

低秩四元数矩阵分解模型的优化理论及其应用

郑佳莉¹, 吴玉玲¹, 柯艺芬^{1,2,3*}, 许小芳¹

(1. 福建师范大学 数学与统计学院, 福建 福州 350117;

2. 福建师范大学 分析数学及应用教育部重点实验室, 福建 福州 350117;

3. 福建省应用数学中心, 福建 福州 350117)

摘要: 本文研究低秩四元数矩阵分解模型的优化理论及其在彩色图像处理中的应用, 提出四元数梯度下降法及其改进算法来求解低秩四元数矩阵分解模型并分析算法的收敛性。通过将模型应用于彩色图像处理上, 数值结果表明所提算法的可行性和有效性。

关键词: 四元数; 低秩四元数矩阵分解模型; 四元数梯度下降法

中图分类号: O241.6 **文献标志码:** A **文章编号:** 0253-2395(2024)06-1115-12

Optimization Theory of Low Rank Quaternion Matrix Decomposition Model and Its Application

ZHENG Jiali¹, WU Yuling¹, KE Yifen^{1,2,3*}, XU Xiaofang¹

(1. School of Mathematics and Statistics, Fujian Normal University, Fuzhou 350117, China;

2. Key Laboratory of Analytical Mathematics and Applications (Ministry of Education),

Fujian Normal University, Fuzhou 350117, China;

3. Center for Applied Mathematics of Fujian Province (FJNU), Fuzhou 350117, China)

Abstract: This paper studies the low rank quaternion matrix decomposition model and its application in color image processing. Quaternion gradient descent algorithm and its improved version are proposed to solve the low rank quaternion matrix decomposition model. And the convergence of the algorithms is considered. Numerical results on color images show the feasibility and effectiveness of the proposed algorithms.

Key words: quaternion; low rank quaternion matrix decomposition model; quaternion gradient descent algorithm

0 引言

四元数矩阵在彩色图像处理中得到了广泛的应用, 包括: 彩色图像去噪与修复、彩色人脸识别以及彩色图像修复等方面^[1-4]。此外, 四元数矩阵在其他工程领域也得到了广泛地应用^[5-12]。这些研究工作中的数学模型可以转化为以四元数矩阵为决策变量的优化问题。

1996年, 文献[13]提出了一种编码纯四元数, 其三个虚部分别表示RGB(红-绿-蓝)图像的三个通道分量。在文献[14]中给出了一些关于四元数和四元数矩阵理论的简要证明。此后, 文献

收稿日期: 2023-12-06; 接受日期: 2024-03-20

基金项目: 国家自然科学基金(12371378; 11901098); 福建省自然科学基金(2023J011127; 2023J01955; 2020J05034)

作者简介: 郑佳莉(2000-), 女, 福建莆田人, 硕士研究生, 研究方向为数值代数及其应用。E-mail: zhengjl202203@163.com

* 通信作者: 柯艺芬(KE Yifen), E-mail: keyifen@fjnu.edu.cn

引文格式: 郑佳莉, 吴玉玲, 柯艺芬, 等. 低秩四元数矩阵分解模型的优化理论及其应用[J]. 山西大学学报(自然科学版), 2024, 47(6): 1115-1126. DOI: 10.13451/j.sxu.ns.2024051

[8, 15]提出了关于四元数变量实值函数的梯度。另外,文献[16]将复数域理论推广到了四元数域。Wang等证明了在复数域中,梯度下降算法是可行的^[17]。所有这些丰富的数学工具为使用四元数表示彩色图像提供了理论保证。

最近,文献[3]将彩色图像修复问题转化为一个低秩四元数矩阵分解(Low rank quaternion matrix decomposition, LRQD)模型。该模型旨在解决四元数矩阵变量中实值函数的等式约束优化问题,并对其收敛性进行了分析。另一方面,文献[18]提出一个关键的工具,即R积(R-product),通过将两个四元数矩阵的R积等价于这两个四元数内积的实部来定义。它还指出了低秩四元数矩阵分解模型满足二阶连续可微,其可视为四元数矩阵优化问题(Quaternion matrix optimization problem)的一个特例。然而,它并未对低秩四元数矩阵分解模型进行深入详细地分析。

本文针对低秩四元数矩阵分解模型进行了最优性分析,并引入了四元数矩阵变量实值函数的一阶导数。将几个重要的优化理论推广到了四元数域,并提出了四元数梯度下降法及其改进版本,给出算法的收敛性分析。

本文剩余的组织结构如下。第1节回顾了四元数和四元数矩阵的一些必要知识,并介绍了低秩四元数矩阵分解模型。第2节引入了关键工具R积和R-线性无关,用来定义四元数矩阵变量的实值函数的一阶导数,并给出了低秩四元数矩阵分解模型的一阶最优性条件。第3节介绍了四元数梯度下降法及改进算法并分析了其算法的收敛性。第4节给出了数值实验。最后一节总结了本文的研究工作。

本文用 \mathbb{R} 、 \mathbb{C} 和 \mathbb{Q} 分别表示实数域、复数域和四元数域;用 $\mathbb{R}^{m \times n}$ 、 $\mathbb{C}^{m \times n}$ 和 $\mathbb{Q}^{m \times n}$ 分别表示 $m \times n$ 阶实矩阵、复矩阵和四元数矩阵的集合。记 $\mathbb{H} := \mathbb{Q}^{m_1 \times n_1} \times \mathbb{Q}^{m_2 \times n_2} \times \mathbb{Q}^{m_3 \times n_3}$,若 $A \in \mathbb{Q}^{m_1 \times n_1}$, $H \in \mathbb{Q}^{m_2 \times n_2}$ 和 $W \in \mathbb{Q}^{m_3 \times n_3}$,称 $X = (A, H, W) \in \mathbb{H}$ 为矩阵分量向量,用 0 和 O 分别表示具有适当维数的零向量和零矩阵。若 $A = (a_{st}) \in \mathbb{Q}^{n \times n}$,用 $\text{tr}(A)$ 表示 A 的迹,其定义为 $\text{tr}(A) = \sum_{s=1}^n a_{ss}$ 。

1 预备知识

本节主要回顾四元数、四元数矩阵以及介绍低秩四元数矩阵分解模型的相关结论。

1.1 四元数及四元数矩阵

四元数 x 可以表示为

$$x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k},$$

其中 $x_0, x_1, x_2, x_3 \in \mathbb{R}$, $\mathbf{i}, \mathbf{j}, \mathbf{k}$ 为三个虚单位,其满足下列条件

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

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

这些规则和分布规律决定了四元数的乘法是非交换的。

四元数 x 的实部记为 $\text{Re}(x) = x_0$,三个虚部分别记为 $\text{Im}_i(x) = x_1$ 、 $\text{Im}_j(x) = x_2$ 和 $\text{Im}_k(x) = x_3$ 。

四元数 x 的共轭为

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

四元数 x 的模为

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

若 $x \neq 0$,则 $x^{-1} = \frac{x^*}{|x|^2}$ 。

四元数矩阵 $A \in \mathbb{Q}^{m \times n}$ 可表示为

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

其中 $A_0, A_1, A_2, A_3 \in \mathbb{R}^{m \times n}$ 。四元数矩阵 A 的转置定义为

$$A^T = A_0^T + A_1^T i + A_2^T j + A_3^T k,$$

四元数矩阵 A 的共轭转置定义为

$$A^* = A_0^T - A_1^T i - A_2^T j - A_3^T k.$$

设 $A \in \mathbb{Q}^{m \times n}$, 若存在矩阵 $B \in \mathbb{Q}^{n \times m}$ 满足 $AB = BA = I_n$, 则称 A 是可逆的, 记 $A^{-1} = B$.

