Konuralp Journal of Mathematics, 12 (2) (2024) 99-105

Konuralp Journal of Mathematics

Research Paper

Journal Homepage: *www.dergipark.gov.tr/konuralpjournalmath*

e-ISSN: 2147-625X

Solution of a Sylvester Quaternion Matrix Equation by a Block Krylov Subspace Method

Gökay Karabacak[1*] and Neslihan Bektaşoğlu[1]

[1]*Department of Mathematics, Faculty of Science and Arts, Kırklareli University, Kırklareli, Turkey*
[*]*Corresponding author*

Abstract

We focus on the solution of the Sylvester quaternion matrix equation $AX - XB = C$, where $A \in \mathbb{H}^{m\times m}$, $B \in \mathbb{H}^{n\times n}$, $C \in \mathbb{H}^{m\times n}$ and m is very large such that $m \gg n$. Non-commutative nature of quaternion scalars under multiplication is a hurdle in solving such a matrix equation. Thus, instead of directly dealing with the quaternion matrix equation, we make use of the complex matrix representations of quaternion matrices, and turn the quaternion matrix equation into a complex matrix equation of size twice as big. Since the resulting complex matrix equation involves large matrices, assuming m is large, in particular $m \gg n$, we present a block Generalized Minimal Residual (GMRES) method that seeks the solution of the complex matrix equation in small affine spaces defined in terms of Krylov subspaces. The solution in such a small affine space can equivalently be posed as the solution of a small complex matrix equation, which can be solved directly, for instance, by rewriting it as a linear system. At every iteration of our block GMRES method, the Krylov subspaces are expanded by adding new vectors, and the small complex matrix equations are altered accordingly. Our block GMRES method eventually produces the complex representation of an approximate solution of the original Sylvester quaternion matrix equation. Finally, this complex matrix representation is transformed back into the corresponding quaternion matrix, which is an approximate solution of the original quaternion matrix equation $AX - XB = C$.

Keywords: *Block Arnoldi process; Block GMRES method; Complex representation; Krylov subspace; Sylvester quaternion matrix equation*
2010 Mathematics Subject Classification: *15A06; 15A24; 15B33; 65F10*

1. Introduction

Matrix equations are mathematical tools that arise when modelling real-life situations. More specifically, many problems that appear in applied sciences can be expressed in terms of linear or non-linear matrix equations. A Sylvester matrix equation is a particular linear matrix equation of the form

$$AX - XB = C, \tag{1.1}$$

where A, B, C are known matrices of appropriate sizes, and X is an unknown matrix of appropriate size. The homogeneous version of a Sylvester matrix equation was first introduced by J. J. Sylvester in 1884 [15]. In the symmetric case, that is when $B = A^T$ and $C = C^T$, equation (1.1) is called a Lyapunov equation, used in the stability analysis of dynamical systems, as well as designing controllers for them [6]. Sylvester matrix equations and its variations, as well as special cases arise in various fields, including control theory, model reduction, matrix decompositions, image processing, and numerical ordinary differential equations; for details, see [29].

In the literature, the solution of a real or a complex Sylvester matrix equation has been widely studied by means of either direct approaches or iterative methods; see, e.g., [3, 4, 5, 9, 12, 17, 20, 21, 23, 32]. In this study, we attempt to solve a Sylvester matrix equation as in (1.1) for $m\times m$, $n\times n$, $m\times n$ quaternion matrices $A \in \mathbb{H}^{m\times m}$, $B \in \mathbb{H}^{n\times n}$, $C \in \mathbb{H}^{m\times n}$, respectively, assuming A is large and sparse, as well as $m \gg n$. Quaternions, discovered by W. R. Hamilton in 1843, are generalizations of real and complex numbers [30]. The multiplication of two quaternion numbers is not commutative, which makes them a skew-field rather than a field. The quaternion matrix equation in (1.1) that we deal with here has applications in various fields, including computer science and signal processing [16].

Motivated by the applications in the fields listed above, solving (1.1) and its variations over the quaternion skew-field has recently attracted attention. There are several approaches in the literature for solving these type of matrix equations [1, 2, 7, 8, 10, 11, 13, 14, 18, 19, 22, 24, 25, 26, 27, 28, 31, 33, 34, 35, 36]. One of the most notable among these approaches is the transformation of the quaternion matrix equations into a real or a complex matrix equation by using the real or complex representations of quaternion matrices [7, 11, 19, 22, 27]. After performing such transformations, some researchers have derived necessary and sufficient conditions for the solvability of Sylvester-type

Email addresses: gokayykarabacak@gmail.com (Gökay Karabacak), neslihanbektas07@gmail.com (Neslihan Bektaşoğlu)

