Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 18, No. 1, pp. 55–75, 2022

COLOR IMAGE COMPLETION USING A LOW-RANK QUATERNION MATRIX APPROXIMATION*

Yannan Chen, Liqun Qi and Xinzhen Zhang†

Abstract: In this paper, we propose a lower rank quaternion matrix approximation algorithm and apply it to color image completion. We introduce a concise form for the gradient of a real function in quaternion matrix variables. The optimality conditions of our quaternion least squares problem have a simple expression with this form. The convergence and convergence rate of our algorithm are established with this tool. Numerical tests on a set of color images demonstrate the efficiency of the algorithm.

Key words: *quaternion matrix, gradient, lower rank approximation, color image completion*

Mathematics Subject Classification: *15A23, 15A83, 65F55, 90C90*

---

1 Introduction

In 1996, Sangwine [9] proposed to encode three channel components of an RGB (red-green-blue) image on the three imaginary parts of a pure quaternion matrix. In 1997, Zhang [19] established the quaternion singular value decomposition (QSVD) of a quaternion matrix. Also see [6, 10]. Since then, quaternion matrices and their QSVD are used widely in color image processing, including color image denoising, completion and inpainting. Recently, Chen, Xiao and Zhou [4] proposed the low-rank quaternion approximation (LRQA) for color image denoising and inpainting. They proposed to use the quaternion rank or some rank surrogates of a quaternion variable matrix X as a regularized term in the approximation of X to a given quaternion data matrix Y.

In this paper, we propose a lower rank quaternion matrix approximation (LRQMA) algorithm and apply it to color image completion. To conduct convergence analysis for our algorithm, we need to consider optimization problems of real functions in quaternion matrix variables. To handle this, we introduce a concise form for the gradient of a real function in quaternion matrix variables. The optimality conditions of our quaternion least squares problem have a simple expression with this form. With this tool, convergence and convergence rate of our algorithm are established. Numerical tests on practical datasets demonstrate the efficiency of the algorithm.

*This work was supported by the National Natural Science Foundation of China (No. 11771405,11871369, 12071159 and 12171168) and the Guangdong Basic and Applied Basic Research Foundation (2020A1515010489).

†Cesponding author.

The rest of the paper is organized as follows. Some necessary preliminary knowledge on quaternions and quaternion matrices are given in the next section. A lower rank quaternion matrix approximation algorithm is presented in Section 3. In Section 4, we introduce a concise form for the gradient of a real function in quaternion matrix variables. The gradients of some real functions in quaternion matrix variables, which are useful for our least squares minimization, have simple expressions. The convergence analysis of our algorithm is presented in Section 5. We show that the Kurdyka-Łojasiewicz inequality [1, 2, 3] holds for the critical points of our algorithm there. The LRQMA algorithm is applied to color image completion. Numerical tests on practical datasets are reported in Section 6. In Section 7, we make some concluding remarks.

2 Preliminary

2.1 Quaternions

In this paper, the real field, the complex field and the quaternion field are denoted by $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{Q}$, respectively. Furthermore, scalars, vectors, matrices and tensors are denoted by small letters, bold small letters, capital letters and calligraphic letters, respectively. We denote vectors with matrix components by bold capital letters. For example, we denote $\mathbf{Z} = (A, B, X)$. We use $\mathbf{0}, O$ and $\mathcal{O}$ to denote zero vector, zero matrix and zero tensor with adequate dimensions. An exception is that $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ denote the three imaginary units of quaternions. We use the notation in [5, 18, 19]. We have

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

which means

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

These rules, along with the distribution law, determine the product of two quaternions. Hence, the multiplication of quaternions is noncommutative.

Let $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} \in \mathbb{Q}$, where $x_0, x_1, x_2, x_3 \in \mathbb{R}$. We define $\mathbf{Re}(x) = x_0$ the real part of x and $\mathbf{Im}(x) = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ the imaginary part of x. Then the conjugate of x is

$$\bar{x} \equiv x^* = x_0 - x_1\mathbf{i} - x_2\mathbf{j} - x_3\mathbf{k},$$

the modulus of x is

$$|x| = |x^*| = \sqrt{xx^*} = \sqrt{x^*x} = \sqrt{x_0^2 + x_1^2 + x_2^2 + x_3^2},$$

and if $x \neq 0$, then $x^{-1} = \frac{x^*}{|x|^2}$.

2.2 Quaternion Matrices

Denote the collections of real, complex and quaternion $m \times n$ matrices by $\mathbb{R}^{m\times n}$, $\mathbb{C}^{m\times n}$ and $\mathbb{Q}^{m\times n}$, respectively. Then $A = (a_{ij}) \in \mathbb{Q}^{m\times n}$ can be denoted as

$$A = A_0 + A_1\mathbf{i} + A_2\mathbf{j} + A_3\mathbf{k}, \tag{2.1}$$

where $A_0, A_1, A_2, A_3 \in \mathbb{R}^{m\times n}$. The transpose of A is $A^\top = (a_{ji})$. The conjugate of A is $\bar{A} = (a_{ij}^*)$. The conjugate transpose of A is $A^* = (a_{ji}^*) = \bar{A}^T$. The Frobenius norm of A is

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2}.$$

By [18], we have the following proposition.

Proposition 2.1. *Suppose that* $A \in \mathbb{Q}^{m\times r}$ *and* $B \in \mathbb{Q}^{r\times n}$*. Then*

$$\|AB\|_F \leq \|A\|_F \|B\|_F.$$

A square matrix $A \in \mathbb{Q}^{m\times m}$ is called a unitary matrix if and only if $AA^* = A^*A = I_m$, where $I_m \in \mathbb{R}^{m\times m}$ is the real $m \times m$ identity matrix.

For the complex representation of a quaternion matrix $A \in \mathbb{Q}^{m\times n}$ with the expression (2.1), we follow [18], and denote it as A^C. Let $B_1 = A_0 + A_1\mathbf{i}$ and $B_2 = A_2 + A_3\mathbf{i}$. Then $B_1, B_2 \in \mathbb{C}^{m\times n}$, and $A = B_1 + B_2\mathbf{j}$. The complex representation of A is

$$A^C = \begin{pmatrix} B_1 & B_2 \\ -\bar{B}_2 & \bar{B}_1 \end{pmatrix}.$$

A color image can be expressed as a third order tensor $\mathcal{A} \in \mathbb{R}^{m\times n\times 3}$. On the other hand, we may also represent it by a pure quaternion matrix [9, 12]

$$A = A(:,:,1)\mathbf{i} + A(:,:,2)\mathbf{j} + A(:,:,3)\mathbf{k} \in \mathbb{Q}^{m\times n},$$

where $A(:,:,1), A(:,:,2)$ and $A(:,:,3)$ are the three frontal slices of $\mathcal{A}$.

We have the following theorem on the QSVD of a quaternion matrix [19].

Theorem 2.2. *Any quaternion matrix* $X \in \mathbb{Q}^{m\times n}$ *has the following QSVD form*

$$X = U \begin{pmatrix} \Sigma_r & O \\ O & O \end{pmatrix} V^*, \tag{2.2}$$

where $U \in \mathbb{Q}^{m\times m}$ *and* $V \in \mathbb{Q}^{n\times n}$ *are unitary, and* $\Sigma_r = diag\{\sigma_1, \cdots, \sigma_r\}$ *is a real nonnegative* $r \times r$ *diagonal matrix, with* $\sigma_1 \geq \cdots \geq \sigma_r > 0$ *as the singular values of* X*.*

The quaternion rank of X is the number of its positive singular values, denoted as $\text{rank}(X)$.

corollaryollary 2.3. *Suppose that quaternion matrix* $X \in \mathbb{Q}^{m\times n}$*. Then*

$$\text{rank}(X) \leq \min\{m, n\}. \tag{2.3}$$

For the ranks of quaternion matrices, we have the following theorem. This theorem can be found on Page 295 of [16] and Page 35 of [20]. Since these two references are not in English, we give a proof here for completeness.

Theorem 2.4. *Suppose that* $A \in \mathbb{Q}^{m\times r}$ *and* $B \in \mathbb{Q}^{r\times n}$*. Then*

$$\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}.$$

In particular,

$$\text{rank}(AB) \leq r. \tag{2.4}$$

Proof. (1) We first show that $\text{rank}(AB) \leq \text{rank}(B)$.

For any $x \in \mathbb{Q}^n$ satisfying $Bx = 0$, we have $ABx = 0$. This means that $\mathcal{N}(B) \subseteq \mathcal{N}(AB)$, where $\mathcal{N}(C)$ denotes the null space of matrix C. Hence, $\dim\mathcal{N}(B) \leq \dim\mathcal{N}(AB)$ and $\text{rank}(AB) \leq \text{rank}(B)$ since $\text{rank}(C) = n - \dim\mathcal{N}(C)$.