设 $A = (a_{st}), B = (b_{st}) \in \mathbb{Q}^{m \times n}$, 它们的内积定义为

$$\langle A, B \rangle = \text{tr}(B^* A) \in \mathbb{Q},$$

它们实内积的定义为

$$A \cdot B := \text{Re} \langle A, B \rangle = \text{Re} \left(\sum_{s=1}^m \sum_{t=1}^n b_{st}^* a_{st} \right) \in \mathbb{R}.$$

四元数矩阵 A 的 Frobenius 范数定义为

$$\|A\|_F = \sqrt{\langle A, A \rangle} = \sqrt{\text{tr}(A^* A)} = \sqrt{\sum_{s=1}^m \sum_{t=1}^n |a_{st}|^2}.$$

显然, 有 $A \cdot A = \langle A, A \rangle = \|A\|_F^2$. 四元数矩阵其他范数的定义可参考文献[1].

定义 1^[18] 若存在不全为零的实数 $\alpha_1, \alpha_2, \dots, \alpha_p$ 满足

$$\sum_{j=1}^p \alpha_j A_j = O,$$

则称 $\{A_j \in \mathbb{R}^{m \times n}; j = 1, \dots, p\}$ 是 \mathbb{R} -线性相关的. 否则, 称为 \mathbb{R} -线性无关的.

定义 2^[18-19] 假设 $f: \mathbb{Q}^{m \times n} \rightarrow \mathbb{R}$ 是二阶连续可微, 若其梯度满足

$$\|\nabla f(Y) - \nabla f(\hat{Y})\|_F \leq \eta \|Y - \hat{Y}\|_F, \forall Y, \hat{Y} \in \mathbb{Q}^{m \times n},$$

则称 ∇f 是 Lipschitz 连续的, 其中 $\eta > 0$ 是 Lipschitz 常数.

引理 1^[18-19] 假设 $f: \mathbb{Q}^{m \times n} \rightarrow \mathbb{R}$ 是二阶连续可微, 且梯度 ∇f 是 Lipschitz 连续, 则有

$$f(Y) \leq f(\hat{Y}) + \nabla f(\hat{Y}) \cdot (Y - \hat{Y}) + \frac{\eta}{2} \|Y - \hat{Y}\|_F^2, \forall Y, \hat{Y} \in \mathbb{Q}^{m \times n},$$

其中 Lipschitz 常数 $\eta > 0$.

引理 2^[3] 假设序列 $\{Z^k = (Z_1^k, Z_2^k, Z_3^k)\}_{k=0}^{+\infty} \subset \mathbb{H}$, 若序列满足

$$\sum_{k=0}^{+\infty} \|Z^k - Z^{k+1}\|_F^2 < +\infty,$$

即:

$$\sum_{k=0}^{+\infty} \|Z^k - Z^{k+1}\|_F^2$$

收敛, 且序列 $\{Z^k\}_{k=1}^{+\infty}$ 是有界的, 则有 $\lim_{k \rightarrow +\infty} \|Z^k - Z^{k+1}\|_F^2 = 0$.

1.2 低秩四元数矩阵分解模型

考虑如下低秩四元数矩阵分解模型^[3]:

$$\min_{A, H, W} \frac{1}{2} \|HW - A\|_F^2, \text{ s.t. } A_\Omega = D_\Omega, \tag{1}$$

其中 $A \in \mathbb{Q}^{m \times n}, H \in \mathbb{Q}^{m \times r}, W \in \mathbb{Q}^{r \times n}$ 是决策变量, D 是给定的四元数数据集矩阵, Ω 为 D 观测部分的索引集, $A_\Omega = D_\Omega$ 表示 A 和 D 对应项在索引集上是相等的.

显然, 若选取 $X = (A, H, W), f(X) = \frac{1}{2} \|HW - A\|_F^2, h_l(X) = a_{st} - d_{st}$, 其中 (s, t) 为 Ω 的第 l 指标 ($l = 1, \dots, p$), $p = |\Omega|$ 为 Ω 基数的大小. 因此, 模型(1)为二阶连续可微的四元数矩阵优化问题的一个特例^[18].

文献[3]提出了求解低秩四元数矩阵分解模型(1)的数值算法, 即: 给定 $H^0 \in \mathbb{Q}^{m \times r}$,

$W^0 \in \mathbb{Q}^{r \times n}$, 其中 $r < \min\{m, n\}$, 记 Ω^c 为 Ω 的补集, 则对于第 k 步迭代有

$$\begin{cases} A_\Omega^k = D_\Omega^k, A_{\Omega^c}^k = (H^k W^k)_{\Omega^c}, \\ W^{k+1} = [(H^k)^* H^k]^{-1} (H^k)^* A^k, \\ H^{k+1} = A^k (W^{k+1})^* [W^{k+1} (W^{k+1})^*]^{-1}. \end{cases} \quad (2)$$

迭代格式(2)可求得模型(1)的最优解。然而, 迭代格式(2)涉及两个四元数矩阵方程组的求解。此外, 文献[18]给出了低秩四元数矩阵分解模型的部分最优性理论结果, 没有给出迭代格式的收敛性分析。

本文将从避免求解四元数矩阵方程组的角度, 提出新的数值算法用于求解模型(1)。同时, 给出模型(1)的最优性条件。进一步, 提出四元数梯度下降算法并分析所提算法的收敛性。

2 低秩四元数矩阵分解模型的最优化理论

在本节中, 借助于 \mathbb{R} 积、 \mathbb{R} -线性无关和梯度的定义, 用它们来定义四元数矩阵变量中实值函数的一阶导数, 建立模型(1)的一阶最优性条件。

定义 3^[3] 设 $f: \mathbb{Q}^{m \times n} \rightarrow \mathbb{R}$, $Y = Y_0 + Y_1 \mathbf{i} + Y_2 \mathbf{j} + Y_3 \mathbf{k}$, 这里 $Y_0, Y_1, Y_2, Y_3 \in \mathbb{R}^{m \times n}$ 。若 $\frac{\partial f}{\partial Y_i}$ 在 Y_i 处存在 ($i = 0, 1, 2, 3$), 则称 f 在 Y 处可微, 记

$$\nabla f(Y) = \frac{\partial f}{\partial Y_0} + \frac{\partial f}{\partial Y_1} \mathbf{i} + \frac{\partial f}{\partial Y_2} \mathbf{j} + \frac{\partial f}{\partial Y_3} \mathbf{k}.$$