matrix equations or a system consisting of Sylvester-type matrix equations. Additionally, they have presented expressions for general or special solutions to these matrix equations using the Moore-Penrose pseudoinverse [11, 19]. Some other researchers have dealt with the solution of the equation by using an iterative method after transforming the quaternion equation into its real or complex counterpart [7, 11, 22, 27]. In contrast to these approaches, it is also possible to solve the quaternion matrix equation without transforming it to a real or a complex matrix equation [8, 10, 13, 18, 24, 25, 26, 31, 33, 34, 35, 36]. In this direction, Kyrchei has derived explicit expressions in terms of determinants for the solutions (analogous to Cramer's Rule) of a two-sided quaternion generalized Sylvester matrix equation [13]. Beik and Ahmadi-Asl have developed conjugate gradient least-squares (CGLS) methods for finding the η-Hermitian and η-anti-Hermitian solutions of the least-squares problem associated with the quaternion matrix equation $AXB+CYD=E$ by making use of Sylvester operators and a real inner product defined over the skew-field of quaternions [10]. Şimşek has provided explicit formulas for the general solutions, as well as the perhermitian, skew-perhermitian solutions of the least-squares problems associated with $AXB+CYD=E$ and $AXB+CXD=E$ over the skew-field of quaternions in terms of Kronecker products and Moore-Penrose pseudoinverse [24]. In this last study, applications of the derived formulas to color image restoration are also presented. In another study, Şimsek has proposed an approach for the numerical solution of a large-scale Sylvester matrix equation by means of a global GMRES method based on a real inner product defined on the space of quaternion matrices [26]. In addition to these studies, some studies attempt to solve Sylvester quaternion matrix equations by employing matrix decompositions [33, 36]; for example, in [36] the solutions of systems of Sylvester-type matrix equations are obtained by computing simultaneous decompositions of quaternion coefficient matrices of the matrix equations. In this study, we solve Sylvester quaternion matrix equations by exploiting their complex representations, particularly by applying the well-known block GMRES method in the complex setting to these complex representations. Thus, we first transform the quaternion matrix equation (1.1) into a complex matrix equation by utilizing complex representations of quaternion matrices. This transformation allows us to work fully in the complex algebra. However, the sizes of the matrices in the resulting complex matrix equation are twice as large as their counterparts in the original quaternion matrix equation. The block GMRES method in the complex setting, as it seeks the solution of the problem in a low-dimensional (Krylov) subspace, is a suitable approach to solve the resulting complex matrix equation. Finally, we form the solution of the original quaternion matrix equation from the solution of the complex equation by applying a simple transformation.

2. Background

We first introduce some basic concepts and definitions related to the skew-field of quaternion scalars, as well as quaternion matrices. The set of quaternion scalars is given by

$$\mathbb{H}=\left\{a=a_0+a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}\;:\;\mathbf{i^2}=\mathbf{j^2}=\mathbf{k^2}=-1,\;\mathbf{ij}=-\mathbf{ji}=\mathbf{k},\;a_0,a_1,a_2,a_3\in\mathbb{R}\right\}.$$

The conjugate and modulus of $a=\;a_0+a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}\in\mathbb{H}$ with $a_0,a_1,a_2,a_3\in\mathbb{R}$ are defined as

$$\bar{a}=a_0-a_1\mathbf{i}-a_2\mathbf{j}-a_3\mathbf{k}\quad\text{and}\quad|a|=\sqrt{a\bar{a}}=\sqrt{a_0^2+a_1^2+a_2^2+a_3^2},$$

respectively. Throughout this study, the set of vectors of size n and the set of $m\times n$ matrices with quaternion entries are denoted by $\mathbb{H}^{n\times 1}$ and $\mathbb{H}^{m\times n}$, respectively.

Since the multiplication of the units $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ with each other is not commutative, the product of two quaternion scalars does also not commute. As a result, the set of quaternions is a skew-field, but not a field. The multiplication of vectors in $\mathbb{H}^{n\times 1}$ with scalars in the skew-field $\mathbb{H}$ can be defined from left or right, and $\mathbb{H}^{n\times 1}$ equipped with such a multiplication with scalars can be viewed as a right vector space or a left vector space accordingly. In this work, we avoid the use of left or right vector spaces, and instead, we benefit from an isomorphism that converts $\mathbb{H}^{m\times n}$ into a complex subspace of $\mathbb{C}^{2m\times 2n}$ explained next.

Any quaternion matrix $A\in\mathbb{H}^{m\times n}$ can be expressed as $A=A_1+jA_2\in\mathbb{H}^{m\times n}$ for same $A_1,A_2\in\mathbb{C}^{m\times n}$ in a unique way. For this representation of a quaternion matrix A, consider the map $\Omega:\mathbb{H}^{m\times n}\to\mathbb{C}^{2m\times 2n}$ defined as

$$\Omega(A)=\begin{pmatrix} A_1 & -\overline{A}_2\\ A_2 & \overline{A}_1\end{pmatrix}\in\mathbb{C}^{2m\times 2n},\tag{2.1}$$

where the notation $\overline{M}$ stands for the complex conjugate of a complex matrix M, which is the matrix obtained by taking the complex conjugate of each entry of M. We remark that the map Ω is on isomorphism between $\mathbb{H}^{m\times n}$ and the subspace of $\mathbb{C}^{2m\times 2n}$ consisting of notices of the form $\begin{pmatrix} Z & -\overline{U}\\ U & \overline{Z}\end{pmatrix}$ for same $Z,U\in\mathbb{C}^{m\times n}$. We now give same basic definitions related to complex and more generally quaternion vectors, matrices. The Euclidian norm of $u\in\mathbb{H}^n$ is defined by $\|u\|_2=\sqrt{\sum_{i=1}^n|v_i|^2}$, and the Frobenius norm of $A\in\mathbb{H}^{m\times n}$ is defined by $\|A\|_F=\sqrt{\sum_{i=1}^m\sum_{j=1}^n|a_{ij}|^2}$. The symbols A^T and A^* stand for transpose and conjugate transpose, respectively, of a quaternion matrix A of size $m\times n$. A matrix $A\in\mathbb{H}^{n\times n}$ is called Hermitian if $A^*=A$. Moreaver, if the conditions $AB=BA=I$ are satisfied, then B is called the inverse of A and denoted by A^{-1}. In particular if $A^*A=AA^*=I$ are satisfied, then the matrix A is said to be unitary. A matrix $A\in\mathbb{H}^{m\times n}$ is called upper Hessenberg if $a_{ij}=0$ for all i such that is $i>j+1$. Finally, the *vec*-operator is a linear transformation that maps a quaternion matrix into a column vector by vertically stacking the columns of the quaternion matrix, while the Kronecker product of two complex matrices A and B is denoted by $A\otimes B$.

