Linear Algebra and its Applications 475 (2015) 45–61

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

The Moore–Penrose inverses of matrices over quaternion polynomial rings

Liji Huang [b], Qing-Wen Wang [a,*], Yang Zhang [b]

[a] Department of Mathematics, Shanghai University, Shanghai, PR China
[b] Department of Mathematics, University of Manitoba, Winnipeg, MB R3T 2N2, Canada

ARTICLE INFO

Article history:
Received 8 May 2014
Accepted 1 February 2015
Available online 20 February 2015
Submitted by W.B. Wu

MSC:
15A09
68W30

Keywords:
Quaternion polynomial matrix
Moore–Penrose inverse
Leverrier–Faddeev algorithm

ABSTRACT

In this paper, we define and discuss the Moore–Penrose inverses of matrices with quaternion polynomial entries. When the Moore–Penrose inverses exist, we prove that Leverrier–Faddeev algorithm works for these matrices by using generalized characteristic polynomials. Furthermore, after studying interpolations for quaternion polynomials, we give an efficient algorithm to compute the Moore–Penrose inverses. We developed a Maple package for quaternion polynomial matrices. All algorithms in this paper are implemented, and tested on examples.

1. Introduction

In 1843 Sir Rowan Hamilton discovered the algebra $\mathbb{H}$ of real quaternion, which is a four-dimensional non-commutative algebra over real number field $\mathbb{R}$ with canonical basis $1, \mathbf{i}, \mathbf{j}, \mathbf{k}$ satisfying the conditions:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1,$$

* Corresponding author.
E-mail addresses: wqwshu9@126.com (Q.-W. Wang), yang.zhang@umanitoba.ca (Y. Zhang).

http://dx.doi.org/10.1016/j.laa.2015.02.004
0024-3795/

that implies

$$\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \ \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \quad \text{and} \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}.$$

The elements in $\mathbb{H}$ can be written in a unique way: $\alpha = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, where a, b, c and d are real numbers, that is, $\mathbb{H} = \{a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \mid a, b, c, d \in \mathbb{R}\}$. The conjugate of α is defined as $\bar{\alpha} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$, and the norm $|\alpha|$ is $|\alpha| = \sqrt{\alpha\bar{\alpha}}$. It is well-known that $\mathbb{H}$ is a skew field, that is, for $0 \neq \alpha \in \mathbb{H}$, $\alpha^{-1} = \bar{\alpha}/|\alpha|^2$.

The study of polynomials with quaternion coefficients may go back to Niven [17] and [18] in the early 1940's. Due to the non-commutativity of $\mathbb{H}$, there are several forms of quaternion polynomials depending on the positions of coefficients. In this paper, we will use the following Definition 1.1, which puts the coefficients on the left side of a variable x. Also they are called regular quaternion polynomials in [5] and quaternion simple polynomials in [19]. Some properties have been discussed (see, for example, [14] and [20]).

Definition 1.1. A quaternion polynomial $f(x)$ over $\mathbb{H}$ is defined as

$$f(x) = a_n x^n + \cdots + a_1 x + a_0, \quad a_i \in \mathbb{H}, \ i = 0, \ldots, n,$$

where x commutes element-wise with $\mathbb{H}$.

The set of all quaternion polynomials in x is denoted by $\mathbb{H}[x]$. The degrees, leading terms and leading coefficients are defined in a common way. Also the degree of 0 is defined as $-\infty$. For $f(x) = \sum_{i=0}^{n} a_i x^i$, $g(x) = \sum_{j=0}^{m} b_j x^j \in \mathbb{H}[x]$, the addition and multiplication are defined as:

$$f(x) + g(x) = \sum_{k=0}^{\max\{n,m\}} (a_k + b_k)x^k, \quad f(x)g(x) = \sum_{k=0}^{n+m} \sum_{i+j=k} a_i b_j x^k,$$

where some a_i, b_j may be assumed to be zero for simplicity.

The conjugate of $f(x) = a_n x^n + \cdots + a_0 \in \mathbb{H}[x]$ is defined as $\bar{f}(x) = \bar{a}_n x^n + \cdots + \bar{a}_0$, and has the following properties:

Lemma 1.2. *(See [21].) Let $f, g \in \mathbb{H}[x]$. Then (i) $\overline{fg} = \bar{g}\bar{f}$ (ii) $f\bar{f} = \bar{f}f \in \mathbb{R}[x]$, where $\mathbb{R}$ are reals, (iii) if $fg \in \mathbb{R}[x]$, then $fg = gf$.*

The quaternion polynomials and matrices with quaternion polynomial entries have been widely studied with many applications in past decades. For example, in [27], the Fast Fourier Transform for the product of two quaternion polynomial has been discussed with the complexity analysis. In [5], Damiano et al. studied Gröbner basis theory for the ring of quaternion polynomials and explored how to compute the module syzygy. Smith–McMilian forms of quaternion polynomial matrices are defined in [22] and some applications to dynamical systems are given.

For matrices over commutative rings, it is well-known that the Moore–Penrose inverses have been defined and explored for many years (see, for example, [3,12,13,23]). This motivates us to consider the Moore–Penrose inverses of quaternion polynomial matrices.

The structure of this paper is as follows. In Section 2, we give the definition of the Moore–Penrose inverse for quaternion polynomial matrices and discuss some basic properties including a sufficient and necessary condition for the existence of the Moore–Penrose inverse. Leverrier–Faddeev algorithm is extended to quaternion polynomial matrices in Section 3 by using generalized characteristic polynomials. In Section 4, we discuss the interpolation problems for quaternion polynomials and give an efficient algorithm to compute the Moore–Penrose inverses. Finally using algorithms in this paper and our Maple package some examples are given in Section 5.

2. Definitions and basic properties

In this section, we first define the Moore–Penrose inverse for matrices over $\mathbb{H}[x]$, and then discuss some basic properties as well as the sufficient and necessary conditions for the existence of the Moore–Penrose inverses.

Let $\mathbb{H}[x]^{m\times n}$ denote the set of all $m \times n$ matrices with entries from $\mathbb{H}[x]$. For $A \in \mathbb{H}[x]^{m\times n}$, the conjugate $\overline{A}$ of A is defined as $\overline{A} = (\overline{A_{ij}})$. If $A = P + Q\mathbf{j}$ with $P, Q \in \mathbb{C}[x]^{m\times n}$, then $\chi_A = \begin{pmatrix} P & Q \\ -\overline{Q} & \overline{P} \end{pmatrix} \in \mathbb{C}[x]^{2m\times 2n}$ denotes the complex adjoint of A. Moreover, $A^T, A^* \in \mathbb{H}[x]^{n\times m}$ denote the transpose and the conjugate transpose of A, respectively. In particular, for all $f \in \mathbb{H}[x]$, $f^* = \overline{f}$. More properties of $\mathbb{H}[x]^{m\times n}$ can be found, for example, in [21,22].

Definition 2.1. $A^\dagger \in \mathbb{H}[x]^{n\times m}$ is called a Moore–Penrose inverse of $A \in \mathbb{H}[x]^{m\times n}$ if it is a solution of the following system of equations:

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

Note that we require that $A^\dagger$ must be in $\mathbb{H}[x]^{n\times m}$. Therefore unlike the matrices over fields, the Moore–Penrose inverses for some quaternion polynomial matrices might not exist.

Recall that $A \in \mathbb{H}^{m\times m}$ is unitary if $AA^* = A^*A = I_m$. Based on the properties of quaternions, one can easily prove the following elementary properties that will be often used throughout this paper.