若 f 在 Y_i ($i = 0, 1, 2, 3$) 的邻域内有定义且在 Y_i 处连续, 则称 f 在 Y 处连续可微。

记 $X = (A, H, W) \in \mathbb{H}$, 其中 $A = A_0 + A_1 \mathbf{i} + A_2 \mathbf{j} + A_3 \mathbf{k} \in \mathbb{Q}^{m_1 \times m_1}$, $H = H_0 + H_1 \mathbf{i} + H_2 \mathbf{j} + H_3 \mathbf{k} \in \mathbb{Q}^{m_2 \times n_2}$, $W = W_0 + W_1 \mathbf{i} + W_2 \mathbf{j} + W_3 \mathbf{k} \in \mathbb{Q}^{m_3 \times n_3}$ 。假设 $f: \mathbb{H} \rightarrow \mathbb{R}$ 是连续可微的, 则 f 在 X 处对 A_i, H_i, W_i ($i = 0, 1, 2, 3$) 是可微, 进而定义 f 在 X 处的偏导数和梯度分别为:

$$\begin{cases} \nabla_A f(X) = \frac{\partial}{\partial A_0} f(X) + \frac{\partial}{\partial A_1} f(X) \mathbf{i} + \frac{\partial}{\partial A_2} f(X) \mathbf{j} + \frac{\partial}{\partial A_3} f(X) \mathbf{k}, \\ \nabla_H f(X) = \frac{\partial}{\partial H_0} f(X) + \frac{\partial}{\partial H_1} f(X) \mathbf{i} + \frac{\partial}{\partial H_2} f(X) \mathbf{j} + \frac{\partial}{\partial H_3} f(X) \mathbf{k}, \\ \nabla_W f(X) = \frac{\partial}{\partial W_0} f(X) + \frac{\partial}{\partial W_1} f(X) \mathbf{i} + \frac{\partial}{\partial W_2} f(X) \mathbf{j} + \frac{\partial}{\partial W_3} f(X) \mathbf{k}, \\ \nabla f(X) = [\nabla_A f(X), \nabla_H f(X), \nabla_W f(X)]. \end{cases}$$

显然 $\nabla f(X) \in \mathbb{H}$ 。

下面给出模型(1)的拉格朗日函数。

记 $p = |\Omega|$ 为 Ω 基数的大小。令

$$f(X) = \frac{1}{2} \|HW - A\|_F^2, \quad (3)$$

$$h_l(X) = h_{0,l}(X) + h_{1,l}(X) \mathbf{i} + h_{2,l}(X) \mathbf{j} + h_{3,l}(X) \mathbf{k} = a_{st} - d_{st}, \quad (4)$$

其中 $h_{0,l}(X) = \text{Re}(a_{st} - d_{st})$, $h_{1,l}(X) = \text{Im}_1(a_{st} - d_{st})$, $h_{2,l}(X) = \text{Im}_2(a_{st} - d_{st})$ 和 $h_{3,l}(X) = \text{Im}_3(a_{st} - d_{st})$, 这里 (s, t) 对应 Ω 的第 l 指标 ($l = 1, \dots, p$)。则模型(1)的拉格朗日函数为

$$\mathcal{L}(X) = f(X) + \sum_{i=0}^3 \sum_{l=1}^p \lambda_{i,l} h_{i,l}(X).$$

定理 1 假设函数 $f, h_{i,l}: \mathbb{H} \rightarrow \mathbb{R}$ ($i = 0, 1, 2, 3; l = 1, \dots, p$) 如式(3)和式(4)所定义, 且满足在 X 处是一阶连续可微的。设 $X^\# = (A^\#, H^\#, W^\#) \in \mathbb{H}$ 是模型(1)的最优解。若 $\nabla h_{i,l}(X^\#)$ 是 \mathbb{R} -线性无关, 则存在拉格朗日乘子 $\lambda_{i,l} \in \mathbb{R}$ ($i = 0, 1, 2, 3; l = 1, \dots, p$) 满足

$$\nabla f(\mathbf{X}^\#) + \sum_{i=0}^3 \sum_{l=1}^p \lambda_{i,l} \nabla h_{i,l}(\mathbf{X}^\#) = \mathbf{O}, \tag{5}$$

$$h_l(\mathbf{X}^\#) = \mathbf{O}. \tag{6}$$

证明 将上式函数中的四元数矩阵变量转化为由实矩阵 $A_i, H_i, W_i (i = 0, 1, 2, 3)$ 为变量的实值函数。由于 $\nabla h_{i,l}(\mathbf{X}^\#) (i = 0, 1, 2, 3; l = 1, \dots, p)$ 是 \mathbb{R} -线性无关, 根据一阶实优化条件理论, 可得式(5)和式(6)成立。

定义4 设 $f: \mathbb{H} \rightarrow \mathbb{R}$ 是二阶连续可微, 若 $Y^\# \in \mathbb{H}$ 是模型(1)的最优解, 且满足

$$\nabla f(Y^\#) = \mathbf{O},$$

则称 $Y^\#$ 是模型(1)的稳定点。

接下来考虑模型(1)的目标函数 $f(\mathbf{X}) = \frac{1}{2} \|\mathbf{HW} - \mathbf{A}\|_F^2$ 。由文献[3], 可得

$$\nabla_A f(\mathbf{X}) = \mathbf{A} - \mathbf{HW}, \tag{7}$$

$$\nabla_H f(\mathbf{X}) = (\mathbf{HW} - \mathbf{A})\mathbf{W}^*, \tag{8}$$

$$\nabla_W f(\mathbf{X}) = \mathbf{H}^*(\mathbf{HW} - \mathbf{A}). \tag{9}$$

并有以下定理成立。

定理2 假设 $f(\mathbf{X}) = \frac{1}{2} \|\mathbf{HW} - \mathbf{A}\|_F^2$, 则偏导数 $\nabla_H f, \nabla_W f, \nabla_A f$ 满足 Lipschitz 连续, 即有下列不等式

$$_F, \forall H_1, H_2,$$

$$\|\nabla_W f(\mathbf{A}, H, \mathbf{W}_2) - \nabla_W f(\mathbf{A}, H, \mathbf{W}_1)\|_F \leq \eta \|\mathbf{W}_2 - \mathbf{W}_1\|_F, \forall \mathbf{W}_1, \mathbf{W}_2,$$

$$\|\nabla_A f(\mathbf{A}_2, H, \mathbf{W}) - \nabla_A f(\mathbf{A}_1, H, \mathbf{W})\|_F \leq \eta \|\mathbf{A}_2 - \mathbf{A}_1\|_F, \forall \mathbf{A}_1, \mathbf{A}_2,$$