3. Problem Definition

Here, we consider the Sylvester quaternion matrix equation

$$AX-XB=C,\tag{3.1}$$

for given large and sparse quaternion matrices $A \in \mathbb{H}^{m\times m}$, $B \in \mathbb{H}^{n\times n}$, $C \in \mathbb{H}^{m\times n}$, and the unknown matrix $X \in \mathbb{H}^{m\times n}$ for the setting when $m \gg n$. A linear quaternion matrix equation as in (3.1) can be transformed into a linear complex matrix equation using the complex matrix representations of quaternion matrices in (2.1). It is worth noting that there are alternative complex representations of a quaternion matrix employed in the literature, but these complex representations can be obtained from each other by permuting rows and/or columns. We will use the following properties related to the complex matrix representation of a quaternion matrix.

Theorem 3.1. *The following properties hold for every $A,B \in \mathbb{H}^{m\times n}$, $C \in \mathbb{H}^{n\times p}$, and $k \in \mathbb{R}$:*

1. $A = B \;\Leftrightarrow\; \Omega(A) = \Omega(B)$.
2. $\Omega(A\pm B) = \Omega(A)\pm\Omega(B)$.
3. $\Omega(AC) = \Omega(A)\Omega(C)$.
4. $\Omega(kA) = k\Omega(A)$.
5. $\Omega(A^*) = \Omega(A)^*$.
6. $A \in \mathbb{H}^{n\times n}$ *is invertible* $\Leftrightarrow \Omega(A)$ *is invertible, and if* $A \in \mathbb{H}^{n\times n}$ *is invertible, then* $\Omega\left(A^{-1}\right) = \Omega(A)^{-1}$.
7. $A \in \mathbb{H}^{n\times n}$ *is unitary* $\Leftrightarrow \Omega(A)$ *is unitary.*
8. $\|A\|_F = \frac{1}{\sqrt{2}}\|\Omega(A)\|_F$.

By making use of the complex representations of $A \in \mathbb{H}^{m\times m}$, $B \in \mathbb{H}^{n\times n}$, $C \in \mathbb{H}^{m\times n}$, $X \in \mathbb{H}^{m\times n}$ in (3.1), as well as items (1), (2) and (3) in Theorem 3.1, we rewrite the Sylvester equation in (3.1) as

$$\Omega(A)\Omega(X) - \Omega(X)\Omega(B) = \Omega(C), \tag{3.2}$$

where $\Omega(A) = \begin{pmatrix} A_1 & -\overline{A_2} \\ A_2 & \overline{A_1} \end{pmatrix} \in \mathbb{C}^{2m\times 2m}$, $\Omega(B) = \begin{pmatrix} B_1 & -\overline{B_2} \\ B_2 & \overline{B_1} \end{pmatrix} \in \mathbb{C}^{2n\times 2n}$, $\Omega(C) = \begin{pmatrix} C_1 & -\overline{C_2} \\ C_2 & \overline{C_1} \end{pmatrix} \in \mathbb{C}^{2m\times 2n}$, and $\Omega(X) = \begin{pmatrix} X_1 & -\overline{X_2} \\ X_2 & \overline{X_1} \end{pmatrix} \in \mathbb{C}^{2m\times 2n}$. To ease the notation, we represent $\Omega(A)$, $\Omega(B)$, $\Omega(C)$, $\Omega(X)$ with $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{X}$, respectively. Equation (3.2) with these shorter representations can be expressed as

$$\tilde{A}\tilde{X} - \tilde{X}\tilde{B} = \tilde{C}. \tag{3.3}$$

Equation (3.3) can further be expressed in terms of the Sylvester operator $S : \mathbb{C}^{2m\times 2n} \to \mathbb{C}^{2m\times 2n}$, $S\left(\tilde{X}\right) = \tilde{A}\tilde{X} - \tilde{X}\tilde{B}$ for given matrices $\tilde{A} \in \mathbb{C}^{2m\times 2m}$ and $\tilde{B} \in \mathbb{C}^{2n\times 2n}$ as

$$S\left(\tilde{X}\right) = \tilde{C}. \tag{3.4}$$

Throughout the rest of this paper, we focus on the solution of the Sylvester quaternion matrix equation (3.1) by setting up a block GMRES method based on its representation in (3.4) in terms of the Sylvester operator. When setting up the block GMRES method, we always assume that (3.1) has a unique solution, equivalently that the complex matrices $\tilde{A}$ and $\tilde{B}$ have no common eigenvalues. We, however, remark that in case the matrix equation does not have any solution, a least-squares solution for (3.1) can be obtained by applying the method to the normal equations corresponding to the least-squares problem associated with (3.1) in a similar way. In the next section, we describe a block Arnoldi process for constructing orthonormal bases for Krylov subspaces, as well as a block GMRES method for finding the best solutions in the least-squares sense of (3.4) in affine spaces associated with these Krylov subspaces.