Proposition 2.2. *Let $A \in \mathbb{H}[x]^{m\times n}$ and $B \in \mathbb{H}[x]^{n\times l}$. Then*

(i) $(AB)^* = B^*A^*$ *and* $AA^* = (AA^*)^*$.
(ii) *If A has a Moore–Penrose inverse $A^\dagger$, then* $(A^*)^\dagger = \left(A^\dagger\right)^*$, $A^\dagger \left(A^\dagger\right)^* A^* = A^\dagger = A^* \left(A^\dagger\right)^* A^\dagger$ *and* $A^\dagger AA^* = A^* = A^*AA^\dagger$.
(iii) *If A has a Moore–Penrose inverse $A^\dagger$, then $A^\dagger$ is unique.*

(iv) *Let A have the Moore–Penrose inverse $A^\dagger$. If $U \in \mathbb{H}^{m\times m}$ is a unitary matrix, then $(UA)^\dagger = A^\dagger U^*$.*

Next we will give conditions for quaternion polynomial matrices to have Moore–Penrose inverses. We need to prove some lemmas first.

It is easy to see that $A \in \mathbb{H}^{m\times n}[x]$ induces an additive homomorphism from $\mathbb{H}[x]^{n\times 1}$ to $\mathbb{H}[x]^{m\times 1}$, that is, for all $P, Q \in \mathbb{H}[x]^{n\times 1}$, $A(P+Q) = AP + AQ \in \mathbb{H}[x]^{m\times 1}$. By the definition of Moore–Penrose inverses and Proposition 2.2, it is easy to prove the following lemma:

Lemma 2.3. *Let $A \in \mathbb{H}[x]^{m\times n}$ have the Moore–Penrose inverse $A^\dagger$. Consider A as a homomorphism from $\mathbb{H}[x]^{n\times 1}$ to $\mathbb{H}[x]^{m\times 1}$. Then $Image(A) = Image(AA^*) = Image(AA^\dagger)$ and $Image(A^*) = Image(A^*A) = Image(A^\dagger A)$.*

Lemma 2.4. *If $E \in \mathbb{H}[x]^{m\times m}$ is a symmetric projection, that is, $E = E^2 = E^*$, then $E \in \mathbb{H}^{m\times m}$.*

Proof. Let $f_1, \ldots, f_m$ be the entries on the first row of E. From $E = E^*$, we may assume that $f_1 = \overline{f_1} \neq 0$. Then by $E = E^2$, we have

$$f_1 = f_1\overline{f_1} + \sum_{i=2}^{m} f_i\overline{f_i} = f_1^2 + \sum_{i=2}^{m} f_i\overline{f_i}.$$

Since $f_1 = \overline{f_1}$, the leading coefficient of f_1^2 is a positive real number. Note that the leading coefficient of $\sum_{i=2}^{m} f_i\overline{f_i}$ is also a positive real number. Thus,

$$\begin{aligned}\deg\left(f_1^2\right) \geq \deg f_1 &= \deg\left(f_1^2 + \sum_{i=2}^{m} f_i\overline{f_i}\right)\\ &= \max\left\{\deg\left(f_1^2\right), \deg\left(\sum_{i=2}^{m} f_i\overline{f_i}\right)\right\} \geq \deg\left(f_1^2\right).\end{aligned}$$

This shows that $f_1 \in \mathbb{H}$. Furthermore, $0 = \deg f_1 = \deg\left(\sum_{i=2}^{m} f_i\overline{f_i}\right)$ and the leading coefficients of $\{f_i\bar{f_i}\}$ $(f_i \neq 0)$ are positive reals imply that $f_i \in \mathbb{H}$ for all $1 \leq i \leq m$. The same discussions can be done for the other rows of E. Therefore, $E \in \mathbb{H}^{m\times m}$. □

Due to the non-commutativity of quaternions, there are two types of eigenvalues: right eigenvalues and left eigenvalues. Right eigenvalues have been studied extensively (see, for example, [1,4,15]). We will work with right eigenvalues towards our main result. For convenience, we will just use the term “eigenvalue” from now on. The following result is well-known and very useful.

Lemma 2.5. *(See [26].) $A \in \mathbb{H}^{m\times m}$ is hermitian, that is, $A = A^*$, if and only if there exists a unitary matrix $U \in \mathbb{H}^{m\times m}$ such that $U^*AU = diag(d_1, \dots, d_m)$, where d_i are the eigenvalues of A.*

Now we are ready to give conditions that quaternion polynomial matrices must satisfy in order to have Moore–Penrose inverses. The following theorem is well-known in some cases, see, for example, [3,24]. Here is an analog version for quaternion polynomial matrices.

Theorem 2.6. *Let $A \in \mathbb{H}[x]^{m\times n}$. Then A has the Moore–Penrose inverse $A^\dagger$ if and only if $A = U\begin{pmatrix} A_1 & A_2 \\ 0 & 0\end{pmatrix}$ with $U \in \mathbb{H}^{m\times m}$ unitary and $A_1A_1^* + A_2A_2^*$ a unit in $\mathbb{H}[x]^{r\times r}$ with $r \le \min\{m, n\}$. Moreover,*

$$A^\dagger = \begin{pmatrix} A_1^*\left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \end{pmatrix} U^*.$$

Proof. ($\Longrightarrow$) If A has the Moore–Penrose inverse $A^\dagger$, then

$$AA^\dagger = AA^\dagger AA^\dagger = \left(AA^\dagger\right)^2 = \left(AA^\dagger\right)^*.$$

By Lemma 2.4, $AA^\dagger \in \mathbb{H}^{m\times m}$. $AA^\dagger$ is hermitian and hence, by Lemma 2.5, there exists a unitary matrix $U \in \mathbb{H}^{m\times m}$ such that $U^*AA^\dagger U = D$ where D is diagonal. Since

$$D^2 = (U^*AA^\dagger U)(U^*AA^\dagger U) = U^*AA^\dagger AA^\dagger U = U^*AA^\dagger U = D,$$

the diagonal entries of D are either 1 or 0. Therefore, we can rearrange the rows of U so that $D = \begin{pmatrix} I_r & 0 \\ 0 & 0\end{pmatrix}$ with $r \le \min\{m, n\}$.

Set $A' = U^*A$. By Lemma 2.2, A' has its own generalized inverse $A'^\dagger$ and $A'A'^\dagger = \begin{pmatrix} I_r & 0 \\ 0 & 0\end{pmatrix}$. Set $A' = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4\end{pmatrix}$, for arbitrary quaternion polynomial matrices $A_1 \in \mathbb{H}[x]^{r\times r}$, $A_2 \in \mathbb{H}[x]^{r\times(n-r)}$, $A_3 \in \mathbb{H}[x]^{(m-r)\times r}$ and $A_4 \in \mathbb{H}[x]^{(m-r)\times(n-r)}$. Since $A' = A'A'^\dagger A' = \begin{pmatrix} I_r & 0 \\ 0 & 0\end{pmatrix}\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4\end{pmatrix} = \begin{pmatrix} A_1 & A_2 \\ 0 & 0\end{pmatrix}$, we must have $A' = \begin{pmatrix} A_1 & A_2 \\ 0 & 0\end{pmatrix}$ and therefore $A'A'^* = \begin{pmatrix} A_1A_1^* + A_2A_2^* & 0 \\ 0 & 0\end{pmatrix}$. Similarly, $A'^\dagger = \begin{pmatrix} B_1 & 0 \\ B_2 & 0\end{pmatrix}$ for some B_1 and B_2. By Lemma 2.3,

$$\mathrm{Image}(A'A'^*) = \mathrm{Image}(A') = \mathrm{Image}(A'A'^\dagger) = \mathrm{Image}\begin{pmatrix} I_r & 0 \\ 0 & 0\end{pmatrix}.$$