其中 $\eta = \max\{1, \|\mathbf{WW}^*\|_F, \|\mathbf{H}^*H\|_F, c\}$ 是 Lipschitz 常数, 这里 c 为给定的适当的常数。

证明 由定义2和式(8)可得

$$\|\nabla_H f(\mathbf{A}, H_2, \mathbf{W}) - \nabla_H f(\mathbf{A}, H_1, \mathbf{W})\|_F =$$

$$\|(\mathbf{H}_2\mathbf{W} - \mathbf{A})(\mathbf{W})^* - (\mathbf{H}_1\mathbf{W} - \mathbf{A})\mathbf{W}^*\|_F =$$

$$\|(\mathbf{H}_2 - \mathbf{H}_1)\mathbf{WW}^*\|_F \leq \|\mathbf{WW}^*\|_F \|\mathbf{H}^{k+1} - \mathbf{H}^k\|_F \leq \eta \|\mathbf{H}^{k+1} - \mathbf{H}^k\|_F.$$

同理, 可得另外两个不等式。

定理3 设 $\mathbf{X}^\# = (\mathbf{A}^\#, \mathbf{H}^\#, \mathbf{W}^\#) \in \mathbb{H}$ 是模型(1)的稳定点, 则有

$$(\mathbf{H}^\#\mathbf{W}^\# - \mathbf{A}^\#)(\mathbf{W}^\#)^* = \mathbf{O}_{m \times r},$$

$$(\mathbf{H}^\#)^*(\mathbf{H}^\#\mathbf{W}^\# - \mathbf{A}^\#) = \mathbf{O}_{r \times n},$$

$$\mathbf{A}^\#_\Omega = \mathbf{D}_\Omega, \mathbf{A}^\#_{\Omega^c} = (\mathbf{H}^\#\mathbf{W}^\#)_{\Omega^c}$$

成立, 其中 $r < \min\{m, n\}$, Ω^c 是 Ω 的补集。

证明 由于 $\mathbf{X}^\#$ 是模型(1)的稳定点, 则 $\mathbf{X}^\#$ 需满足约束条件 $\mathbf{A}^\#_\Omega = \mathbf{D}_\Omega$, 且有

$$\nabla f(\mathbf{X}^\#) = (\nabla_A f(\mathbf{X}^\#), \nabla_H f(\mathbf{X}^\#), \nabla_W f(\mathbf{X}^\#)) = \mathbf{O}.$$

再由式(7)、式(8)和式(9)可得

$$\mathbf{A}^\#_{\Omega^c} - (\mathbf{H}^\#\mathbf{W}^\#)_{\Omega^c} = \mathbf{O}_{m \times n},$$

$$(\mathbf{H}^\#\mathbf{W}^\# - \mathbf{A}^\#)(\mathbf{W}^\#)^* = \mathbf{O}_{m \times r},$$

$$(\mathbf{H}^\#)^*(\mathbf{H}^\#\mathbf{W}^\# - \mathbf{A}^\#) = \mathbf{O}_{r \times n}.$$

即证明完成。

3 四元数梯度下降法及其收敛性分析

令

$$\mathbb{S} = \{Z \in \mathbb{Q}^{m \times n} \mid h_l(Z) = 0, l = 1, \dots, p, p = |\Omega|\}.$$

记 $\mathbb{H}_1 = \mathbb{S} \times \mathbb{Q}^{m \times r} \times \mathbb{Q}^{r \times n}$, 则模型(1)可改写为如下优化问题:

$$\min_{X=(A, H, W) \in \mathbb{H}_1} f(X) = \frac{1}{2} \|HW - A\|_{F^\circ}^2 \quad (10)$$

受梯度下降法思想的启发, 本文提出一种四元数梯度下降法求解问题(10)。记 $X^k = (A^k, H^k, W^k)$ 。对于第 k 步迭代, 有

$$X^{k+1} = X^k - \alpha \nabla f(X^k),$$

其中 α 是步长因子且 $\nabla f(X^k) = ((A^k - H^k W^k, (H^k W^k - A^k)(W^k)^*, (H^k)^*(H^k W^k - A^k))$ 。

具体实施步骤见算法1。

算法1 (四元数梯度下降法求解低秩四元数矩阵分解模型)

步1: 给定 ϵ , 选取 $X^0 = (A^0, H^0, W^0) \in \mathbb{H}_1$ 。

步2: 更新 $H^{k+1} = H^k - \alpha(H^k W^k - A^k)(W^k)^*$ 。

步3: 更新 $W^{k+1} = W^k - \alpha(H^{k+1})^*(H^{k+1} W^k - A^k)$ 。

步4: 更新 $A_\Omega^{k+1} = D_\Omega, A_\Omega^{k+1} = (A^k - \alpha(A^k - H^{k+1} W^{k+1}))_\Omega$ 。

步5: 如果 $\|X^{k+1} - X^k\|_F < \epsilon$, 停止; 否则, 置 $k := k + 1$, 转步2。

下面分析算法1的收敛性。

定理4 假设 $X^k, X^{k+1} \in \mathbb{H}_1$ 是由算法1生成, 当步长因子 $\alpha \in \left(0, \frac{2}{\eta}\right)$, 算法1收敛且有以下不等式成立

$$f(X^{k+1}) \leq f(X^k) + \left(\frac{\eta}{2} - \frac{1}{\alpha}\right) \|X^{k+1} - X^k\|_F^2, \quad (11)$$

其中 Lipschitz 常数 $\eta > 0$ 。

证明 由算法1可得

$$\begin{aligned} H^{k+1} &= H^k - \alpha \nabla_H f(A^k, H^k, W^k), \\ W^{k+1} &= W^k - \alpha \nabla_W f(A^k, H^{k+1}, W^k), \\ A_\Omega^{k+1} &= (A^k - \alpha \nabla_A f(A^k, H^{k+1}, W^{k+1}))_\Omega. \end{aligned} \quad (12)$$

由定理2, 可得偏导数 $\nabla_H f, \nabla_W f, \nabla_A f$ 满足 Lipschitz 连续。根据引理1和式(12)可得

$$\begin{aligned} & f(A^k, H^{k+1}, W^k) \leq \\ & f(A^k, H^k, W^k) + \nabla_H f(A^k, H^k, W^k) \cdot (H^{k+1} - H^k) + \frac{\eta}{2} \|H^{k+1} - H^k\|_F^2 \leq \\ & f(A^k, H^k, W^k) - \frac{1}{\alpha} (H^{k+1} - H^k) \cdot (H^{k+1} - H^k) + \frac{\eta}{2} \|H^{k+1} - H^k\|_F^2 \leq \\ & f(A^k, H^k, W^k) + \left(\frac{\eta}{2} - \frac{1}{\alpha}\right) \|H^{k+1} - H^k\|_F^2, \end{aligned} \quad (13)$$

$$\begin{aligned} & f(A^k, H^{k+1}, W^{k+1}) \leq \\ & f(A^k, H^{k+1}, W^k) + \nabla_W f(A^k, H^{k+1}, W^k) \cdot (W^{k+1} - W^k) + \frac{\eta}{2} \|W^{k+1} - W^k\|_F^2 \leq \\ & f(A^k, H^{k+1}, W^k) + \left(\frac{\eta}{2} - \frac{1}{\alpha}\right) \|W^{k+1} - W^k\|_F^2 \end{aligned} \quad (14)$$

