_{X \in SSO\mathbb{R}^{2m \times 2m}(SO\mathbb{R}^{n \times n})} \|AX - B\|^2 + \|XC - D\|^2. \tag{6.1}$$

With the help of the definition of the Frobenius norm and the properties of the skew-symmetric orthogonal matrix, we get that

$$\|AX - B\|^2 + \|XC - D\|^2 = \|A\|^2 + \|B\|^2 + \|C\|^2 + \|D\|^2 - 2\,\mathrm{tr}\left(X^T\left(A^TB + DC^T\right)\right). \tag{6.2}$$

Let

$$\begin{aligned} A^TB + DC^T &= K, \\ \frac{K^T - K}{2} &= T. \end{aligned} \tag{6.3}$$

Then, it follows from the skew-symmetric matrix X that

$$\mathrm{tr}\left(X^T\left(A^TB + DC^T\right)\right) = \mathrm{tr}\left(X^TK\right) = \mathrm{tr}(-XK) = \mathrm{tr}\left(X\left(\frac{K^T - K}{2}\right)\right) = \mathrm{tr}(XT). \tag{6.4}$$

Therefore, (6.1) holds if and only if (6.4) reaches its maximum. Now, we pay our attention to find the maximum value of (6.4). Assume the eigenvalue decomposition of T is

$$T = E\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}E^T \tag{6.5}$$

with

$$\Lambda = \mathrm{diag}(\Lambda_1, \ldots, \Lambda_l), \qquad \Lambda_i = \begin{pmatrix} 0 & \alpha_i \\ -\alpha_i & 0 \end{pmatrix}, \quad \alpha_i > 0, \ i = 1, \ldots, l; \qquad 2l = \mathrm{rank}(T). \tag{6.6}$$

Denote

$$E^TXE = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}, \tag{6.7}$$

partitioned according to

$$\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}, \tag{6.8}$$

then (6.4) has the following form:

$$\mathrm{tr}(XT) = \mathrm{tr}\left(E^TXE\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}\right) = \mathrm{tr}(X_{11}\Lambda). \tag{6.9}$$

Thus, by

$$X_{11} = \widetilde{I} = \operatorname{diag}\left(\widetilde{I}_1, \ldots, \widetilde{I}_l\right), \tag{6.10}$$

where

$$\widetilde{I}_i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad i = 1, \ldots, l. \tag{6.11}$$

Equation (6.9) gets its maximum. Since $E^T X E$ is skew-symmetric, it follows from

$$X = E \begin{pmatrix} \widetilde{I} & 0 \\ 0 & G \end{pmatrix} E^T, \tag{6.12}$$

where $G \in SSO\mathbb{R}^{(2m-2l)\times(2m-2l)}$ is arbitrary, that (6.1) obtains its minimum. Hence we have the following theorem.

Theorem 6.1. *Given $A, B \in \mathbb{R}^{n\times 2m}$ and $C, D \in \mathbb{R}^{2m\times n}$, denote*

$$A^T B + DC^T = K, \qquad \frac{K^T - K}{2} = T, \tag{6.13}$$

and let the spectral decomposition of T be (6.5)*. Then the least squares skew-symmetric orthogonal solutions of the system* (1.1) *can be expressed as* (6.12)*.*

If X in (6.1) is a symmetric orthogonal matrix, then by the definition of the Frobenius norm and the properties of the symmetric orthogonal matrix, (6.2) holds. Let

$$A^T B + DC^T = H, \qquad \frac{H^T + H}{2} = N. \tag{6.14}$$

Then we get that

$$\min_{X\in SO\mathbb{R}^{n\times n}} \|AX - B\|^2 + \|XC - D\|^2 = \min_{X\in SO\mathbb{R}^{n\times n}} \left[\|A\|^2 + \|B\|^2 + \|C\|^2 + \|D\|^2 - 2\operatorname{tr}(XN)\right]. \tag{6.15}$$

Thus (6.15) reaches its minimum if and only if $\operatorname{tr}(XN)$ obtains its maximum. Now, we focus on finding the maximum value of $\operatorname{tr}(XN)$. Let the spectral decomposition of the symmetric matrix N be

$$N = M \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} M^T, \tag{6.16}$$

where

$$\Sigma = \begin{pmatrix} \Sigma^+ & 0 \\ 0 & \Sigma^- \end{pmatrix}, \quad \Sigma^+ = \operatorname{diag}(\lambda_1, \ldots, \lambda_s),$$
$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s > 0, \quad \Sigma^- = \operatorname{diag}(\lambda_{s+1}, \ldots, \lambda_t), \tag{6.17}$$
$$\lambda_t \leq \lambda_{t-1} \leq \cdots \leq \lambda_{s+1} < 0, \quad t = \operatorname{rank}(N).$$

Denote

$$M^T X M = \begin{pmatrix} \overline{X_{11}} & \overline{X_{12}} \\ \overline{X_{21}} & \overline{X_{22}} \end{pmatrix} \tag{6.18}$$

being compatible with

$$\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}. \tag{6.19}$$

Then

$$\operatorname{tr}(XN) = \operatorname{tr}\left(M^T X M \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}\right) = \operatorname{tr}\left(\begin{pmatrix} \overline{X_{11}} & \overline{X_{12}} \\ \overline{X_{21}} & \overline{X_{22}} \end{pmatrix}\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}\right) = \operatorname{tr}\left(\overline{X_{11}}\Sigma\right). \tag{6.20}$$

Therefore, it follows from

$$\overline{X_{11}} = \widehat{I} = \begin{pmatrix} I_s & 0 \\ 0 & -I_{t-s} \end{pmatrix} \tag{6.21}$$

that (6.20) reaches its maximum. Since $M^T X M$ is a symmetric orthogonal matrix, then when X has the form

$$X = M \begin{pmatrix} \widehat{I} & 0 \\ 0 & L \end{pmatrix} M^T, \tag{6.22}$$

where $L \in SO\mathbb{R}^{(n-t)\times(n-t)}$ is arbitrary, (6.15) gets its minimum. Thus we obtain the following theorem.

Theorem 6.2. *Given $A, B \in \mathbb{R}^{n\times m}$, $C, D \in \mathbb{R}^{m\times n}$, denote*

$$A^T B + DC^T = H, \qquad \frac{H^T + H}{2} = N, \tag{6.23}$$

and let the eigenvalue decomposition of N be (6.16). *Then the least squares symmetric orthogonal solutions of the system* (1.1) *can be described as* (6.22).

Algorithm 6.3. Consider the following.

Step 1. Input $A, B \in \mathbb{R}^{n\times m}$ and $C, D \in \mathbb{R}^{m\times n}$.

Step 2. Compute

$$A^T B + DC^T = H,$$
$$N = \frac{H^T + H}{2}. \tag{6.24}$$

Step 3. Compute the spectral decomposition of N with the form (6.16).

Step 4. Compute the least squares symmetric orthogonal solutions of (1.1) according to (6.22).

Example 6.4. Assume

$$A = \begin{pmatrix} 12.2 & 8.4 & -5.6 & 6.3 & 9.4 & 10.7 \\ 11.8 & 2.9 & 8.5 & 6.9 & 9.6 & -7.8 \\ 10.6 & 2.3 & 11.5 & 7.8 & 6.7 & 8.9 \\ 3.6 & 7.8 & 4.9 & 11.9 & 9.4 & 5.9 \\ 4.5 & 6.7 & 7.8 & 3.1 & 5.6 & 11.6 \end{pmatrix}, \quad B = \begin{pmatrix} 8.5 & 9.4 & 3.6 & 7.8 & 6.3 & 4.7 \\ 2.7 & 3.6 & 7.9 & 9.4 & 5.6 & 7.8 \\ 3.7 & 6.7 & 8.6 & 9.8 & 3.4 & 2.9 \\ -4.3 & 6.2 & 5.7 & 7.4 & 5.4 & 9.5 \\ 2.9 & 3.9 & -5.2 & 6.3 & 7.8 & 4.6 \end{pmatrix},$$
$$C = \begin{pmatrix} 5.4 & 6.4 & 3.7 & 5.6 & 9.7 \\ 3.6 & 4.2 & 7.8 & -6.3 & 7.8 \\ 6.7 & 3.5 & -4.6 & 2.9 & 2.8 \\ -2.7 & 7.2 & 10.8 & 3.7 & 3.8 \\ 1.9 & 3.9 & 8.2 & 5.6 & 11.2 \\ 8.9 & 7.8 & 9.4 & 7.9 & 5.6 \end{pmatrix}, \quad D = \begin{pmatrix} 7.9 & 9.5 & 5.4 & 2.8 & 8.6 \\ 8.7 & 2.6 & 6.7 & 8.4 & 8.1 \\ 5.7 & 3.9 & -2.9 & 5.2 & 1.9 \\ 4.8 & 5.8 & 1.8 & -7.2 & 5.8 \\ 2.8 & 7.9 & 4.5 & 6.7 & 9.6 \\ 9.5 & 4.1 & 3.4 & 9.8 & 3.9 \end{pmatrix}. \tag{6.25}$$

It can be verified that the given matrices A, B, C, and D do not satisfy the solvability conditions in Theorem 4.1 or Theorem 4.2. So we intend to derive the least squares symmetric orthogonal solutions of the system (1.1). By Algorithm 6.3, we have the following results:

(1) the least squares symmetric orthogonal solution

$$X = \begin{pmatrix} 0.01200 & 0.06621 & -0.24978 & 0.27047 & 0.91302 & -0.16218 \\ 0.06621 & 0.65601 & 0.22702 & -0.55769 & 0.24587 & 0.37715 \\ -0.24978 & 0.22702 & 0.80661 & 0.40254 & 0.04066 & -0.26784 \\ 0.27047 & -0.55769 & 0.40254 & 0.06853 & 0.23799 & 0.62645 \\ 0.91302 & 0.24587 & 0.04066 & 0.23799 & -0.12142 & -0.18135 \\ -0.16218 & 0.37715 & -0.26784 & 0.62645 & -0.181354 & 0.57825 \end{pmatrix}, \tag{6.26}$$

(2)

$$\min_{X\in SO\mathbb{R}^{m\times m}} \|AX - B\|^2 + \|XC - D\|^2 = 1.98366,$$
$$\left\| X^T X - I_6 \right\| = 2.84882 \times 10^{-15}, \qquad \left\| X^T - X \right\| = 0.00000. \tag{6.27}$$

Remark 6.5. (1) There exists a unique symmetric orthogonal solution such that (6.1) holds if and only if the matrix

$$N = \frac{H^T + H}{2}, \tag{6.28}$$

where

$$H = A^T B + DC^T, \tag{6.29}$$

is invertible. Example 6.4 just illustrates it.

(2) The algorithm about computing the least squares skew-symmetric orthogonal solutions of the system (1.1) can be shown similarly; we omit it here.

7. Conclusions

This paper is devoted to giving the solvability conditions of and the expressions of the orthogonal solutions, the symmetric orthogonal solutions, and the skew-symmetric orthogonal solutions to the system (1.1), respectively, and meanwhile obtaining the least squares symmetric orthogonal and skew-symmetric orthogonal solutions of the system (1.1). In addition, an algorithm and an example have been provided to compute its least squares symmetric orthogonal solutions.

Acknowledgments

This research was supported by the grants from Innovation Program of Shanghai Municipal Education Commission (13ZZ080), the National Natural Science Foundation of China (11171205), the Natural Science Foundation of Shanghai (11ZR1412500), the Discipline Project at the corresponding level of Shanghai (A. 13-0101-12-005), and Shanghai Leading Academic Discipline Project (J50101).

References

[1] C. J. Meng and X. Y. Hu, "The inverse problem of symmetric orthogonal matrices and its optimal approximation," *Mathematica Numerica Sinica*, vol. 28, pp. 269–280, 2006 (Chinese).

[2] C. J. Meng, X. Y. Hu, and L. Zhang, "The skew-symmetric orthogonal solutions of the matrix equation $AX = B$," *Linear Algebra and Its Applications*, vol. 402, no. 1–3, pp. 303–318, 2005.

[3] D. L. Chu, H. C. Chan, and D. W. C. Ho, "Regularization of singular systems by derivative and proportional output feedback," *SIAM Journal on Matrix Analysis and Applications*, vol. 19, no. 1, pp. 21–38, 1998.

[4] K. T. Joseph, "Inverse eigenvalue problem in structural design," *AIAA Journal*, vol. 30, no. 12, pp. 2890–2896, 1992.

[5] A. Jameson and E. Kreinder, "Inverse problem of linear optimal control," *SIAM Journal on Control*, vol. 11, no. 1, pp. 1–19, 1973.

[6] Q. W. Wang, "Bisymmetric and centrosymmetric solutions to systems of real quaternion matrix equations," *Computers and Mathematics with Applications*, vol. 49, no. 5-6, pp. 641–650, 2005.

[7] Q. W. Wang, X. Liu, and S. W. Yu, "The common bisymmetric nonnegative definite solutions with extreme ranks and inertias to a pair of matrix equations," *Applied Mathematics and Computation*, vol. 218, pp. 2761–2771, 2011.

[8] C. G. Khatri and S. K. Mitra, "Hermitian and nonnegative definite solutions of linear matrix equations," *SIAM Journal on Applied Mathematics*, vol. 31, pp. 578–585, 1976.

[9] Q. Zhang and Q. W. Wang, "The (P, Q)-(skew)symmetric extremal rank solutions to a system of quaternion matrix equations," *Applied Mathematics and Computation*, vol. 217, no. 22, pp. 9286–9296, 2011.

[10] F. L. Li, X. Y. Hu, and L. Zhang, "Least-squares mirrorsymmetric solution for matrix equations $(AX = B, XC = D)$," *Numerical Mathematics*, vol. 15, pp. 217–226, 2006.

[11] Y. X. Yuan, "Least-squares solutions to the matrix equations $AX = B$ and $XC = D$," *Applied Mathematics and Computation*, vol. 216, no. 10, pp. 3120–3125, 2010.

[12] A. Dajić and J. J. Koliha, "Positive solutions to the equations $AX = C$ and $XB = D$ for Hilbert space operators," *Journal of Mathematical Analysis and Applications*, vol. 333, no. 2, pp. 567–576, 2007.

[13] F. L. Li, X. Y. Hu, and L. Zhang, "The generalized reflexive solution for a class of matrix equations $(AX = B, XC = D)$," *Acta Mathematica Scientia*, vol. 28, no. 1, pp. 185–193, 2008.

[14] S. Kumar Mitra, "The matrix equations $AX = C, XB = D$," *Linear Algebra and Its Applications*, vol. 59, pp. 171–181, 1984.

[15] Y. Qiu, Z. Zhang, and J. Lu, "The matrix equations $AX = B, XC = D$ with $PX = sXP$ constraint," *Applied Mathematics and Computation*, vol. 189, no. 2, pp. 1428–1434, 2007.

[16] Y. Y. Qiu and A. D. Wang, "Least squares solutions to the equations $AX = B, XC = B$ with some constraints," *Applied Mathematics and Computation*, vol. 204, no. 2, pp. 872–880, 2008.

[17] Q. W. Wang, X. Zhang, and Z. H. He, "On the Hermitian structures of the solution to a pair of matrix equations," *Linear Multilinear Algebra*, vol. 61, no. 1, pp. 73–90, 2013.