This implies the surjectivity of $A_1A_1^* + A_2A_2^*$ on $\mathbb{H}[x]^{r\times 1}$. Therefore, $A_1A_1^* + A_2A_2^*$ is a unit in $\mathbb{H}[x]^{r\times r}$ and

$$A = UA' = U\begin{pmatrix} A_1 & A_2 \\ 0 & 0\end{pmatrix}.$$

Next, we have that:

$$\begin{aligned} A'^{\dagger} &= A'^{\dagger}\left(A'^{\dagger}\right)^{*} A'^{*} = A'^{\dagger}\left(A'^{*}\right)^{\dagger} A'^{*} = A'^{*}\left(A' A'^{*}\right)^{\dagger} \\ &= \begin{pmatrix} A_1^* & 0 \\ A_2^* & 0 \end{pmatrix} \begin{pmatrix} \left(A_1 A_1^* + A_2 A_2^*\right)^{-1} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} A_1^*\left(A_1 A_1^* + A_2 A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1 A_1^* + A_2 A_2^*\right)^{-1} & 0 \end{pmatrix}, \end{aligned}$$

which gives:

$$A^{\dagger} = \begin{pmatrix} A_1^*\left(A_1 A_1^* + A_2 A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1 A_1^* + A_2 A_2^*\right)^{-1} & 0 \end{pmatrix} U^*.$$

($\Longleftarrow$) The converse can be proved by direct computation. □

3. Leverrier–Faddeev algorithm

Leverrier–Faddeev algorithm has been used to compute the Moore–Penrose inverses for many years. We refer the reader to [2,6,11,25] for more details. In this section, we define the characteristic polynomial for quaternion polynomial matrix A by using AA^*. In particular, we prove that the coefficients of this characteristic polynomial are reals. Then we show that Leverrier–Faddeev algorithm works very well for quaternion polynomial matrices.

Given $A \in \mathbb{H}[x]^{m\times m}$. An element $\lambda \in \mathbb{H}$ is called an *eigenvalue* of A if there exists a vector $X \in \mathbb{H}^{m\times 1}[x]$ such that $AX = X\lambda$.

Lemma 3.1. *Let $A \in \mathbb{H}[x]^{m\times n}$. Then eigenvalues of AA^* are real.*

Proof. Let $B = AA^*$ and $\lambda \in \mathbb{H}$ be an eigenvalue of B with corresponding eigenvector $X = \left(\, x_1 \cdots x_m \,\right)^T \neq 0$ such that $BX = X\lambda$. Then $X^*BX = X^*X\lambda$. Note that $B = B^*$. We have that $X^*BX = \lambda^* X^* X$, and thus

$$X^* X\lambda = \lambda^* X^* X = \left(X^* X\lambda\right)^*,$$

that is,

$$(\bar{x}_1, \cdots \bar{x}_m) \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \lambda = \left(\sum \bar{x}_i x_i\right)\lambda = \left(\left(\sum \bar{x}_i x_i\right)\lambda\right)^* = \lambda^*\left(\sum \bar{x}_i x_i\right)^* = \lambda^*\left(\sum \bar{x}_i x_i\right).$$

By Lemma 1.2, $0 \neq \sum \bar{x}_i x_i \in \mathbb{R}\,[x]$. The above equation gives $\lambda = \lambda^*$, which implies $\lambda \in \mathbb{R}$. □

Cayley–Hamilton theorem for quaternion matrices has been extensively discussed. A survey can be found in [26]. Next we define the characteristic polynomial for a given quaternion polynomial matrix.

Definition 3.2. For $A \in \mathbb{H}[x]^{m\times n}$, let $B = AA^*$ and χ_B be its complex adjoint. Then $f_B(\lambda) = \det(\lambda I_{2m} - \chi_B)$ is called the characteristic polynomial of A.

Remark 3.3. By Lemma 3.1, λ can be assumed to be a real indeterminate that enjoys the following: $\lambda = \overline{\lambda}$ and λ commutes element-wise with $\mathbb{H}[x]$.

Theorem 3.4. *Let $A \in \mathbb{H}[x]^{m\times n}$ and $B = AA^*$. Then $f_B(\lambda) = g(\lambda)^2$ where $g(\lambda) \in (\mathbb{R}[x])[\lambda]$.*

Proof. We first show that $f_B(\lambda) \in (\mathbb{R}[x])[\lambda]$. Note that $B = AA^*$, we have

$$\det\left((\lambda I_{2m} - \chi_B)^T\right) = \det(\lambda I_{2m} - \chi_B) = \det\left((\lambda I_{2m} - \chi_B)^*\right),$$

and thus

$$\det(\lambda I_{2m} - \chi_B) = \det\overline{(\lambda I_{2m} - \chi_B)} = \overline{\det(\lambda I_{2m} - \chi_B)}.$$

Therefore

$$\det(\lambda I_{2m} - \chi_B) = f_B(\lambda) \in (\mathbb{R}[x])[\lambda]. \tag{1}$$

Next, we show that $f_B(\lambda) = g(\lambda)^2$ where $g(\lambda) \in (\mathbb{C}[x])[\lambda]$. Let $B = P + Q\mathbf{j}$. For any fixed $1 \le i, j \le m$, we have $B_{ij} = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ where a, b, c and $d \in \mathbb{R}[x]$. Since B is hermitian, $B_{ji} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$ and therefore $P_{ij} = a + b\mathbf{i}$, $P_{ji} = a - b\mathbf{i}$ and $Q_{ij} = c + d\mathbf{i}$, $Q_{ji} = -c - d\mathbf{i}$. So $P^T = \overline{P}$ and $Q = -Q^T$. Therefore,

$$\chi_B = \begin{pmatrix} P & Q \\ -\overline{Q} & \overline{P} \end{pmatrix} = \begin{pmatrix} P & Q \\ -\overline{Q} & P^T \end{pmatrix} \implies \lambda I_{2m} - \chi_B = \begin{pmatrix} \lambda I_m - P & Q \\ -\overline{Q} & \lambda I_m - P^T \end{pmatrix}.$$

Next, we have:

$$\begin{aligned} &\begin{pmatrix} I_m & -I_m \\ 0 & I_m \end{pmatrix}\begin{pmatrix} I_m & 0 \\ I_m & I_m \end{pmatrix}\begin{pmatrix} I_m & -I_m \\ 0 & I_m \end{pmatrix}\begin{pmatrix} \lambda I_m - P & Q \\ -\overline{Q} & \lambda I_m - P^T \end{pmatrix} \\ &= \begin{pmatrix} \overline{Q} & P^T - \lambda I_m \\ \lambda I_m - P & Q \end{pmatrix}. \end{aligned}$$

Therefore,

$$f_B(\lambda) = \det\begin{pmatrix} \lambda I_m - P & Q \\ -\overline{Q} & \lambda I_m - P^T \end{pmatrix} = \det\begin{pmatrix} \overline{Q} & P^T - \lambda I_m \\ \lambda I_m - P & Q \end{pmatrix}.$$

Note that

$$\begin{pmatrix} \overline{Q} & P^T - \lambda I_m \\ \lambda I_m - P & Q \end{pmatrix}^T = -\begin{pmatrix} \overline{Q} & P^T - \lambda I_m \\ \lambda I_m - P & Q \end{pmatrix},$$

which implies that $\begin{pmatrix} \overline{Q} & P^T - \lambda I_m \\ \lambda I_m - P & Q \end{pmatrix}$ is skew-symmetric. By [16], the determinant of $\begin{pmatrix} \overline{Q} & P^T - \lambda I_m \\ \lambda I_m - P & Q \end{pmatrix}$, also called its Pfaffian, can be written as the square of a polynomial in its entries. Therefore, $f_B(\lambda) = g(\lambda)^2$ where $g(\lambda) \in (\mathbb{C}[x])[\lambda]$.