令

$$\begin{aligned} f_\Omega(A, H^{k+1}, W^{k+1}) &= \frac{1}{2} \|(H^{k+1} W^{k+1} - A)_\Omega\|_F^2, \\ f_\Omega(A, H^{k+1}, W^{k+1}) &= \frac{1}{2} \|(H^{k+1} W^{k+1} - A)_{\Omega^c}\|_F^2, \end{aligned}$$

则

$$f(A^{k+1}, H^{k+1}, W^{k+1}) = \frac{1}{2} \|H^{k+1}W^{k+1} - A^{k+1}\|_F^2 = \frac{1}{2} \|(H^{k+1}W^{k+1} - A^{k+1})_\Omega\|_F^2 + \frac{1}{2} \|(H^{k+1}W^{k+1} - A^{k+1})_{\Omega^c}\|_F^2 := f_\Omega(A^{k+1}, H^{k+1}, W^{k+1}) + f_{\Omega^c}(A^{k+1}, H^{k+1}, W^{k+1}).$$

注意到

$$A_\Omega^{k+1} = A_\Omega^k = D_\Omega,$$

则有

$$f_\Omega(A^{k+1}, H^{k+1}, W^{k+1}) = f_\Omega(A^k, H^{k+1}, W^{k+1}). \tag{15}$$

又因为

$$\begin{aligned} f_{\Omega^c}(A^{k+1}, H^{k+1}, W^{k+1}) &\leq f_{\Omega^c}(A^k, H^{k+1}, W^{k+1}) + (\nabla_{A^k} f(A^k, H^{k+1}, W^{k+1}))_{\Omega^c} \cdot (A^{k+1} - A^k)_{\Omega^c} + \frac{\eta}{2} \| (A^{k+1} - A^k)_{\Omega^c} \|_F^2 \leq \\ & f_{\Omega^c}(A^k, H^{k+1}, W^{k+1}) - \frac{1}{\alpha} (A^{k+1} - A^k)_{\Omega^c} \cdot (A^{k+1} - A^k)_{\Omega^c} + \frac{\eta}{2} \| (A^{k+1} - A^k)_{\Omega^c} \|_F^2 \leq \\ & f_{\Omega^c}(A^k, H^{k+1}, W^{k+1}) + \left(\frac{\eta}{2} - \frac{1}{\alpha} \right) \| (A^{k+1} - A^k)_{\Omega^c} \|_F^2, \end{aligned} \tag{16}$$

将式(15)和式(16)相加可得

$$\begin{aligned} f(A^{k+1}, H^{k+1}, W^{k+1}) &\leq f(A^k, H^{k+1}, W^{k+1}) + \left(\frac{\eta}{2} - \frac{1}{\alpha} \right) \| (A^{k+1} - A^k)_{\Omega^c} \|_F^2 = \\ & f(A^k, H^{k+1}, W^{k+1}) + \left(\frac{\eta}{2} - \frac{1}{\alpha} \right) \| A^{k+1} - A^k \|_{F^{\circ}}^2. \end{aligned} \tag{17}$$

令 $X^k = (A^k, H^k, W^k)$, 将式(13)、式(14)和式(17)相加可得

$$\begin{aligned} f(X^{k+1}) &\leq f(X^k) + \left(\frac{\eta}{2} - \frac{1}{\alpha} \right) [\| H^{k+1} - H^k \|_F^2 + \| W^{k+1} - W^k \|_F^2 + \| A^{k+1} - A^k \|_F^2]. \end{aligned}$$

即:

$$f(X^{k+1}) \leq f(X^k) + \left(\frac{\eta}{2} - \frac{1}{\alpha} \right) \| X^{k+1} - X^k \|_{F^{\circ}}^2$$

因此, 当 $\frac{\eta}{2} - \frac{1}{\alpha} < 0$ 时, 即证明式(11)。

定理5 记 $\{X^k = (A^k, H^k, W^k)\}_{k=0}^{+\infty} \subseteq \mathbb{H}_1$ 是算法1所生成的序列, 则有以下式子成立

$$\sum_{k=0}^{+\infty} \| X^k - X^{k+1} \|_F^2 < \infty,$$

从而有

$$\lim_{k \rightarrow +\infty} \| X^k - X^{k+1} \|_F^2 = 0,$$

即:

$$\begin{aligned} \lim_{k \rightarrow +\infty} \| A^k - A^{k+1} \|_F &= 0, \\ \lim_{k \rightarrow +\infty} \| H^k - H^{k+1} \|_F &= 0, \\ \lim_{k \rightarrow +\infty} \| W^k - W^{k+1} \|_F &= 0. \end{aligned}$$

证明 由定理4可得, 当 $\alpha \in \left(0, \frac{\eta}{2}\right)$ 时, 有

$$\begin{aligned} & \left(\frac{1}{\alpha} - \frac{\eta}{2}\right) \|X^k - X^{k+1}\|_F^2 \leq f(X^k) - f(X^{k+1}) \Leftrightarrow \\ & \frac{2 - \alpha\eta}{2\alpha} \|X^k - X^{k+1}\|_F^2 \leq f(X^k) - f(X^{k+1}) \Leftrightarrow \\ & \|X^k - X^{k+1}\|_F^2 \leq \frac{2\alpha}{2 - \alpha\eta} (f(X^k) - f(X^{k+1})). \end{aligned}$$

对上式 k 进行从 0 累加到 K 可得

$$\begin{aligned} \sum_{k=0}^K \|X^k - X^{k+1}\|_F^2 & \leq \sum_{k=0}^K \frac{2\alpha}{2 - \alpha\eta} (f(X^k) - f(X^{k+1})) \leq \frac{2\alpha}{2 - \alpha\eta} (f(X^0) - f(X^1) + \dots + f(X^K) - f(X^{K+1})) = \\ & \frac{2\alpha}{2 - \alpha\eta} (f(X^0) - f(X^{K+1})) \leq \frac{2\alpha}{2 - \alpha\eta} f(X^0) < \infty. \end{aligned}$$

令 $K \rightarrow +\infty$, 再由引理 2 可得

$$\lim_{k \rightarrow +\infty} \|X^k - X^{k+1}\|_F^2 = 0.$$

又由于

$$\begin{aligned} \|X^k - X^{k+1}\|_F^2 & = \|(A^k, H^k, W^k) - (A^{k+1}, H^{k+1}, W^{k+1})\|_F^2 = \\ & \|(A^k - A^{k+1}, H^k - H^{k+1}, W^k - W^{k+1})\|_F^2 = \\ & \|A^k - A^{k+1}\|_F^2 + \|H^k - H^{k+1}\|_F^2 + \|W^k - W^{k+1}\|_F^2, \end{aligned}$$