[18] Q. X. Xu, "Common Hermitian and positive solutions to the adjointable operator equations $AX = C, XB = D$," *Linear Algebra and Its Applications*, vol. 429, no. 1, pp. 1–11, 2008.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2012, Article ID 398085, 14 pages
doi:10.1155/2012/398085

Research Article

On the Hermitian R-Conjugate Solution of a System of Matrix Equations

Chang-Zhou Dong,[1] Qing-Wen Wang,[2] and Yu-Ping Zhang[3]

[1] *School of Mathematics and Science, Shijiazhuang University of Economics, Shijiazhuang, Hebei 050031, China*
[2] *Department of Mathematics, Shanghai University, Shanghai, Shanghai 200444, China*
[3] *Department of Mathematics, Ordnance Engineering College, Shijiazhuang, Hebei 050003, China*

Correspondence should be addressed to Qing-Wen Wang, wqw858@yahoo.com.cn

Received 3 October 2012; Accepted 26 November 2012

Academic Editor: Yang Zhang

Let R be an n by n nontrivial real symmetric involution matrix, that is, $R = R^{-1} = R^T \neq I_n$. An $n \times n$ complex matrix A is termed R-conjugate if $\overline{A} = RAR$, where $\overline{A}$ denotes the conjugate of A. We give necessary and sufficient conditions for the existence of the Hermitian R-conjugate solution to the system of complex matrix equations $AX = C$ and $XB = D$ and present an expression of the Hermitian R-conjugate solution to this system when the solvability conditions are satisfied. In addition, the solution to an optimal approximation problem is obtained. Furthermore, the least squares Hermitian R-conjugate solution with the least norm to this system mentioned above is considered. The representation of such solution is also derived. Finally, an algorithm and numerical examples are given.

1. Introduction

Throughout, we denote the complex $m \times n$ matrix space by $\mathbb{C}^{m\times n}$, the real $m \times n$ matrix space by $\mathbb{R}^{m\times n}$, and the set of all matrices in $\mathbb{R}^{m\times n}$ with rank r by $\mathbb{R}_r^{m\times n}$. The symbols I, $\overline{A}$, A^T, A^*, $A^\dagger$, and $\|A\|$ stand for the identity matrix with the appropriate size, the conjugate, the transpose, the conjugate transpose, the Moore-Penrose generalized inverse, and the Frobenius norm of $A \in \mathbb{C}^{m\times n}$, respectively. We use V_n to denote the $n \times n$ backward matrix having the elements 1 along the southwest diagonal and with the remaining elements being zeros.

Recall that an $n \times n$ complex matrix A is centrohermitian if $\overline{A} = V_n A V_n$. Centrohermitian matrices and related matrices, such as k-Hermitian matrices, Hermitian Toeplitz matrices, and generalized centrohermitian matrices, appear in digital signal processing and others areas (see, [1–4]). As a generalization of a centrohermitian matrix and related matrices, Trench [5] gave the definition of R-conjugate matrix. A matrix $A \in \mathbb{C}^{n\times n}$ is R-conjugate if $\overline{A} = RAR$, where R is a nontrivial real symmetric involution matrix, that is, $R = R^{-1} = R^T$ and $R \neq I_n$. At the same time, Trench studied the linear equation $Az = w$ for R-conjugate matrices in [5], where z, w are known column vectors.

Investigating the matrix equation

$$AX = B \tag{1.1}$$

with the unknown matrix X being symmetric, reflexive, Hermitian-generalized Hamiltonian, and repositive definite is a very active research topic [6–14]. As a generalization of (1.1), the classical system of matrix equations

$$AX = C, \qquad XB = D \tag{1.2}$$

has attracted many author's attention. For instance, [15] gave the necessary and sufficient conditions for the consistency of (1.2), [16, 17] derived an expression for the general solution by using singular value decomposition of a matrix and generalized inverses of matrices, respectively. Moreover, many results have been obtained about the system (1.2) with various constraints, such as bisymmetric, Hermitian, positive semidefinite, reflexive, and generalized reflexive solutions (see, [18–28]). To our knowledge, so far there has been little investigation of the Hermitian R-conjugate solution to (1.2).

Motivated by the work mentioned above, we investigate Hermitian R-conjugate solutions to (1.2). We also consider the optimal approximation problem

$$\left\|\widehat{X} - E\right\| = \min_{X\in S_X}\|X - E\|, \tag{1.3}$$

where E is a given matrix in $\mathbb{C}^{n\times n}$ and S_X the set of all Hermitian R-conjugate solutions to (1.2). In many cases the system (1.2) has not Hermitian R-conjugate solution. Hence, we need to further study its least squares solution, which can be described as follows: Let $RH\mathbb{C}^{n\times n}$ denote the set of all Hermitian R-conjugate matrices in $\mathbb{C}^{n\times n}$:

$$S_L = \left\{X \mid \min_{X\in RH\mathbb{C}^{n\times n}}\left(\|AX - C\|^2 + \|XB - D\|^2\right)\right\}. \tag{1.4}$$

Find $\widetilde{X} \in \mathbb{C}^{n\times n}$ such that

$$\left\|\widetilde{X}\right\| = \min_{X\in S_L}\|X\|. \tag{1.5}$$

In Section 2, we present necessary and sufficient conditions for the existence of the Hermitian R-conjugate solution to (1.2) and give an expression of this solution when the solvability conditions are met. In Section 3, we derive an optimal approximation solution to (1.3). In Section 4, we provide the least squares Hermitian R-conjugate solution to (1.5). In Section 5, we give an algorithm and a numerical example to illustrate our results.

2. *R*-Conjugate Hermitian Solution to (1.2)

In this section, we establish the solvability conditions and the general expression for the Hermitian R-conjugate solution to (1.2).

We denote $R\mathbb{C}^{n\times n}$ and $RH\mathbb{C}^{n\times n}$ the set of all R-conjugate matrices and Hermitian R-conjugate matrices, respectively, that is,

$$\begin{aligned} R\mathbb{C}^{n\times n} &= \left\{A \mid \overline{A} = RAR\right\}, \\ HR\mathbb{C}^{n\times n} &= \left\{A \mid \overline{A} = RAR, A = A^*\right\}, \end{aligned} \tag{2.1}$$

where R is $n\times n$ nontrivial real symmetric involution matrix.

Chang et al. in [29] mentioned that for nontrivial symmetric involution matrix $R \in \mathbb{R}^{n\times n}$, there exist positive integer r and $n\times n$ real orthogonal matrix $[P,\ Q]$ such that

$$R = \begin{bmatrix} P, & Q \end{bmatrix} \begin{bmatrix} I_r & 0 \\ 0 & -I_{n-r} \end{bmatrix} \begin{bmatrix} P^T \\ Q^T \end{bmatrix}, \tag{2.2}$$

where $P \in \mathbb{R}^{n\times r}$, $Q \in \mathbb{R}^{n\times(n-r)}$. By (2.2),

$$RP = P, \qquad RQ = -Q, \qquad P^TP = I_r, \qquad Q^TQ = I_{n-r}, \qquad P^TQ = 0, \qquad Q^TP = 0. \tag{2.3}$$

Throughout this paper, we always assume that the nontrivial symmetric involution matrix R is fixed which is given by (2.2) and (2.3). Now, we give a criterion of judging a matrix is R-conjugate Hermitian matrix.

Theorem 2.1. *A matrix $K \in HR\mathbb{C}^{n\times n}$ if and only if there exists a symmetric matrix $H \in \mathbb{R}^{n\times n}$ such that $K = \Gamma H \Gamma^*$, where*

$$\Gamma = \begin{bmatrix} P, & iQ \end{bmatrix}, \tag{2.4}$$

with P, Q being the same as (2.2).

Proof. If $K \in HR\mathbb{C}^{n\times n}$, then $\overline{K} = RKR$. By (2.2),

$$\overline{K} = RKR = \begin{bmatrix} P, & Q \end{bmatrix} \begin{bmatrix} I_r & 0 \\ 0 & -I_{n-r} \end{bmatrix} \begin{bmatrix} P^T \\ Q^T \end{bmatrix} K \begin{bmatrix} P, & Q \end{bmatrix} \begin{bmatrix} I_r & 0 \\ 0 & -I_{n-r} \end{bmatrix} \begin{bmatrix} P^T \\ Q^T \end{bmatrix}, \tag{2.5}$$

which is equivalent to

$$\begin{bmatrix} P^T \\ Q^T \end{bmatrix} \overline{K} [P,\ Q] = \begin{bmatrix} I_r & 0 \\ 0 & -I_{n-r} \end{bmatrix} \begin{bmatrix} P^T \\ Q^T \end{bmatrix} K [P,\ Q] \begin{bmatrix} I_r & 0 \\ 0 & -I_{n-r} \end{bmatrix}. \tag{2.6}$$

Suppose that

$$\begin{bmatrix} P^T \\ Q^T \end{bmatrix} K [P,\ Q] = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}. \tag{2.7}$$

Substituting (2.7) into (2.6), we obtain

$$\begin{bmatrix} \overline{K_{11}} & \overline{K_{12}} \\ \overline{K_{21}} & \overline{K_{22}} \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & -I_{n-r} \end{bmatrix} \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} I_r & 0 \\ 0 & -I_{n-r} \end{bmatrix} = \begin{bmatrix} K_{11} & -K_{12} \\ -K_{21} & K_{22} \end{bmatrix}. \tag{2.8}$$

Hence, $\overline{K_{11}} = K_{11}$, $\overline{K_{12}} = -K_{12}$, $\overline{K_{21}} = -K_{21}$, $\overline{K_{22}} = K_{22}$, that is, K_{11}, iK_{12}, iK_{21}, K_{22} are real matrices. If we denote $M = iK_{12}$, $N = -iK_{21}$, then by (2.7)

$$K = [P,\ Q] \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} P^T \\ Q^T \end{bmatrix} = [P,\ iQ] \begin{bmatrix} K_{11} & M \\ N & K_{22} \end{bmatrix} \begin{bmatrix} P^T \\ -iQ^T \end{bmatrix}. \tag{2.9}$$

Let $\Gamma = [P, iQ]$, and

$$H = \begin{bmatrix} K_{11} & M \\ N & K_{22} \end{bmatrix}. \tag{2.10}$$

Then, K can be expressed as $\Gamma H \Gamma^*$, where Γ is unitary matrix and H is a real matrix. By $K = K^*$

$$\Gamma H^T \Gamma^* = K^* = K = \Gamma H \Gamma^*, \tag{2.11}$$

we obtain $H = H^T$.

Conversely, if there exists a symmetric matrix $H \in \mathbb{R}^{n\times n}$ such that $K = \Gamma H \Gamma^*$, then it follows from (2.3) that

$$RKR = R\Gamma H\Gamma^* R = R\left[P,\ iQ\right] H \begin{bmatrix} P^T \\ -iQ^T \end{bmatrix} R = \left[P,\ -iQ\right] H \begin{bmatrix} P^T \\ iQ^T \end{bmatrix} = \overline{\Gamma} H \overline{\Gamma^*} = \overline{K},$$
$$K^* = \Gamma H^T \Gamma^* = \Gamma H \Gamma^* = K, \tag{2.12}$$

that is, $K \in HR\mathbb{C}^{n\times n}$. □

Theorem 2.1 implies that an arbitrary complex Hermitian R-conjugate matrix is equivalent to a real symmetric matrix.

Lemma 2.2. *For any matrix $A \in \mathbb{C}^{m\times n}$, $A = A_1 + iA_2$, where*

$$A_1 = \frac{A + \overline{A}}{2}, \qquad A_2 = \frac{A - \overline{A}}{2i}. \tag{2.13}$$

Proof. For any matrix $A \in \mathbb{C}^{m\times n}$, it is obvious that $A = A_1 + iA_2$, where A_1, A_2 are defined as (2.13). Now, we prove that the decomposition $A = A_1 + iA_2$ is unique. If there exist B_1, B_2 such that $A = B_1 + iB_2$, then

$$A_1 - B_1 + i(B_2 - A_2) = 0. \tag{2.14}$$

It follows from A_1, A_2, B_1, and B_2 are real matrix that

$$A_1 = B_1, \qquad A_2 = B_2. \tag{2.15}$$

Hence, $A = A_1 + iA_2$ holds, where A_1, A_2 are defined as (2.13). □

By Theorem 2.1, for $X \in HR\mathbb{C}^{n\times n}$, we may assume that

$$X = \Gamma Y \Gamma^*, \tag{2.16}$$

where Γ is defined as (2.4) and $Y \in \mathbb{R}^{n\times n}$ is a symmetric matrix.

Suppose that $A\Gamma = A_1 + iA_2 \in \mathbb{C}^{m\times n}$, $C\Gamma = C_1 + iC_2 \in \mathbb{C}^{m\times n}$, $\Gamma^* B = B_1 + iB_2 \in \mathbb{C}^{n\times l}$, and $\Gamma^* D = D_1 + iD_2 \in \mathbb{C}^{n\times l}$, where

$$A_1 = \frac{A\Gamma + \overline{A\Gamma}}{2}, \qquad A_2 = \frac{A\Gamma - \overline{A\Gamma}}{2i}, \qquad C_1 = \frac{C\Gamma + \overline{C\Gamma}}{2}, \qquad C_2 = \frac{C\Gamma - \overline{C\Gamma}}{2i},$$
$$B_1 = \frac{\Gamma^* B + \overline{\Gamma^* B}}{2}, \qquad B_2 = \frac{\Gamma^* B - \overline{\Gamma^* B}}{2i}, \qquad D_1 = \frac{\Gamma^* D + \overline{\Gamma^* D}}{2}, \qquad D_2 = \frac{\Gamma^* D - \overline{\Gamma^* D}}{2i}. \tag{2.17}$$

Then, system (1.2) can be reduced into

$$(A_1 + iA_2)Y = C_1 + iC_2, \qquad Y(B_1 + iB_2) = D_1 + iD_2, \tag{2.18}$$

which implies that

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} Y = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, \qquad Y\left[B_1,\ B_2\right] = \left[D_1,\ D_2\right]. \tag{2.19}$$

Let

$$\begin{gathered} F = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \qquad G = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, \qquad K = \left[B_1,\ B_2\right], \\ L = \left[D_1,\ D_2\right], \qquad M = \begin{bmatrix} F \\ K^T \end{bmatrix}, \qquad N = \begin{bmatrix} G \\ L^T \end{bmatrix}. \end{gathered} \tag{2.20}$$

Then, system (1.2) has a solution X in $HR\mathbb{C}^{n\times n}$ if and only if the real system

$$MY = N \tag{2.21}$$

has a symmetric solution Y in $\mathbb{R}^{n\times n}$.

Lemma 2.3 (Theorem 1 in [7]). *Let $A \in \mathbb{R}^{m\times n}$. The SVD of matrix A is as follows*

$$A = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^T, \tag{2.22}$$

where $U = [U_1, U_2] \in \mathbb{R}^{m\times m}$ and $V = [V_1, V_2] \in \mathbb{R}^{n\times n}$ are orthogonal matrices, $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$, $\sigma_i > 0$ $(i = 1, \ldots, r)$, $r = \operatorname{rank}(A)$, $U_1 \in \mathbb{R}^{m\times r}$, $V_1 \in \mathbb{R}^{n\times r}$. Then, (1.1) *has a symmetric solution if and only if*

$$AB^T = BA^T, \qquad U_2^T B = 0. \tag{2.23}$$

In that case, it has the general solution

$$X = V_1\Sigma^{-1}U_1^T B + V_2V_2^T B^T U_1 \Sigma^{-1} V_1^T + V_2 G V_2^T, \tag{2.24}$$

where G is an arbitrary $(n-r)\times(n-r)$ symmetric matrix.