Finally we show that $g(\lambda) \in (\mathbb{R}[x])[\lambda]$. Suppose otherwise. Then $g(\lambda) = a(\lambda) + b(\lambda)\mathbf{i}$ where $a)\lambda)$ and $b(\lambda) \in (\mathbb{R}[x])[\lambda]$ with $b(\lambda) \neq 0$. By (1), $g(\lambda)^2 = a(\lambda)^2 - b(\lambda)^2 + 2a(\lambda)b(\lambda)\mathbf{i} \in (\mathbb{R}[x])[\lambda]$. Hence $a(\lambda) = 0$ and $f_B(\lambda) = g(\lambda)^2 = (b(\lambda)\mathbf{i})^2 = -b(\lambda)^2$ where $b(\lambda) \in (\mathbb{R}[x])[\lambda]$. For a fixed $x \in \mathbb{R}$, let $\lambda' \in \mathbb{R}$ be large enough such that $\lambda' I_{2m} - \chi_B \in \mathbb{C}^{2m\times 2m}$ is diagonally dominant with non-negative diagonal entries and that $(b(x))(\lambda') \neq 0$. Since $\lambda' I_{2m} - \chi_B$ is also hermitian, $\lambda' I_{2m} - \chi_B$ is positive definite [9]. But $\det(\lambda' I_{2m} - \chi_B) = -((b(x))(\lambda'))^2 < 0$, a contradiction. Therefore, $b = 0$ and thus $f_B(\lambda) = g(\lambda)^2$ where $g(\lambda) \in (\mathbb{R}[x])[\lambda]$. □

Corollary 3.5. *Let $A \in \mathbb{H}[x]^{m\times n}$, $B = AA^*$ and $f_B(\lambda) = g(\lambda)^2$. Then $g(B) = 0$. We will call $g(\lambda)$ the generalized characteristic polynomial of A.*

Proof. Note that $g(\lambda) \in (\mathbb{R}[x])[\lambda]$ by Theorem 3.4. Then $\chi_{g(B)} = g(\chi_B)$. Next, $f_B(\chi_B) = 0$ by the Cayley–Hamilton theorem for complex polynomial matrices [9]. Therefore $g(\chi_B) = 0$, and $0 = g(\chi_B) = \chi_{g(B)}$, that is, $g(B) = 0$. □

From the definition, it is easy to check the following lemma, which have the analogues in the complex case.

Lemma 3.6. *Let $A \in \mathbb{H}[x]^{m\times n}$ have the Moore–Penrose inverse $A^\dagger$. Set $B = AA^*$. Then*

(i) $B^\dagger = (A^*)^\dagger A^\dagger$ *and* $B^\dagger B = AA^\dagger$.
(ii) $B^\dagger B = BB^\dagger$ *and* $(B^\dagger B)^2 = B^\dagger B$.
(iii) $(B^\dagger)^k = (B^k)^\dagger$ *and* $(B^{n-k})^\dagger B^{n-k} = B^\dagger B$, *for any* $k \in \mathbb{N}$.

The following result is well-known for quaternion case and it is easy to check that the result also holds for quaternion polynomials.

Lemma 3.7. *Let $A \in \mathbb{H}[x]^{m\times n}$, $B \in \mathbb{H}[x]^{p\times q}$ and $C \in \mathbb{H}[x]^{m\times q}$. If $A^\dagger$ and $B^\dagger$ both exist, then the quaternion polynomial matrix equation $AXB = C$ has a solution in $\mathbb{H}[x]^{n\times p}$ if and only if $AA^\dagger CB^\dagger B = C$, in which case the general solution is*

$$X = A^\dagger CB^\dagger + Y - A^\dagger AYBB^\dagger,$$

where $Y \in \mathbb{H}[x]^{n\times p}$ is arbitrary.

Theorem 3.8. *Let $A \in \mathbb{H}[x]^{m\times n}$ have the Moore–Penrose inverse $A^\dagger$ and $B = AA^*$. Suppose the generalized characteristic polynomial of A is*

$$g(\lambda) = \lambda^m + a_1\lambda^{m-1} + \cdots + a_k\lambda^{m-k} + \cdots + a_{m-1}\lambda + a_m,$$

where $a_i \in \mathbb{R}[x]$. If k is the largest integer such that $a_k \neq 0$, then the generalized inverse of A is given by

$$A^{\dagger} = -\frac{1}{a_k} A^* \left[B^{k-1} + a_1 B^{k-2} + \cdots + a_{k-1} I\right].$$

If $a_i = 0$ for all $1 \leq i \leq m$, then $A^{\dagger} = 0$.

Proof. The proof is similar to the complex case in [6] by using Corollary 3.5, Lemmas 3.6, 3.7 and 2.2. $\square$

From the above theorem, we can find the Moore–Penrose inverse $A^{\dagger}$ of A by computing its generalized characteristic polynomials. Faddeev [7] modified Leverrier's method and gave one algorithm to compute $\{a_i\}$ without computing $g(\lambda)$. Next we extend this algorithm to quaternion polynomial matrices.

Lemma 3.9. *Let $A \in \mathbb{H}[x]^{m \times n}$ have the Moore–Penrose inverse $A^{\dagger}$ and set $B = AA^*$. Then for $1 \leq k \leq m$,*

$$tr\left[\left(B^k + a_1 B^{k-1} + \cdots + a_{k-1} B\right)\right] = -k a_k,$$

where the a_i arise from the generalized characteristic polynomial of A:

$$g(\lambda) = \lambda^m + a_1 \lambda^{m-1} + \cdots + a_k \lambda^{m-k} + \cdots + a_{m-1}\lambda + a_m.$$

Proof. Let $Y = yI$ where $y \in \mathbb{R}$. We can write $g(Y)$ as:

$$\begin{aligned} g(Y) &= g(Y) - g(B) \\ &= (Y - B)\left[Y^{m-1} + (B + a_1 I) Y^{m-2} + \cdots + \left(B^{m-1} + a_1 B^{m-2} + \cdots + a_m I\right)\right]. \end{aligned}$$

As long as y is not an eigenvalue of B, $(yI - B) = Y - B$ is non-singular, so we can write:

$$\begin{aligned} (Y - B)^{-1} g(Y) = Y^{m-1} &+ (B + a_1 I) Y^{m-2} + \left(B^2 + a_1 B + a_2 I\right) Y^{m-3} + \cdots \\ &+ \left(B^{m-1} + a_1 B^{m-2} + \cdots + a_m I\right). \end{aligned}$$

Taking the traces gives:

$$\begin{aligned} \operatorname{tr}\left[(Y - B)^{-1} g(Y)\right] &= m y^{m-1} + \operatorname{tr}\left[(B + a_1 I)\right] y^{m-2} + \cdots \\ &\quad + \operatorname{tr}\left(B^{m-1} + a_1 B^{m-2} + \cdots + a_m I\right). \end{aligned}$$

Let $C = (Y - B)^{-1} g(Y)$. Since $g(Y) = g(yI) = g(y) I$, $C = g(y)(Y - B)^{-1}$. Therefore,

$$\mathrm{tr}C = g(y)\,\mathrm{tr}\left[(Y-B)^{-1}\right].$$

Let $\lambda_1, \cdots, \lambda_{m'}$, where $m' \leq m$, be all the non-zero eigenvalues of B. $\mathrm{tr}\left[(Y-B)^{-1}\right]$ is the sum of the eigenvalue of $(Y-B)^{-1}$. We will show that these eigenvalues are the fractions $\frac{1}{y-\lambda_1}, \cdots, \frac{1}{y-\lambda_{m'}}$.

Let ζ be an eigenvalue of $(Y-B)^{-1}$ with corresponding eigenvector Z such that:

$$(Y-B)^{-1} Z = Z\zeta.$$

ζ is real by Lemma 3.1, and hence

$$(Y-B)\,Z = Z\frac{1}{\zeta} \implies BZ = Z\left(y - \frac{1}{\zeta}\right).$$

Therefore, $y - \frac{1}{\zeta} = \lambda_i \Rightarrow \zeta = \frac{1}{y-\lambda_i}$ for some $1 \leq i \leq m'$.