令上式 $k \rightarrow +\infty$ 时, 有

$$\lim_{k \rightarrow +\infty} \|A^k - A^{k+1}\|_F^2 = 0, \quad \lim_{k \rightarrow +\infty} \|H^k - H^{k+1}\|_F^2 = 0, \quad \lim_{k \rightarrow +\infty} \|W^k - W^{k+1}\|_F^2 = 0$$

成立, 即证明完成。

根据定理 5 可知, $\{X^k\}_{k=0}^{+\infty}$ 是有界的且 S 是闭集, 则存在子序列 $\{X^{k_j}\}$ 和 $X^\#$, 当 $j \rightarrow +\infty$ 时, 满足 $X^{k_j} \rightarrow X^\#$ 。由于 f 在 $S \times \mathbb{Q}^{m \times r} \times \mathbb{Q}^{r \times n}$ 上是连续的, 则当 $j \rightarrow +\infty$ 时, 有 $f(X^{k_j}) \rightarrow f(X^\#)$ 成立^[20]。

结合低秩四元数矩阵分解模型一阶最优性分析中的定理 2, 四元数梯度下降法基础上进行改进, 提出改进的四元数梯度下降法, 如算法 2 所示。

算法 2 (改进的四元数梯度下降法求解低秩四元数矩阵分解模型)

步 1: 给定 ϵ , 选取 $X^0 = (A^0, H^0, W^0) \in \mathbb{H}_1$ 。

步 2: 更新 $H^{k+1} = H^k - \alpha(H^k W^k - A^k)(W^k)^*$ 。

步 3: 更新 $W^{k+1} = W^k - \alpha(H^{k+1})^*(H^{k+1} W^k - A^k)$ 。

步 4: 更新 $A_\Omega^{k+1} = D_\Omega, A_\Omega^{k+1} = (H^{k+1} W^{k+1})_\Omega$ 。

步 5: 如果 $\|X^{k+1} - X^k\|_F < \epsilon$, 停止; 否则, 置 $k := k + 1$, 转步 2。

算法 1 与算法 2 的主要区别在于变量 A 的更新, 然而算法 2 利用最新得到的变量 H^{k+1} 和 W^{k+1} 来更新变量 A 。

下面分析算法 2 的收敛性。

定理 6 假设 $X^k, X^{k+1} \in \mathbb{H}_1$ 是由算法 2 生成, 当步长因子 $\alpha \in \left(0, \frac{2}{\eta}\right)$, 算法 2 收敛且有以下不等式成立

$$f(X^{k+1}) \leq f(X^k) + \lambda_0 \|X^{k+1} - X^k\|_F^2, \tag{18}$$

其中 $\lambda_0 = \max\left\{-\frac{1}{2}, -\left(\frac{1}{\alpha} - \frac{\eta}{2}\right)\right\}$, $\eta > 0$ 是 Lipschitz 常数。

证明 由定理 2, 可得偏导数 $\nabla_H f, \nabla_W f, \nabla_A f$ 满足 Lipschitz 连续, 则根据引理 1 和算法 2 可得

$$\begin{aligned} & f(A^k, H^{k+1}, W^k) \leq \\ & f(A^k, H^k, W^k) + \nabla_H f(A^k, H^k, W^k) \cdot (H^{k+1} - H^k) + \frac{\eta}{2} \|H^{k+1} - H^k\|_F^2 \leq \end{aligned}$$

$$\begin{aligned}
 f(A^k, H^k, W^k) - \frac{1}{\alpha} (H^{k+1} - H^k) \cdot (H^{k+1} - H^k) + \frac{\eta}{2} \|H^{k+1} - H^k\|_F^2 &\leq \\
 f(A^k, H^k, W^k) + \left(\frac{\eta}{2} - \frac{1}{\alpha}\right) \|H^{k+1} - H^k\|_F^2, & \tag{19}
 \end{aligned}$$

$$\begin{aligned}
 f(A^k, H^{k+1}, W^k) + \nabla_w f(A^k, H^{k+1}, W^k) \cdot (W^{k+1} - W^k) + \frac{\eta}{2} \|W^{k+1} - W^k\|_F^2 &\leq \\
 f(A^k, H^{k+1}, W^k) + \left(\frac{\eta}{2} - \frac{1}{\alpha}\right) \|W^{k+1} - W^k\|_{F^\circ}^2 & \tag{20}
 \end{aligned}$$

又由算法 2 可得

$$A_{\Omega}^{k+1} = A_{\Omega}^k = D_{\Omega},$$

则有

$$\begin{aligned}
 f(A^{k+1}, H^{k+1}, W^{k+1}) - f(A^k, H^{k+1}, W^{k+1}) &= \\
 \frac{1}{2} \|H^{k+1} W^{k+1} - A^{k+1}\|_F^2 - \frac{1}{2} \|H^{k+1} W^{k+1} - A^k\|_F^2 &= \\
 \frac{1}{2} \|(H^{k+1} W^{k+1} - A^{k+1})_{\Omega}\|_F^2 + \frac{1}{2} \|(H^{k+1} W^{k+1} - A^{k+1})_{\Omega^c}\|_F^2 - \frac{1}{2} \|(H^{k+1} W^{k+1} - A^k)_{\Omega}\|_F^2 - \\
 \frac{1}{2} \|(H^{k+1} W^{k+1} - A^k)_{\Omega^c}\|_F^2 &= \\
 \frac{1}{2} \|(H^{k+1} W^{k+1} - A^{k+1})_{\Omega}\|_F^2 - \frac{1}{2} \|(H^{k+1} W^{k+1} - A^k)_{\Omega}\|_F^2 &= \\
 -\frac{1}{2} \|(A^{k+1} - A^k)_{\Omega}\|_F^2 = -\frac{1}{2} \|A^{k+1} - A^k\|_{F^\circ}^2 & \tag{21}
 \end{aligned}$$

令 $X^k = (A^k, H^k, W^k)$, 将式 (19)、式 (20) 和式 (21) 相加可得

$$\begin{aligned}
 W^{k+1} &\leq f(A^k, H^k, W^k) - \left(\frac{1}{\alpha} - \frac{\eta}{2}\right) [\|H^{k+1} - H^k\|_F^2 + \|W^{k+1} - W^k\|_F^2] - \frac{1}{2} \|A^{k+1} - A^k\|_F^2 \leq \\
 f(A^k, H^k, W^k) + \lambda_0 [\|H^{k+1} - H^k\|_F^2 + \|W^{k+1} - W^k\|_F^2 + \|A^{k+1} - A^k\|_F^2], &
 \end{aligned}$$

这里 $\lambda_0 = \max\{-\frac{1}{2}, -(\frac{1}{\alpha} - \frac{\eta}{2})\}$, 即:

$$f(X^{k+1}) \leq f(X^k) + \lambda_0 \|X^{k+1} - X^k\|_F^2,$$

即证明式 (18)。

与定理 5 证明类似, 可得 $\|A^k - A^{k+1}\|_F \rightarrow 0, \|H^k - H^{k+1}\|_F \rightarrow 0, \|W^k - W^{k+1}\|_F \rightarrow 0$ 。