By Lemma 2.3, we have the following theorem.

Theorem 2.4. *Given $A \in \mathbb{C}^{m\times n}$, $C \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times l}$, and $D \in \mathbb{C}^{n\times l}$. Let A_1, A_2, C_1, C_2, B_1, B_2, D_1, D_2, F, G, K, L, M, and N be defined in* (2.17), (2.20), *respectively. Assume that the SVD of $M \in \mathbb{R}^{(2m+2l)\times n}$ is as follows*

$$M = U\begin{bmatrix} M_1 & 0 \\ 0 & 0 \end{bmatrix}V^T, \tag{2.25}$$

where $U = [U_1, U_2] \in \mathbb{R}^{(2m+2l)\times(2m+2l)}$ and $V = [V_1, V_2] \in \mathbb{R}^{n\times n}$ are orthogonal matrices, $M_1 = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$, $\sigma_i > 0$ $(i = 1, \ldots, r)$, $r = \operatorname{rank}(M)$, $U_1 \in \mathbb{R}^{(2m+2l)\times r}$, $V_1 \in \mathbb{R}^{n\times r}$. Then, system (1.2) *has a solution in $HR\mathbb{C}^{n\times n}$ if and only if*

$$MN^T = NM^T, \qquad U_2^T N = 0. \tag{2.26}$$

In that case, it has the general solution

$$X = \Gamma\left(V_1 M_1^{-1} U_1^T N + V_2 V_2^T N^T U_1 M_1^{-1} V_1^T + V_2 G V_2^T\right)\Gamma^*, \tag{2.27}$$

where G is an arbitrary $(n-r)\times(n-r)$ symmetric matrix.

3. The Solution of Optimal Approximation Problem (1.3)

When the set S_X of all Hermitian R-conjugate solution to (1.2) is nonempty, it is easy to verify S_X is a closed set. Therefore, the optimal approximation problem (1.3) has a unique solution by [30].

Theorem 3.1. *Given $A \in \mathbb{C}^{m\times n}$, $C \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times l}$, $D \in \mathbb{C}^{n\times l}$, $E \in \mathbb{C}^{n\times n}$, and $E_1 = (1/2)(\Gamma^* E\Gamma + \overline{\Gamma^* E\Gamma})$. Assume S_X is nonempty, then the optimal approximation problem* (1.3) *has a unique solution $\widehat{X}$ and*

$$\widehat{X} = \Gamma\left(V_1 M_1^{-1} U_1^T N + V_2 V_2^T N^T U_1 M_1^{-1} V_1^T + V_2 V_2^T E_1 V_2 V_2^T\right)\Gamma^*. \tag{3.1}$$

Proof. Since S_X is nonempty, $X \in S_X$ has the form of (2.27). By Lemma 2.2, $\Gamma^* E\Gamma$ can be written as

$$\Gamma^* E\Gamma = E_1 + iE_2, \tag{3.2}$$

where

$$E_1 = \frac{1}{2}\left(\Gamma^* E\Gamma + \overline{\Gamma^* E\Gamma}\right), \qquad E_2 = \frac{1}{2i}\left(\Gamma^* E\Gamma - \overline{\Gamma^* E\Gamma}\right). \tag{3.3}$$

According to (3.2) and the unitary invariance of Frobenius norm

$$\begin{aligned}\|X - E\| &= \left\|\Gamma\left(V_1 M^{-1} U_1^T N + V_2 V_2^T N^T U_1 M^{-1} V_1^T + V_2 G V_2^T\right)\Gamma^* - E\right\| \\ &= \left\|\left(E_1 - V_1 M^{-1} U_1^T N - V_2 V_2^T N^T U_1 M^{-1} V_1^T - V_2 G V_2^T\right) + iE_2\right\|.\end{aligned} \tag{3.4}$$

We get

$$\|X - E\|^2 = \left\|E_1 - V_1 M^{-1} U_1^T N - V_2 V_2^T N^T U_1 M^{-1} V_1^T - V_2 G V_2^T\right\|^2 + \|E_2\|^2. \tag{3.5}$$

Then, $\min_{X \in S_X}\|X - E\|$ is consistent if and only if there exists $G \in \mathbb{R}^{(n-r)\times(n-r)}$ such that

$$\min\left\|E_1 - V_1 M^{-1} U_1^T N - V_2 V_2^T N^T U_1 M^{-1} V_1^T - V_2 G V_2^T\right\|. \tag{3.6}$$

For the orthogonal matrix V

$$\begin{aligned}&\left\|E_1 - V_1 M^{-1} U_1^T N - V_2 V_2^T N^T U_1 M^{-1} V_1^T - V_2 G V_2^T\right\|^2 \\ &\quad= \left\|V^T (E_1 - V_1 M^{-1} U_1^T N - V_2 V_2^T N^T U_1 M^{-1} V_1^T - V_2 G V_2^T) V\right\|^2 \\ &\quad= \left\|V_1^T\left(E_1 - V_1 M^{-1} U_1^T N\right)V_1\right\|^2 + \left\|V_1^T (E_1 - V_1 M^{-1} U_1^T N) V_2\right\|^2 \\ &\qquad+ \left\|V_2^T\left(E_1 - V_2 V_2^T N^T U_1 M^{-1} V_1^T\right)V_1\right\|^2 + \left\|V_2^T\left(E_1 - V_2 G V_2^T\right)V_2\right\|.\end{aligned} \tag{3.7}$$

Therefore,

$$\min\left\|E_1 - V_1 M^{-1} U_1^T N - V_2 V_2^T N^T U_1 M^{-1} V_1^T - V_2 G V_2^T\right\| \tag{3.8}$$

is equivalent to

$$G = V_2^T E_1 V_2. \tag{3.9}$$

Substituting (3.9) into (2.27), we obtain (3.1). □

4. The Solution of Problem (1.5)

In this section, we give the explicit expression of the solution to (1.5).

Theorem 4.1. *Given $A \in \mathbb{C}^{m\times n}$, $C \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times l}$, and $D \in \mathbb{C}^{n\times l}$. Let A_1, A_2, C_1, C_2, B_1, B_2, D_1, D_2, F, G, K, L, M, and N be defined in* (2.17), (2.20), *respectively. Assume that the*

SVD of $M \in \mathbb{R}^{(2m+2l)\times n}$ *is as* (2.25) *and system* (1.2) *has not a solution in* $HR\mathbb{C}^{n\times n}$. *Then,* $X \in S_L$ *can be expressed as*

$$X = \Gamma V \begin{bmatrix} M_1^{-1}U_1^T N V_1 & M_1^{-1}U_1^T N V_2 \\ V_2^T N^T U_1 M_1^{-1} & Y_{22} \end{bmatrix} V^T \Gamma^*, \tag{4.1}$$

where $Y_{22} \in \mathbb{R}^{(n-r)\times(n-r)}$ *is an arbitrary symmetric matrix.*

Proof. It yields from (2.17)–(2.21) and (2.25) that

$$\begin{aligned} \|AX - C\|^2 + \|XB - D\|^2 &= \|MY - N\|^2 \\ &= \left\| U \begin{bmatrix} M_1 & 0 \\ 0 & 0 \end{bmatrix} V^T Y - N \right\|^2 \\ &= \left\| \begin{bmatrix} M_1 & 0 \\ 0 & 0 \end{bmatrix} V^T Y V - U^T N V \right\|^2. \end{aligned} \tag{4.2}$$

Assume that

$$V^T Y V = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}, \quad Y_{11} \in \mathbb{R}^{r\times r},\ Y_{22} \in \mathbb{R}^{(n-r)\times(n-r)}. \tag{4.3}$$

Then, we have

$$\begin{aligned} &\|AX - C\|^2 + \|XB - D\|^2 \\ &\quad = \left\| M_1 Y_{11} - U_1^T N V_1 \right\|^2 + \left\| M_1 Y_{12} - U_1^T N V_2 \right\|^2 \\ &\qquad + \left\| U_2^T N V_1 \right\|^2 + \left\| U_2^T N V_2 \right\|^2. \end{aligned} \tag{4.4}$$

Hence,

$$\min\left(\|AX - C\|^2 + \|XB - D\|^2 \right) \tag{4.5}$$

is solvable if and only if there exist Y_{11}, Y_{12} such that

$$\begin{aligned} \left\| M_1 Y_{11} - U_1^T N V_1 \right\|^2 &= \min, \\ \left\| M_1 Y_{12} - U_1^T N V_2 \right\|^2 &= \min. \end{aligned} \tag{4.6}$$

It follows from (4.6) that

$$Y_{11} = M_1^{-1} U_1^T N V_1,$$
$$Y_{12} = M_1^{-1} U_1^T N V_2. \tag{4.7}$$

Substituting (4.7) into (4.3) and then into (2.16), we can get that the form of elements in S_L is (4.1). □

Theorem 4.2. *Assume that the notations and conditions are the same as Theorem 4.1. Then,*

$$\left\| \widetilde{X} \right\| = \min_{X \in S_L} \|X\| \tag{4.8}$$

if and only if

$$\widetilde{X} = \Gamma V \begin{bmatrix} M_1^{-1} U_1^T N V_1 & M_1^{-1} U_1^T N V_2 \\ V_2^T N^T U_1 M_1^{-1} & 0 \end{bmatrix} V^T \Gamma^*. \tag{4.9}$$

Proof. In Theorem 4.1, it implies from (4.1) that $\min_{X \in S_L} \|X\|$ is equivalent to X has the expression (4.1) with $Y_{22} = 0$. Hence, (4.9) holds. □

5. An Algorithm and Numerical Example

Base on the main results of this paper, we in this section propose an algorithm for finding the solution of the approximation problem (1.3) and the least squares problem with least norm (1.5). All the tests are performed by MATLAB 6.5 which has a machine precision of around 10^{-16}.

Algorithm 5.1. (1) Input $A \in \mathbb{C}^{m\times n}$, $C \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times l}$, $D \in \mathbb{C}^{n\times l}$.

(2) Compute $A_1, A_2, C_1, C_2, B_1, B_2, D_1, D_2, F, G, K, L, M$, and N by (2.17) and (2.20).

(3) Compute the singular value decomposition of M with the form of (2.25).

(4) If (2.26) holds, then input $E \in \mathbb{C}^{n\times n}$ and compute the solution $\widehat{X}$ of problem (1.3) according (3.1), else compute the solution $\widetilde{X}$ to problem (1.5) by (4.9).

To show our algorithm is feasible, we give two numerical example. Let an nontrivial symmetric involution be

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{5.1}$$

We obtain $[P, Q]$ in (2.2) by using the spectral decomposition of R, then by (2.4)

$$\Gamma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & i \end{bmatrix}. \tag{5.2}$$

Example 5.2. Suppose $A \in \mathbb{C}^{2\times 4}$, $C \in \mathbb{C}^{2\times 4}$, $B \in \mathbb{C}^{4\times 3}$, $D \in \mathbb{C}^{4\times 3}$, and

$$A = \begin{bmatrix} 3.33 - 5.987i & 45i & 7.21 & -i \\ 0 & -0.66i & 7.694 & 1.123i \end{bmatrix},$$
$$C = \begin{bmatrix} 0.2679 - 0.0934i & 0.0012 + 4.0762i & -0.0777 - 0.1718i & -1.2801i \\ 0.2207 & -0.1197i & 0.0877 & 0.7058i \end{bmatrix},$$
$$B = \begin{bmatrix} 4 + 12i & 2.369i & 4.256 - 5.111i \\ 4i & 4.66i & 8.21 - 5i \\ 0 & 4.83i & 56 + i \\ 2.22i & -4.666 & 7i \end{bmatrix}, \tag{5.3}$$
$$D = \begin{bmatrix} 0.0616 + 0.1872i & -0.0009 + 0.1756i & 1.6746 - 0.0494i \\ 0.0024 + 0.2704i & 0.1775 + 0.4194i & 0.7359 - 0.6189i \\ -0.0548 + 0.3444i & 0.0093 - 0.3075i & -0.4731 - 0.1636i \\ 0.0337i & 0.1209 - 0.1864i & -0.2484 - 3.8817i \end{bmatrix}.$$

We can verify that (2.26) holds. Hence, system (1.2) has an Hermitian R-conjugate solution. Given

$$E = \begin{bmatrix} 7.35i & 8.389i & 99.256 - 6.51i & -4.6i \\ 1.55 & 4.56i & 7.71 - 7.5i & i \\ 5i & 0 & -4.556i & -7.99 \\ 4.22i & 0 & 5.1i & 0 \end{bmatrix}. \tag{5.4}$$

Applying Algorithm 5.1, we obtain the following:

$$\widehat{X}=\begin{bmatrix}1.5597 & 0.0194i & 2.8705 & 0.0002i\\ -0.0194i & 9.0001 & 0.2005i & -3.9997\\ 2.8705 & -0.2005i & -0.0452 & 7.9993i\\ -0.0002i & -3.9997 & -7.9993i & 5.6846\end{bmatrix}. \tag{5.5}$$

Example 5.2 illustrates that we can solve the optimal approximation problem with Algorithm 5.1 when system (1.2) have Hermitian R-conjugate solutions.

Example 5.3. Let A, B, and C be the same as Example 5.2, and let D in Example 5.2 be changed into

$$D=\begin{bmatrix}0.0616+0.1872i & -0.0009+0.1756i & 1.6746+0.0494i\\ 0.0024+0.2704i & 0.1775+0.4194i & 0.7359-0.6189i\\ -0.0548+0.3444i & 0.0093-0.3075i & -0.4731-0.1636i\\ 0.0337i & 0.1209-0.1864i & -0.2484-3.8817i\end{bmatrix}. \tag{5.6}$$

We can verify that (2.26) does not hold. By Algorithm 5.1, we get

$$\widetilde{X}=\begin{bmatrix}0.52 & 2.2417i & 0.4914 & 0.3991i\\ -2.2417i & 8.6634 & 0.1921i & -2.8232\\ 0.4914 & -0.1921i & 0.1406 & 1.3154i\\ -0.3991i & -2.8232 & -1.3154i & 6.3974\end{bmatrix}. \tag{5.7}$$

Example 5.3 demonstrates that we can get the least squares solution with Algorithm 5.1 when system (1.2) has not Hermitian R-conjugate solutions.

Acknowledgments

This research was supported by the Grants from the Key Project of Scientific Research Innovation Foundation of Shanghai Municipal Education Commission (13ZZ080), the National Natural Science Foundation of China (11171205), the Natural Science Foundation of Shanghai (11ZR1412500), the Ph.D. Programs Foundation of Ministry of Education of China (20093108110001), the Discipline Project at the corresponding level of Shanghai (A. 13-0101-12-005), Shanghai Leading Academic Discipline Project (J50101), the Natural Science Foundation of Hebei province (A2012403013), and the Natural Science Foundation of Hebei province (A2012205028). The authors are grateful to the anonymous referees for their helpful comments and constructive suggestions.

References

[1] R. D. Hill, R. G. Bates, and S. R. Waters, "On centro-Hermitian matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 11, no. 1, pp. 128–133, 1990.