To show that $\text{rank}(AB) \le \text{rank}(A)$. Let $A_1, A_2, \cdots, A_m$ be all rows of matrix A. Then $A_1B, A_2B, \cdots, A_mB$ are all rows of matrix AB. Let $\text{rank}(AB) = r$. Without loss of generality, we assume that $A_1B, \ldots, A_rB$ are left linearly independent. Now we show that $A_1, \ldots, A_r$ are linearly independent by contradiction.

Assume that there exist $k_1, \ldots, k_r \in \mathbb{Q}$ such that $k_1A_1 + \cdots + k_rA_r = 0$. Then

$$k_1A_1B + \cdots + k_rA_rB = (k_1A_1 + \cdots + k_rA_r)B = 0.$$

From $A_1B, \ldots, A_rB$ are left linearly independent [18], we have $k_1 = \cdots = k_r = 0$ and hence $A_1, \ldots, A_r$ are left linearly independent. This means that $\text{rank}(A) \ge r$ since $\text{rank}(A)$ is the maximum number of rows that are left linearly independent. That is, $\text{rank}(A) \ge \text{rank}(AB)$.

In all, $\text{rank}(AB) \le \min\{\text{rank}(A), \text{rank}(B)\}$.

From the fact that $\text{rank}(A) \le r$, $\text{rank}(AB) \le r$ is clear. □

3 An LRQMA Algorithm

Suppose that we have a quaternion data matrix $D \in \mathbb{Q}^{m\times n}$, which is only partially observed. Let Ω be the set of observed entries of D. The low-rank quaternion matrix approximation (LRQMA) model for color image completion is as follows:

$$\min_{X\in\mathbb{Q}^{m\times n}} \left\{ \frac{1}{2}\|(X-D)_\Omega\|_F^2 : \text{rank}(X) \le r, \mathbf{Re}(X) = O \right\}, \tag{3.1}$$

where $r < \min\{m, n\}$.

We derive a low rank quaternion decomposition theorem for a quaternion matrix X.

Theorem 3.1. *Suppose that $X \in \mathbb{Q}^{m\times n}$. Let positive integer $r < \min\{m, n\}$. Then rank$(X) \le r$ if and only if there are $A \in \mathbb{Q}^{m\times r}$ and $B \in \mathbb{Q}^{r\times n}$ such that $X = AB$.*

Proof. If $X = AB$, $A \in \mathbb{Q}^{m\times r}$ and $B \in \mathbb{Q}^{r\times n}$, then by Theorem 2.4, we have $\text{rank}(X) \le r$.

On the other hand, suppose that $\text{rank}(X) \le r$. Then by Theorem 2.2, we have (2.2). Let

$$U = (U_1\ U_2)$$

and

$$V^* = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$$

where $U_1 \in \mathbb{Q}^{m\times r}$, $U_2 \in \mathbb{Q}^{m\times(m-r)}$, $V_1 \in \mathbb{Q}^{r\times n}$ and $V_2 \in \mathbb{Q}^{(n-r)\times n}$. Let $\Sigma = \text{diag}\{\sqrt{\sigma_1}, \cdots, \sqrt{\sigma_r}\}$, $A = U_1\Sigma$ and $B = \Sigma V_1$. Then we have $X = AB$, $A \in \mathbb{Q}^{m\times r}$ and $B \in \mathbb{Q}^{r\times n}$. □

The LRQMA model (3.1) can be rewritten as

$$\min_{A\in\mathbb{Q}^{m\times r}, B\in\mathbb{Q}^{r\times n}, X\in\mathbb{Q}^{m\times n}} \left\{ f(A,B,X) \equiv \frac{1}{2}\|AB - X\|_F^2 : \mathbf{Re}(X) = O, \mathbf{Im}(X_\Omega) = \mathbf{Im}(D_\Omega) \right\}. \tag{3.2}$$

Here, $r < \min\{m, n\}$, $X \in \mathbb{Q}^{m\times n}$ is a surrogate quaternion matrix. We may use the following alternative scheme to find A, B and X.

At the beginning, we choose $A^{(0)} \in \mathbb{Q}^{m\times r}$ and $B^{(0)} \in \mathbb{Q}^{r\times n}$. Set $k \leftarrow 0$.

At the kth iteration, we first calculate $X^{(k)}$. Let $X^{(k)}$ be the solution of

$$\min_{X\in\mathbb{Q}^{m\times n}} \left\{ p(X) \equiv \frac{1}{2}\|A^{(k)}B^{(k)} - X\|_F^2 : \mathbf{Re}(X) = O, \mathbf{Im}(X_\Omega) = \mathbf{Im}(D_\Omega) \right\}. \tag{3.3}$$

Then, it is straightforward to know that

$$\mathbf{Re}(X^{(k)}) = O,\ \mathbf{Im}(X^{(k)}_{\Omega}) = \mathbf{Im}(D_{\Omega}),\ \mathbf{Im}(X^{(k)}_{\Omega_C}) = \mathbf{Im}\left((A^{(k)}B^{(k)})_{\Omega_C}\right), \tag{3.4}$$

where Ω_C is the complement of Ω.

Second, we find $A^{(k+1)}$ as the solution of

$$\min_{A\in\mathbb{Q}^{m\times r}} g(A) \equiv \frac{1}{2}\left\|AB^{(k)} - X^{(k)}\right\|_F^2 + \frac{\lambda}{2}\left\|A - A^{(k)}\right\|_F^2, \tag{3.5}$$

where $\lambda > 0$ is a regularization parameter. Note that this least squares problem has a closed-form solution

$$A^{(k+1)} = \left[X^{(k)}\left(B^{(k)}\right)^* + \lambda A^{(k)}\right]\left[B^{(k)}\left(B^{(k)}\right)^* + \lambda I_r\right]^{-1}. \tag{3.6}$$

We can see this by the following argument. Let

$$\begin{aligned}
A =&\ A_0 + A_1\mathbf{i} + A_2\mathbf{j} + A_3\mathbf{k},\\
A^{(k)} =&\ A_0^{(k)} + A_1^{(k)}\mathbf{i} + A_2^{(k)}\mathbf{j} + A_3^{(k)}\mathbf{k},\\
B^{(k)} =&\ B_0^{(k)} + B_1^{(k)}\mathbf{i} + B_2^{(k)}\mathbf{j} + B_3^{(k)}\mathbf{k},\\
X^{(k)} =&\ X_0^{(k)} + X_1^{(k)}\mathbf{i} + X_2^{(k)}\mathbf{j} + X_3^{(k)}\mathbf{k},
\end{aligned}$$

where $A_i, A_i^{(k)}, B_i^{(k)}$ and $X_i^{(k)}$ for $k = 0, 1, 2, 3$, are corresponding real matrices. Then

$$\begin{aligned}
g(A) =&\ \tfrac{1}{2}\left\|A_0B_0^{(k)} - A_1B_1^{(k)} - A_2B_2^{(k)} - A_3B_3^{(k)} - X_0^{(k)}\right\|_F^2\\
+&\ \tfrac{1}{2}\left\|A_0B_1^{(k)} + A_1B_0^{(k)} + A_2B_3^{(k)} - A_3B_2^{(k)} - X_1^{(k)}\right\|_F^2\\
+&\ \tfrac{1}{2}\left\|A_0B_2^{(k)} + A_2B_0^{(k)} + A_3B_1^{(k)} - A_1B_3^{(k)} - X_2^{(k)}\right\|_F^2\\
+&\ \tfrac{1}{2}\left\|A_0B_3^{(k)} + A_3B_0^{(k)} + A_1B_2^{(k)} - A_2B_1^{(k)} - X_3^{(k)}\right\|_F^2\\
+&\ \tfrac{\lambda}{2}\sum_{i=0}^{3}\left\|A_i - A_i^{(k)}\right\|_F^2.
\end{aligned}$$

By derivation, we see that the solution of (3.5) is (3.6).

Finally, we find $B^{(k+1)}$ as the solution of

$$\min_{B\in\mathbb{Q}^{r\times n}} h(B) \equiv \frac{1}{2}\left\|A^{(k+1)}B - X^{(k)}\right\|_F^2 + \frac{\lambda}{2}\left\|B - B^{(k)}\right\|_F^2. \tag{3.7}$$

Similarly, this least squares problem has a closed-form solution

$$B^{(k+1)} = \left[\left(A^{(k+1)}\right)^* A^{(k+1)} + \lambda I_r\right]^{-1}\left[\left(A^{(k+1)}\right)^* X^{(k)} + \lambda B^{(k)}\right]. \tag{3.8}$$

We thus have the following algorithm to solve (3.2).

4 The Gradient of A Real Function in Quaternion Matrix Variables

To conduct convergence analysis for our algorithm, we need to consider optimization problems of real functions in quaternion matrix variables. To handle this, we introduce a concise form for the gradient of a real function in quaternion matrix variables.