4 数值实验

在本节中, 为了证明所提出算法的有效性, 给出四元数梯度下降法和改进的四元数梯度下降法在图像恢复上的应用并去验证所提算法的可行性。

所有实验都是在 MATLAB (R2022a), Intel (R) Core (TM) i5-7200U CPU @2.50 GHz 的计算机下进行。设采样率 $\rho \in (0, 1)$ 为给定的数。指标集 Ω 在 MATLAB 中随机生成。图像大小统一设置为 $256 \times 256 \times 3$ 或 $135 \times 198 \times 3$ 。此外, 实验中的原始彩色图像是不含噪声。设定算法 1 和算法 2 的初始值 A^0 为

$$A_{ij}^0 = \begin{cases} D_{ij}, & (i, j) \in \Omega, \\ \text{随机数}, & (i, j) \notin \Omega, \end{cases}$$

这里 D 是取采用率为 ρ 的退化图像。设定步长因子 $\alpha = 0.0025, \rho = 0.7$ 和 $\epsilon = 10^{-4}$ 。设置其终止条件为 $\|X^k - X^{k+1}\|_F < 10^{-4}$, 其中 k 为迭代次数。设 $A^\#$ 和 A 分别表示恢复后的彩色图像和原始彩色图像, 定量指标选择峰值信噪比 (R_{PSN}) 和结构相识度 (M_{SSI})^[21], 分别定义为

$$R_{\text{PSN}}(A^\#, A) = 20 \log_{10} \frac{255}{\frac{1}{mn} \|A^\# - A\|_F},$$

$$M_{\text{SSI}}(A^\#, A) = \frac{(2\mu_A \mu_{A^\#} + c_1)(2\sigma_{AA^\#} + c_2)}{(\mu_A^2 + \mu_{A^\#}^2 + c_1)(\sigma_A^2 + \sigma_{A^\#}^2 + c_2)},$$

其中 m, n 为 A 的维数, $\mu_A, \mu_{A^\#}, \sigma_A, \sigma_{A^\#}$ 分别为 A 和 $A^\#$ 的均值和标准差, 常数 c_1 和 c_2 用来避免分母为零。

在实验中使用 MATLAB 内置函数“psnr”和“ssim”来计算 R_{PSN} 和 M_{SSI} 结果。 R_{PSN} 和 M_{SSI} 的值越高, 表示恢复图像的质量和效果越好, 但 R_{PSN} 值在 $(0, 1)$ 之间。利用本文提出四元数梯度下降法和改进的四元数梯度下降法对退化图像进行图像恢复任务。具体如图 1 所示。

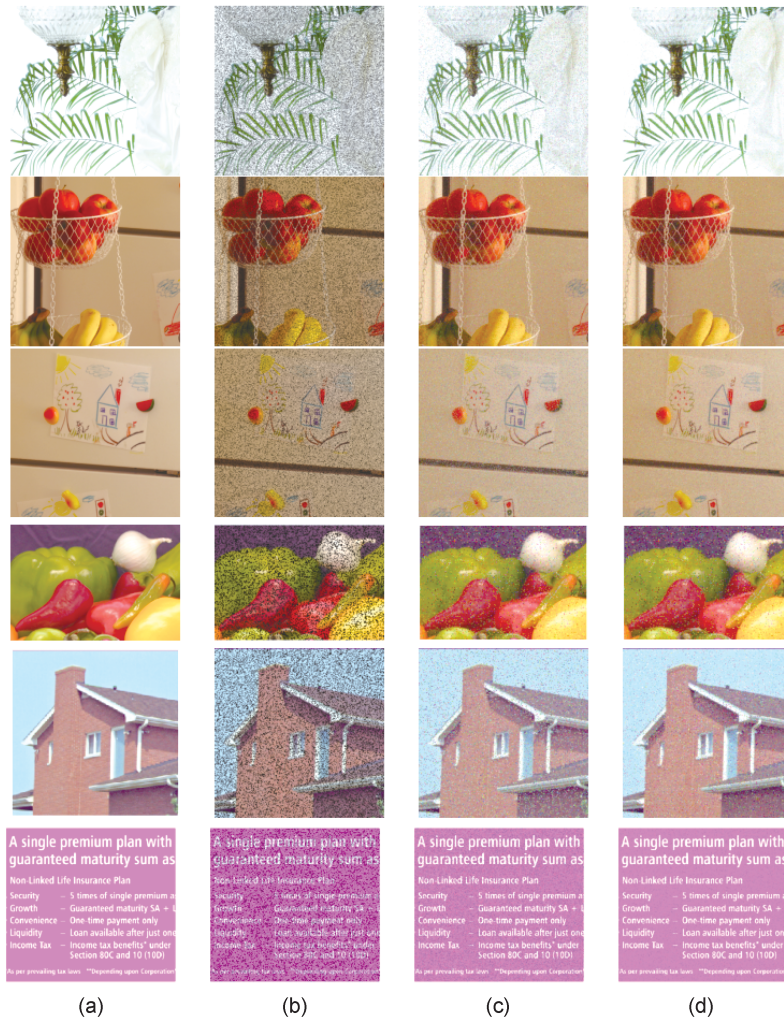


图 1 在图像上恢复具有视觉质量和数值结构的彩色图像
 (a)原始图像; (b)采样率 $\rho = 0.7$ 生成的退化图像; (c)利用四元数梯度下降法恢复的图像;
 (d)利用改进的四元数梯度下降法算法恢复的图像。

Fig. 1 Color images with visual quality and numerical structure are restored on the image
 (a) The original images; (b) Degraded images generated by sampling rate $\rho = 0.7$; (c) The images restored by quaternion gradient descent method; (d) The images restored by the improved quaternion gradient descent algorithm.

本节主要比较所提出的四元数梯度下降法和改进的四元数梯度下降法对测试图像整体恢复效果。当采样率 $\rho = 0.7$ 时, 恢复图像的 R_{PSN} 和 M_{SSI} 值见表 1。从表 1 可看出, 当步长因子 $\alpha = 0.0025$ 时, 改进的四元数梯度下降法的数值表现较好。图 1 给出了四元数梯度下降法和改进的四

元数梯度下降法恢复图像的效果图,以展示所提出方法的视觉质量。从视觉结果来看,当步长因子 $\alpha = 0.0025$ 时,两种算法均能够较好地保留颜色通道的颜色结构。然而,当步长因子过大时,会适当地破坏颜色结构。此外,改进的四元数梯度下降法在恢复彩色图像上效果较好,其可以更好地防止图像平滑,能够较好地恢复图像的细节结构。