[2] R. Kouassi, P. Gouton, and M. Paindavoine, "Approximation of the Karhunen-Loeve tranformation and Its application to colour images," *Signal Processing: Image Communication*, vol. 16, pp. 541–551, 2001.

[3] A. Lee, "Centro-Hermitian and skew-centro-Hermitian matrices," *Linear Algebra and Its Applications*, vol. 29, pp. 205–210, 1980.

[4] Z.-Y. Liu, H.-D. Cao, and H.-J. Chen, "A note on computing matrix-vector products with generalized centrosymmetric (centrohermitian) matrices," *Applied Mathematics and Computation*, vol. 169, no. 2, pp. 1332–1345, 2005.

[5] W. F. Trench, "Characterization and properties of matrices with generalized symmetry or skew symmetry," *Linear Algebra and Its Applications*, vol. 377, pp. 207–218, 2004.

[6] K.-W. E. Chu, "Symmetric solutions of linear matrix equations by matrix decompositions," *Linear Algebra and Its Applications*, vol. 119, pp. 35–50, 1989.

[7] H. Dai, "On the symmetric solutions of linear matrix equations," *Linear Algebra and Its Applications*, vol. 131, pp. 1–7, 1990.

[8] F. J. H. Don, "On the symmetric solutions of a linear matrix equation," *Linear Algebra and Its Applications*, vol. 93, pp. 1–7, 1987.

[9] I. Kyrchei, "Explicit representation formulas for the minimum norm least squares solutions of some quaternion matrix equations," *Linear Algebra and Its Applications*, vol. 438, no. 1, pp. 136–152, 2013.

[10] Y. Li, Y. Gao, and W. Guo, "A Hermitian least squares solution of the matrix equation $AXB = C$ subject to inequality restrictions," *Computers & Mathematics with Applications*, vol. 64, no. 6, pp. 1752–1760, 2012.

[11] Z.-Y. Peng and X.-Y. Hu, "The reflexive and anti-reflexive solutions of the matrix equation $AX = B$," *Linear Algebra and Its Applications*, vol. 375, pp. 147–155, 2003.

[12] L. Wu, "The re-positive definite solutions to the matrix inverse problem $AX = B$," *Linear Algebra and Its Applications*, vol. 174, pp. 145–151, 1992.

[13] S. F. Yuan, Q. W. Wang, and X. Zhang, "Least-squares problem for the quaternion matrix equation $AXB + CYD = E$ over different constrained matrices," Article ID 722626, *International Journal of Computer Mathematics*. In press.

[14] Z.-Z. Zhang, X.-Y. Hu, and L. Zhang, "On the Hermitian-generalized Hamiltonian solutions of linear matrix equations," *SIAM Journal on Matrix Analysis and Applications*, vol. 27, no. 1, pp. 294–303, 2005.

[15] F. Cecioni, "Sopra operazioni algebriche," *Annali della Scuola Normale Superiore di Pisa*, vol. 11, pp. 17–20, 1910.

[16] K.-W. E. Chu, "Singular value and generalized singular value decompositions and the solution of linear matrix equations," *Linear Algebra and Its Applications*, vol. 88-89, pp. 83–98, 1987.

[17] S. K. Mitra, "A pair of simultaneous linear matrix equations $A_1XB_1 = C_1, A_2XB_2 = C_2$ and a matrix programming problem," *Linear Algebra and Its Applications*, vol. 131, pp. 107–123, 1990.

[18] H.-X. Chang and Q.-W. Wang, "Reflexive solution to a system of matrix equations," *Journal of Shanghai University*, vol. 11, no. 4, pp. 355–358, 2007.

[19] A. Dajić and J. J. Koliha, "Equations $ax = c$ and $xb = d$ in rings and rings with involution with applications to Hilbert space operators," *Linear Algebra and Its Applications*, vol. 429, no. 7, pp. 1779–1809, 2008.

[20] A. Dajić and J. J. Koliha, "Positive solutions to the equations $AX = C$ and $XB = D$ for Hilbert space operators," *Journal of Mathematical Analysis and Applications*, vol. 333, no. 2, pp. 567–576, 2007.

[21] C.-Z. Dong, Q.-W. Wang, and Y.-P. Zhang, "The common positive solution to adjointable operator equations with an application," *Journal of Mathematical Analysis and Applications*, vol. 396, no. 2, pp. 670–679, 2012.

[22] X. Fang, J. Yu, and H. Yao, "Solutions to operator equations on Hilbert C^*-modules," *Linear Algebra and Its Applications*, vol. 431, no. 11, pp. 2142–2153, 2009.

[23] C. G. Khatri and S. K. Mitra, "Hermitian and nonnegative definite solutions of linear matrix equations," *SIAM Journal on Applied Mathematics*, vol. 31, no. 4, pp. 579–585, 1976.

[24] I. Kyrchei, "Analogs of Cramer's rule for the minimum norm least squares solutions of some matrix equations," *Applied Mathematics and Computation*, vol. 218, no. 11, pp. 6375–6384, 2012.

[25] F. L. Li, X. Y. Hu, and L. Zhang, "The generalized reflexive solution for a class of matrix equations $(AX = C, XB = D)$," *Acta Mathematica Scientia*, vol. 28B, no. 1, pp. 185–193, 2008.

[26] Q.-W. Wang and C.-Z. Dong, "Positive solutions to a system of adjointable operator equations over Hilbert C^*-modules," *Linear Algebra and Its Applications*, vol. 433, no. 7, pp. 1481–1489, 2010.

[27] Q. Xu, "Common Hermitian and positive solutions to the adjointable operator equations $AX = C$, $XB = D$," *Linear Algebra and Its Applications*, vol. 429, no. 1, pp. 1–11, 2008.

[28] F. Zhang, Y. Li, and J. Zhao, "Common Hermitian least squares solutions of matrix equations $A_1XA_1^* = B_1$ and $A_2XA_2^* = B_2$ subject to inequality restrictions," *Computers & Mathematics with Applications* , vol. 62, no. 6, pp. 2424–2433, 2011.

[29] H.-X. Chang, Q.-W. Wang, and G.-J. Song, "(R, S)-conjugate solution to a pair of linear matrix equations," *Applied Mathematics and Computation*, vol. 217, no. 1, pp. 73–82, 2010.

[30] E. W. Cheney, *Introduction to Approximation Theory*, McGraw-Hill, New York, NY, USA, 1966.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2012, Article ID 784620, 13 pages
doi:10.1155/2012/784620

Research Article

Perturbation Analysis for the Matrix Equation $X - \sum_{i=1}^{m} A_i^* X A_i + \sum_{j=1}^{n} B_j^* X B_j = I$

Xue-Feng Duan[1,2] and Qing-Wen Wang[3]

[1] *School of Mathematics and Computational Science, Guilin University of Electronic Technology, Guilin 541004, China*
[2] *Research Center on Data Science and Social Computing, Guilin University of Electronic Technology, Guilin 541004, China*
[3] *Department of Mathematics, Shanghai University, Shanghai 200444, China*

Correspondence should be addressed to Qing-Wen Wang, wqw858@yahoo.com.cn

Received 18 September 2012; Accepted 19 November 2012

Academic Editor: Yang Zhang

We consider the perturbation analysis of the matrix equation $X - \sum_{i=1}^{m} A_i^* X A_i + \sum_{j=1}^{n} B_j^* X B_j = I$. Based on the matrix differentiation, we first give a precise perturbation bound for the positive definite solution. A numerical example is presented to illustrate the sharpness of the perturbation bound.

1. Introduction

In this paper, we consider the matrix equation

$$X - \sum_{i=1}^{m} A_i^* X A_i + \sum_{j=1}^{n} B_j^* X B_j = I, \tag{1.1}$$

where $A_1, A_2, \ldots, A_m, B_1, B_2, \ldots, B_n$ are $n \times n$ complex matrices, I is an $n \times n$ identity matrix, m, n are nonnegative integers and the positive definite solution X is practical interest. Here, A_i^* and B_i^* denote the conjugate transpose of the matrices A_i and B_i, respectively. Equation (1.1) arises in solving some nonlinear matrix equations with Newton method. See, for example, the nonlinear matrix equation which appears in Sakhnovich [1]. Solving these nonlinear matrix equations gives rise to (1.1). On the other hand, (1.1) is the general case of the generalized Lyapunov equation

$$MYS^* + SYM^* + \sum_{k=1}^{t} N_k Y N_k^* + CC^* = 0, \tag{1.2}$$

whose positive definite solution is the controllability Gramian of the bilinear control system (see [2, 3] for more details)

$$M\dot{x}(t) = Sx(t) + \sum_{k=1}^{t} N_k x(t) u_k(t) + Cu(t). \tag{1.3}$$

Set

$$\begin{gathered} E = \frac{1}{\sqrt{2}}(M - S + I), \qquad F = \frac{1}{\sqrt{2}}(M + S + I), \qquad G = S - I, \qquad Q = CC^*, \\ X = Q^{-1/2} Y Q^{-1/2}, \qquad A_i = Q^{-1/2} N_i Q^{1/2}, \quad i = 1, 2, \ldots, t, \\ A_{t+1} = Q^{-1/2} F Q^{1/2}, \qquad A_{t+2} = Q^{-1/2} G Q^{1/2}, \\ B_1 = Q^{-1/2} E Q^{1/2}, \qquad B_2 = Q^{-1/2} C Q^{1/2}, \end{gathered} \tag{1.4}$$

then (1.2) can be equivalently written as (1.1) with $m = t + 2$ and $n = 2$.

Some special cases of (1.1) have been studied. Based on the kronecker product and fixed point theorem in partially ordered sets, Reurings [4] and Ran and Reurings [5, 6] gave some sufficient conditions for the existence of a unique positive definite solution of the linear matrix equations $X - \sum_{i=1}^{m} A_i^* X A_i = I$ and $X + \sum_{j=1}^{n} B_i^* X B_i = I$. And the expressions for these unique positive definite solutions were also derived under some constraint conditions. For the general linear matrix equation (1.1), Reurings [[4], Page 61] pointed out that it is hard to find sufficient conditions for the existence of a positive definite solution, because the map

$$G(X) = I + \sum_{i=1}^{m} A_i^* X A_i - \sum_{j=1}^{n} B_j^* X B_j \tag{1.5}$$

is not monotone and does not map the set of $n \times n$ positive definite matrices into itself. Recently, Berzig [7] overcame these difficulties by making use of Bhaskar-Lakshmikantham coupled fixed point theorem and gave a sufficient condition for (1.1) existing a unique positive definite solution. An iterative method was constructed to compute the unique positive definite solution, and the error estimation was given too.

Recently, the matrix equations of the form (1.1) have been studied by many authors (see [8–14]). Some numerical methods for solving the well-known Lyapunov equation $X + A^T X A = Q$, such as Bartels-Stewart method and Hessenberg-Schur method, have been proposed in [12]. Based on the fixed point theorem, the sufficient and necessary conditions for the existence of a positive definite solution of the matrix equation $X^s \pm A^* X^{-t} A = Q$, $s, t \in N$ have been given in [8, 9]. The fixed point iterative method and inversion-free iterative method were developed for solving the matrix equations $X \pm A^*X^{-\alpha}A = Q,\ \alpha > 0$ in [13, 14]. By making use of the fixed point theorem of mixed monotone operator, the matrix equation $X - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i = Q,\ \ 0 < |\delta_i| < 1$ was studied in [10], and derived a sufficient condition for the existence of a positive definite solution. Assume that F maps positive definite matrices either into positive definite matrices or into negative definite matrices, the general nonlinear matrix equation $X + A^*F(X)A = Q$ was studied in [11], and the fixed point iterative method was constructed to compute the positive definite solution under some additional conditions.

Motivated by the works and applications in [2–6], we continue to study the matrix equation (1.1). Based on a new mathematical tool (i.e., the matrix differentiation), we firstly give a differential bound for the unique positive definite solution of (1.1), and then use it to derive a precise perturbation bound for the unique positive definite solution. A numerical example is used to show that the perturbation bound is very sharp.

Throughout this paper, we write $B > 0\ \ (B \geq 0)$ if the matrix B is positive definite (semidefinite). If $B - C$ is positive definite (semidefinite), then we write $B > C\ \ (B \geq C)$. If a positive definite matrix X satisfies $B \leq X \leq C$, we denote that $X \in [B, C]$. The symbols $\lambda_1(B)$ and $\lambda_n(B)$ denote the maximal and minimal eigenvalues of an $n \times n$ Hermitian matrix B, respectively. The symbol $H^{n\times n}$ stands for the set of $n \times n$ Hermitian matrices. The symbol $\|B\|$ denotes the spectral norm of the matrix B.

2. Perturbation Analysis for the Matrix Equation (1.1)

Based on the matrix differentiation, we firstly give a differential bound for the unique positive definite solution X_U of (1.1), and then use it to derive a precise perturbation bound for X_U in this section.

Definition 2.1 ([15], Definition 3.6). Let $F = (f_{ij})_{m\times n}$, then the matrix differentiation of F is $dF = (df_{ij})_{m\times n}$. For example, let

$$F = \begin{pmatrix} s+t & s^2-2t \\ 2s+t^3 & t^2 \end{pmatrix}. \tag{2.1}$$

Then

$$dF = \begin{pmatrix} ds+dt & 2sds-2dt \\ 2ds+3t^2dt & 2tdt \end{pmatrix}. \tag{2.2}$$

Lemma 2.2 ([15], Theorem 3.2). *The matrix differentiation has the following properties:*

(1) $d(F_1 \pm F_2) = dF_1 \pm dF_2$;

(2) $d(kF) = k(dF)$, *where k is a complex number;*

(3) $d(F^*) = (dF)^*$;

(4) $d(F_1F_2F_3) = (dF_1)F_2F_3 + F_1(dF_2)F_3 + F_1F_2(dF_3)$;

(5) $dF^{-1} = -F^{-1}(dF)F^{-1}$;

(6) $dF = 0$, *where F is a constant matrix.*

Lemma 2.3 ([7], Theorem 3.1). *If*

$$\sum_{i=1}^{m} A_i^* A_i < \frac{1}{2} I, \qquad \sum_{j=1}^{n} B_j^* B_j < \frac{1}{2} I, \tag{2.3}$$

then (1.1) *has a unique positive definite solution* X_U *and*

$$X_U \in \left[I - 2\sum_{j=1}^{n} B_j^* B_j,\ I + 2\sum_{i=1}^{m} A_i^* A_i \right]. \tag{2.4}$$

Theorem 2.4. *If*

$$\sum_{i=1}^{m} \|A_i\|^2 < \frac{1}{2}, \qquad \sum_{j=1}^{n} \|B_j\|^2 < \frac{1}{2}, \tag{2.5}$$

then (1.1) *has a unique positive definite solution* X_U, *and it satisfies*

$$\|dX_U\| \le \frac{2\left(1 + 2\sum_{i=1}^{m} \|A_i\|^2\right)\left[\sum_{i=1}^{m} (\|A_i\|\|dA_i\|) + \sum_{j=1}^{n} \left(\|B_j\|\|dB_j\|\right)\right]}{1 - \sum_{i=1}^{m} \|A_i\|^2 - \sum_{j=1}^{n} \|B_j\|^2}. \tag{2.6}$$