4. Results

4.1. The Block Arnoldi Process

In this section, we first present the block Arnoldi process for solving the complex matrix equation in (3.4) expressed in terms of Sylvester operators. Let $\tilde{X}_0 \in \mathbb{C}^{2m\times 2n}$ be an initial approximation of the solution of (3.4), and $\tilde{R}_0 := \tilde{C} - \left(\tilde{A}\tilde{X}_0 - \tilde{X}_0\tilde{B}\right) = \tilde{C} - S\left(\tilde{X}_0\right)$ be the initial residual related to the error of this approximation. For a positive integer k, the kth block Krylov subspace $\mathscr{K}_k\left(S,\tilde{R}_0\right)$ associated with $\tilde{R}_0$ and the Sylvester operator S is defined by

$$\mathscr{K}_k\left(S,\tilde{R}_0\right) := \text{blockspan}\left\{\tilde{R}_0, S\left(\tilde{R}_0\right), \ldots, S\left(\tilde{R}_0\right)^{k-1}\right\} = \left\{\tilde{R}_0\beta_1 + S\left(\tilde{R}_0\right)\beta_2, \ldots, S\left(\tilde{R}_0\right)^{k-1}\beta_k \;:\beta_1,\ldots,\beta_k \in \mathbb{C}^{2n\times 2n}\right\}. \tag{4.1}$$

Lemma 2.1 in [17] shows that the Krylov subspaces constructed using the matrix $\tilde{A}$ and using the Sylvester operator S are the same, i.e., for all $k \geq 1$, we have

$$\mathscr{K}_k\left(S,\tilde{R}_0\right) = \mathscr{K}_k\left(\tilde{A},\tilde{R}_0\right) := \text{blockspan}\left\{\tilde{R}_0, \tilde{A}\tilde{R}_0, \ldots, \tilde{A}^{k-1}\tilde{R}_0\right\}.$$

Thanks to this lemma, throughout the rest of this paper, we work on $\mathscr{K}_k\left(\tilde{A},\tilde{R}_0\right)$ rather than $\mathscr{K}_k\left(S,\tilde{R}_0\right)$.
The block Arnoldi process is formally described below in Algorithm 1. It starts with $Q_0 \in \mathbb{C}^{2m\times 2n}$, whose columns form an orthonormal basis for the columns space of $\tilde{R}_0$. At the kth iteration, the Krylov subspace is expanded by adding $2n$ new orthonormal basis vectors stored inside Q_k due to the inclusion of the column space of $\tilde{A}^k\tilde{R}_0$ in the subspace. This is achieved by multiplying $\tilde{A}$ with Q_{k-1}, the orthonormal basis matrix from the previous iteration, then orthonormalizing the resulting vectors against all previous basis vectors by the Gram–Schmidt procedure. At the end of the kth iteration, Algorithm 1 yields the recurrence

$$\tilde{A}\tilde{Q}_k = \tilde{Q}_{k+1}\tilde{H}_k, \tag{4.2}$$

where $\tilde{Q}_k := \left[\begin{array}{cccc} Q_0 & Q_1 & \cdots & Q_{k-1} \end{array}\right] \in \mathbb{C}^{2m\times 2nk}$ and $\tilde{Q}_{k+1} := \left[\begin{array}{ccccc} Q_0 & Q_1 & \cdots & Q_{k-1} & Q_k \end{array}\right] \in \mathbb{C}^{2m\times 2n(k+1)}$ with columns forming an orthonormal basis for $\mathscr{K}_k\left(\tilde{A},\tilde{R}_0\right)$ and $\mathscr{K}_{k+1}\left(\tilde{A},\tilde{R}_0\right)$, respectively, while $\tilde{H}_k \in \mathbb{C}^{2n(k+1)\times 2nk}$ is an upper block Hessenberg matrix given by

$$\tilde{H}_k = \begin{pmatrix} H_{11} & H_{12} & H_{13} & \cdots & H_{1k} \\ H_{21} & H_{22} & H_{23} & \cdots & H_{2k} \\ 0 & H_{32} & H_{33} & \cdots & H_{3k} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & H_{k(k-1)} & H_{kk} \\ 0 & 0 & 0 & 0 & H_{(k+1)k} \end{pmatrix}. \tag{4.3}$$

We remark that the submatrices $H_{(i+1)i}$ for $i = 1, \ldots, k$ in (4.3) are upper triangular matrices.

Algorithm 1 The Block Arnoldi Process

$\tilde{R}_0 \leftarrow \tilde{C} - \left(\tilde{A}\tilde{X}_0 - \tilde{X}_0\tilde{B}\right)$
Obtain Q_0 from the QR factorization $\tilde{R}_0 = Q_0U_0$
for $i = 1,\ldots,k$ **do**
    $W_i \leftarrow \tilde{A}Q_{i-1}$
    for $j = 1,\ldots,i$ **do**
        $H_{ji} \leftarrow Q_{j-1}^{*}W_i$
        $W_i \leftarrow W_i - Q_{j-1}H_{ji}$
    end for
    Form Q_i and $H_{(i+1)i}$ by computing the QR factorization $W_i = Q_iH_{(i+1)i}$
end for
$\tilde{Q}_k \leftarrow \left[\begin{array}{cccc} Q_0 & Q_1 & \cdots & Q_{k-1} \end{array}\right]$, $\tilde{Q}_{k+1} \leftarrow \left[\begin{array}{ccccc} Q_0 & Q_1 & \cdots & Q_{k-1} & Q_k \end{array}\right]$ and $\tilde{H}$ is as in (4.3).