表1 采样率 $\rho = 0.7$ 的 R_{PSN} 和 M_{SSI} 值

Table 1 The values of R_{PSN} and M_{SSI} for the sampling rate $\rho = 0.7$

Image serial number	Degraded image		Algorithm 1		Algorithm 2	
	R_{PSN}	M_{SSI}	R_{PSN}	M_{SSI}	R_{PSN}	M_{SSI}
01	7.28	0.152 6	13.13	0.339 7	19.11	0.796 6
02	10.03	0.332 7	17.32	0.682 6	20.52	0.821 6
03	10.05	0.302 7	18.68	0.679 0	19.24	0.803 5
04	9.85	0.208 1	16.32	0.481 8	18.57	0.695 4
05	11.85	0.308 1	16.71	0.638 0	20.51	0.796 0
06	6.71	0.489 5	13.90	0.815 9	21.24	0.947 5

从测试图像的数值结果可以看出,所提出的四元数梯度下降法和改进的四元数梯度下降法均能够有效地处理图像恢复问题,且改进的四元数梯度下降法在视觉质量和数值结果上有较好的表现。

5 结论

本文研究低秩四元数矩阵分解模型及其在彩色图像处理中的应用。利用了一个关键工具R-积,将优化理论推广到四元数域上,给出低秩四元数矩阵分解模型的一阶优化条件,同时提出了四元数梯度下降法和改进的四元数梯度下降法,并分析了其算法的收敛性。数值实验表明了所提出算法的鲁棒性。特别地,改进的四元数梯度下降法能够更好地保留彩色图像颜色的结构。然而,所提出的策略也有其局部性,步长因子太大时会破坏图像的结构。

参考文献:

- [1] JIA Z G, NG M K, SONG G J. Robust Quaternion Matrix Completion with Applications to Image Inpainting [J]. *Numer Linear Algebra Appl*, 2019, **26**(4): e2245. DOI: 10.1002/nla.2245.
- [2] KE Y F, MA C F, JIA Z G, et al. Quasi Non-negative Quaternion Matrix Factorization with Application to Color Face Recognition[J]. *J Sci Comput*, 2023, **95**(2): 38. DOI: 10.1007/s10915-023-02157-x.
- [3] CHEN Y N, QI L Q, ZHANG X Z, et al. A Low Rank Quaternion Decomposition Algorithm and Its Application in Color Image Inpainting[EB/OL]. (2020-09-25) [2023-09-10]. <https://arxiv.org/abs/2009.12203>.
- [4] WU P L, KOU K I, MIAO J F. Efficient Low-rank Quaternion Matrix Completion Under the Learnable Transforms for Color Image Recovery[J]. *Appl Math Lett*, 2024, **148**: 108880. DOI: 10.1016/j.aml.2023.108880.
- [5] SUN S N, DIAO Q K, XU D P, et al. Convex Quaternion Optimization for Signal Processing: Theory and Applications[J]. *IEEE Trans Signal Process*, 2023, **71**: 4106-4115. DOI: 10.1109/TSP.2023.3328053.
- [6] ZHENG M M, NI G Y. Approximation Strategy Based on the T-product for Third-order Quaternion Tensors with Application to Color Video Compression[J]. *Appl Math Lett*, 2023, **140**: 108587. DOI: 10.1016/j.aml.2023.108587.
- [7] CHEN Y Y, XIAO X L, ZHOU Y C. Low-rank Quaternion Approximation for Color Image Processing[J]. *IEEE Trans Image Process*, 2019: **29**: 1426-1439. DOI: 10.1109/TIP.2019.2941319.
- [8] XU D P, JAHANCHAH C, TOOK C C, et al. Enabling Quaternion Derivatives: The Generalized HR Calculus [J]. *R Soc Open Sci*, 2015, **2**(8): 150255. DOI: 10.1098/rsos.150255.
- [9] JIA Z G, JIN Q Y, NG M K, et al. Non-local Robust Quaternion Matrix Completion for Large-scale Color Image and Video Inpainting[J]. *IEEE Trans Image Process*,

- 2022, **31**: 3868–3883. DOI: 10.1109/TIP.2022.3176133.
- [10] MIAO J F, KOU K I. Color Image Recovery Using Low-rank Quaternion Matrix Completion Algorithm[J]. *IEEE Trans Image Process*, 2022, **31**: 190–201. DOI: 10.1109/TIP.2021.3128321.
- [11] MIAO J F, KOU K I. Quaternion-based Bilinear Factor Matrix Norm Minimization for Color Image Inpainting [J]. *IEEE Trans Signal Process*, 2020, **68**: 5617–5631. DOI: 10.1109/TSP.2020.3025519.
- [12] JIA Z G, WEI M S, LING S T. A New Structure-preserving Method for Quaternion Hermitian Eigenvalue Problems[J]. *J Comput Appl Math*, 2013, **239**: 12–24. DOI: 10.1016/j.cam.2012.09.018.
- [13] SANGWINE S J. Fourier Transforms of Colour Images Using Quaternion or Hyper Complex, Numbers[J]. *Electron Lett*, 1996, **32**(21): 1979. DOI: 10.1049/el:19961331.
- [14] ZHANG F Z. Quaternions and Matrices of Quaternions[J]. *Linear Algebra Appl*, 1997, **251**: 21–57. DOI: 10.1016/0024-3795(95)00543-9.
- [15] MANDIC D P, JAHANCHAH C, TOOK C C. A Quaternion Gradient Operator and Its Applications[J]. *IEEE Signal Process Lett*, 2011, **18**(1): 47–50. DOI: 10.1109/LSP.2010.2091126.
- [16] QI L Q, ZHANG X Z. Constrained Optimization of Real Functions in Quaternion Matrix Variables[J/OL]. (2020–09–29)[2023–09–10]. <https://arxiv.org/abs/2009.13884>.
- [17] WANG S, GONG Y. Low-complexity Iterative Methods for Complex-variable Matrix Optimization Problems in Frobenius Norm[J/OL]. (2023–03–14) [2023–09–10]. <https://doi.org/10.48550/arXiv.2303.0761>.
- [18] QI L Q, LUO Z Y, WANG Q W, *et al.* Quaternion Matrix Optimization: Motivation and Analysis[J]. *J Optim Theory Appl*, 2022, **193**(1/2/3): 621–648. DOI: 10.1007/s10957-021-01906-y.
- [19] BAUSCHKE H H, BOLTE J, TEBOULLE M. A Descent Lemma Beyond Lipschitz Gradient Continuity: First-order Methods Revisited and Applications[J]. *Math Oper Res*, 2017, **42**(2): 330–348. DOI: 10.1287/moor.2016.0817.
- [20] CHEN J F, WANG Q W, SONG G J, *et al.* Quaternion Matrix Factorization for Low-rank Quaternion Matrix Completion[J]. *Mathematics*, 2023, **11**(9): 2144. DOI: 10.3390/math11092144.
- [21] WANG Z, BOVIK A C, SHEIKH H R, *et al.* Image Quality Assessment: From Error Visibility to Structural Similarity[J]. *IEEE Trans Image Process*, 2004, **13**(4): 600–612. DOI: 10.1109/tip.2003.819861.