Proof. Since

$$\begin{aligned} \lambda_1(A_i^* A_i) &\le \|A_i^* A_i\| \le \|A_i\|^2, \quad i = 1, 2, \ldots, m, \\ \lambda_1\left(B_j^* B_j\right) &\le \left\|B_j^* B_j\right\| \le \|B_j\|^2, \quad j = 1, 2, \ldots, n, \end{aligned} \tag{2.7}$$

then

$$\begin{aligned} A_i^* A_i &\le \lambda_1(A_i^* A_i) I \le \|A_i^* A_i\| I \le \|A_i\|^2 I, \quad i = 1, 2, \ldots, m, \\ B_j^* B_j &\le \lambda_1\left(B_j^* B_j\right) I \le \left\|B_j^* B_j\right\| \le \|B_j\|^2 I, \quad j = 1, 2, \ldots, n, \end{aligned} \tag{2.8}$$

consequently,

$$\begin{aligned} \sum_{i=1}^{m} A_i^* A_i &\le \sum_{i=1}^{m} \|A_i\|^2 I, \\ \sum_{j=1}^{n} B_j^* B_j &\le \sum_{j=1}^{n} \|B_j\|^2 I. \end{aligned} \tag{2.9}$$

Combining (2.5)–(2.9) we have

$$\sum_{i=1}^{m} A_i^* A_i \le \sum_{i=1}^{m} \|A_i\|^2 I < \frac{1}{2} I,$$
$$\sum_{j=1}^{n} B_j^* B_j \le \sum_{j=1}^{n} \|B_j\|^2 I < \frac{1}{2} I. \tag{2.10}$$

Then by Lemma 2.3 we obtain that (1.1) has a unique positive definite solution X_U, which satisfies

$$X_U \in \left[I - 2\sum_{j=1}^{n} B_j^* B_j,\ I + 2\sum_{i=1}^{m} A_i^* A_i \right]. \tag{2.11}$$

Noting that X_U is the unique positive definite solution of (1.1), then

$$X_U - \sum_{i=1}^{m} A_i^* X_U A_i + \sum_{j=1}^{n} B_j^* X_U B_j = I. \tag{2.12}$$

It is known that the elements of X_U are differentiable functions of the elements of A_i and B_i. Differentiating (2.12), and by Lemma 2.2, we have

$$\begin{aligned} dX_U &- \sum_{i=1}^{m} \left[(dA_i^*) X_U A_i + A_i^* (dX_U) A_i + A_i^* X_U (dA_i) \right] \\ &+ \sum_{j=1}^{n} \left[\left(dB_j^*\right) X_U B_j + B_j^* (dX_U) B_j + B_j^* X_U \left(dB_j\right) \right] = 0, \end{aligned} \tag{2.13}$$

which implies that

$$\begin{aligned} dX_U &- \sum_{i=1}^{m} A_i^* (dX_U) A_i + \sum_{j=1}^{n} B_j^* (dX_U) B_j \\ &= \sum_{i=1}^{m} (dA_i^*) X_U A_i + \sum_{i=1}^{m} A_i^* X_U (dA_i) - \sum_{j=1}^{n} \left(dB_j^*\right) X_U B_j - \sum_{j=1}^{n} B_j^* X_U \left(dB_j\right). \end{aligned} \tag{2.14}$$

By taking spectral norm for both sides of (2.14), we obtain that

$$\begin{aligned}
&\left\| dX_U - \sum_{i=1}^{m} A_i^*(dX_U)A_i + \sum_{j=1}^{n} B_j^*(dX_U)B_j \right\| \\
&= \left\| \sum_{i=1}^{m} (dA_i^*)X_U A_i + \sum_{i=1}^{m} A_i^* X_U (dA_i) - \sum_{j=1}^{n} \left(dB_j^*\right) X_U B_j - \sum_{j=1}^{n} B_j^* X_U (dB_j) \right\| \\
&\leq \left\| \sum_{i=1}^{m} (dA_i^*)X_U A_i \right\| + \left\| \sum_{i=1}^{m} A_i^* X_U (dA_i) \right\| + \left\| \sum_{j=1}^{n} \left(dB_j^*\right) X_U B_j \right\| + \left\| \sum_{j=1}^{n} B_j^* X_U (dB_j) \right\| \\
&\leq \sum_{i=1}^{m} \left\| (dA_i^*)X_U A_i \right\| + \sum_{i=1}^{m} \left\| A_i^* X_U (dA_i) \right\| + \sum_{j=1}^{n} \left\| \left(dB_j^*\right) X_U B_j \right\| + \sum_{j=1}^{n} \left\| B_j^* X_U (dB_j) \right\| \\
&\leq \sum_{i=1}^{m} \|dA_i^*\| \|X_U\| \|A_i\| + \sum_{i=1}^{m} \|A_i^*\| \|X_U\| \|dA_i\| + \sum_{j=1}^{n} \left\|dB_j^*\right\| \|X_U\| \|B_j\| + \sum_{j=1}^{n} \left\|B_j^*\right\| \|X_U\| \|dB_j\| \\
&= 2\sum_{i=1}^{m} (\|A_i\| \|X_U\| \|dA_i\|) + 2\sum_{j=1}^{n} \left(\|B_j\| \|X_U\| \|dB_j\|\right) \\
&= 2\left[\sum_{i=1}^{m} (\|A_i\| \|dA_i\|) + \sum_{j=1}^{n} \left(\|B_j\| \|dB_j\|\right) \right] \|X_U\|,
\end{aligned} \tag{2.15}$$

and noting (2.11) we obtain that

$$\|X_U\| \leq \left\| I + 2\sum_{i=1}^{m} A_i^* A_i \right\| \leq 1 + 2\sum_{i=1}^{m} \|A_i\|^2. \tag{2.16}$$

Then

$$\begin{aligned}
&\left\| dX_U - \sum_{i=1}^{m} A_i^*(dX_U)A_i + \sum_{j=1}^{n} B_j^*(dX_U)B_j \right\| \\
&\leq 2\left[\sum_{i=1}^{m} (\|A_i\| \|dA_i\|) + \sum_{j=1}^{n} \left(\|B_j\| \|dB_j\|\right) \right] \|X_U\| \\
&\leq 2\left(1 + 2\sum_{i=1}^{m} \|A_i\|^2 \right) \left[\sum_{i=1}^{m} (\|A_i\| \|dA_i\|) + \sum_{j=1}^{n} \left(\|B_j\| \|dB_j\|\right) \right],
\end{aligned} \tag{2.17}$$

$$\begin{aligned}
&\left\| dX_U - \sum_{i=1}^{m} A_i^*(dX_U)A_i + \sum_{j=1}^{n} B_j^*(dX_U)B_j \right\| \\
&\quad \geq \|dX_U\| - \left\| \sum_{i=1}^{m} A_i^*(dX_U)A_i \right\| - \left\| \sum_{j=1}^{n} B_j^*(dX_U)B_j \right\| \\
&\quad \geq \|dX_U\| - \sum_{i=1}^{m} \left\| A_i^*(dX_U)A_i \right\| - \sum_{j=1}^{n} \left\| B_j^*(dX_U)B_j \right\| \\
&\quad \geq \|dX_U\| - \sum_{i=1}^{m} \left\| A_i^* \right\| \|dX_U\| \|A_i\| - \sum_{j=1}^{n} \left\| B_j^* \right\| \|dX_U\| \left\| B_j \right\| \\
&\quad = \left(1 - \sum_{i=1}^{m} \|A_i\|^2 - \sum_{j=1}^{n} \left\| B_j \right\|^2 \right) \|dX_U\|.
\end{aligned} \tag{2.18}$$

Due to (2.5) we have

$$1 - \sum_{i=1}^{m} \|A_i\|^2 - \sum_{j=1}^{n} \left\| B_j \right\|^2 > 0. \tag{2.19}$$

Combining (2.17), (2.18) and noting (2.19), we have

$$\begin{aligned}
&\left(1 - \sum_{i=1}^{m} \|A_i\|^2 - \sum_{j=1}^{n} \left\| B_j \right\|^2 \right) \|dX_U\| \\
&\quad \leq \left\| dX_U - \sum_{i=1}^{m} A_i^*(dX_U)A_i + \sum_{j=1}^{n} B_j^*(dX_U)B_j \right\| \\
&\quad \leq 2\left(1 + 2\sum_{i=1}^{m} \|A_i\|^2 \right) \left[\sum_{i=1}^{m} (\|A_i\| \|dA_i\|) + \sum_{j=1}^{n} \left(\left\| B_j \right\| \left\| dB_j \right\| \right) \right],
\end{aligned} \tag{2.20}$$

which implies that

$$\|dX_U\| \leq \frac{2\left(1 + 2\sum_{i=1}^{m} \|A_i\|^2 \right) \left[\sum_{i=1}^{m} (\|A_i\| \|dA_i\|) + \sum_{j=1}^{n} \left(\left\| B_j \right\| \left\| dB_j \right\| \right) \right]}{1 - \sum_{i=1}^{m} \|A_i\|^2 - \sum_{j=1}^{n} \left\| B_j \right\|^2}. \tag{2.21}$$

□

Theorem 2.5. *Let* $\widetilde{A}_1, \widetilde{A}_2, \ldots, \widetilde{A}_m, \widetilde{B}_1, \widetilde{B}_2, \ldots, \widetilde{B}_n$ *be perturbed matrices of* $A_1, A_2, \ldots, A_m$, $B_1,\ B_2, \ldots, B_n$ *in* (1.1) *and* $\Delta_i = \widetilde{A}_i - A_i,\ i = 1, 2, \ldots, m,\ \Delta_j = \widetilde{B}_j - B_j,\quad j = 1, 2, \ldots, n$*. If*

$$\sum_{i=1}^{m} \|A_i\|^2 < \frac{1}{2}, \qquad \sum_{j=1}^{n} \|B_j\|^2 < \frac{1}{2}, \tag{2.22}$$

$$2\sum_{i=1}^{m} (\|A_i\| \|\Delta_i\|) + \sum_{i=1}^{m} \|A_i\|^2 < \frac{1}{2} - \sum_{i=1}^{m} \|A_i\|^2, \tag{2.23}$$

$$2\sum_{j=1}^{n} (\|B_j\| \|\Delta_j\|) + \sum_{j=1}^{n} \|\Delta_j\|^2 < \frac{1}{2} - \sum_{j=1}^{n} \|B_j\|^2, \tag{2.24}$$

then (1.1) *and its perturbed equation*

$$\widetilde{X} - \sum_{i=1}^{m} \widetilde{A}_i^* \widetilde{X} \widetilde{A}_i + \sum_{j=1}^{n} \widetilde{B}_j^* \widetilde{X} \widetilde{B}_j = I \tag{2.25}$$

have unique positive definite solutions X_U *and* $\widetilde{X}_U$*, respectively, which satisfy*

$$\left\| \widetilde{X}_U - X_U \right\| \leq S_{\mathrm{err}}, \tag{2.26}$$

where

$$S_{\mathrm{err}} = \frac{2\left[1 + 2\sum_{i=1}^{m} (\|A_i\| + \|\Delta_i\|)^2\right] \left[\sum_{i=1}^{m} (\|A_i\| + \|\Delta_i\|)^2 \|\Delta_i\| + \sum_{j=1}^{n} \left(\|B_j\| + \|\Delta_j\|\right)^2 \|\Delta_j\|\right]}{1 - \sum_{i=1}^{m} (\|A_i\| + \|\Delta_i\|)^2 - \sum_{j=1}^{n} \left(\|B_j\| + \|\Delta_j\|\right)^2}. \tag{2.27}$$

Proof. By (2.22) and Theorem 2.4, we know that (1.1) has a unique positive definite solution X_U. And by (2.23) we have

$$\begin{aligned}
\sum_{i=1}^{m} \left\| \widetilde{A}_i \right\|^2 &= \sum_{i=1}^{m} \|A_i + \Delta_i\|^2 \leq \sum_{i=1}^{m} (\|A_i\| + \|\Delta_i\|)^2 \\
&= \sum_{i=1}^{m} \left(\|A_i\|^2 + 2\|A_i\| \|\Delta_i\| + \|\Delta_i\|^2 \right) \\
&= \sum_{i=1}^{m} \|A_i\|^2 + 2\sum_{i=1}^{m} (\|A_i\| \|\Delta_i\|) + \sum_{i=1}^{m} \|\Delta_i\|^2 \\
&< \sum_{i=1}^{m} \|A_i\|^2 + \frac{1}{2} - \sum_{i=1}^{m} \|A_i\|^2 = \frac{1}{2},
\end{aligned} \tag{2.28}$$

similarly, by (2.24) we have

$$\sum_{j=1}^{n}\left\|\widetilde{B}_j\right\|^2 < \frac{1}{2}. \tag{2.29}$$

By (2.28), (2.29), and Theorem 2.4 we obtain that the perturbed equation (2.25) has a unique positive definite solution $\widetilde{X}_U$. □

Set

$$A_i(t) = A_i + t\Delta_i, \quad B_j(t) = B_j + t\Delta_j, \quad t \in [0,1], \tag{2.30}$$

then by (2.23) we have

$$\begin{aligned}
\sum_{i=1}^{m}\|A_i(t)\|^2 &= \sum_{i=1}^{m}\|A_i + t\Delta_i\|^2 \le \sum_{i=1}^{m}(\|A_i\| + t\|\Delta_i\|)^2 \\
&= \sum_{i=1}^{m}\left(\|A_i\|^2 + 2t\|A_i\|\|\Delta_i\| + t^2\|\Delta_i\|^2\right) \\
&\le \sum_{i=1}^{m}\left(\|A_i\|^2 + 2\|A_i\|\|\Delta_i\| + \|\Delta_i\|^2\right) \\
&= \sum_{i=1}^{m}\|A_i\|^2 + 2\sum_{i=1}^{m}(\|A_i\|\|\Delta_i\|) + \sum_{i=1}^{m}\|\Delta_i\|^2 \\
&< \sum_{i=1}^{m}\|A_i\|^2 + \frac{1}{2} - \sum_{i=1}^{m}\|A_i\|^2 = \frac{1}{2},
\end{aligned} \tag{2.31}$$

similarly, by (2.24) we have

$$\sum_{j=1}^{n}\left\|B_j(t)\right\|^2 = \sum_{i=1}^{m}\left\|B_j + t\Delta_j\right\|^2 < \frac{1}{2}. \tag{2.32}$$

Therefore, by (2.31), (2.32), and Theorem 2.4 we derive that for arbitrary $t \in [0,1]$, the matrix equation

$$X - \sum_{i=1}^{m}A_i^*(t)XA_i(t) + \sum_{j=1}^{n}B_j^*(t)XB_j(t) = I \tag{2.33}$$

has a unique positive definite solution $X_U(t)$, especially,

$$X_U(0) = X_U, \qquad X_U(1) = \widetilde{X}_U. \tag{2.34}$$

From Theorem 2.4 it follows that

$$\begin{aligned}
\left\|\widetilde{X}_U - X_U\right\| &= \|X_U(1) - X_U(0)\| = \left\|\int_0^1 dX_U(t)\right\| \le \int_0^1 \|dX_U(t)\| \\
&\le \int_0^1 \frac{2\left(1 + 2\sum_{i=1}^m \|A_i(t)\|^2\right)\left[\sum_{i=1}^m (\|A_i(t)\|\|dA_i(t)\|) + \sum_{j=1}^n \left(\left\|B_j(t)\right\|\left\|dB_j(t)\right\|\right)\right]}{1 - \sum_{i=1}^m \|A_i(t)\|^2 - \sum_{j=1}^n \left\|B_j(t)\right\|^2} \\
&\le \int_0^1 \frac{2\left(1 + 2\sum_{i=1}^m \|A_i(t)\|^2\right)\left[\sum_{i=1}^m (\|A_i(t)\|\|\Delta_i\|dt) + \sum_{j=1}^n \left(\left\|B_j(t)\right\|\left\|\Delta_j\right\|dt\right)\right]}{1 - \sum_{i=1}^m \|A_i(t)\|^2 - \sum_{j=1}^n \left\|B_j(t)\right\|^2} \\
&= \int_0^1 \frac{2\left(1 + 2\sum_{i=1}^m \|A_i(t)\|^2\right)\left[\sum_{i=1}^m (\|A_i(t)\|\|\Delta_i\|) + \sum_{j=1}^n \left(\left\|B_j(t)\right\|\left\|\Delta_j\right\|\right)\right]}{1 - \sum_{i=1}^m \|A_i(t)\|^2 - \sum_{j=1}^n \left\|B_j(t)\right\|^2} dt.
\end{aligned} \tag{2.35}$$