Since $g(y) = (y-\lambda_1)(y-\lambda_2)\cdots(y-\lambda_{m'})$, we have that $g'(y) = g(y)\left(\frac{1}{y-\lambda_1} + \cdots + \frac{1}{y-\lambda_{m'}}\right)$ and $\mathrm{tr}C = g'(y)$. The derivative of g is also equal to:

$$g'(y) = my^{m-1} + a_1(m-1)\,y^{m-2} + \cdots + a_{m-1}.$$

Therefore,

$$\begin{aligned} &my^{m-1} + a_1(m-1)\,y^{m-2} + \cdots + a_{m-1} \\ &\quad = my^{m-1} + \mathrm{tr}\,(B + a_1 I)\,y^{m-2} + \cdots + \mathrm{tr}\left(B^{m-1} + a_1 B^{m-2} + \cdots + a_m I\right). \end{aligned}$$

Comparing the coefficient of y^{m-k-1} on both sides, we obtain

$$\begin{aligned} a_k(m-k) &= \mathrm{tr}\left(B^k + a_1 B^{k-1} + \cdots + a_{k-1}B + a_k I\right) \\ &= \mathrm{tr}\left(B^k + a_1 B^{k-1} + \cdots + a_{k-1}B\right) + \mathrm{tr}(a_k I), \end{aligned}$$

and then

$$-ka_k = \mathrm{tr}\left(B^k + a_1 B^{k-1} + \cdots + a_{k-1}B\right). \qquad \square$$

Next we give the Leverrier–Faddeev algorithm for finding Moore–Penrose inverses of quaternion polynomial matrices.

Proposition 3.10. *Let $A \in \mathbb{H}[x]^{m\times n}$ have the generalized inverse $A^\dagger$ and $B = AA^*$. Suppose the generalized characteristic polynomial of A is*

$$g(\lambda) = \lambda^m + a_1\lambda^{m-1} + \cdots + a_k\lambda^{m-k} + \cdots + a_{m-1}\lambda + a_m,$$

where $a_i \in \mathbb{R}[x]$. Define $a_0 = 1$. If p is the largest integer such that $a_p \neq 0$ and we construct the sequence $A_0, \cdots, A_p$ as follows:

$$\begin{array}{llllll} A_0 & = 0 & -1 & = q_0 & B_0 & = I \\ A_1 & = AA^*B_0 & \frac{trA_1}{1} & = q_1 & B_1 & = A_1 - q_1 I \\ & \vdots & & \vdots & & \vdots \\ A_{p-1} & = AA^*B_{p-2} & \frac{trA_{p-1}}{p-1} & = q_{p-1} & B_{p-1} & = A_{p-1} - q_{p-1}I \\ A_p & = AA^*B_{p-1} & \frac{trA_p}{p} & = q_p & B_p & = A_p - q_p I \end{array}$$

then $q_i(x) = -a_i(x)$, $i = 0, \cdots, p$.

Proof. We will show $q_i(x) = -a_i(x)$ by mathematical induction. By the definition, clearly $q_0 = -a_0$ holds.

Now we assume that $q_i(x) = -a_i(x)$ holds for all $0 \leq i \leq k-1$. Then

$$\begin{aligned} A_k &= AA^*B_{k-1} \\ &= BB_{k-1} \\ &= B(A_{k-1} - q_{k-1}I) \\ &= B(B(A_{k-2} - q_{k-2}I) - q_{k-1}I) \\ &\quad \cdots \\ &= B^k - q_1 B^{k-1} - q_2 B^{k-2} - \cdots - q_{k-1}B \\ &= B^k + a_1 B^{k-1} + a_2 B^{k-2} + \cdots + a_{k-1}B, \end{aligned}$$

and thus

$$\operatorname{tr}(A_k) = \operatorname{tr}\left(B^k + a_1 B^{k-1} + \cdots + a_{k-1}B\right),$$

which, by Lemma 3.9, is equal to $-ka_k$. So $q_k = \frac{\operatorname{tr}A_k}{k} = -a_k$. Therefore, $q_i(x) = -a_i(x)$ for all $p \geq i \geq 0$. $\square$

Now combining Theorem 3.8 and Proposition 3.10, we have the following algorithm to compute Moore–Penrose inverses.

Algorithm 3.11. Leverrier–Faddeev algorithm for quaternion polynomial matrices

Input: $A \in \mathbb{H}[x]^{m\times n}$
Output: The Moore–Penrose inverse $A^\dagger$ of A in $\mathbb{H}[x]^{n\times m}$ if exists.

1. $B_0 \leftarrow I_m$, $a_0 \leftarrow 1$
2. for $i = 1, \ldots, m$ do
   $A_i \leftarrow AA^*B_{i-1}$, $a_i \leftarrow -\frac{\operatorname{tr}A_i}{i}$, $B_i \leftarrow A_i + a_i I_m$

3. Find the maximal index p such that $a_p \neq 0$.
4. Return $A^\dagger = \begin{cases} -\frac{1}{a_p} A^* B_{p-1}, & p > 0, \\ 0, & p = 0. \end{cases}$

Note that we have to compute many matrix products in Proposition 3.10, which means that Leverrier–Faddeev method is not efficient. In the next section, we will give an efficient way by combining Theorem 3.8 and interpolation methods.

4. Finding Moore–Penrose inverses by interpolation

In this section we present an efficient method to obtain the Moore–Penrose inverse of a quaternion polynomial matrix by interpolation at data points of real numbers.

Since $\mathbb{H}$ is a skew field, the roots of quaternion polynomials in $\mathbb{H}[x]$ and interpolations are quite complicated comparing with complex number case. First we recall some basic definitions and properties.

An element $r \in \mathbb{H}$ is a root of a nonzero polynomial $f \in \mathbb{H}[x]$ iff $x - r$ is a right divisor of f. The set of polynomials in $\mathbb{H}[x]$ having r as a root is the left ideal $\mathbb{H}[x] \cdot (x - r)$. It is worth mentioning that the evaluations of quaternion polynomials are quite different from commutative case. Let f, g and $h \in \mathbb{H}[x]$, $f = gh$ and $r \in \mathbb{H}$. If $h(r) = 0$, then $f(r) = 0$. Otherwise, set $\beta = h(r) \neq 0$. Then the evaluation of $f(x)$ at $x = r$ is defined as

$$f(r) = g\left(\beta r \beta^{-1}\right) h(r). \tag{2}$$

In particular, if r is a root of f but not of h, then $\beta r \beta^{-1}$ is a root of g. We refer the reader to [14] for more details.

In [8], Gordon and Motzkin proved that if $f \in \mathbb{H}[x]$ is of degree n, then the roots of f lie in at most n conjugacy classes of $\mathbb{H}$.

It is well-known that Newton's interpolation and Lagrange's interpolation play important roles in studying polynomials over fields. Unfortunately one cannot get similar nice formulas in quaternion case. But we still can find a quaternion polynomial from a given set of points.

Lemma 4.1. *Let $c_1, \ldots, c_n$ be n pairwise non-conjugate elements of $\mathbb{H}$. Then there is a unique monic polynomial $g_n \in \mathbb{H}[x]$ of degree n such that $g_n(c_1) = \cdots = g_n(c_n) = 0$. Moreover, $c_1, \ldots, c_n$ are the only roots (up to conjugacy classes) of g_n in $\mathbb{H}$.*

Proof. We first show the existence of g_n for all $n \geq 1$ by mathematical induction. For $n = 1$, it is trivially true as $g = x - c_1$.

Suppose the claim holds for all $1 \leq n \leq k - 1$. Let $c_1, \cdots, c_k \in \mathbb{H}$ be pairwise non-conjugate. Invoking the inductive hypothesis, there exists a monic polynomial g_{k-1} of degree $k - 1$ with $c_2, \cdots, c_k$ as its only roots (up to conjugacy classes), that is, $g_{k-1}(c_1) \neq 0$. Construct g_k as follows,

$$g_k(x) = \left(x - g_{k-1}(c_1)\, c_1 g_{k-1}(c_1)^{-1}\right) \cdot g_{k-1}(x).$$

By Eq. (2), $g_k(c_1) = 0$. Thus, the claim holds for k. Therefore this claim holds for all $n \geq 1$.