Algorithm 1 An LRQMA algorithm for solving (3.2).

1: Choose $A^{(0)} \in \mathbb{Q}^{m\times r}$ and $B^{(0)} \in \mathbb{Q}^{r\times n}$. Set $k \leftarrow 0$.
2: Use (3.4) to find $X^{(k)}$.
3: Use (3.6) to find $A^{(k+1)}$.
4: Use (3.8) to find $B^{(k+1)}$.
5: If $A^{(k+1)} = A^{(k)}$ and $B^{(k+1)} = B^{(k)}$, stop. Otherwise, $k \leftarrow k+1$ and goto Step 2.

Consider the following optimization problem

$$\min\{f(X) : X \in \mathbb{Q}^{m\times n}, g_j(X) = 0, j = 1, \dots, p\}, \tag{4.1}$$

where $f, g_j : \mathbb{Q}^{m\times n} \to \mathbb{R}$ for $j = 1, \dots, p$. We need to have a form for $\nabla f(X)$ and $\nabla g_j(X)$ for $j = 1, \dots, p$.

Definition 4.1. Let $f : \mathbb{Q}^{m\times n} \to \mathbb{R}$. Let $X = X_0+X_1\mathbf{i}+X_2\mathbf{j}+X_3\mathbf{k}$, where $X_0, X_1, X_2, X_3 \in \mathbb{R}^{m\times n}$. Then we say that f is differentiable at X if $\frac{\partial f}{\partial X_i}$ exists at X_i for $i = 0, 1, 2, 3$, and we denote

$$\nabla f(X) = \frac{\partial f}{\partial X_0} + \frac{\partial f}{\partial X_1}\mathbf{i} + \frac{\partial f}{\partial X_2}\mathbf{j} + \frac{\partial f}{\partial X_3}\mathbf{k}. \tag{4.2}$$

If $\frac{\partial f}{\partial X_i}$ exists in a neighborhood of X_i, and is continuous at X_i, for $i = 0, 1, 2, 3$, then we say that f is continuously differentiable at X. If f is continuously differentiable for any $X \in \mathbb{Q}^{m\times n}$, then we say that f is continuously differentiable.

If f has more variables, then we may change $\nabla f(X)$ in (4.2) to $\frac{\partial f}{\partial X}$.

This form is different from the generalized HR calculus studied in [17], which is similar to the approach in optimization of real functions with complex variables [13].

Based from this definition, we have the following theorem.

Theorem 4.2. *Suppose that $f, g_j : \mathbb{Q}^{m\times n} \to \mathbb{R}$, for $j = 1, \dots, p$, are continuously differentiable, and $X^{\#} \in \mathbb{Q}^{m\times n}$ is an optimal solution of (4.1). Then there are Lagrangian multipliers $\pi_j \in \mathbb{Q}$ for $j = 1, \dots, p$, such that*

$$\nabla f(X^{\#}) + \sum_{j=1}^{p} \pi_j \nabla g_j(X^{\#}) = O.$$

Proof. Let $X = X_0 + X_1\mathbf{i} + X_2\mathbf{j} + X_3\mathbf{k}$, where $X_0, X_1, X_2, X_3 \in \mathbb{R}^{m\times n}$. Then (4.1) is converted to an optimization problem with real matrix variables. By optimization theory [14] and Definition 4.1, the conclusion holds. □

This theorem can be extended to other optimization problems involving continuously differentiable real functions in quaternion matrix variables. The Lagrangian multipliers π_j are quaternion numbers. We will see this more clearly in the next section.

With Definition 4.1, the gradients of some functions, which are useful for our model, have simple expressions.

Theorem 4.3. *Suppose that $f : \mathbb{Q}^{m\times r} \to \mathbb{R}$ is defined by $f(X) = \frac{1}{2}\|XB + C\|_F^2$, where $B \in \mathbb{Q}^{r\times n}$ and $C \in \mathbb{Q}^{m\times n}$. Then*

$$\nabla f(X) = (XB + C)B^*.$$

Proof. We have $XB + C = M_0 + M_1\mathbf{i} + M_2\mathbf{j} + M_3\mathbf{k}$, where

$$M_0 = X_0B_0 - X_1B_1 - X_2B_2 - X_3B_3 + C_0,$$

$$M_1 = X_0B_1 + X_1B_0 + X_2B_3 - X_3B_2 + C_1,$$

$$M_2 = X_0B_2 + X_2B_0 + X_1B_3 - X_3B_1 + C_2,$$

$$M_3 = X_0B_3 + X_3B_0 + X_1B_2 - X_2B_1 + C_3.$$

Then,

$$f(X) = \frac{1}{2}\sum_{i=0}^{3}\|M_i\|_F^2,$$

$$\frac{\partial f}{\partial X_0} = M_0B_0^\top + M_1B_1^\top + M_2B_2^\top + M_3B_3^\top,$$

$$\frac{\partial f}{\partial X_1} = -M_0B_1^\top + M_1B_0^\top + M_2B_3^\top + M_3B_2^\top,$$

$$\frac{\partial f}{\partial X_2} = -M_0B_2^\top + M_1B_3^\top + M_2B_0^\top - M_3B_1^\top,$$

$$\frac{\partial f}{\partial X_3} = -M_0B_3^\top - M_1B_2^\top - M_2B_1^\top + M_3B_0^\top.$$

Thus, we have

$$\nabla f(X) = \frac{\partial f}{\partial X_0} + \frac{\partial f}{\partial X_1}\mathbf{i} + \frac{\partial f}{\partial X_2}\mathbf{j} + \frac{\partial f}{\partial X_3}\mathbf{k} = (XB+C)(B_0 - B_1^\top\mathbf{i} - B_2^\top\mathbf{j} - B_3^\top\mathbf{k}) = (XB+C)B^*.$$

□

Similarly, we have the following theorem.

Theorem 4.4. *Suppose that* $f : \mathbb{Q}^{r\times n} \to \mathbb{R}$ *is defined by* $f(X) = \frac{1}{2}\|AX + C\|_F^2$, *where* $A \in \mathbb{Q}^{m\times r}$ *and* $C \in \mathbb{Q}^{m\times n}$. *Then*

$$\nabla f(X) = A^*(AX + C).$$

In our convergence analysis, the Kurdyka-Łojasiewicz property [1, 2, 3] plays a critical role. Since we regard functions in this paper as functions defined on an abstract vector space with real coefficients, the Kurdyka-Łojasiewicz property also holds for functions related with the optimization problems studied in this paper.

5 Convergence Analysis

In this section, we present convergence analysis for the LRQMA algorithm. As stated in Section 3, we may regard the objective function f in (3.2) as a function defined in the abstract vector space of dimension $4(mr + rn + mn)$ with real coefficients. Then we use the gradient of f to study the stationary points and the first order optimality conditions of (3.2) without ambiguity.

In the following, we always denote $\mathbf{Z} \equiv (A, B, X)$. Thus, $\mathbf{Z}^{(k)} \equiv (A^{(k)}, B^{(k)}, X^{(k)})$, $\mathbf{Z}^{\#} \equiv (A^{\#}, B^{\#}, X^{\#})$, so on. We have the following theorem.

Theorem 5.1. *Let $A^\# \in \mathbb{Q}^{m\times r}, B^\# \in \mathbb{Q}^{r\times n}$ and $X^\# \in \mathbb{Q}^{m\times n}$. Suppose that $X^\#$ satisfies*

$$\mathbf{Re}(X^\#) = O, \quad \mathbf{Im}(X^\#_\Omega) = \mathbf{Im}(D_\Omega), \tag{5.1}$$

i.e., $X^\#$ is a feasible point of (3.2). If $\mathbf{Z}^\# \equiv \left(A^\#, B^\#, X^\#\right)$ is an optimal solution of (3.2), then we have

$$\left(A^\# B^\# - X^\#\right)\left(B^\#\right)^* = O_{m\times r}, \tag{5.2}$$

$$\left(A^\#\right)^*\left(A^\# B^\# - X^\#\right) = O_{r\times n}, \tag{5.3}$$

$$\mathbf{Im}\left(X^\#_{\Omega_C}\right) = \mathbf{Im}\left((A^\# B^\#)_{\Omega_C}\right), \tag{5.4}$$

i.e., $\mathbf{Z}^\# \equiv \left(A^\#, B^\#, X^\#\right)$ is a stationary point of (3.2).

Proof. By Theorem 4.2, $\mathbf{Z}^\# \equiv \left(A^\#, B^\#, X^\#\right)$ should satisfy

$$\begin{aligned} \frac{\partial}{\partial A} f(A^\#, B^\#, X^\#) &= O_{m\times r}, \\ \frac{\partial}{\partial B} f(A^\#, B^\#, X^\#) &= O_{r\times n}, \\ \frac{\partial}{\partial X} f(A^\#, B^\#, X^\#) &= -\Pi, \end{aligned}$$

where $\Pi \in \mathbb{Q}^{m\times n}$ with $\mathbf{Im}(\Pi_{\Omega_C}) = 0$ is a matrix of Lagrangian multipliers.