Noting that

$$\begin{aligned}
\|A_i(t)\| &= \|A_i + t\Delta_i\| \le \|A_i\| + t\|\Delta_i\|, \quad i = 1, 2, \ldots, m, \\
\left\|B_j(t)\right\| &= \left\|B_j + t\Delta_i\right\| \le \left\|B_j\right\| + t\|\Delta_i\|, \quad j = 1, 2, \ldots, n,
\end{aligned} \tag{2.36}$$

and combining Mean Value Theorem of Integration, we have

$$\begin{aligned}
&\left\|\widetilde{X}_U - X_U\right\| \\
&\le \int_0^1 \frac{2\left(1 + 2\sum_{i=1}^m \|A_i(t)\|^2\right)\left[\sum_{i=1}^m (\|A_i(t)\|\|\Delta_i\|) + \sum_{j=1}^n \left(\left\|B_j(t)\right\|\left\|\Delta_j\right\|\right)\right]}{1 - \sum_{i=1}^m \|A_i(t)\|^2 - \sum_{j=1}^n \left\|B_j(t)\right\|^2} dt \\
&\le \int_0^1 \frac{2\left[1 + 2\sum_{i=1}^m (\|A_i\| + t\|\Delta_i\|)^2\right]\left[\sum_{i=1}^m (\|A_i\| + t\|\Delta_i\|)^2\|\Delta_i\| + \sum_{j=1}^n \left(\left\|B_j\right\| + t\left\|\Delta_j\right\|\right)^2\left\|\Delta_j\right\|\right]}{1 - \sum_{i=1}^m (\|A_i\| + t\|\Delta_i\|)^2 - \sum_{j=1}^n \left(\left\|B_j\right\| + t\left\|\Delta_j\right\|\right)^2} dt \\
&= \frac{2\left[1 + 2\sum_{i=1}^m (\|A_i\| + \xi\|\Delta_i\|)^2\right]\left[\sum_{i=1}^m (\|A_i\| + \xi\|\Delta_i\|)^2\|\Delta_i\| + \sum_{j=1}^n \left(\left\|B_j\right\| + \xi\left\|\Delta_j\right\|\right)^2\left\|\Delta_j\right\|\right]}{1 - \sum_{i=1}^m (\|A_i\| + \xi\|\Delta_i\|)^2 - \sum_{j=1}^n \left(\left\|B_j\right\| + \xi\left\|\Delta_j\right\|\right)^2} \\
&\quad \times (1 - 0), \quad \xi \in [0, 1] \\
&\le \frac{2\left[1 + 2\sum_{i=1}^m (\|A_i\| + \|\Delta_i\|)^2\right]\left[\sum_{i=1}^m (\|A_i\| + \|\Delta_i\|)^2\|\Delta_i\| + \sum_{j=1}^n \left(\left\|B_j\right\| + \left\|\Delta_j\right\|\right)^2\left\|\Delta_j\right\|\right]}{1 - \sum_{i=1}^m (\|A_i\| + \|\Delta_i\|)^2 - \sum_{j=1}^n \left(\left\|B_j\right\| + \left\|\Delta_j\right\|\right)^2} = S_{\text{err}}.
\end{aligned} \tag{2.37}$$

3. Numerical Experiments

In this section, we use a numerical example to confirm the correctness of Theorem 2.5 and the precision of the perturbation bound for the unique positive definite solution X_U of (1.1).

Example 3.1. Consider the symmetric linear matrix equation

$$X - A_1^* X A_1 - A_2^* X A_2 + B_1^* X B_1 + B_2^* X B_2 = I, \tag{3.1}$$

and its perturbed equation

$$\widetilde{X} - \widetilde{A}_1^* \widetilde{X} \widetilde{A}_1 - \widetilde{A}_2^* \widetilde{X} \widetilde{A}_2 + \widetilde{B}_1^* \widetilde{X} \widetilde{B}_1 + \widetilde{B}_2^* \widetilde{X} \widetilde{B}_2 = I, \tag{3.2}$$

where

$$\begin{gathered}
A_1 = \begin{pmatrix} 0.02 & -0.10 & -0.02 \\ 0.08 & -0.10 & 0.02 \\ -0.06 & -0.12 & 0.14 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0.08 & -0.10 & -0.02 \\ 0.08 & -0.10 & 0.02 \\ -0.06 & -0.12 & 0.14 \end{pmatrix}, \\
B_1 = \begin{pmatrix} 0.47 & 0.02 & 0.04 \\ -0.10 & 0.36 & -0.02 \\ -0.04 & 0.01 & 0.47 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} 0.10 & 0.10 & 0.05 \\ 0.15 & 0.275 & 0.075 \\ 0.05 & 0.05 & 0.175 \end{pmatrix}, \\
\widetilde{A}_1 = A_1 + \begin{pmatrix} 0.5 & 0.1 & -0.2 \\ -0.4 & 0.2 & 0.6 \\ -0.2 & 0.1 & -0.1 \end{pmatrix} \times 10^{-j}, \qquad \widetilde{A}_2 = A_2 + \begin{pmatrix} -0.4 & 0.1 & -0.2 \\ 0.5 & 0.7 & -1.3 \\ 1.1 & 0.9 & 0.6 \end{pmatrix} \times 10^{-j}, \\
\widetilde{B}_1 = B_1 + \begin{pmatrix} 0.8 & 0.2 & 0.05 \\ -0.2 & 0.12 & 0.14 \\ -0.25 & -0.2 & 0.26 \end{pmatrix} \times 10^{-j}, \quad \widetilde{B}_2 = B_2 + \begin{pmatrix} 0.2 & 0.2 & 0.1 \\ -0.3 & 0.15 & -0.15 \\ 0.1 & -0.1 & 0.25 \end{pmatrix} \times 10^{-j}, \quad j \in N.
\end{gathered} \tag{3.3}$$

It is easy to verify that the conditions (2.22)–(2.24) are satisfied, then (3.1) and its perturbed equation (3.2) have unique positive definite solutions X_U and $\widetilde{X}_U$, respectively. From Berzig [7] it follows that the sequences $\{X_k\}$ and $\{Y_k\}$ generated by the iterative method

$$\begin{gathered}
X_0 = 0, \qquad Y_0 = 2I, \\
X_{k+1} = I + A_1^* X_k A_1 + A_2^* X_k A_2 - B_1^* Y_k B_1 - B_2^* Y_k B_2, \\
Y_{k+1} = I + A_1^* Y_k A_1 + A_2^* Y_k A_2 - B_1^* X_k B_1 - B_2^* X_k B_2, \quad k = 0, 1, 2, \ldots
\end{gathered} \tag{3.4}$$

both converge to X_U. Choose $\tau = 1.0 \times 10^{-15}$ as the termination scalar, that is,

$$R(X) = \left\| X - A_1^* X A_1 - A_2^* X A_2 + B_1^* X B_1 + B_2^* X B_2 - I \right\| \le \tau = 1.0 \times 10^{-15}. \tag{3.5}$$

By using the iterative method (3.4) we can get the computed solution X_k of (3.1). Since $R(X_k) < 1.0 \times 10^{-15}$, then the computed solution X_k has a very high precision. For simplicity,

Table 1: Numerical results for the different values of j.

j	2	3	4	5	6
$\Vert\widetilde{X}_U - X_U\Vert/\Vert X_U\Vert$	2.13×10^{-4}	6.16×10^{-6}	5.23×10^{-8}	4.37×10^{-10}	8.45×10^{-12}
$S_{\mathrm{err}}/\Vert X_U\Vert$	3.42×10^{-4}	7.13×10^{-6}	6.63×10^{-8}	5.12×10^{-10}	9.77×10^{-12}

we write the computed solution as the unique positive definite solution X_U. Similarly, we can also get the unique positive definite solution $\widetilde{X}_U$ of the perturbed equation (3.2).

Some numerical results on the perturbation bounds for the unique positive definite solution X_U are listed in Table 1.

From Table 1, we see that Theorem 2.5 gives a precise perturbation bound for the unique positive definite solution of (3.1).

4. Conclusion

In this paper, we study the matrix equation (1.1) which arises in solving some nonlinear matrix equations and the bilinear control system. A new method of perturbation analysis is developed for the matrix equation (1.1). By making use of the matrix differentiation and its elegant properties, we derive a precise perturbation bound for the unique positive definite solution of (1.1). A numerical example is presented to illustrate the sharpness of the perturbation bound.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (11101100; 11171205; 11261014; 11226323), the Natural Science Foundation of Guangxi Province (2012GXNSFBA053006; 2011GXNSFA018138), the Key Project of Scientific Research Innovation Foundation of Shanghai Municipal Education Commission (13ZZ080), the Natural Science Foundation of Shanghai (11ZR1412500), the Ph.D. Programs Foundation of Ministry of Education of China (20093108110001), the Discipline Project at the corresponding level of Shanghai (A. 13-0101-12-005), and Shanghai Leading Academic Discipline Project (J50101). The authors wish to thank Professor Y. Zhang and the anonymous referees for providing very useful suggestions for improving this paper. The authors also thank Dr. Xiaoshan Chen for discussing the properties of matrix differentiation.

References

[1] L. A. Sakhnovich, *Interpolation Theory and Its Applications, Mathematics and Its Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.

[2] T. Damm, "Direct methods and ADI-preconditioned Krylov subspace methods for generalized Lyapunov equations," *Numerical Linear Algebra with Applications*, vol. 15, no. 9, pp. 853–871, 2008.

[3] L. Zhang and J. Lam, "On H_2 model reduction of bilinear systems," *Automatica*, vol. 38, no. 2, pp. 205–216, 2002.

[4] M. C. B. Reurings, *Symmetric matrix equations [Ph.D. thesis]*, Vrije Universiteit, Amsterdam, The Netherlands, 2003.

[5] A. C. M. Ran and M. C. B. Reurings, "The symmetric linear matrix equation," *Electronic Journal of Linear Algebra*, vol. 9, pp. 93–107, 2002.

[6] A. C. M. Ran and M. C. B. Reurings, "A fixed point theorem in partially ordered sets and some applications to matrix equations," *Proceedings of the American Mathematical Society*, vol. 132, no. 5, pp. 1435–1443, 2004.

[7] M. Berzig, "Solving a class of matrix equation via Bhaskar-Lakshmikantham coupled fixed point theorem," *Applied Mathematics Letters*, vol. 25, no. 11, pp. 1638–1643, 2012.

[8] J. Cai and G. Chen, "Some investigation on Hermitian positive definite solutions of the matrix equation $X^s + A^*X^{-t}A = Q$," *Linear Algebra and Its Applications*, vol. 430, no. 8-9, pp. 2448–2456, 2009.

[9] X. Duan and A. Liao, "On the existence of Hermitian positive definite solutions of the matrix equation $X^s + A^*X^{-t}A = Q$," *Linear Algebra and its Applications*, vol. 429, no. 4, pp. 673–687, 2008.

[10] X. Duan, A. Liao, and B. Tang, "On the nonlinear matrix equation $X - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i = Q$," *Linear Algebra and its Applications*, vol. 429, no. 1, pp. 110–121, 2008.

[11] S. M. El-Sayed and A. C. M. Ran, "On an iteration method for solving a class of nonlinear matrix equations," *SIAM Journal on Matrix Analysis and Applications*, vol. 23, no. 3, pp. 632–645, 2001.

[12] Z. Gajic and M. T. J. Qureshi, *Lyapunov Matrix Equation in System Stability and Control*, Academic Press, London, UK, 1995.

[13] V. I. Hasanov, "Notes on two perturbation estimates of the extreme solutions to the equations $X \pm A^*X^{-1}A = Q$," *Applied Mathematics and Computation*, vol. 216, no. 5, pp. 1355–1362, 2010.

[14] Z.-Y. Peng, S. M. El-Sayed, and X.-l. Zhang, "Iterative methods for the extremal positive definite solution of the matrix equation $X + A^*X^{-\alpha}A = Q$," *Journal of Computational and Applied Mathematics*, vol. 200, no. 2, pp. 520–527, 2007.

[15] B. R. Fang, J. D. Zhou, and Y. M. Li, *Matrix Theory*, Tsinghua University Press, Beijing, China, 2006.

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2012, Article ID 483624, 11 pages
doi:10.1155/2012/483624

Research Article

Refinements of Kantorovich Inequality for Hermitian Matrices

Feixiang Chen

School of Mathematics and Statistics, Chongqing Three Gorges University, Wanzhou, Chongqing 404000, China

Correspondence should be addressed to Feixiang Chen, cfx2002@126.com

Received 28 August 2012; Accepted 5 November 2012

Academic Editor: K. C. Sivakumar

Some new Kantorovich-type inequalities for Hermitian matrix are proposed in this paper. We consider what happens to these inequalities when the positive definite matrix is allowed to be invertible and provides refinements of the classical results.

1. Introduction and Preliminaries

We first state the well-known Kantorovich inequality for a positive definite Hermite matrix (see [1, 2]), let $A \in M_n$ be a positive definite Hermitian matrix with real eigenvalues $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$. Then

$$1 \le x^* A x x^* A^{-1} x \le \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1 \lambda_n}, \tag{1.1}$$

for any $x \in \mathbb{C}^n$, $\|x\| = 1$, where A^* denotes the conjugate transpose of matrix A. A matrix $A \in M_n$ is Hermitian if $A = A^*$. An equivalent form of this result is incorporated in

$$0 \le x^* A x x^* A^{-1} x - 1 \le \frac{(\lambda_n - \lambda_1)^2}{4\lambda_1 \lambda_n}, \tag{1.2}$$

for any $x \in \mathbb{C}^n$, $\|x\| = 1$.

Attributed to Kantorovich, the inequality has built up a considerable literature. This typically comprises generalizations. Examples are [3–5] for matrix versions. Operator versions are developed in [6, 7]. Multivariate versions have been useful in statistics to assess the robustness of least squares, see [8, 9] and the references therein.

Due to the important applications of the original Kantorovich inequality for matrices [10] in Statistics [8, 11, 12] and Numerical Analysis [13, 14], any new inequality of this type will have a flow of consequences in the areas of applications.

Motivated by the interest in both pure and applied mathematics outlined above we establish in this paper some improvements of Kantorovich inequalities. The classical Kantorovich-type inequalities are modified to apply not only to positive definite but also to invertible Hermitian matrices. As natural tools in deriving the new results, the recent Grüss-type inequalities for vectors in inner product in [6, 15–19] are utilized.