We next show that g_n is unique. For a fixed n, let $g' \neq g_n$ be a monic polynomial of degree n such that $g'(c_1) = \cdots = g'(c_n) = 0$ too. Then $\deg(g_n - g') \leq n - 1$ but $g_n - g'$ has roots $c_1, \cdots, c_n$ which lie in n different conjugacy classes of $\mathbb{H}$, a contradiction. Therefore, g_n is unique for all $n \geq 1$. □

Proposition 4.2. *Let $c_1, \cdots, c_{n+1} \in \mathbb{H}$ be pairwise non-conjugate and let $d_1, \cdots, d_{n+1} \in \mathbb{H}$. Then there exists a unique lowest degree polynomial $f \in \mathbb{H}[x]$, of degree $p \leq n$, such that $f(c_i) = d_i$ for all $1 \leq i \leq n+1$.*

Proof. For any $1 \leq s' \leq n+1$, let $S = \{1, \cdots, n+1\} \setminus \{s'\}$. By Lemma 4.1, we can find a unique monic $h_S \in \mathbb{H}[x]$ of degree n such that $h_S(c_s) = 0,\ s \in S$ and that $\{c_s \mid s \in S\}$ are the only roots (up to conjugacy class) of h_S in $\mathbb{H}$. Then $h_S(c_{s'}) \neq 0$, and thus we can construct a quaternion polynomial g_S of degree n such that $g_S(c_\alpha) = \begin{cases} 0 & \alpha \in S, \\ 1 & \alpha = s', \end{cases}$ as follows:

$$g_S(x) = h_S(c_{s'})^{-1} h_S(x).$$

Furthermore we construct a quaternion polynomial f of degree at most n such that $f(c_i) = d_i$ for all $1 \leq i \leq n+1$ as follows:

$$f = \sum_{s'=1}^{n+1} d_{s'} g_S.$$

Finally we show that f is unique. Suppose we have $f' \in \mathbb{H}[x]$ of degree $p' \leq n$ such that $f' \neq f$ and that $f'(c_i) = d_i$ for all $1 \leq i \leq n+1$ too. Then $f - f' \neq 0$ is of degree at most $\leq n$. But $f - f'$ has roots $c_1, \cdots, c_{n+1}$ which lie in $n+1$ conjugacy classes of $\mathbb{H}$, a contradiction. Therefore, f is unique. □

Next we consider the interpolation for quaternion polynomial matrices. Recall that the degree of a given $A \in \mathbb{H}[x]^{m\times n}$ is equal to $\deg A = \max\{\deg(A_{ij}) \mid 1 \leq i \leq m,\ 1 \leq j \leq n\}$. Then the upper bound of the degree of its Moore–Penrose inverse (if it exists) $A^\dagger$ is given by the following lemma.

Lemma 4.3. *Let $A \in \mathbb{H}[x]^{m\times n}$ have the Moore–Penrose inverse $A^\dagger$. Then*

$$\deg A^\dagger \leq (2m-1)\deg A.$$

Proof. By Theorem 3.8, $\deg A^\dagger \leq \deg\left(A^*\left(AA^*\right)^{m-2}\right) \leq \deg\left(A^{2m-1}\right) \leq (2m-1)\deg A$. $\square$

For $A = (A_{ij}) \in \mathbb{H}[x]$ and $c \in \mathbb{H}$, the evaluation of A at c can be defined as entry-wise in a common sense, that is, $A(c) = (A_{ij}(c))$. One has to pay an attention that the evaluations of quaternion polynomials have some special rules as we explained at the beginning of this section.

Proposition 4.4. *Let $c_1, \cdots, c_{k+1} \in \mathbb{H}$ be pairwise non-conjugate and let $A_1, \cdots, A_{k+1} \in \mathbb{H}^{n\times m}$. Then there is a unique lowest degree matrix $A \in \mathbb{H}\left[x\right]^{n\times m}$ of degree $p \leq k$, such that $A\left(c_i\right) = A_i$ for all $1 \leq i \leq k+1$.*

Proof. For any $1 \leq n_1 \leq n$ and $1 \leq m_1 \leq m$, by Proposition 4.2, there is a lowest degree polynomial $A_{n_1m_1}(x)$ determined by the values $c_1, \cdots, c_{k+1}$ and $\left(A_1\right)_{n_1m_1}, \cdots, \left(A_{k+1}\right)_{n_1m_1}$. In fact, for any $1 \leq s' \leq k+1$, let $S = \{1, \cdots, k+1\} \setminus \{s'\}$. Then

$$A_{n_1m_1}(x) = \sum_{s'=1}^{k+1} \left(A_{s'}\right)_{n_1m_1} g_S(x),$$

where $g_S\left(c_\alpha\right) = \begin{cases} 0 & \alpha \in S, \\ 1 & \alpha = s'. \end{cases}$ Since n_1 and m_1 are chosen randomly, a lowest degree matrix A that satisfies $A\left(c_i\right) = A_i$ for all $1 \leq i \leq k+1$ is determined by $A = (\sum_{s'=1}^{k+1} A_{s'} g_S)$.

Next we show that A is unique. Suppose $A' \neq A$ of degree $p' \leq p$ also satisfies $A'\left(c_i\right) = A_i$ for all $1 \leq i \leq k+1$. Then for some $1 \leq n_2 \leq n$ and $1 \leq m_2 \leq m$, $\left(A - A'\right)_{n_2m_2} \neq 0$. But $\left(A - A'\right)_{n_2m_2}$, of degree at most $p \leq k$, has roots $c_1, \cdots, c_{k+1}$ which lie in $k+1$ conjugacy classes of $\mathbb{H}$, a contradiction. Therefore, A is unique. $\square$

Let $A \in \mathbb{H}\left[x\right]^{m\times n}$ have the Moore–Penrose inverse $A^\dagger$, and set $B = AA^*$. Let p be the largest integer such that $a_p \neq 0$. We can construct the sequence $A_0, \cdots, A_p$ as in Proposition 3.10. The next theorem gives the interpolation version of Leverrier–Faddeev algorithm.

Theorem 4.5. *In the above setting, let $k = (2m-1)\deg A$ and $c_1, \cdots, c_{k+1} \in \mathbb{R}$ be $k+1$ distinct real numbers such that $q_p\left(c_{s'}\right) \neq 0$ for any $1 \leq s' \leq k+1$. Let $S = \{1, \cdots, k+1\} \setminus \{s'\}$. Then*

$$A^\dagger = \sum_{s'=1}^{k+1} A\left(c_{s'}\right)^\dagger g_S$$

where

$$A\left(c_{s'}\right)^\dagger = \frac{1}{q_p\left(c_{s'}\right)} A\left(c_{s'}\right)^* \left[B\left(c_{s'}\right)^{p-1} - q_1\left(c_{s'}\right) B\left(c_{s'}\right)^{p-2} - \cdots - q_{p-1}\left(c_{s'}\right) I\right]$$

and

$$g_S(c_\alpha) = \begin{cases} 0 & \alpha \in S, \\ 1 & \alpha = s'. \end{cases}$$

Proof. It follows from Theorem 3.8, Proposition 3.10 and Proposition 4.4. □

5. Implementation and examples

The calculations of quaternions are very complicated and time-consuming. It is almost impossible to do some calculations for quaternion polynomials and quaternion polynomial matrices with a little big size by hand. There are only few quaternion packages in computer algebra system Maple. But no one has commends for quaternion polynomials and quaternion polynomial matrices. In [10], we developed a Maple package which includes all basic operations for quaternion polynomials and quaternion polynomial matrices. In particular, we implemented all algorithms in this paper. Next we will give some examples by using our Maple package.