4.2. A Block GMRES Method

Using the initial approximation $\tilde{X}_0 \in \mathbb{C}^{2m\times 2n}$ for the solution of (3.4), our block quaternion GMRES method determines $\tilde{X}_k$ minimizing $\left\|\tilde{C} - S\left(\tilde{X}\right)\right\|_F$ over all $\tilde{X} \in \tilde{X}_0 + \mathscr{K}_k\left(\tilde{A},\tilde{R}_0\right)$. The minimizing $\tilde{X}_k$ can be expressed as $\tilde{X}_k = \tilde{X}_0 + \tilde{Q}_k\tilde{Y}_k$ for some $\tilde{Y}_k \in \mathbb{C}^{2nk\times 2n}$, where $\tilde{Q}_k = \left[\begin{array}{cccc} Q_0 & Q_1 & \cdots & Q_{k-1} \end{array}\right] \in \mathbb{C}^{2m\times 2nk}$ is the matrix with columns forming an orthonormal basis for $\mathscr{K}_k\left(\tilde{A},\tilde{R}_0\right)$ generated by the block Arnoldi process, i.e., by Algorithm 1. The matrix $\tilde{Y}_k$ such that $\tilde{X}_k = \tilde{X}_0 + \tilde{Q}_k\tilde{Y}_k$ is optimal for the least-squares problem must satisfy

$$\min_{\tilde{X}\in\tilde{X}_0+\mathscr{K}_k\left(\tilde{A},\tilde{R}_0\right)} \left\|\tilde{C} - S\left(\tilde{X}\right)\right\|_F = \min_{\tilde{Y}\in\mathbb{C}^{2nk\times 2n}} \left\|\tilde{C} - S\left(\tilde{X}_0 + \tilde{Q}_k\tilde{Y}\right)\right\|_F = \left\|\tilde{C} - S\left(\tilde{X}_0 + \tilde{Q}_k\tilde{Y}_k\right)\right\|_F.$$

Recalling the definition of the operator S, as well as $\tilde{R}_0 = \tilde{C} - S\left(\tilde{X}_0\right)$, and exploiting the linearity of S, letting $\underline{m} := \min_{\tilde{X}\in\tilde{X}_0+\mathscr{K}_k\left(\tilde{A},\tilde{R}_0\right)} \left\|\tilde{C} - S\left(\tilde{X}\right)\right\|_F$, we obtain

$$\underline{m} = \min_{\tilde{Y}\in\mathbb{C}^{2nk\times 2n}} \left\|\tilde{R}_0 - \left(\tilde{A}\left(\tilde{Q}_k\tilde{Y}\right) - \left(\tilde{Q}_k\tilde{Y}\right)\tilde{B}\right)\right\|_F. \tag{4.4}$$

Since $\tilde{R}_0 - \left(\tilde{A}\left(\tilde{Q}_k\tilde{Y}\right) - \left(\tilde{Q}_k\tilde{Y}\right)\tilde{B}\right)$ is contained in $\mathscr{K}_{k+1}\left(\tilde{A},\tilde{R}_0\right)$ for which the columns of $\tilde{Q}_k$ form an orthonormal basis, its Frobenius norm does not change when it is multiplied with $\tilde{Q}_{k+1}^{*}$ from right. Hence, we have

$$\underline{m} = \min_{\tilde{Y}\in\mathbb{C}^{2nk\times 2n}} \left\|\tilde{Q}_{k+1}^{*}\tilde{R}_0 - \tilde{Q}_{k+1}^{*}\left(\tilde{A}\left(\tilde{Q}_k\tilde{Y}\right) - \left(\tilde{Q}_k\tilde{Y}\right)\tilde{B}\right)\right\|_F. \tag{4.5}$$

As $\tilde{R}_0 = Q_0U_0$ is the QR factorization of $\tilde{R}_0$, and Q_0 is the first block column of the matrix $\tilde{Q}_{k+1}$ with orthonormal columns, it follows that $\tilde{Q}_{k+1}^{*}\tilde{R}_0 = \begin{bmatrix} U_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$. Finally, letting $\tilde{U} = \begin{bmatrix} U_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, and recalling the Arnoldi recurrence in (4.2), the minimization problem in (4.5) can equivalently be written as

$$\underline{m} = \min_{\tilde{Y}\in\mathbb{C}^{2nk\times 2n}} \left\|\tilde{U} - \left(\tilde{H}_k\tilde{Y} - \begin{bmatrix} I \\ 0 \end{bmatrix}\tilde{Y}\tilde{B}\right)\right\|_F. \tag{4.6}$$

Consequently, $\tilde{Y}_k$ is the best solution in the least-squares sense of the Sylvester matrix equation

$$\tilde{H}_k\tilde{Y} - \begin{bmatrix} I \\ 0 \end{bmatrix}\tilde{Y}\tilde{B} = \tilde{U}, \tag{4.7}$$

which has a smaller size than the original Sylvester matrix equation in (3.4). The least-squares solution $\tilde{Y}_k$ of the matrix equation in (4.7) satisfies the linear system

$$\left(I\otimes\tilde{H}_k-\tilde{B}^T\otimes\begin{bmatrix} I \\ 0 \end{bmatrix}\right) vec\left(\tilde{Y}\right)=vec\left(\tilde{U}\right) \tag{4.8}$$

with the help of the vec operator. Denoting $\left(I\otimes\tilde{H}_k-\tilde{B}^T\otimes\begin{bmatrix} I \\ 0 \end{bmatrix}\right)$, $vec\left(\tilde{Y}\right)$, $vec\left(\tilde{U}\right)$ with $\mathscr{A}$, y, g, respectively, we represent it as

$$\underline{m}=\min_{y\in\mathbb{C}^{4n^2k\times 1}}\|g-\mathscr{A}y\|_2 . \tag{4.9}$$