By Theorem 4.3,

$$\frac{\partial}{\partial A} f(A^\#, B^\#, X^\#) = \left(A^\# B^\# - X^\#\right)\left(B^\#\right)^*.$$

We have (5.2).

By Theorem 4.4,

$$\frac{\partial}{\partial B} f(A^\#, B^\#, X^\#) = \left(A^\#\right)^*\left(A^\# B^\# - X^\#\right).$$

We have (5.3).

By Theorem 4.3 or 4.4,

$$\frac{\partial}{\partial X} f(A^\#, B^\#, X^\#) = X^\# - A^\# B^\# = \Pi.$$

We have (5.4) and

$$\mathbf{Re}(\Pi) = \mathbf{Re}(A^\# B^\# - X^\#), \quad \mathbf{Im}(\Pi_\Omega) = \mathbf{Im}\left((A^\# B^\# - X^\#)_\Omega\right).$$

This shows that the Lagrangian multipliers are quaternion numbers.

The theorem is proved. □

We now consider the case that the LRQMA algorithm stops in a finite number of iterations.

Proposition 5.2. *If LRQMA algorithm stops in a finite number of iterations, then $\mathbf{Z}^\# \equiv \left(A^\#, B^\#, X^\#\right) = \left(A^{(k)}, B^{(k)}, X^{(k)}\right)$ is a stationary point of (3.2).*

Proof. If LRQMA algorithm stops in a finite number of iterations, then $A^{(k+1)} = A^{(k)}$, $B^{(k+1)} = B^{(k)}$ and $X^{(k+1)} = X^{(k)}$. Denote $A^{\#} = A^{(k)}$, $B^{\#} = B^{(k)}$ and $X^{\#} = X^{(k)}$. By (3.4), we have (5.1) and (5.4). By (3.6), we have

$$A^{\#}\left(B^{\#}(B^{\#})^* + \lambda I_r\right) = X^{\#}(B^{\#})^* + \lambda A^{\#}.$$

This leads to (5.2). By (3.8), we have

$$\left((A^{\#})^* A^{\#} + \lambda I_r\right) B^{\#} = (A^{\#})^* X^{\#} + \lambda B^{\#}.$$

This leads to (5.3). These imply that $\mathbf{Z}^{\#} \equiv \left(A^{\#}, B^{\#}, X^{\#}\right) \equiv \left(A^{(k)}, B^{(k)}, X^{(k)}\right)$ is a stationary point of (3.2). □

We then consider the case that the LRQMA algorithm generates an infinite sequence $\left\{\mathbf{Z}^{(k)} \equiv \left(A^{(k)}, B^{(k)}, X^{(k)}\right) : k = 0, 1, 2, \ldots\right\}$.

Lemma 5.3. *Suppose that* $\left\{\mathbf{Z}^{(k)} \equiv \left(A^{(k)}, B^{(k)}, X^{(k)}\right) : k = 0, 1, 2, \ldots\right\}$ *is the sequence generated by the LRQMA algorithm. Then*

$$\begin{aligned} & f\left(A^{(k)}, B^{(k)}, X^{(k))}\right) - f\left(A^{(k+1)}, B^{(k+1)}, X^{(k+1)}\right) \\ \geq\ & \frac{\lambda}{2}\left[\|A^{(k)} - A^{(k+1)}\|_F^2 + \|B^{(k)} - B^{(k+1)}\|_F^2 + \|X^{(k)} - X^{(k+1)}\|_F^2\right]. \end{aligned}$$

Proof. First, we have

$$\begin{aligned} & f\left(A^{(k)}, B^{(k)}, X^{(k))}\right) - f\left(A^{(k+1)}, B^{(k)}, X^{(k)}\right) \\ =\ & \frac{1}{2}\left\|A^{(k)}B^{(k)} - X^{(k)}\right\|_F^2 - \frac{1}{2}\left\|A^{(k+1)}B^{(k)} - X^{(k)}\right\|_F^2 \\ =\ & g(A^{(k)}) - g(A^{(k+1)}) + \frac{\lambda}{2}\left\|A^{(k+1)} - A^{(k)}\right\|_F^2 \\ \geq\ & \frac{\lambda}{2}\left\|A^{(k+1)} - A^{(k)}\right\|_F^2. \end{aligned}$$

Here, g is the objective function of (3.5). Since $A^{(k+1)}$ minimizes (3.5), we have $g(A^{(k)}) \geq g(A^{(k+1)})$. This leads to the last inequality.

Similarly, we have

$$\begin{aligned} & f\left(A^{(k+1)}, B^{(k)}, X^{(k))}\right) - f\left(A^{(k+1)}, B^{(k+1)}, X^{(k)}\right) \\ =\ & \frac{1}{2}\left\|A^{(k+1)}B^{(k)} - X^{(k)}\right\|_F^2 - \frac{1}{2}\left\|A^{(k+1)}B^{(k+1)} - X^{(k)}\right\|_F^2 \\ =\ & h(B^{(k)}) - h(B^{(k+1)}) + \frac{\lambda}{2}\left\|B^{(k+1)} - B^{(k)}\right\|_F^2 \\ \geq\ & \frac{\lambda}{2}\left\|B^{(k+1)} - B^{(k)}\right\|_F^2. \end{aligned}$$

Here, h is the objective function of (3.7). Since $B^{(k+1)}$ minimizes (3.7), we have $h(B^{(k)}) \geq h(B^{(k+1)})$. This leads to the last inequality.

Finally, we have

$$\begin{aligned}
& f\left(A^{(k+1)}, B^{(k+1)}, X^{(k))}\right) - f\left(A^{(k+1)}, B^{(k+1)}, X^{(k+1)}\right) \\
= \; & \frac{1}{2}\left\|A^{(k+1)}B^{(k+1)} - X^{(k)}\right\|_F^2 - \frac{1}{2}\left\|A^{(k+1)}B^{(k+1)} - X^{(k+1)}\right\|_F^2 \\
= \; & \frac{1}{2}\left\|\mathbf{Re}(A^{(k+1)}B^{(k+1)} - X^{(k)})\right\|_F^2 - \frac{1}{2}\left\|\mathbf{Re}(A^{(k+1)}B^{(k+1)} - X^{(k+1)})\right\|_F^2 \\
& + \frac{1}{2}\left\|\mathbf{Im}((A^{(k+1)}B^{(k+1)} - X^{(k)})_\Omega)\right\|_F^2 - \frac{1}{2}\left\|\mathbf{Im}((A^{(k+1)}B^{(k+1)} - X^{(k+1)})_\Omega)\right\|_F^2 \\
& + \frac{1}{2}\left\|\mathbf{Im}((A^{(k+1)}B^{(k+1)} - X^{(k)})_{\Omega_C})\right\|_F^2 - \frac{1}{2}\left\|\mathbf{Im}((A^{(k+1)}B^{(k+1)} - X^{(k+1)})_{\Omega_C})\right\|_F^2 \\
= \; & \frac{1}{2}\left\|\mathbf{Im}((A^{(k+1)}B^{(k+1)} - X^{(k)})_{\Omega_C})\right\|_F^2 \\
= \; & \frac{1}{2}\left\|\mathbf{Im}((X^{(k+1)} - X^{(k)})_{\Omega_C})\right\|_F^2 \\
= \; & \frac{1}{2}\|X^{(k)} - X^{(k+1)}\|_F^2.
\end{aligned}$$

Summing up these three parts, we have the conclusion. □

Lemma 5.4. *Suppose that* $\left\{\mathbf{Z}^{(k)} \equiv \left(A^{(k)}, B^{(k)}, X^{(k)}\right) : k = 0, 1, 2, \dots\right\}$ *is the sequence generated by the LRQMA algorithm. Then*

$$\sum_{k=0}^{\infty}\left[\left\|A^{(k)} - A^{(k+1)}\right\|_F^2 + \left\|B^{(k)} - B^{(k+1)}\right\|_F^2 + \left\|X^{(k)} - X^{(k+1)}\right\|_F^2\right] < \infty,$$

$$\lim_{k\to\infty}\left[A^{(k)} - A^{(k+1)}\right] = O_{m\times r},$$

$$\lim_{k\to\infty}\left[B^{(k)} - B^{(k+1)}\right] = O_{r\times n}$$

and

$$\lim_{k\to\infty}\left[X^{(k)} - X^{(k+1)}\right] = O_{m\times n}.$$

Proof. By Lemma 5.3, for $k = 0, 1, 2, \dots$, we have

$$\begin{aligned}
& \left\|A^{(k)} - A^{(k+1)}\right\|_F^2 + \left\|B^{(k)} - B^{(k+1)}\right\|_F^2 + \left\|X^{(k)} - X^{(k+1)}\right\|_F^2 \\
\leq \; & \frac{2}{\lambda}\left[f\left(A^{(k)}, B^{(k)}, X^{(k))}\right) - f\left(A^{(k+1)}, B^{(k+1)}, X^{(k+1)}\right)\right].
\end{aligned}$$

