中国科学: 数学 2026 年 第 56 卷 第 4 期: 833～846
SCIENTIA SINICA Mathematica
《中国科学》杂志社
SCIENCE CHINA PRESS
论 文

# 基于 Tucker 秩张量补全的双层模型及算法

王川龙*, 钟林江, 闫喜红

智能优化计算与区块链技术山西省重点实验室, 晋中 030619
E-mail: clwang1964@163.com, 2403242166@qq.com, 415562629@qq.com

收稿日期: 2023-04-23; 接受日期: 2024-09-22; 网络出版日期: 2024-12-20; * 通信作者
国家自然科学基金 (批准号: 12371381) 和山西省科技创新团队项目 (批准号: 202204051002018) 资助项目

**摘要** 基于张量的 Tucker 秩, 本文提出张量按模展开矩阵核范数的极小极小、极小极大以及它们组合的优化模型. 首先研究这三种新模型的零空间性质和约束等距性质 (restricted isometry property, RIP). 然后采用增广 Lagrange 乘子方法框架, 设计求解张量补全的三种优化算法, 分析求解极小极小与极小极大组合优化模型算法的收敛性. 最后进行数值实验, 将本文提出的三种算法与求解传统模型的高精度低秩填充算法进行比较, 结果显示: 极小极小与极小极大组合的优化模型与传统的核范数优化模型精度基本相同, CPU (central processing unit) 时间明显减少.

**关键词** 张量补全 双层模型 收敛性 零空间性质 RIP 条件

**MSC (2020) 主题分类** 90C26, 90C47, 15A69

## 1 引言

张量是向量和矩阵的高阶推广, 已成为高维数据处理的一个重要工具. 随着信息科学的飞速发展, 大量高维的、复杂的数据需要被采集和处理, 张量及其应用研究越来越受到重视. 由于张量补全可以应用于图像恢复 [2,16]、数据挖掘 [1]、机器学习 [18] 和信号处理 [23] 等领域, 因此张量补全已成为张量研究的热点之一. 张量补全是通过极小化张量的秩来补全具有缺失数据的张量, 其优化模型为

$$\begin{aligned} &\min_{\mathcal{X}} \ \operatorname{rank}(\mathcal{X}) \\ &\text{s.t.} \ \ P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{T}), \end{aligned} \tag{1.1}$$

其中 $\mathcal{X}$ 是需要补全的张量, $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ 是 $N$- 阶采样张量, $\Omega$ 表示已知元素的下标集合, $P_\Omega(\cdot)$ 是集合 $\Omega$ 上的正交投影.

在矩阵补全问题中, 极小化秩的优化模型是非凸非连续的, 通常将其松弛为极小化矩阵核范数来近似求解 (参见文献 [3,4]). 在张量补全中, 张量的秩不唯一, 如 CP (CANDECOMP/PARAFAC) 秩、Tucker 秩、TT (tensor train) 秩和 TR (tensor ring) 秩等 (参见文献 [13,22,27,31]), 这就增加了张

**英文引用格式:** Wang C L, Zhong L J, Yan X H. Bi-level models and algorithms based on Tucker rank for tensor completion (in Chinese). Sci Sin Math, 2026, 56: 833–846, doi: 10.1360/SSM-2023-0095

© 2024《中国科学》杂志社 

量补全的复杂性. 本文仅研究 Tucker 秩张量补全问题. Liu 等 [20] 将矩阵核范数优化模型推广到张量补全上, 模型 (1.1) 松弛为下面的凸优化模型:

$$\begin{aligned} &\min_{\mathcal{X}} \ \sum_{i=1}^{N} \alpha_i \|\mathcal{X}_i\|_* \\ &\text{s.t.} \ \ P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{T}), \end{aligned} \tag{1.2}$$