Note that the coefficient matrix $\mathscr{A}$ is of size $4n^2(k+1)\times 4n^2k$ independent of m. Thus, assuming n, k are small, $\mathscr{A}$ is of small size, and the least-squares solution to (4.9) can be retrieved efficiently by computing a QR decomposition of $\mathscr{A}$. Let $y_*\in\mathbb{C}^{4n^2k\times 1}$ be the solution of (4.9) retrieved this way. Then, $\tilde{Y}_k\in\mathbb{C}^{2nk\times 2n}$ is such that $y_*=vec(\tilde{Y}_k)$, so $\tilde{Y}_k$ can be constructed from y_* by just reshaping it as a matrix without performing any computation. In the end, the matrix $\tilde{X}_k=\tilde{X}_0+\tilde{Q}_k\tilde{Y}_k$ is an approximate solution to the complex Sylvester matrix equation (3.4) and must be of the form

$$\tilde{X}_k=\Omega\left(X_k\right)=\begin{bmatrix} X_{k,1} & -\overline{X_{k,2}} \\ X_{k,2} & \overline{X_{k,1}} \end{bmatrix}\in\mathbb{C}^{2m\times 2n}$$

for some X_k that is an approximate solution to the quaternion matrix equation (3.1). The approximate solution X_k to (3.1) is then given by

$$X_k=X_{k,1}+\mathbf{j}X_{k,2}\in\mathbb{H}^{m\times n}.$$

5. Numerical Experiments

Here, we provide a numerical example to demonstrate the accuracy of the block GMRES method proposed for the Sylvester quaternion matrix equation $AX-XB=C$, as well as its convergence. We randomly generate the known matrices $A\in\mathbb{H}^{m\times m}$, $B\in\mathbb{H}^{n\times n}$, $C\in\mathbb{H}^{m\times n}$ with $m=50$ and $n=5$ in MATLAB. The precise data is made available on the web[1]. By utilizing the complex matrix representations of the matrices A,B,C, we transform the Sylvester quaternion matrix equation $AX-XB=C$ into the corresponding Sylvester complex matrix equation $\tilde{A}\tilde{X}-\tilde{X}\tilde{B}=\tilde{C}$ for $\tilde{A}\in\mathbb{C}^{100\times 100}$, $\tilde{B}\in\mathbb{C}^{10\times 10}$, $\tilde{C}\in\mathbb{C}^{100\times 10}$. This transformation allows us to deal with the problem in the usual complex space by using the complex arithmetic implemented in computers. In particular, we apply the proposed block GMRES method to the Sylvester complex matrix equation $\tilde{A}\tilde{X}-\tilde{X}\tilde{B}=\tilde{C}$. For this experiment, we initialize the block Arnoldi Algorithm with the matrix $\tilde{X}_0$ set equal to the 100×10 zeros matrix. As a result, the initial residual matrix $\tilde{R}_0$ satisfies $\tilde{R}_0=\tilde{C}$. When carrying out the block Arnoldi algorithm, the maximal number of iterations is set equal to 10, so the block Arnoldi algorithm yields $\tilde{Q}_{10}$ and $\tilde{Q}_{11}$, which are orthonormal bases for $\mathscr{K}_{10}\left(\tilde{A},\tilde{C}\right)$ and $\mathscr{K}_{11}\left(\tilde{A},\tilde{C}\right)$, respectively. We then apply the block GMRES method to find the best solution $\tilde{X}_{10}$ of the complex matrix equation $\tilde{A}\tilde{X}-\tilde{X}\tilde{B}=\tilde{C}$ in the Krylov subspace $\mathscr{K}_{10}\left(\tilde{A},\tilde{C}\right)$. Finally, an approximate solution X_{10} of the original problem $AX-XB=C$ is obtained by converting $\tilde{X}_{10}$ into a quaternion matrix as explained at the end of the previous section.
Figure 5.1 below depicts the residual norm $\left\|\tilde{R}_k\right\|_F=\left\|\tilde{C}-S\left(\tilde{X}_k\right)\right\|_F$ for the Sylvester complex matrix equation $\tilde{A}\tilde{X}-\tilde{X}\tilde{B}=\tilde{C}$ in this example as a function of k, that is the number of iterations.