To simplify the proof, we first introduce some lemmas.

2. Lemmas

Let B be a Hermitian matrix with real eigenvalues $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_n$, if $A - B$ is positive semidefinite, we write

$$A \geq B, \tag{2.1}$$

that is, $\lambda_i \geq \mu_i$, $i = 1, 2, \ldots, n$. On $\mathbb{C}^n$, we have the standard inner product defined by $\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y}_i$, where $x = (x_1, \ldots, x_n)^* \in \mathbb{C}^n$ and $y = (y_1, \ldots, y_n)^* \in \mathbb{C}^n$.

Lemma 2.1. *Let a, b, c, and d be real numbers, then one has the following inequality:*

$$\left(a^2 - b^2\right)\left(c^2 - d^2\right) \leq (ac - bd)^2. \tag{2.2}$$

Lemma 2.2. *Let A and B be Hermitian matrices, if $AB = BA$, then*

$$AB \leq \frac{(A + B)^2}{4}. \tag{2.3}$$

Lemma 2.3. *Let $A \geq 0$, $B \geq 0$, if $AB = BA$, then*

$$AB \geq 0. \tag{2.4}$$

3. Some Results

The following lemmas can be obtained from [16–19] by replacing Hilbert space $(H, \langle\cdot,\cdot\rangle)$ with inner product spaces $\mathbb{C}^n$, so we omit the details.

Lemma 3.1. *Let u, v, and e be vectors in $\mathbb{C}^n$, and $\|e\| = 1$. If α, β, δ, and γ are real or complex numbers such that*

$$\operatorname{Re}\langle \beta e - u, u - \alpha e\rangle \geq 0, \qquad \operatorname{Re}\langle \delta e - v, v - \gamma e\rangle \geq 0, \tag{3.1}$$

then

$$|\langle u,v\rangle - \langle u,e\rangle\langle e,v\rangle| \le \frac{1}{4}(\beta-\alpha)(\delta-\gamma) - \left[\operatorname{Re}\langle \beta e - u, u - \alpha e\rangle \operatorname{Re}\langle \delta e - v, v - \gamma e\rangle\right]^{1/2}. \quad (3.2)$$

Lemma 3.2. *With the assumptions in Lemma 3.1, one has*

$$|\langle u,v\rangle - \langle u,e\rangle\langle e,v\rangle| \le \frac{1}{4}(\beta-\alpha)(\gamma-\delta) - \left|\langle u,e\rangle - \frac{\alpha+\beta}{2}\right|\left|\langle v,e\rangle - \frac{\gamma+\delta}{2}\right|. \quad (3.3)$$

Lemma 3.3. *With the assumptions in Lemma 3.1, if* $\operatorname{Re}(\beta\overline{\alpha}) \ge 0$, $\operatorname{Re}(\delta\overline{\gamma}) \ge 0$, *one has*

$$|\langle u,v\rangle - \langle u,e\rangle\langle e,v\rangle| \le \frac{|\beta-\alpha||\delta-\gamma|}{4\left[\operatorname{Re}(\beta\overline{\alpha})\operatorname{Re}(\delta\overline{\gamma})\right]^{1/2}}|\langle u,e\rangle\langle e,v\rangle|. \quad (3.4)$$

Lemma 3.4. *With the assumptions in Lemma 3.3, one has*

$$\begin{aligned} &|\langle u,v\rangle - \langle u,e\rangle\langle e,v\rangle| \\ &\quad \le \left\{\left(|\alpha+\beta| - 2\left[\operatorname{Re}(\beta\overline{\alpha})\right]^{1/2}\right)\left(|\delta+\gamma| - 2\left[\operatorname{Re}(\delta\overline{\gamma})\right]^{1/2}\right)\right\}^{1/2}\left[|\langle u,e\rangle\langle e,v\rangle|\right]^{1/2}. \end{aligned} \quad (3.5)$$

4. New Kantorovich Inequalities for Hermitian Matrices

For a Hermitian matrix A, as in [6], we define the following transform:

$$C(A) = (\lambda_n I - A)(A - \lambda_1 I). \quad (4.1)$$

When A is invertible, if $\lambda_1\lambda_n > 0$, then,

$$C\left(A^{-1}\right) = \left(\frac{1}{\lambda_1}I - A^{-1}\right)\left(A^{-1} - \frac{1}{\lambda_n}I\right). \quad (4.2)$$

Otherwise, $\lambda_1\lambda_n < 0$, then,

$$C\left(A^{-1}\right) = \left(\frac{1}{\lambda_{k+1}}I - A^{-1}\right)\left(A^{-1} - \frac{1}{\lambda_k}I\right), \quad (4.3)$$

where

$$\lambda_1 \le \cdots \le \lambda_k < 0 < \lambda_{k+1} \le \cdots \le \lambda_n. \quad (4.4)$$

From Lemma 2.3 we can conclude that $C(A) \ge 0$ and $C(A^{-1}) \ge 0$.

For two Hermitian matrices A and B, and $x \in \mathbb{C}^n$, $\|x\| = 1$, we define the following functional:

$$G(A, B; x) = \langle Ax, Bx \rangle - \langle Ax, x \rangle \langle x, Bx \rangle. \tag{4.5}$$

When $A = B$, we denote

$$G(A; x) = \|Ax\|^2 - \langle Ax, x \rangle^2, \tag{4.6}$$

for $x \in \mathbb{C}^n$, $\|x\| = 1$.

Lemma 4.1. *With notations above, and for $x \in \mathbb{C}^n$, $\|x\| = 1$, then*

$$0 \le \langle C(A)x, x \rangle \le \frac{(\lambda_n - \lambda_1)^2}{4}, \tag{4.7}$$

if $\lambda_n \lambda_1 > 0$,

$$0 \le \left\langle C\left(A^{-1}\right)x, x \right\rangle \le \frac{(\lambda_n - \lambda_1)^2}{4(\lambda_n \lambda_1)^2}, \tag{4.8}$$

if $\lambda_n \lambda_1 < 0$,

$$0 \le \left\langle C\left(A^{-1}\right)x, x \right\rangle \le \frac{(\lambda_{k+1} - \lambda_k)^2}{4(\lambda_{k+1} \lambda_k)^2}. \tag{4.9}$$

Proof. From $C(A) \ge 0$, then

$$\langle C(A)x, x \rangle \ge 0. \tag{4.10}$$

While, from Lemma 2.2, we can get

$$C(A) = (\lambda_n I - A)(A - \lambda_1 I) \le \frac{(\lambda_n - \lambda_1)^2}{4} I. \tag{4.11}$$

Then $\langle C(A)x, x \rangle \le (\lambda_n - \lambda_1)^2 / 4$ is straightforward. The proof for $C(A^{-1})$ is similar. □

Lemma 4.2. *With notations above, and for $x \in \mathbb{C}^n$, $\|x\| = 1$, then*

$$\left| x^* A x x^* A^{-1} x - 1 \right|^2 \le G(A; x) G\left(A^{-1}; x\right). \tag{4.12}$$

Proof. Thus,

$$\begin{aligned}\left|x^*Axx^*A^{-1}x-1\right|^2 &= \left|x^*\left(\left(x^*A^{-1}x\right)I-A^{-1}\right)((x^*Ax)I-A)x\right|^2\\ &\le \left\|\left(\left(x^*A^{-1}x\right)I-A^{-1}\right)x\right\|^2\|((x^*Ax)I-A)x\|^2,\end{aligned} \tag{4.13}$$

while

$$\begin{aligned}\|((x^*Ax)I-A)x\|^2 &= x^*\left((x^*Ax)^2I-2(x^*Ax)A+A^2\right)x\\ &= x^*A^2x-(x^*Ax)^2\\ &= \|Ax\|^2-\langle Ax,x\rangle^2\\ &= G(A;x).\end{aligned} \tag{4.14}$$

Similarly, we can get $\|((x^*A^{-1}x)I-A^{-1})x\|^2 = G(A^{-1};x)$, then we complete the proof. □

Theorem 4.3. *Let A, B be two Hermitian matrices, and $C(A)\ge 0$, $C(B)\ge 0$ are defined as above, then*

$$\begin{aligned}|G(A,B;x)| &\le \frac{1}{4}(\lambda_n-\lambda_1)(\mu_n-\mu_1)-[\langle C(A)x,x\rangle\langle C(B)x,x\rangle]^{1/2},\\ |G(A,B;x)| &\le \frac{1}{4}(\lambda_n-\lambda_1)(\mu_n-\mu_1)-\left|\left\langle\left(A-\frac{\lambda_1+\lambda_n}{2}\right)x,x\right\rangle\right|\left|\left\langle\left(B-\frac{\mu_1+\mu_n}{2}\right)x,x\right\rangle\right|,\end{aligned} \tag{4.15}$$

for any $x\in\mathbb{C}^n$, $\|x\|=1$.

If $\lambda_1\lambda_n>0$, $\mu_1\mu_n>0$, then

$$\begin{aligned}|G(A,B;x)| &\le \frac{(\lambda_n-\lambda_1)(\mu_n-\mu_1)}{4[(\lambda_n\lambda_1)(\mu_n\mu_1)]^{1/2}}|\langle Ax,x\rangle\langle Bx,x\rangle|,\\ |G(A,B;x)| &\le \left\{\left(|\lambda_1+\lambda_n|-2(\lambda_1\lambda_n)^{1/2}\right)\left(|\mu_1+\mu_n|-2(\mu_1\mu_n)^{1/2}\right)\right\}^{1/2}[|\langle Ax,x\rangle\langle Bx,x\rangle|]^{1/2},\end{aligned} \tag{4.16}$$

for any $x\in\mathbb{C}^n$, $\|x\|=1$.

Proof. The proof follows by Lemmas 3.1, 3.2, 3.3, and 3.4 on choosing $u=Ax$, $v=Bx$, and $e=x$, $\beta=\lambda_n$, $\alpha=\lambda_1$, $\delta=\mu_n$, and $\gamma=\mu_1$, $x\in\mathbb{C}^n$, $\|x\|=1$, respectively. □

Corollary 4.4. *Let A be a Hermitian matrices, and $C(A) \geq 0$ is defined as above, then*

$$|G(A;x)| \leq \frac{1}{4}(\lambda_n - \lambda_1)^2 - \langle C(A)x, x\rangle, \tag{4.17}$$

$$|G(A;x)| \leq \frac{1}{4}(\lambda_n - \lambda_1)^2 - \left|\left\langle\left(A - \frac{\lambda_1 + \lambda_n}{2}I\right)x, x\right\rangle\right|^2, \tag{4.18}$$

for any $x \in \mathbb{C}^n$, $\|x\| = 1$.

If $\lambda_1\lambda_n > 0$, then

$$|G(A;x)| \leq \frac{(\lambda_n - \lambda_1)^2}{4(\lambda_n\lambda_1)}|\langle Ax, x\rangle|^2, \tag{4.19}$$

$$|G(A;x)| \leq \left(|\lambda_1 + \lambda_n| - 2\sqrt{\lambda_1\lambda_n}\right)|\langle Ax, x\rangle|, \tag{4.20}$$

for any $x \in \mathbb{C}^n$, $\|x\| = 1$.

Proof. The proof follows by Theorem 4.3 on choosing $A = B$, respectively. □

Corollary 4.5. *Let A be a Hermitian matrices and $C(A^{-1}) \geq 0$ is defined as above, then one has the following.*

If $\lambda_1\lambda_n > 0$, then

$$\left|G\left(A^{-1};x\right)\right| \leq \frac{(\lambda_n - \lambda_1)^2}{4(\lambda_n\lambda_1)^2} - \left\langle C\left(A^{-1}\right)x, x\right\rangle, \tag{4.21}$$

$$\left|G\left(A^{-1};x\right)\right| \leq \frac{(\lambda_n - \lambda_1)^2}{4(\lambda_n\lambda_1)^2} - \left|\left\langle\left(A^{-1} - \frac{\lambda_1 + \lambda_n}{2\lambda_n\lambda_1}I\right)x, x\right\rangle\right|^2, \tag{4.22}$$

$$\left|G\left(A^{-1};x\right)\right| \leq \frac{(\lambda_n - \lambda_1)^2}{4\lambda_n\lambda_1}\left|\left\langle A^{-1}x, x\right\rangle\right|^2, \tag{4.23}$$

$$\left|G\left(A^{-1};x\right)\right| \leq \left(\frac{|\lambda_1 + \lambda_n|}{\lambda_1\lambda_n} - 2\frac{1}{\sqrt{\lambda_1\lambda_n}}\right)\left|\left\langle A^{-1}x, x\right\rangle\right|, \tag{4.24}$$

for any $x \in \mathbb{C}^n$, $\|x\| = 1$.

If $\lambda_1\lambda_n < 0$, then

$$\left|G\left(A^{-1};x\right)\right| \leq \frac{(\lambda_k - \lambda_{k+1})^2}{4(\lambda_k\lambda_{k+1})^2} - \left\langle C\left(A^{-1}\right)x, x\right\rangle, \tag{4.25}$$

where $C(A^{-1}) = ((1/\lambda_{k+1})I - A^{-1})(A^{-1} - (1/\lambda_k)I)$, *and*

$$\left|G\left(A^{-1};x\right)\right| \leq \frac{(\lambda_k - \lambda_{k+1})^2}{4(\lambda_k\lambda_{k+1})^2} - \left|\left\langle\left(A^{-1} - \frac{\lambda_{k+1}+\lambda_k}{2\lambda_k\lambda_{k+1}}I\right)x, x\right\rangle\right|^2, \tag{4.26}$$

for any $x \in \mathbb{C}^n$, $\|x\| = 1$.

Proof. The proof follows by Corollary 4.4 by replacing A with A^{-1}, respectively. □

Theorem 4.6. *Let A be an $n\times n$ invertible Hermitian matrix with real eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$, then one has the following.*

If $\lambda_1\lambda_n > 0$, then

$$\left|x^*Axx^*A^{-1}x - 1\right| \leq \frac{(\lambda_n - \lambda_1)^2}{4(\lambda_n\lambda_1)} - \sqrt{\langle C(A)x, x\rangle\langle C(A^{-1})x, x\rangle}, \tag{4.27}$$

$$\left|x^*Axx^*A^{-1}x - 1\right| \leq \frac{(\lambda_n - \lambda_1)^2}{4(\lambda_n\lambda_1)} - \left|\left\langle\left(A - \frac{\lambda_1+\lambda_n}{2}I\right)x, x\right\rangle\right|\left|\left\langle\left(A^{-1} - \frac{\lambda_1+\lambda_n}{2\lambda_n\lambda_1}I\right)x, x\right\rangle\right|, \tag{4.28}$$

$$\left|x^*Axx^*A^{-1}x - 1\right| \leq \frac{(\lambda_n - \lambda_1)^2}{4(\lambda_n\lambda_1)^2}\langle Ax, x\rangle\left\langle A^{-1}x, x\right\rangle, \tag{4.29}$$

$$\left|x^*Axx^*A^{-1}x - 1\right| \leq \frac{\left(\sqrt{|\lambda_n|} - \sqrt{|\lambda_1|}\right)^2}{\sqrt{\lambda_n\lambda_1}}\sqrt{\langle Ax, x\rangle\langle A^{-1}x, x\rangle}, \tag{4.30}$$

for any $x \in \mathbb{C}^n$, $\|x\| = 1$.