Summarizing with respect to k, we have

$$\begin{aligned}
& \sum_{k=0}^{\infty}\left[\left\|A^{(k)} - A^{(k+1)}\right\|_F^2 + \left\|B^{(k)} - B^{(k+1)}\right\|_F^2 + \left\|X^{(k)} - X^{(k+1)}\right\|_F^2\right] \\
\leq \; & \sum_{k=0}^{\infty}\frac{2}{\lambda}\left[f\left(A^{(k)}, B^{(k)}, X^{(k))}\right) - f\left(A^{(k+1)}, B^{(k+1)}, X^{(k+1)}\right)\right] \\
\leq \; & \frac{2}{\lambda}f\left(A^{(0)}, B^{(0)}, X^{(0))}\right) < \infty.
\end{aligned}$$

Hence, $\left\|A^{(k)} - A^{(k+1)}\right\|_F^2 \to 0$, $\left\|B^{(k)} - B^{(k+1)}\right\|_F^2 \to 0$ and $\left\|X^{(k)} - X^{(k+1)}\right\|_F^2 \to 0$. The second conclusion of the lemma follows. □

Theorem 5.5. *Suppose that* $\left\{\mathbf{Z}^{(k)} \equiv \left(A^{(k)}, B^{(k)}, X^{(k)}\right) : k = 0, 1, 2, \dots\right\}$ *is the sequence generated by the LRQMA algorithm and it is bounded. Then every limiting point of this sequence is a stationary point of (3.2).*

Proof. Since $\left\{\left(A^{(k)}, B^{(k)}, X^{(k)}\right) : k = 0, 1, 2, \dots\right\}$ is bounded, it must have a subsequence $\left\{\left(A^{(k_i)}, B^{(k_i)}, X^{(k_i)}\right) : i = 0, 1, 2, \dots\right\}$ that converges to a limiting point $\left(A^{\#}, B^{\#}, X^{\#}\right)$. By Lemma 5.4, the subsequence $\left\{\left(A^{(k_i+1)}, B^{(k_i+1)}, X^{(k_i+1)}\right) : i = 0, 1, 2, \dots\right\}$ converges to the same limiting point.

In (3.4), (3.6) and (3.8), replace k by k_i and let $i \to \infty$. Then, with an argument similar to the proof of Proposition 5.2, we conclude that $\left(A^{\#}, B^{\#}, X^{\#}\right)$ is a stationary point of (3.2). □

Consider optimization problem (3.2). As we regard the objective function f as a function defined on an abstract vector space with real coefficients, it is a semi-algebraic function in the sense of [1, 2, 3]. Then, for any critical point $\left(A^{\#}, B^{\#}, X^{\#}\right)$ of f, there are a neighborhood N of this critical point, an exponent $\theta \in [\frac{1}{2}, 1)$ and a positive constant μ such that the Kurdyka-Łojasiewicz inequality [1, 2, 3] below holds, i.e., for any $(A, B, X) \in N$, we have

$$\left|f(A, B, X) - f\left(A^{\#}, B^{\#}, X^{\#}\right)\right|^{\theta} \le \mu \left\|\Pi_\Sigma(\nabla f(A, B, X))\right\|_F, \tag{5.5}$$

where Σ is the feasible set of (3.2), $\Pi_\Sigma(\nabla f)$ is the projected gradient of f with respect to Σ. Then, we have

$$\Pi_\Sigma(\nabla f(A, B, X)) = \begin{pmatrix} (AB - X)B^* \\ A^*(AB - X) \\ \mathbf{Im}((X - AB)_{\Omega_C}) \end{pmatrix}. \tag{5.6}$$

We may further confirm (5.5) by the following argument in the real field. Let

$$\begin{aligned} A &= A_0 + A_1\mathbf{i} + A_2\mathbf{j} + A_3\mathbf{k}, \\ B &= B_0 + B_1\mathbf{i} + B_2\mathbf{j} + B_3\mathbf{k}, \\ X &= X_0 + X_1\mathbf{i} + X_2\mathbf{j} + X_3\mathbf{k}, \\ D &= D_0 + D_1\mathbf{i} + D_2\mathbf{j} + D_3\mathbf{k}, \\ \mathbf{W} &= (A_0, A_1, A_2, A_3, B_0, B_1, B_2, B_3, X_0, X_1, X_2, X_3), \end{aligned}$$

where $A_i \in \mathbb{R}^{m\times r}, B_i \in \mathbb{R}^{r\times n}$, $X_i \in \mathbb{R}^{m\times n}$ and $D_i \in \mathbb{R}^{m\times n}$ for $i = 0, 1, 2, 3$. Define $\phi(\mathbf{W}) \equiv f(A, B, X)$. Then

$$\begin{aligned} \phi(\mathbf{W}) =& \frac{1}{2}\, \|(A_0B_0 - A_1B_1 - A_2B_2 - A_3B_3 - X_0) + (A_0B_1 + A_1B_0 + A_2B_3 - A_3B_2 - X_1)\mathbf{i} \\ &+ (A_0B_2 + A_2B_0 + A_3B_1 - A_1B_3 - X_2)\mathbf{j} + (A_0B_3 + A_3B_0 + A_1B_2 - A_2B_1 - X_3)\mathbf{k}\|_F^2 \\ =& \frac{1}{2}\left[\|A_0B_0 - A_1B_1 - A_2B_2 - A_3B_3 - X_0\|_F^2 + \|A_0B_1 + A_1B_0 + A_2B_3 - A_3B_2 - X_1\|_F^2\right. \\ &\left. + \|A_0B_2 + A_2B_0 + A_3B_1 - A_1B_3 - X_2\|_F^2 + \|A_0B_3 + A_3B_0 + A_1B_2 - A_2B_1 - X_3\|_F^2\right]. \end{aligned}$$

This shows that ϕ is a semi-algebraic function. The constraints of (3.2) in the real field are $X_0 = O$ and $(X_i)_\Omega = (D_i)_\Omega$ for $i = 1, 2, 3$. Applying the Kurdyka-Łojasiewicz inequality in the real field [1, 2, 3], we may still obtain (5.5).

In the following, we use (5.5) to show that the infinite sequence generated by the LRQMA algorithm converges to a stationary point. We first to prove two lemmas.

Lemma 5.6. *Suppose that* $\left\{\mathbf{Z}^{(k)} \equiv \left(A^{(k)}, B^{(k)}, X^{(k)}\right) : k = 0, 1, 2, \ldots\right\}$ *is the sequence generated by the LRQMA algorithm and it is bounded. Then there is a positive constant* $\eta \in (0, 1]$ *such that for* $k = 0, 1, 2, 3, \ldots,$

$$\left\|\left(A^{(k)}, B^{(k)}, X^{(k)})\right) - \left(A^{(k+1)}, B^{(k+1)}, X^{(k+1)}\right)\right\|_F \geq \eta \left\|\Pi_\Sigma(\nabla f(A^{(k)}, B^{(k)}, X^{(k)}))\right\|_F.$$

Proof. By (3.6) and Proposition 2.1, we have

$$\begin{aligned}
&\left\|\left(A^{(k)}B^{(k)} - X^{(k)}\right)\left(B^{(k)}\right)^*\right\|_F^2 \\
&= \left\|\left(A^{(k)} - A^{(k+1)}\right)\left[B^{(k)}\left(B^{(k)}\right)^* + \lambda I_r\right]\right\|_F^2 \\
&\leq \left\|A^{(k)} - A^{(k+1)}\right\|_F^2 \left\|B^{(k)}\left(B^{(k)}\right)^* + \lambda I_r\right\|_F^2 \\
&\leq \frac{1}{2\eta^2}\left\|A^{(k+1)} - A^{(k)}\right\|_F^2.
\end{aligned}$$

Here $\eta \in (0, 1]$ is a positive constant. Such a positive constant exists, as $\left\{\left(A^{(k)}, B^{(k)}, X^{(k)}\right)\right\}$ is bounded. Similarly, by (3.8) and Proposition 2.1, we have

$$\left\|\left(A^{(k)}\right)^*\left(A^{(k)}B^{(k)} - X^{(k)}\right)\right\|_F^2 \leq \frac{1}{2\eta^2}\left\|A^{(k+1)} - A^{(k)}\right\|_F^2 + \frac{1}{\eta^2}\left\|B^{(k+1)} - B^{(k)}\right\|_F^2.$$

We also have

$$\left\|\mathbf{Im}\left((X^{(k)} - A^{(k)}B^{(k)})_{\Omega_C}\right)\right\|_F^2 \leq \left\|X^{(k+1)} - X^{(k)}\right\|_F^2.$$