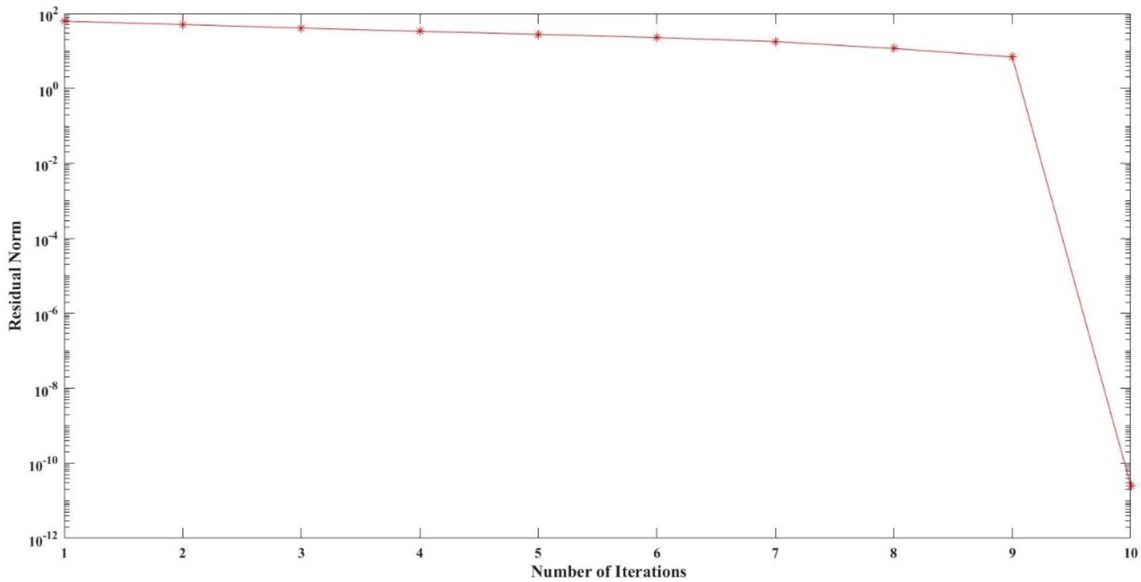

Figure 5.1: The residual norm $\left\|\tilde{R}_k\right\|_F$ is plotted as a function of k (the number of iterations) for Algorithm 4.1

Additionally, the residual norm up to four decimal digits at every iteration until termination is reported in Table 1 below. Observe that the residual norms decrease monotonically. After 10 iterations $\left\|\tilde{R}_{10}\right\|_F$ is zero up to rounding errors as expected in theory.

[1] https://drive.google.com/drive/folders/11pUmV-O-285nRgCaNn9XA01eMoFZW61g?u

Table 1: Residual norm $\|\tilde{R}_k\|$ up to four decimal digits with respect to k

k	1	2	3	4	5	6	7	8	9	10
$\|\tilde{R}_k\|_F$	62.2118	50.6591	40.7386	33.3585	27.8083	22.7574	17.8061	11.6200	6.8859	0.0000

6. Conclusion

In this paper, we present a block GMRES method for solving the Sylvester quaternion matrix equation $AX - XB = C$ by employing the complex representations of quaternion matrices. Our method is suitable for large and sparse quaternion matrices $A \in \mathbb{H}^{m\times m}$, $B \in \mathbb{H}^{n\times n}$, $C \in \mathbb{H}^{m\times n}$ with the unknown matrix $X \in \mathbb{H}^{m\times n}$, in particular for the case when $m \gg n$. We have also provided an example to illustrate the accuracy and convergence of the proposed method.
The method could possibly be improved by taking the structures of the complex representations when dealing with the resulting complex matrix equations, which may in turn reduce the number of operations. Moreover, the proposed approach could potentially be adapted to solve other types of matrix equations.

Article Information

Acknowledgements: The authors would like to express their sincere thanks to the editor and the anonymous reviewers for their helpful comments and suggestions.

Author's contributions: All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Conflict of Interest Disclosure: No potential conflict of interest was declared by the author.

Supporting/Supporting Organizations: No grants were received from any public, private or non-profit organizations for this research.

Ethical Approval and Participant Consent: It is declared that during the preparation process of this study, scientific and ethical principles were followed and all the studies benefited from are stated in the bibliography.

Plagiarism Statement: This article was scanned by the plagiarism program. No plagiarism detected.