If $\lambda_1\lambda_n < 0$, then

$$\left|x^*Axx^*A^{-1}x - 1\right| \leq \frac{(\lambda_n - \lambda_1)(\lambda_{k+1} - \lambda_k)}{4|\lambda_k\lambda_{k+1}|} - \sqrt{\langle C(A)x, x\rangle\langle C(A^{-1})x, x\rangle}, \tag{4.31}$$

where $C(A^{-1}) = ((1/\lambda_{k+1})I - A^{-1})(A^{-1} - (1/\lambda_k)I)$,

$$\begin{aligned}\left|x^*Axx^*A^{-1}x - 1\right| \leq{}& \frac{(\lambda_n - \lambda_1)(\lambda_{k+1} - \lambda_k)}{4|\lambda_k\lambda_{k+1}|} \\ &- \left|\left\langle\left(A - \frac{\lambda_1+\lambda_n}{2}I\right)x, x\right\rangle\right|\left|\left\langle\left(A^{-1} - \frac{\lambda_k+\lambda_{k+1}}{2\lambda_1\lambda_{k+1}}I\right)x, x\right\rangle\right|.\end{aligned} \tag{4.32}$$

Proof. Considering

$$\left|x^*Axx^*A^{-1}x - 1\right|^2 \leq G(A;x)G\left(A^{-1};x\right). \tag{4.33}$$

When $\lambda_1\lambda_n > 0$, from (4.17) and (4.21), we get

$$\left|x^*Axx^*A^{-1}x-1\right|^2 \leq \left\{\frac{1}{4}(\lambda_n-\lambda_1)^2-\langle C(A)x,x\rangle\right\}\left\{\frac{(\lambda_n-\lambda_1)^2}{4(\lambda_n\lambda_1)^2}-\left\langle C\left(A^{-1}\right)x,x\right\rangle\right\}. \tag{4.34}$$

From $(a^2-b^2)(c^2-d^2) \leq (ac-bd)^2$, we have

$$\left|x^*Axx^*A^{-1}x-1\right|^2 \leq \left\{\frac{(\lambda_n-\lambda_1)^2}{4(\lambda_n\lambda_1)}-\sqrt{[\langle C(A)x,x\rangle\langle C(A^{-1})x,x\rangle]}\right\}^2, \tag{4.35}$$

then, the conclusion (4.27) holds.

Similarly, from (4.18) and (4.22), we get

$$\begin{aligned}\left|x^*Axx^*A^{-1}x-1\right|^2 &\leq \left\{\frac{1}{4}(\lambda_n-\lambda_1)^2-\left|\left\langle\left(A-\frac{\lambda_1+\lambda_n}{2}I\right)x,x\right\rangle\right|^2\right\}\\ &\quad\times\left\{\frac{(\lambda_n-\lambda_1)^2}{4(\lambda_n\lambda_1)^2}-\left|\left\langle\left(A^{-1}-\frac{\lambda_1+\lambda_n}{2\lambda_n\lambda_1}I\right)x,x\right\rangle\right|^2\right\}\\ &\leq \left\{\frac{(\lambda_n-\lambda_1)^2}{4(\lambda_n\lambda_1)}-\left|\left\langle\left(A-\frac{\lambda_1+\lambda_n}{2}I\right)x,x\right\rangle\right|\left|\left\langle\left(A^{-1}-\frac{\lambda_1+\lambda_n}{2\lambda_n\lambda_1}I\right)x,x\right\rangle\right|\right\}^2,\end{aligned} \tag{4.36}$$

then, the conclusion (4.28) holds.

From (4.19) and (4.23), we get

$$\left|x^*Axx^*A^{-1}x-1\right|^2 \leq \frac{(\lambda_n-\lambda_1)^2}{4(\lambda_n\lambda_1)}|\langle Ax,x\rangle|^2\frac{(\lambda_n-\lambda_1)^2}{4(\lambda_n\lambda_1)}\left|\left\langle A^{-1}x,x\right\rangle\right|^2, \tag{4.37}$$

then, the conclusion (4.29) holds.

From (4.20) and (4.24), we get

$$\left|x^*Axx^*A^{-1}x-1\right|^2 \leq \left(|\lambda_1+\lambda_n|-2\sqrt{\lambda_1\lambda_n}\right)|\langle Ax,x\rangle|\left(\frac{|\lambda_1+\lambda_n|}{\lambda_1\lambda_n}-2\frac{1}{\sqrt{\lambda_1\lambda_n}}\right)\left|\left\langle A^{-1}x,x\right\rangle\right|, \tag{4.38}$$

then, the conclusion (4.30) holds.

When $\lambda_1\lambda_n < 0$, from (4.17) and (4.25), we get

$$\left|x^*Axx^*A^{-1}x-1\right|^2 \leq \left\{\frac{1}{4}(\lambda_n-\lambda_1)^2-\langle C(A)x,x\rangle\right\}\left\{\frac{(\lambda_k-\lambda_{k+1})^2}{4(\lambda_k\lambda_{k+1})^2}-\left\langle C\left(A^{-1}\right)x,x\right\rangle\right\}, \tag{4.39}$$

then, the conclusion (4.31) holds.

From (4.18) and (4.26), we get

$$\begin{aligned}\left|x^* A x x^* A^{-1} x-1\right|^2 \leq &\left\{\frac{1}{4}(\lambda_n-\lambda_1)^2-\left|\left\langle\left(A-\frac{\lambda_1+\lambda_n}{2} I\right) x, x\right\rangle\right|^2\right\} \\ &\times\left\{\frac{(\lambda_k-\lambda_{k+1})^2}{4(\lambda_k \lambda_{k+1})^2}-\left|\left\langle\left(A^{-1}-\frac{\lambda_{k+1}+\lambda_k}{2 \lambda_k \lambda_{k+1}} I\right) x, x\right\rangle\right|^2\right\} \\ \leq &\left\{\frac{(\lambda_n-\lambda_1)(\lambda_{k+1}-\lambda_k)}{4|\lambda_k \lambda_{k+1}|}\right. \\ &\left.-\left|\left\langle\left(A-\frac{\lambda_1+\lambda_n}{2} I\right) x, x\right\rangle\right|\left|\left\langle\left(A^{-1}-\frac{\lambda_k+\lambda_{k+1}}{2 \lambda_1 \lambda_{k+1}} I\right) x, x\right\rangle\right|\right\}^2,\end{aligned} \tag{4.40}$$

then, the conclusion (4.32) holds. □

Corollary 4.7. *With the notations above, for any $x \in \mathbb{C}^n$, $\|x\|=1$, one lets*

$$\begin{aligned}|G_1(A;x)| &= \frac{1}{4}(\lambda_n-\lambda_1)^2-\langle C(A)x,x\rangle, \\ |G_2(A;x)| &= \frac{1}{4}(\lambda_n-\lambda_1)^2-\left|\left\langle\left(A-\frac{\lambda_1+\lambda_n}{2} I\right) x, x\right\rangle\right|^2.\end{aligned} \tag{4.41}$$

If $\lambda_1\lambda_n > 0$, one lets

$$\begin{aligned}|G_3(A;x)| &= \frac{(\lambda_n-\lambda_1)^2}{4(\lambda_n\lambda_1)}|\langle Ax,x\rangle|^2, \\ |G_4(A;x)| &= \left(|\lambda_1+\lambda_n|-2\sqrt{\lambda_1\lambda_n}\right)|\langle Ax,x\rangle|.\end{aligned} \tag{4.42}$$

If $\lambda_1\lambda_n > 0$

$$\begin{aligned}\left|G^1\left(A^{-1};x\right)\right| &= \frac{(\lambda_n-\lambda_1)^2}{4(\lambda_n\lambda_1)^2}-\left\langle C\left(A^{-1}\right)x,x\right\rangle, \\ \left|G^2\left(A^{-1};x\right)\right| &= \frac{(\lambda_n-\lambda_1)^2}{4(\lambda_n\lambda_1)^2}-\left|\left\langle\left(A^{-1}-\frac{\lambda_1+\lambda_n}{2\lambda_n\lambda_1} I\right) x, x\right\rangle\right|^2, \\ \left|G^3\left(A^{-1};x\right)\right| &= \frac{(\lambda_n-\lambda_1)^2}{4\lambda_n\lambda_1}\left|\left\langle A^{-1}x,x\right\rangle\right|^2, \\ \left|G^4\left(A^{-1};x\right)\right| &= \left(\frac{|\lambda_1+\lambda_n|}{\lambda_1\lambda_n}-2\frac{1}{\sqrt{\lambda_1\lambda_n}}\right)\left|\left\langle A^{-1}x,x\right\rangle\right|.\end{aligned} \tag{4.43}$$

If $\lambda_1\lambda_n < 0$, *then*

$$\left|G^5\left(A^{-1};x\right)\right| = \frac{(\lambda_k - \lambda_{k+1})^2}{4(\lambda_k\lambda_{k+1})^2} - \left\langle C\left(A^{-1}\right)x,x\right\rangle, \tag{4.44}$$

where $C(A^{-1}) = ((1/\lambda_{k+1})I - A^{-1})(A^{-1} - (1/\lambda_k)I)$, *and*

$$\left|G^6\left(A^{-1};x\right)\right| = \frac{(\lambda_k - \lambda_{k+1})^2}{4(\lambda_k\lambda_{k+1})^2} - \left|\left\langle\left(A^{-1} - \frac{\lambda_{k+1}+\lambda_k}{2\lambda_k\lambda_{k+1}}I\right)x,x\right\rangle\right|^2. \tag{4.45}$$

Then, one has the following.

If $\lambda_1\lambda_n > 0$,

$$\left|x^*Axx^*A^{-1}x - 1\right| \le \sqrt{G^*(A;x)G^*(A^{-1};x)}, \tag{4.46}$$

where

$$\begin{gathered}
G^*(A;x) = \min\{G_1(A;x), G_2(A;x), G_3(A;x), G_4(A;x)\},\\
G^*(A;x) = \min\left\{G^1\left(A^{-11};x\right), G^2\left(A^{-1};x\right), G^3\left(A^{-1};x\right), G^4\left(A^{-1};x\right)\right\}.
\end{gathered} \tag{4.47}$$

If $\lambda_1\lambda_n < 0$

$$\left|x^*Axx^*A^{-1}x - 1\right| \le \sqrt{G^\circ(A;x)G^\bullet(A^{-1};x)}, \tag{4.48}$$

where

$$\begin{gathered}
G^\circ(A;x) = \min\{G_1(A;x), G_2(A;x)\},\\
G^\bullet(A;x) = \min\left\{G^5\left(A^{-1};x\right), G^6\left(A^{-1};x\right)\right\}.
\end{gathered} \tag{4.49}$$

Proof. The proof follows from that the conclusions in Corollaries 4.4 and 4.5 are independent. □

Remark 4.8. It is easy to see that if $\lambda_1 > 0$, $\lambda_n > 0$, our result coincides with the inequality of operator versions in [6]. So we conclude that our results give an improvement of the Kantorovich inequality [6] that applies to all invertible Hermite matrices.

5. Conclusion

In this paper, we introduce some new Kantorovich-type inequalities for the invertible Hermitian matrices. Inequalities (4.27) and (4.31) are the same as [4], but our proof is simple. In Theorem 4.6, if $\lambda_1 > 0$, $\lambda_n > 0$, the results are similar to the well-known Kantorovich-type inequalities for operators in [6]. Moreover, for any invertible Hermitian matrix, there exists a similar inequality.

Acknowledgments

The authors are grateful to the associate editor and two anonymous referees whose comments have helped to improve the final version of the present work. This work was supported by the Natural Science Foundation of Chongqing Municipal Education Commission (no. KJ091104).

References

[1] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
[2] A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, NY, USA, 1964.
[3] S. Liu and H. Neudecker, "Several matrix Kantorovich-type inequalities," *Journal of Mathematical Analysis and Applications*, vol. 197, no. 1, pp. 23–26, 1996.
[4] A. W. Marshall and I. Olkin, "Matrix versions of the Cauchy and Kantorovich inequalities," *Aequationes Mathematicae*, vol. 40, no. 1, pp. 89–93, 1990.
[5] J. K. Baksalary and S. Puntanen, "Generalized matrix versions of the Cauchy-Schwarz and Kantorovich inequalities," *Aequationes Mathematicae*, vol. 41, no. 1, pp. 103–110, 1991.
[6] S. S. Dragomir, "New inequalities of the Kantorovich type for bounded linear operators in Hilbert spaces," *Linear Algebra and its Applications*, vol. 428, no. 11-12, pp. 2750–2760, 2008.
[7] P. G. Spain, "Operator versions of the Kantorovich inequality," *Proceedings of the American Mathematical Society*, vol. 124, no. 9, pp. 2813–2819, 1996.
[8] S. Liu and H. Neudecker, "Kantorovich inequalities and efficiency comparisons for several classes of estimators in linear models," *Statistica Neerlandica*, vol. 51, no. 3, pp. 345–355, 1997.
[9] P. Bloomfield and G. S. Watson, "The inefficiency of least squares," *Biometrika*, vol. 62, pp. 121–128, 1975.
[10] L. V. Kantorovič, "Functional analysis and applied mathematics," *Uspekhi Matematicheskikh Nauk*, vol. 3, no. 6(28), pp. 89–185, 1948 (Russian).
[11] C. G. Khatri and C. R. Rao, "Some extensions of the Kantorovich inequality and statistical applications," *Journal of Multivariate Analysis*, vol. 11, no. 4, pp. 498–505, 1981.
[12] C. T. Lin, "Extrema of quadratic forms and statistical applications," *Communications in Statistics A*, vol. 13, no. 12, pp. 1517–1520, 1984.
[13] A. Galántai, "A study of Auchmuty's error estimate," *Computers & Mathematics with Applications*, vol. 42, no. 8-9, pp. 1093–1102, 2001.
[14] P. D. Robinson and A. J. Wathen, "Variational bounds on the entries of the inverse of a matrix," *IMA Journal of Numerical Analysis*, vol. 12, no. 4, pp. 463–486, 1992.
[15] S. S. Dragomir, "A generalization of Grüss's inequality in inner product spaces and applications," *Journal of Mathematical Analysis and Applications*, vol. 237, no. 1, pp. 74–82, 1999.
[16] S. S. Dragomir, "Some Grüss type inequalities in inner product spaces," *Journal of Inequalities in Pure and Applied Mathematics*, vol. 4, no. 2, article 42, 2003.
[17] S. S. Dragomir, "On Bessel and Grüss inequalities for orthonormal families in inner product spaces," *Bulletin of the Australian Mathematical Society*, vol. 69, no. 2, pp. 327–340, 2004.
[18] S. S. Dragomir, "Reverses of Schwarz, triangle and Bessel inequalities in inner product spaces," *Journal of Inequalities in Pure and Applied Mathematics*, vol. 5, no. 3, article 76, 2004.
[19] S. S. Dragomir, "Reverses of the Schwarz inequality generalising a Klamkin-McLenaghan result," *Bulletin of the Australian Mathematical Society*, vol. 73, no. 1, pp. 69–78, 2006.
[20] Z. Liu, L. Lu, and K. Wang, "On several matrix Kantorovich-type inequalities," *Journal of Inequalities and Applications*, vol. 2010, Article ID 571629, 5 pages, 2010.