By (5.6) and the above three inequalities, we have

$$\begin{aligned}
&\left\|\Pi_\Sigma\left(\nabla f(A^{(k)}, B^{(k)}, X^{(k)})\right)\right\|_F^2 \\
&= \left\|\left(A^{(k)}B^{(k)} - X^{(k)}\right)\left(B^{(k)}\right)^*\right\|_F^2 + \left\|\left(A^{(k)}\right)^*\left(A^{(k)}B^{(k)} - X^{(k)}\right)\right\|_F^2 \\
&\quad + \left\|\mathbf{Im}\left((X^{(k)} - A^{(k)}B^{(k)})_{\Omega_C}\right)\right\|_F^2 \\
&\leq \frac{1}{\eta^2}\left\|A^{(k+1)} - A^{(k)}\right\|_F^2 + \frac{1}{\eta^2}\left\|B^{(k+1)} - B^{(k)}\right\|_F^2 + \left\|X^{(k+1)} - X^{(k)}\right\|_F^2 \\
&\leq \frac{1}{\eta^2}\left\|\left(A^{(k)}, B^{(k)}, X^{(k)})\right) - \left(A^{(k+1)}, B^{(k+1)}, X^{(k+1)}\right)\right\|_F^2.
\end{aligned}$$

Hence, we have the conclusion. ☐

Lemma 5.7. *Let* $\mathbf{Z}^\#$ *be one limiting point of* $\{\mathbf{Z}^{(k)}\}$*. Assume that* $\mathbf{Z}^{(0)}$ *satisfies* $\mathbf{Z}^{(0)} \in N$ *and* $\left\|\mathbf{Z}^{(0)} - \mathbf{Z}^\#\right\| \leq \rho$*, where*

$$\rho > \frac{2\mu}{\lambda\eta(1-\theta)}\left|f(\mathbf{Z}^{(0)}) - f(\mathbf{Z}^\#)\right|^{1-\theta} + \left\|\mathbf{Z}^{(0)} - \mathbf{Z}^\#\right\|. \tag{5.7}$$

Then we have the following conclusions:

$$\left\|\mathbf{Z}^{(k)} - \mathbf{Z}^\#\right\| \leq \rho, \quad \text{for } k = 0, 1, 2, \ldots \tag{5.8}$$

and

$$\sum_{k=0}^{\infty}\left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{(k+1)}\right\| \leq \frac{2\mu}{\lambda\eta(1-\theta)}\left|f(\mathbf{Z}^{(0)} - f(\mathbf{Z}^\#)\right|^{1-\theta}. \tag{5.9}$$

Proof. We prove (5.8) by induction. By assumption, (5.8) holds for $k = 0$. We now assume that there is an integer $\hat{k}$ such that (5.8) holds for $0 \le k \le \hat{k}$. This means that the Kurdyka-Łojasiewicz inequality holds at these points. We now prove that (5.8) holds for $k = \hat{k} + 1$. Define a scalar function

$$\psi(\alpha) := \frac{1}{1-\theta}\left|\alpha - f(\mathbf{Z}^{\#})\right|^{1-\theta}.$$

Then ψ is a concave function and $\psi'(\alpha) = |\alpha - f(\mathbf{Z}^{\#})|^{-\theta}$ if $\alpha \ge f(\mathbf{Z}^{\#})$. For $0 \le k \le \hat{k}$, we have

$$\begin{aligned}
&\psi\left(f\left(\mathbf{Z}^{(k)}\right)\right) - \psi\left(f\left(\mathbf{Z}^{(k+1)}\right)\right) \\
\ge\ & \psi'\left(f\left(\mathbf{Z}^{(k)}\right)\right)\left[f\left(\mathbf{Z}^{(k)}\right) - f\left(\mathbf{Z}^{(k+1)}\right)\right] \\
\ge\ & \frac{1}{\left|f\left(\mathbf{Z}^{(k)}\right) - f\left(\mathbf{Z}^{(\#)}\right)\right|^{\theta}}\frac{\lambda}{2}\left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{(k+1)}\right\|_F^2 \quad \text{[by Lemma 5.3]} \\
\ge\ & \frac{\lambda}{2\mu}\frac{\left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{(k+1)}\right\|_F^2}{\left\|\Pi_\Sigma\left(\nabla f\left(\mathbf{Z}^{(k)}\right)\right)\right\|} \quad \text{[by (5.5)]} \\
\ge\ & \frac{\lambda\eta}{2\mu}\frac{\left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{(k+1)}\right\|_F^2}{\left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{(k+1)}\right\|_F} \quad \text{[by Lemma 5.6]} \\
=\ & \frac{\lambda\eta}{2\mu}\left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{(k+1)}\right\|_F.
\end{aligned}$$

Summarizing k from 0 to $\hat{k}$, we have

$$\begin{aligned}
\sum_{k=0}^{\hat{k}}\left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{(k+1)}\right\|_F &\le \frac{2\mu}{\lambda\eta}\sum_{k=0}^{\hat{k}}\left[\psi\left(f\left(\mathbf{Z}^{(k)}\right)\right) - \psi\left(f\left(\mathbf{Z}^{(k+1)}\right)\right)\right] \\
&= \frac{2\mu}{\lambda\eta}\left[\psi\left(f\left(\mathbf{Z}^{(0)}\right)\right) - \psi\left(f\left(\mathbf{Z}^{(\hat{k}+1)}\right)\right)\right] \\
&\le \frac{2\mu}{\lambda\eta}\psi\left(f\left(\mathbf{Z}^{(0)}\right)\right). \qquad (5.10)
\end{aligned}$$

By this and (5.7), we have

$$\begin{aligned}
\left\|\mathbf{Z}^{(\hat{k}+1)} - \mathbf{Z}^{(\#)}\right\|_F &\le \sum_{k=0}^{\hat{k}}\left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{(k+1)}\right\|_F + \left\|\mathbf{Z}^{(0)} - \mathbf{Z}^{\#}\right\| \\
&\le \frac{2\mu}{\lambda\eta}\psi\left(f\left(\mathbf{Z}^{(0)}\right)\right) + \left\|\mathbf{Z}^{(0)} - \mathbf{Z}^{\#}\right\| \\
&< \rho.
\end{aligned}$$

This proves (5.8).

Letting $\hat{k} \to \infty$ in (5.10) and using (5.7), we have (5.9). □

Theorem 5.8. *Suppose that the LRQMA algorithm generates a bounded sequence* $\{\mathbf{Z}^{(k)}\}$. *Then*

$$\sum_{k=0}^{\infty}\left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{(k+1)}\right\| < \infty,$$

which means that the entire sequence $\{\mathbf{Z}^{(k)}\}$ *converges to a limit.*

Proof. Because $\{\mathbf{Z}^{(k)}\}$ is bounded, it must have a limiting point $\mathbf{Z}^{\#}$. Then there is an index k_0 such that

$$\left\|\mathbf{Z}^{k_0} - \mathbf{Z}^{\#}\right\| \leq \rho.$$

We may regard $\mathbf{Z}^{k_0}$ as the initial point. Then Lemma 5.7 holds. The entire sequence satisfies (5.9). The theorem is proved. □

Finally, we establish convergence rates for the convergence of this sequence.

Theorem 5.9. *Suppose that the LRQMA algorithm generates a bounded sequence* $\{\mathbf{Z}^{(k)}\}$. *Then,*

(1) *if* $\theta = \frac{1}{2}$, *then there exist* $\sigma \in [0,1)$ *and* $\beta > 0$ *such that*

$$\left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{\#}\right\| \leq \beta\sigma^k,$$

i.e., the sequence converges R-linearly;

(2) *if* $\frac{1}{2} < \theta < 1$, *then there exists* $\beta > 0$ *such that*

$$\left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{\#}\right\| \leq \beta k^{-\frac{1-\theta}{2\theta-1}}.$$

Proof. Without loss of generality, assume that $\left\|\mathbf{Z}^{(0)} - \mathbf{Z}^{\#}\right\| < \rho$. Let

$$\Delta_k := \sum_{i=k}^{\infty} \left\|\mathbf{Z}^{(i)} - \mathbf{Z}^{(i+1)}\right\| \geq \left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{\#}\right\|. \tag{5.11}$$