Availability of data and materials: Not applicable

References

[1] A. Rehman, I. Kyrchei, M. Z. U. Rahman, V. Leiva and C. Castro, *Solvability and algorithm for Sylvester-type quaternion matrix equations with potential applications*, AIMS Mathematics, 9(8) (2024), 19967-19996. doi: 10.3934/math.2024974
[2] A. Rehman, I. Kyrchei, I. Ali, M. Akram and A. Shakoor, *Explicit formulas and determinantal representation for η-skew Hermitian solution to a system of quaternion matrix equations*, Filomat, 34(8) (2020), 2601-2627.
[3] A. Bouhamidi and K. Jbilou, *A note on the numerical approximate solutions for generalized Sylvester matrix equations with applications*, Applied Mathematics and Computation, 206 (2008), 687–694.
[4] A. El Guennouni, K. Jbilou and A.J. Riquet, *Block Krylov Subspace Methods for Solving Large Sylvester Equations*, Numerical Algorithms, 29 (2002), 75–96.
[5] A. Kaabi, *On the numerical solution of generalized Sylvester matrix equations*, Bulletin of the Iranian Mathematical Society, 40 (2014), 101–113.
[6] A.M. Lypunov, *The general problem of the stability of motion*, International Journal of Control, 55 (3) (1992), 531-534.
[7] A. Wei, Y. Li, W. Ding and J. Zhao, *Three special kinds of least-squares solutions for the quaternion generalized Sylvester matrix equation*, AIMS Mathematics, 7 (4) (2021), 5029-5048.
[8] C.Q Zhang, Q.W. Wang, A. Dmytryshyn and Z.H. He, *Investigation of some Sylvester-type quaternion matrix equations with multiple unknowns*, Computational and Applied Mathematics, 43 (2024), 26 pages.
[9] F.P.A. Beik, *Theoretical Results on the Global GMRES Method for Solving Generalized Sylvester Matrix Equations*, Bulletin of the Iranian Mathematical Society, 40 (5) (2014), 1097–1117.
[10] F.P.A. Beik and S. Ahmadi-Asl, *An Iterative Algorithm for η-(anti)-(Hermitian) Least-Squares Solutions of Quaternion Matrix Equations*, Electronic Journal of Linear Algebra, 30 (2015), 372-401.
[11] F. Zhang, W. Mu, Y. Li and J. Zhao, *Special least-squares solutions of the quaternion matrix equation $AXB + CXD = E$*, Computers and Mathematics with Applications, 72 (5) (2016), 1426-1435.
[12] G.H. Golub, S. Nash and C.V. Loan, *A hessenberg-schur method for the problem $AX + XB = C$*, IEEE Transactions on Automatic Control, 24 (6) (1979), 909-913.
[13] I. Kyrchei, *Cramer's rules for Sylvester quaternion matrix equation and its special cases*, Advances in Applied Clifford Algebras, 28:90 (2018).
[14] I. Kyrchei, *Cramer's Rules of η-(Skew-)Hermitian Solutions to the Quaternion Sylvester-Type Matrix Equations*, Adv. Appl. Clifford Algebras, 29:56 (2019).
[15] J.J. Sylvester, *Sur l'equation en matrices $px = xq$*, C. R. Acad. Sci. Paris., 99 (2) (1884), 67-71, 115-116.
[16] L. Rodman, *Topics in Quaternion Linear Algebra*, Princeton University Press, 2014.
[17] M. Robbe and M. Sadkane, *A convergence analysis of GMRES and FOM methods for Sylvester equations*, Numerical Algorithms, 30 (2002), 71-89.
[18] N. Li and Q. Wang, *Iterative Algorithm for Solving a Class of Quaternion Matrix Equation Over the Generalized (P,Q) Reflexive Matrices*, Abstract and Applied Analysis Article ID 831656 (2013), 15 pages.
[19] Q.W. Wang, Qing-Wen, H.S. Zhang and S.W. Yu, *On solutions to the quaternion matrix equation $AXB + CYD = E$*, The Electronic Journal of Linear Algebra, 17 (2008), 343-358.
[20] R.H. Bartels and G.W. Stewart, *Algorithm 432: Solution of the matrix equation $AX + XB = C$*, Communications of the ACM, 15 (1972), 820-826.
[21] S. Agoujil, A.H. Bentbib, K. Jbilou, and E. Sadek, *A Minimal Residual Norm Method for Large-scale Sylvester Matrix Equations*, Electronic Transactions on Numerical Analysis, 43 (2014), 45-59.
[22] S. Ling and Z. Jia, *Matrix iterative algorithms for least-squares problem in quaternionic quantum theory*, International Journal of Computer Mathematics, 90 (3) (2013), 727-745.
[23] S.K. Li, M.X. Wang and G. Liu, *A global variant of the COCR method for the complex symmetric Sylvester matrix equation $AX + XB = C$*, Computers and Mathematics with Applications, 94 (2021), 104–113.

[24] S. Şimşek, *Least-squares solutions of generalized Sylvester-type quaternion matrix equations*, Advances in Applied Clifford Algebras, 33:28 (2023), 23 pages.
[25] S. Şimşek, *A block quaternion GMRES method and its convergence analysis*, Calcolo, 61 (33) (2024).
[26] S. Şimşek, *A Global Krylov Subspace Method for the Sylvester Quaternion Matrix Equation*, Journal of New Theory, 47 (2024), 39-51.
[27] S. Şimşek and A. Körükçü, *A Block Conjugate Gradient Method for Quaternion Linear Systems*, Yuzuncu Yil University Journal of the Institute of Natural and Applied Sciences, 28 (2) (2023), 394–403.
[28] S. Şimşek, M. Sarduvan and Özdemir, *H. Centrohermitian and Skew-Centrohermitian Solutions to the Minimum Residual and Matrix Nearness Problems of the Quaternion Matrix Equation* $(AXB,DXE)=(C,F)$, Adv. Appl. Clifford Algebras, 27 (2017), 2201–2214.
[29] V. Simoncini, *Computational methods for linear matrix equations*, SIAM Review, 58 (3) (2016), 377-441.
[30] W.R. Hamilton, *Elements of quaternions*, Longmans, Green, Company, London, 1866.
[31] X. Zhang, *A system of generalized Sylvester quaternion matrix equations and its applications*, Applied Mathematics and Computation, 273 (2016), 74-81.
[32] Y. Lin, *Implicitly restarted global FOM and GMRES for nonsymmetric matrix equations and Sylvester equations*, Applied Mathematics and Computation, 167 (2005), 1004-1025.
[33] Z. He, *Pure PSVD approach to Sylvester-type quaternion matrix equations*, Electronic Journal of Linear Algebra, 35 (2019), 266-284.
[34] Z. He, *Some new results on a system of Sylvester-type quaternion matrix equations*, Linear and Multilinear Algebra, 69 (16) (2021), 3069-3091.
[35] Z. He, Q. Wang and Y. Zhang, *A system of quaternary coupled Sylvester-type real quaternion matrix equations*, Automatica, 87 (2018), 25-31.
[36] Z. He, W.L. Qin, J. Tian, X.X. Wang and Y. Zhang, *A new Sylvester-type quaternion matrix equation model for color image data transmission*, Computational and Applied Mathematics, 43 (2024), 30 pages.