Example 5.1. Find, by interpolation, the generalized inverse of the quaternion polynomial matrix

$$A = \begin{pmatrix} 14x+14+76\mathbf{i}+70\mathbf{j}+56\mathbf{k} & 56-28\mathbf{i}-70\mathbf{j}+70\mathbf{k} & 28\mathbf{j}-56\mathbf{k} & 14x-56-8\mathbf{i}-14\mathbf{j}-56\mathbf{k} \\ -2x-2-43\mathbf{i}-10\mathbf{j}-8\mathbf{k} & -8+4\mathbf{i}+10\mathbf{j}-10\mathbf{k} & -4\mathbf{j}+8\mathbf{k} & -2x+8-31\mathbf{i}+2\mathbf{j}+8\mathbf{k} \\ -3x-3+3\mathbf{i}-15\mathbf{j}-12\mathbf{k} & -12+6\mathbf{i}+15\mathbf{j}-15\mathbf{k} & -6\mathbf{j}+12\mathbf{k} & -3x+12+21\mathbf{i}+3\mathbf{j}+12\mathbf{k} \\ -4x-4+4\mathbf{i}-20\mathbf{j}-16\mathbf{k} & -16+8\mathbf{i}+20\mathbf{j}-20\mathbf{k} & -8\mathbf{j}+16\mathbf{k} & -4x+16+28\mathbf{i}+4\mathbf{j}+16\mathbf{k} \end{pmatrix} \in \mathbb{H}^{4\times 3}[x]$$

From Lemma 4.3, we know that the upper bound of the degree of $A^\dagger$ is less than $(2m-1)\deg A = (2\times 4-1)\cdot 1 = 7$. In practice, we don't need to start from the upper bound. For this example, we may guess $\deg A^\dagger = 2$, and choose $c_1 = 0$ and $c_2 = 1$. Then using our Maple package, it is easy to do the following calculations:

$$A(c_1) = \begin{pmatrix} 14+76\mathbf{i}+70\mathbf{j}+56\mathbf{k} & 56-28\mathbf{i}-70\mathbf{j}+70\mathbf{k} & 28\mathbf{j}-56\mathbf{k} & -56-8\mathbf{i}-14\mathbf{j}-56\mathbf{k} \\ -2-43\mathbf{i}-10\mathbf{j}-8\mathbf{k} & -8+4\mathbf{i}+10\mathbf{j}-10\mathbf{k} & -4\mathbf{j}+8\mathbf{k} & 8-31\mathbf{i}+2\mathbf{j}+8\mathbf{k} \\ -3+3\mathbf{i}-15\mathbf{j}-12\mathbf{k} & -12+6\mathbf{i}+15\mathbf{j}-15\mathbf{k} & -6\mathbf{j}+12\mathbf{k} & 12+21\mathbf{i}+3\mathbf{j}+12\mathbf{k} \\ -4+4\mathbf{i}-20\mathbf{j}-16\mathbf{k} & -16+8\mathbf{i}+20\mathbf{j}-20\mathbf{k} & -8\mathbf{j}+16\mathbf{k} & 16+28\mathbf{i}+4\mathbf{j}+16\mathbf{k} \end{pmatrix}$$

and

$$A(c_2) = \begin{pmatrix} 28+76\mathbf{i}+70\mathbf{j}+56\mathbf{k} & 56-28\mathbf{i}-70\mathbf{j}+70\mathbf{k} & 28\mathbf{j}-56\mathbf{k} & -42-8\mathbf{i}-14\mathbf{j}-56\mathbf{k} \\ -4-43\mathbf{i}-10\mathbf{j}-8\mathbf{k} & -8+4\mathbf{i}+10\mathbf{j}-10\mathbf{k} & -4\mathbf{j}+8\mathbf{k} & 6-31\mathbf{i}+2\mathbf{j}+8\mathbf{k} \\ -6+3\mathbf{i}-15\mathbf{j}-12\mathbf{k} & -12+6\mathbf{i}+15\mathbf{j}-15\mathbf{k} & -6\mathbf{j}+12\mathbf{k} & 9+21\mathbf{i}+3\mathbf{j}+12\mathbf{k} \\ -8+4\mathbf{i}-20\mathbf{j}-16\mathbf{k} & -16+8\mathbf{i}+20\mathbf{j}-20\mathbf{k} & -8\mathbf{j}+16\mathbf{k} & 12+28\mathbf{i}+4\mathbf{j}+16\mathbf{k} \end{pmatrix}.$$

By the algorithm stated in Theorem 4.5, we calculate and obtain:

$$A(c_1)^\dagger = A(0)^\dagger = \frac{1}{230175} \times$$

$$\begin{pmatrix} 140-560\mathbf{i}-228\mathbf{j}-342\mathbf{k} & 355+1730\mathbf{i}-96\mathbf{j}+81\mathbf{k} & -255-870\mathbf{i}+126\mathbf{j}+54\mathbf{k} & -340-1160\mathbf{i}+168\mathbf{j}+72\mathbf{k} \\ 276+88\mathbf{i}+426\mathbf{j}-382\mathbf{k} & 282+416\mathbf{i}-93\mathbf{j}-149\mathbf{k} & -252-276\mathbf{i}-72\mathbf{j}+204\mathbf{k} & -336-368\mathbf{i}-96\mathbf{j}+272\mathbf{k} \\ 32+16\mathbf{i}-176\mathbf{j}+292\mathbf{k} & -176-88\mathbf{i}+68\mathbf{j}+194\mathbf{k} & 96+48\mathbf{i}+12\mathbf{j}-204\mathbf{k} & 128+64\mathbf{i}+16\mathbf{j}-272\mathbf{k} \\ -140-122\mathbf{i}+228\mathbf{j}+342\mathbf{k} & -355+2021\mathbf{i}+96\mathbf{j}-81\mathbf{k} & 255-1176\mathbf{i}-126\mathbf{j}-54\mathbf{k} & 340-1568\mathbf{i}-168\mathbf{j}-72\mathbf{k} \end{pmatrix}$$

and

$$A\left(c_2\right)^{\dagger}=A\left(1\right)^{\dagger}=\frac{1}{230175}\times$$
$$\begin{pmatrix} 152-550\mathbf{i}-244\mathbf{j}-330\mathbf{k} & 289+1675\mathbf{i}-8\mathbf{j}+15\mathbf{k} & -219-840\mathbf{i}+78\mathbf{j}+90\mathbf{k} & -292-1120\mathbf{i}+104\mathbf{j}+120\mathbf{k} \\ 268+104\mathbf{i}+406\mathbf{j}-402\mathbf{k} & 326+328\mathbf{i}+17\mathbf{j}-39\mathbf{k} & -276-228\mathbf{i}-132\mathbf{j}+144\mathbf{k} & -368-304\mathbf{i}-176\mathbf{j}+192\mathbf{k} \\ 32+16\mathbf{i}-160\mathbf{j}+300\mathbf{k} & -176-88\mathbf{i}-20\mathbf{j}+150\mathbf{k} & 96+48\mathbf{i}+60\mathbf{j}-180\mathbf{k} & 128+64\mathbf{i}+80\mathbf{j}-240\mathbf{k} \\ -152-132\mathbf{i}+244\mathbf{j}+330\mathbf{k} & -289+2076\mathbf{i}+8\mathbf{j}-15\mathbf{k} & 219-1206\mathbf{i}-78\mathbf{j}-90\mathbf{k} & 292-1608\mathbf{i}-104\mathbf{j}-120\mathbf{k} \end{pmatrix}.$$