By Lemma 5.7 we have

$$\begin{aligned}
\Delta_k &\leq \frac{2\mu}{\lambda\eta(1-\theta)} \left|f(\mathbf{Z}^{(k)} - f(\mathbf{Z}^{\#})\right|^{1-\theta} \\
&= \frac{2\mu}{\lambda\eta(1-\theta)} \left[\left|f(\mathbf{Z}^{(k)} - f(\mathbf{Z}^{\#})\right|^{\theta}\right]^{\frac{1-\theta}{\theta}} \\
&\leq \frac{2\mu}{\lambda\eta(1-\theta)} \mu^{\frac{1-\theta}{\theta}} \left\|\Pi_\Sigma\left(\nabla f\left(\mathbf{Z}^{(k)}\right)\right)\right\|^{\frac{1-\theta}{\theta}} \quad \text{[by (5.5)]} \\
&\leq \frac{2\mu}{\lambda\eta(1-\theta)} \left(\frac{\mu}{\eta}\right)^{\frac{1-\theta}{\theta}} \left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{(k+1)}\right\|^{\frac{1-\theta}{\theta}} \quad \text{[by Lemma 5.6]} \\
&= \frac{2\mu^{\frac{1}{\theta}}}{\lambda\eta^{\frac{1}{\theta}}(1-\theta)} \left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{(k+1)}\right\|^{\frac{1-\theta}{\theta}}.
\end{aligned} \tag{5.12}$$

(1) If $\theta = \frac{1}{2}$, then $\frac{1-\theta}{\theta} = 1$. By (5.12), we have

$$\Delta_k \leq \frac{2\mu^{\frac{1}{\theta}}}{\lambda\eta^{\frac{1}{\theta}}(1-\theta)} \left(\Delta_k - \Delta_{k+1}\right).$$

This implies that

$$\Delta_{k+1} \leq \sigma\Delta_k, \tag{5.13}$$

where

$$\sigma = \frac{2\mu^{\frac{1}{\theta}} - \lambda\eta^{\frac{1}{\theta}}(1-\theta)}{2\mu^{\frac{1}{\theta}}}.$$

By (5.11) and (5.12), we know that

$$\left\|\mathbf{Z}^{(k)} - \mathbf{Z}^{\#}\right\| \le \Delta_k \le \sigma\Delta_{k-1} \le \cdots \le \sigma^k\Delta_0 \equiv \beta\sigma^k,$$

where $\beta \equiv \Delta_0$ is finite by Theorem 5.8. The conclusion follows.

(2) Let

$$\kappa^{\frac{1-\theta}{\theta}} = \frac{2\mu^{\frac{1}{\theta}}}{\lambda\eta^{\frac{1}{\theta}}(1-\theta)}.$$

By (5.12), we get

$$\Delta_k^{\frac{\theta}{1-\theta}} \le \kappa\left(\Delta_k - \Delta_{k+1}\right).$$

Define $\zeta(\alpha) := \alpha^{-\frac{\theta}{1-\theta}}$. Then ζ is monotonically decreasing. We have

$$\begin{aligned} \tfrac{1}{\kappa} \le{}& \zeta(\Delta_k)\left(\Delta_k - \Delta_{k+1}\right) \\ ={}& \int_{\Delta_{k+1}}^{\Delta_k} \zeta(\Delta_k)d\alpha \\ \le{}& \int_{\Delta_{k+1}}^{\Delta_k} \zeta(\alpha)d\alpha \\ ={}& -\tfrac{1-\theta}{2\theta-1}\left(\Delta_k^{-\frac{2\theta-1}{1-\theta}} - \Delta_{k+1}^{-\frac{2\theta-1}{1-\theta}}\right). \end{aligned}$$

Denote $\nu := -\frac{2\theta-1}{1-\theta}$. Then $\nu < 0$ since $\frac{1}{2} < \theta < 1$. We have

$$\Delta_{k+1}^{\nu} - \Delta_k^{\nu} \ge \xi > 0,$$

where $\xi \equiv -\frac{\nu}{\kappa}$. This implies that

$$\Delta_k \le \left[\Delta_0^{\nu} + k\xi\right]^{\frac{1}{\nu}} \le (k\xi)^{\frac{1}{\nu}}.$$

Letting $\beta = \xi^{\frac{1}{\nu}}$, we have the conclusion. □

6 Numerical Experiments

In this section, we are going to evaluate the performance of the proposed low-rank quaternion matrix approximation method for color image completion. Given a color image with possibly missing pixels, we represent it as a quaternion matrix, where red, green, blue values are placed into three imaginary parts of the pure quaternion matrix. For simplicity, missing entries of the quaternion matrix take the mean of known entries. In this way, we get a quaternion matrix $D \in \mathbb{Q}^{m\times n}$. Let $USV^* = D$ be the QSVD decomposition; See Theorem 2.2. Singular values of D are located in the diagonal of S and satisfy $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_{\min(m,n)} \ge 0$. The parameter of rank is determined by choosing the smallest integer r that satisfies

$$\sum_{i=1}^{r} \sigma_i \ge 0.96 \sum_{i=1}^{\min(m,n)} \sigma_i.$$

We define $S_{1:r}$ as the leading $r \times r$ diagonal block of S. Let $U_{1:r}$ and $V_{1:r}$ be the first r columns of U and V, respectively. Then, the initial points are $A^{(0)} = U_{1:r}$ and $B^{(0)} = S_{1:r}V_{1:r}^*$. The LRQMA algorithm terminates if the decrease of the objective function is tiny, $\frac{f(\mathbf{Z}^{(k-1)})-f(\mathbf{Z}^{(k)})}{\max(1,f(\mathbf{Z}^{(k)}))} \le 10^{-5}$, or the number of iterations exceeds 500.

We compare four algorithms as follows.

Figure 1: Original color images are illustrated in the first line. Second to the last lines report associated images with 25%, 50%, and 75% randomly missing pixels, respectively.

- LRQMA: The proposed algorithm in Section 3, where the quaternion toolbox for matlab [11] is used.
- LMaFit: A low-rank matrix factorization model is solved by a nonlinear successive over-relaxation algorithm for the matrix completion problem [15].
- FPCA: A fixed point continuation with approximate SVD (FPCA) algorithm is designed for solving the matrix nuclear norm minimization problem [8].
- LRTC: The low-rank tensor completion approach is proposed for estimating missing values in visual data [7].

Testing color images (named lena, baboon, fruits, peppers, lotus, sailboat, and airplane) are from the public-domain test images dataset, which are illustrated in Figure 1. These color images are all of size $256 \times 256 \times 3$. Incomplete images with 25%, 50%, and 75% randomly missing pixels are shown in the second to the last lines of Figure 1. We calculate the peak signal-to-noise ratio (PSNR) of the recovered images

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{(2^8-1)^2}{\mathrm{MSE}}\right),$$

where the mean square error is

$$\mathrm{MSE} = \frac{1}{256 \times 256 \times 3}\sum_{i=1}^{256}\sum_{j=1}^{256}\sum_{k=1}^{3}\left(X_{ijk}^{rec} - X_{ijk}^{true}\right)^2,$$

and X^{rec} and X^{true} denote the recovered and true color images, respectively.

https://homepages.cae.wisc.edu/ ece533/images/

Methods	lena	baboon	fruits	peppers	lotus	sailboat	airplane
				25% missing pixels			
LRQMA	**34.682**	**29.292**	**33.116**	**37.183**	**34.070**	**33.729**	31.079
LMaFit	28.410	26.911	27.608	27.130	27.565	26.271	27.995
FPCA	29.607	26.263	28.505	29.607	28.392	27.044	29.562
LRTC	31.749	29.020	31.181	32.311	31.987	29.788	**31.961**
				50% missing pixels			
LRQMA	**31.335**	**25.098**	**29.146**	**32.602**	**29.431**	**28.443**	**28.264**
LMaFit	24.963	23.415	24.023	23.772	24.242	22.875	24.552
FPCA	26.353	23.195	25.068	26.140	25.473	23.677	25.823
LRTC	26.414	24.626	26.321	26.777	27.008	24.516	26.822
				75% missing pixels			
LRQMA	**26.887**	**21.873**	**24.998**	**26.956**	**23.052**	**22.873**	22.500
LMaFit	21.559	19.252	18.708	10.685	8.196	18.903	20.003
FPCA	20.942	17.881	19.903	19.746	19.442	17.952	19.724
LRTC	21.765	21.025	22.184	21.610	22.402	20.061	**22.507**

Table 1: PSNRs of recovered color images.

Methods	lena	baboon	fruits	peppers	lotus	sailboat	airplane
LRQMA	**31.533**	**30.223**	**31.698**	**32.912**	**32.442**	**31.652**	28.760
LMaFit	28.814	24.207	25.720	28.157	27.884	27.852	21.881
FPCA	30.108	27.783	30.379	31.078	29.473	28.615	29.351
LRTC	30.399	28.929	31.399	30.880	30.343	29.097	**30.197**

Table 2: PSNRs of color images.

In numerical experiments, we examine four algorithms for completing color images with randomly missing pixels. Resulting images are illustrated in Figures 2–4 and associated PSNRs are reported in Table 1. First, all methods report recovered images. As the pixels missing ratio increases from 25% to 75%, the quality of recovered color images decreases. Second, the tensor-based approach LRTC performs better than real matrix-based approaches LMaFit and FPCA, because the tensor could capture the inherent multilinear structure of color images though the third mode of the tensor is only of 3 dimension. Finally, since we transform the red-green-blue pixel to a pure quaternion as a while, the proposed quaternion-based method works wonderfully and gives higher PSNR color images when compared with other approaches.