其中 $\alpha_i \geqslant 0$ 且 $\sum_{i=1}^{N} \alpha_i = 1$, $\mathcal{X}_i$ 是张量 $\mathcal{X}$ 的模 -$i$ 展开所得到的矩阵, 还给出了张量补全的凸优化算法, 即简单低秩张量补全 (simple low rank tensor completion, SiLRTC) 算法、快速低秩张量补全 (fast low rank tensor completion, FaLRTC) 算法和高精度低秩张量补全 (high accuracy low rank tensor completion, HaLRTC) 算法. Gandy 等 [7] 将此模型应用到具有噪声的张量补全问题, 提出了带有噪声采样数据的张量核范数模型并用 Douglas-Rachford 分裂算法进行求解. 随着维数的增加, 核范数模型求得的解可能是次优的. Han 等 [12] 和 Xue 等 [30] 使用了张量截断核范数模型, 该模型目标函数为核范数减去最大的几个奇异值之和, 更好地逼近了低秩张量. Ng 等 [21] 提出了一种自适应加权张量核范数模型, 研究了每个模式矩阵核范数的自适应加权, 张量补全的更多其他理论和算法可参见文献 [5,6,9,10,15,17,24]. 但是为了极小化 $N$- 模式矩阵核范数的加权组合, $N$ 个模式矩阵 SVD (singular value decomposition) 都需要计算, 计算复杂度较大. 为了减少 SVD 的次数, 本文提出新的极小极小、极小极大以及二者组合核范数优化模型, 具体如下所示:

(1) 张量 $\mathcal{X}$ 按模 -$i$ 展开矩阵最小核范数的极小化模型 (极小极小)

$$\begin{aligned} &\min_{\mathcal{X}} \left\{ \min_{1 \leqslant i \leqslant N} \|\mathcal{X}_i\|_* \right\} \\ &\text{s.t.} \ \ P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{T}); \end{aligned} \tag{1.3}$$

(2) 张量 $\mathcal{X}$ 按模 -$i$ 展开矩阵最大核范数的极小化模型 (极小极大)

$$\begin{aligned} &\min_{\mathcal{X}} \left\{ \max_{1 \leqslant i \leqslant N} \|\mathcal{X}_i\|_* \right\} \\ &\text{s.t.} \ \ P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{T}); \end{aligned} \tag{1.4}$$

(3) 二者组合的优化模型

$$\begin{aligned} &\min_{\mathcal{X}} \left( \alpha_1 \min_{1 \leqslant i \leqslant N} \|\mathcal{X}_i\|_* + \alpha_2 \max_{1 \leqslant i \leqslant N} \|\mathcal{X}_i\|_* \right) \\ &\text{s.t.} \ \ P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{T}), \end{aligned} \tag{1.5}$$

其中 $\alpha_1 \geqslant 0$, $\alpha_2 \geqslant 0$ 且 $\alpha_1 + \alpha_2 = 1$.

极小极小优化模型是寻求核范数最小的模式矩阵, 但没有考虑其他模式矩阵的核范数. 极小极大优化模型是极小化核范数最大的模式矩阵, 从而降低了张量的所有模式矩阵的核范数. 而它们的组合则兼顾了二者的作用. 显然, 新模型不需要所有展开的模矩阵都进行 SVD. 因此, 这些新的优化模型求解时间一定会减少, 这对张量补全是有价值的.

本文余下内容的结构如下. 第 2 节介绍张量的预备知识、相关的概念和定义. 第 3 节详细设计求解 (1.3)–(1.5) 优化模型的增广 Lagrange 乘子迭代算法. 第 4 节给出 3 个优化模型的零空间性质和约束等距性质 (restricted isometry property, RIP) 的充分条件, 并建立非凸优化模型 (1.5) 相应算法的收敛性理论. 第 5 节进行数值实验, 通过随机产生的张量补全问题, 比较 3 个新的优化模型和传统优化模型 (1.2) 的实验效果. 第 6 节给出本文结论.

## 2 预备知识

本文 $\mathbb{R}^{I_1}$ 和 $\mathbb{R}^{I_1\times I_2}$ 分别表示 $I_1$ 维实向量空间及 $I_1\times I_2$ 实矩阵集合. $r(X)$ 代表矩阵 $X$ 的秩, $\sigma_i(X)$ 表示矩阵 $X$ 的第 $i$ 个最大奇异值. 矩阵 $X$ 的核范数记为 $\|X\|_*=\sum_i\sigma_i(X)$. 一个 $N$ 阶实张量表示为 $\mathcal{X}\in\mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$, 其元素为 $\mathcal{X}_{i_1i_2\cdots i_N}$ $(1\leqslant i_k\leqslant I_k,\ 1\leqslant k\leqslant N)$. 张量 $\mathcal{X}$ 和 $\mathcal{Y}$ 的内积定义为 $\langle\mathcal{X},\mathcal{Y}\rangle=\sum_{i_1,i_2,\ldots,i_N}\mathcal{X}_{i_1i_2\cdots i_N}\mathcal{Y}_{i_1i_2\cdots i_N}$. $\|\mathcal{X}\|_F=(\langle\mathcal{X},\mathcal{X}\rangle)^{\frac{1}{2}}=(\sum_{i_1,i_2,\ldots,i_N}|\mathcal{X}_{i_1i_2\cdots i_N}|^2)^{\frac{1}{2}}$ 是张量 $\mathcal{X}$ 的 $F$ 范数. $\bar{\Omega}$ 是 $\Omega$ 的补集. $P_\Omega(\cdot)$ 是集合 $\Omega$ 上的正交投影, 定义为