By Theorem 4.5, we have:

$$\begin{aligned} A^{\dagger} &= \sum_{s'=1}^{2} A\left(c_{s'}\right)^{\dagger} g_S = A\left(0\right)^{\dagger}\left(1-x\right)+A\left(1\right)^{\dagger} x \\ &= \frac{1}{230175}\times \end{aligned}$$
$$\left(\begin{matrix} (12+10\mathbf{i}-16\mathbf{j}+12\mathbf{k})\,x+140-560\mathbf{i}-228\mathbf{j}-342\mathbf{k} & (-66-55\mathbf{i}+88\mathbf{j}-66\mathbf{k})\,x+355+1730\mathbf{i}-96\mathbf{j}+81\mathbf{k} \\ (-8+16\mathbf{i}-20\mathbf{j}-20\mathbf{k})\,x+276+88\mathbf{i}+426\mathbf{j}-382\mathbf{k} & (44-88\mathbf{i}+110\mathbf{j}+110\mathbf{k})\,x+282+416\mathbf{i}-93\mathbf{j}-149\mathbf{k} \\ (16\mathbf{j}+8\mathbf{k})\,x+32+16\mathbf{i}-176\mathbf{j}+292\mathbf{k} & (-88\mathbf{j}-44\mathbf{k})\,x-176-88\mathbf{i}+68\mathbf{j}+194\mathbf{k} \\ (-12-10\mathbf{i}+16\mathbf{j}-12\mathbf{k})\,x-140-122\mathbf{i}+228\mathbf{j}-342\mathbf{k} & (66+55\mathbf{i}-88\mathbf{j}+66\mathbf{k})\,x-355+2021\mathbf{i}+96\mathbf{j}-81\mathbf{k} \end{matrix}\right.$$
$$\left.\begin{matrix} (36+30\mathbf{i}-48\mathbf{j}+36\mathbf{k})\,x-255-870\mathbf{i}+126\mathbf{j}+54\mathbf{k} & (48+40\mathbf{i}-64\mathbf{j}+48\mathbf{k})\,x-340-1160\mathbf{i}+168\mathbf{j}+72\mathbf{k} \\ (-24+48\mathbf{i}-60\mathbf{j}-60\mathbf{k})\,x-252-276\mathbf{i}-72\mathbf{j}+204\mathbf{k} & (-32+64\mathbf{i}-80\mathbf{j}+80\mathbf{k})\,x-366-368\mathbf{i}-96\mathbf{j}+272\mathbf{k} \\ (48\mathbf{j}+24\mathbf{k})\,x+96+48\mathbf{i}+12\mathbf{j}-204\mathbf{k} & (64\mathbf{j}+32\mathbf{k})\,x+128+64\mathbf{i}+16\mathbf{j}-272\mathbf{k} \\ (-36-30\mathbf{i}+48\mathbf{j}-36\mathbf{k})\,x+255-1176\mathbf{i}-126\mathbf{j}-54\mathbf{k} & (-48-40\mathbf{i}+64\mathbf{j}-48\mathbf{k})\,x+340-1568\mathbf{i}-168\mathbf{j}-72\mathbf{k} \end{matrix}\right).$$

Direct calculation shows that $A^{\dagger}$ satisfies the four defining relations of the Moore–Penrose inverse. Therefore $A^{\dagger}$ is the Moore–Penrose inverse of A.

Example 5.2. As special cases, the algorithms in this paper and our Maple package can compute the Moore–Penrose inverses for (real) complex polynomial matrices and quaternion matrices.

For instance, let $A=\begin{pmatrix} 1 & \mathbf{i}+2\mathbf{k} & 3 \\ \mathbf{i} & 6+\mathbf{j} & 7 \end{pmatrix}_{2\times 3}$. Then its Moore–Penrose inverse can be found by using our Maple package as follows:

$$A^{\dagger}=\begin{pmatrix} \frac{47}{347}+\frac{21}{694}\mathbf{i}+\frac{11}{694}\mathbf{j} & -\frac{21}{694}-\frac{11}{347}\mathbf{i}-\frac{11}{694}\mathbf{k} \\ -\frac{63}{347}-\frac{28}{347}\mathbf{i}+\frac{21}{694}\mathbf{j}-\frac{101}{694}\mathbf{k} & \frac{61}{694}+\frac{21}{694}\mathbf{i}-\frac{6}{347}\mathbf{j}+\frac{21}{347}\mathbf{k} \\ \frac{57}{347}+\frac{49}{694}\mathbf{i}+\frac{77}{694}\mathbf{k} & \frac{21}{347}-\frac{21}{694}\mathbf{i}-\frac{33}{694}\mathbf{k} \end{pmatrix}_{3\times 2}.$$

Clearly these examples also show that it could be very time-consuming to find the Moore–Penrose inverses by hand.

Acknowledgments

We would like to thank two anonymous referees for their valuable comments. This research was supported by the grants from the National Natural Science Foundation of China (11171205), the Natural Science Foundation of Shanghai (11ZR1412500), Shanghai Leading Academic Discipline Project (J50101), the National Sciences and Engineering Research Council of Canada (NSERC) and URGP from the University of Manitoba.

References

[1] A. Baker, Right eigenvalues for quaternionic matrices: a topological approach, Linear Algebra Appl. 286 (1–3) (1999) 303–309.

[2] S. Barnett, Leverrier's algorithm: a new proof and extensions, SIAM J. Matrix Anal. Appl. 10 (1989) 551–556.

[3] A. Ben-Israel, T.N.E. Greville, Generalized Inverses: Theory and Applications, second ed., Springer-Verlag, New York, 2003.

[4] J.L. Brenner, Matrices of quaternions, Pacific J. Math. 1 (1951) 329–335.

[5] A. Damiano, G. Gentili, D. Sreuppa, Computations in the ring of quaternionic polynomials, J. Symbolic Comput. 45 (2010) 38–45.

[6] H.P. Decell, An application of the Cayley–Hamilton theorem to generalized matrix inversion, SIAM Rev. 7 (1965) 526–528.

[7] V.N. Faddeeva, Computational Methods of Linear Algebra, Dover Books on Advanced Mathematics, Dover Publications, 1959.

[8] B. Gordon, T.S. Motzkin, On the zeros of polynomials over division rings, Trans. Amer. Math. Soc. 116 (1965) 218–226.

[9] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 1990.

[10] Liji Huang, Yang Zhang, Maple Package for Quaternion Matrices, The University of Manitoba, 2013.

[11] J. Jones, N.P. Karampetakis, A.C. Pugh, The computation and application of the generalized inverse via Maple, J. Symbolic Comput. 25 (1998) 99–124.

[12] I.I. Kyrchei, Determinantal representation of the Moore–Penrose inverse matrix over the quaternion skew field, J. Math. Sci. 180 (1) (2012) 23–33.

[13] I.I. Kyrchei, Cramers rule for some quaternion matrix equations, Appl. Math. Comput. 217 (2010) 2024–2030.

[14] T.Y. Lam, A First Course in Noncommutative Rings, Graduate Texts in Mathematics, vol. 131, Springer, 2001.

[15] H.C. Lee, Eigenvalues and canonical forms of matrices with quaternion coefficients, Proc. R. Ir. Acad. A Math. Phys. Sci. 52 (1949) 253–260.

[16] T. Muir, A Treatise on the Theory of Determinants, Dover Phoenix Editions, Dover Publications, 2003.

[17] I. Niven, Equations in quaternions, Amer. Math. Monthly 48 (1941) 654–661.

[18] I. Niven, The roots of a quaternion, Amer. Math. Monthly 49 (1942) 386–388.

[19] G. Opfer, Polynomials and Vandermonde matrices over the field of quaternions, Electron. Trans. Numer. Anal. 36 (2009) 9–16.

[20] R. Pereira, P. Rocha, On the determinant of quaternionic polynomial matrices and its application to system stability, Math. Methods Appl. Sci. 31 (2008) 99–122.

[21] R. Pereira, P. Rocha, P. Vettori, Algebraic tools for the study of quaternionic behavioral systems, Linear Algebra Appl. 400 (2005) 121–140.

[22] R. Pereira, P. Vettori, Stability of quaternionic linear systems, IEEE Trans. Automat. Control 51 (2006) 518–523.

[23] R. Penrose, A generalized inverse for matrices, Math. Proc. Cambridge Philos. Soc. 51 (1955) 406–413.

[24] R. Puystjens, Moore–Penrose inverses for matrices over some Noetherian rings, J. Pure Appl. Algebra 31 (1–3) (1984) 191–198.

[25] P.S. Stanimirovic, M.D. Petkovic, Computing generalized inverse of polynomial matrices by interpolation, Appl. Math. Comput. 172 (2006) 508–523.

[26] F. Zhang, Quaternions and matrices of quaternions, Linear Algebra Appl. 251 (0) (1997) 21–57.

[27] M. Ziegler, Quasi-optimal arithmetic for quaternion polynomials, ISAAC 2003, pp. 705–715.