Next, we consider the case that color images are of text missing pixels, which are illustrated in the second line of Figure 5. Using four methods, the resulting PSNR values and recovered images are reported in Table 2 and lines 3–6 of Figure 5. Obviously, our quaternion-based method LRQMA outperforms other methods. Hence, we claim that the proposed low-rank quaternion matrix approximation is effective and powerful for color image completion.

7 Conclusion

In a certain sense, the gradient of a real function in quaternion matrix variables, studied in Section 4, is a kind of the pseudo-derivative mentioned in [17]. What we presented in (4.2) is a concise form of such a pseudo-derivative. It turns out to give us simple expressions for

Figure 2: Recovered images by various algorithms from 25% randomly missing pixels.

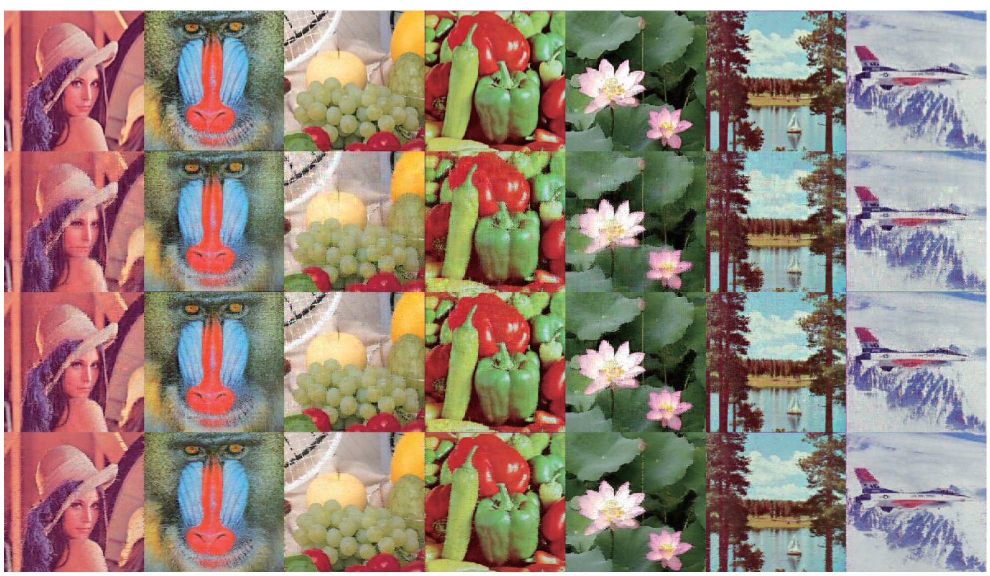

Figure 3: Recovered images by various algorithms from 50% randomly missing pixels.

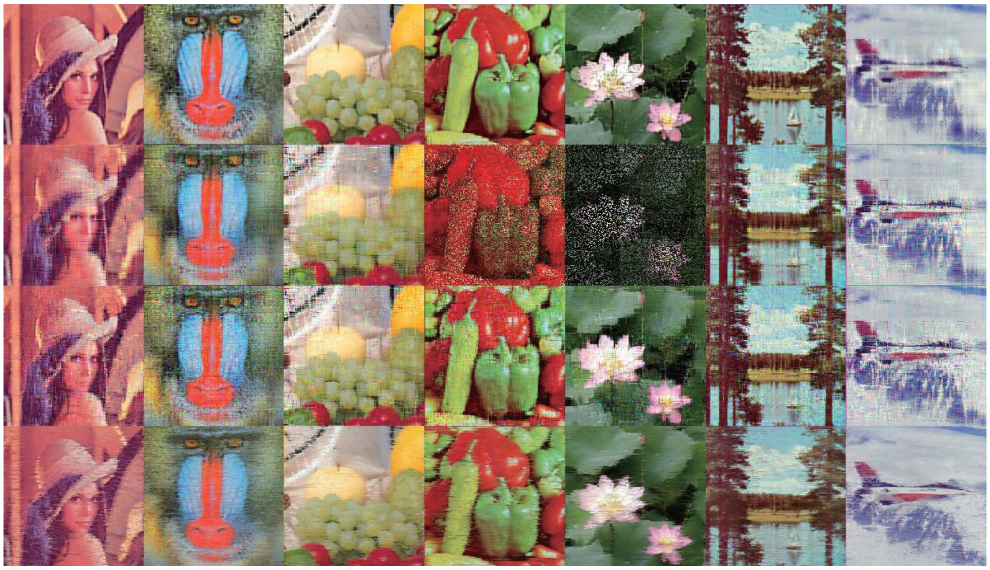

Figure 4: Recovered images by various algorithms from 75% randomly missing pixels.

Figure 5: Text missing images and recovered images by various methods.

gradients of functions, and optimality conditions of optimization problems, involved in our convergence analysis. It may be worth further exploring the use of (4.2).

Acknowledgment

We are thankful to Prof. Qingwen Wang for his help on quaternion matrices.

References

[1] H. Attouch and J. Bolte, On the convergence of the proximal algorithm for nonsmooth functions involving analytic features, *Math. Program.* 116 (2009) 5–16.

[2] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Łojasiewicz inequality, *Math. Oper. Res.* 35 (2010) 438–457.

[3] J. Bolte, A. Daniilidis, and A. Lewis, The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems, *SIAM J. Optim.* 17 (2007) 1205–1223.

[4] Y. Chen, X. Xiao and Y. Zhou, Low-rank quaternion approximation for color image processing, *IEEE Trans. Image Process.* 29 (2019) 1057–7149.

[5] Z. Jia, *The Eigenvalue Problem of Quaternion Matrix: Structure-Preserving Algorithms and Applications*, Science Press, Beijing, 2019.

[6] N. Le Bihan and S.J. Sangwine, Quaternion principal component analysis of color images, *Proceedings 2003 International Conference on Image Processing* 1 (2003) 809–812.

[7] J. Liu, P. Musialski, P. Wonka, and Jieping Ye, Tensor completion for estimating missing values in visual data, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (2013) 208–220.

[8] S. Ma, D. Goldfarb, and L. Chen, Fixed point and Bregman iterative methods for matrix rank minimization, *Math. Program.*, 128 (2011) 321–353.

[9] S.J. Sangwine, Fourier transforms of colour images using quaternion or hypercomplex numbers, *Electron. Lett.* 32 (1996) 1979–1980.

[10] S.J. Sangwine and N. Le Bihan, Quaternion singular value decomposition based on bidiagonalization to a real or complex matrix using quaternion Householder transformations, *Appl. Math. Comput.* 182 (2006) 727–738.

[11] S.J. Sangwine and N. Le Bihan, Quaternion toolbox for Matlab, Version 3.0, (2020). `http://qtfm.sourceforge.net/`.

[12] G. Song, W. Ding, and M. K. Ng, Low rank pure quaternion approximation for pure quaternion matrices, *SIAM J. Matrix Anal. Appl.* 42 (2021) 58–82.

[13] L. Sorber, M. van Barel and L. De Lathauwer, Unconstrianed optimization of real functions in complex variables, *SIAM J. Optim.* 22 (2012) 879–898.

[14] W. Sun and Y. Yuan, *Optimization Theory and Methods: Nonlinear Programming*, Springer, New York, 2006.

[15] Z. Wen, W. Yin, and Y. Zhang, Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm, *Math. Program. Comput.* 4 (2012) 333–361.

[16] B. Xie, *Abstract Algebra*, Shanghai Science Publisher, Shenghai, 1982 (in Chinese).

[17] D. Xu, C. Jahanchachi, C.C. Took and D.P. Mandic, Enabling quaternion derivatives: the generalized HR calculus, *R. Soc. open sci.* 2 (2015): 150255.

[18] M. Wei, Y. Li, F. Zhang and J. Zhao, *Quaternion Matrix Computations*, Nova Science Publisher, New York, 2018.

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

[20] W. Zhuang, *Introduction to the Theory of Matrices over the Skew Field*, Science Publisher, Beijing, 2006 (in Chinese).

Manuscript received 1 June 2021
revised 3 September 2021
accepted for publication 8 September 2021

Yannan Chen
School of Mathematical Sciences
South China Normal University
Guangzhou, China
E-mail address: ynchen@scnu.edu.cn

Liqun Qi
Future Network Theory Lab
2012 Labs Huawei Tech. Investment Co., Ltd, Shatin
New Territory, Hong Kong, China;
Department of Applied Mathematics
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong, China
E-mail address: maqilq@polyu.edu.hk

Xinzhen Zhang
School of Mathematics, Tianjin University, Tianjin, China
E-mail address: xzzhang@tju.edu.cn