$$P_\Omega(\mathcal{X})(i_1,\ldots,i_N)=\begin{cases}\mathcal{X}_{i_1i_2\cdots i_N}, & \text{若 } (i_1,\ldots,i_N)\in\Omega,\\ 0, & \text{若 } (i_1,\ldots,i_N)\notin\Omega.\end{cases}$$

张量 $\mathcal{X}$ 沿模 -$k$ 展开的运算记为 $\mathcal{X}_k=\mathrm{unfold}_k(\mathcal{X})\in\mathbb{R}^{I_k\times\prod_{i\neq k}I_i}$, 其逆运算折叠定义为

$$\mathrm{Fold}_k(\mathcal{X}_k)=\mathcal{X}.$$

显然 $\|\mathcal{X}\|_F=\|\mathcal{X}_i\|_F$ $(1\leqslant i\leqslant N)$. 为第 4 节收敛性证明方便, 令

$$\mathcal{N}=\{1,2,\ldots,N\},\quad \mathrm{Null}(\Omega)=\{H\in\mathbb{R}^{I_1\times I_2\times\cdots\times I_N}\mid P_\Omega(H)=0\}.$$

**定义 2.1** (奇异值分解 (SVD)[8]) 对于一个秩为 $r$ 的矩阵 $Y\in\mathbb{R}^{I_1\times I_2}$, 存在两个矩阵 $U\in\mathbb{R}^{I_1\times r}$ 和 $V\in\mathbb{R}^{I_2\times r}$ 满足 $Y=U\Sigma_rV^{\mathrm{T}}$, 其中 $\Sigma_r=\mathrm{diag}(\sigma_1,\sigma_2,\ldots,\sigma_r)$, $\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_r>0$, $U$ 和 $V$ 分别是具有正交列的左奇异矩阵和右奇异矩阵.

**定义 2.2** (奇异值阈值算子[3]) 秩为 $r$ 矩阵 $Y\in\mathbb{R}^{I_1\times I_2}$ 的 SVD 为

$$Y=U\Sigma_rV^{\mathrm{T}},\quad \Sigma_r=\mathrm{diag}(\{\sigma_i\}_{1\leqslant i\leqslant r}),$$

其中 $U$, $V$ 和 $\Sigma_r$ 由上述定义给出. 对于任意 $\tau\geqslant 0$, 奇异值阈值算子 $D_\tau(Y)$ 被定义为

$$D_\tau(Y)=UD_\tau(\Sigma)V^{\mathrm{T}},\quad D_\tau(\Sigma)=\mathrm{diag}(\max\{\sigma_i-\tau,0\}).$$

对于任意 $\tau\geqslant 0$, 奇异值阈值算满足 $D_\tau(Y)=\arg\min_{X\in\mathbb{R}^{I_1\times I_2}}\tau\|X\|_*+\frac{1}{2}\|X-Y\|_F^2$.

## 3 算法设计

为了方便求解张量补全问题, 将极小极小优化模型 (1.3) 转化为等价的可分离优化模型:

$$\begin{aligned}&\min_{\mathcal{X},\mathcal{Z}}\Big\{\min_{1\leqslant i\leqslant N}\|\mathcal{Z}_i\|_*\Big\}\\ &\text{s.t. }\ \mathcal{Z}=\mathcal{X},\\ &\qquad P_\Omega(\mathcal{X})=P_\Omega(\mathcal{T}).\end{aligned}\tag{3.1}$$

则优化模型 (3.1) 的约束增广 Lagrange 函数优化模型为

$$\begin{aligned}&\min_{\mathcal{X},\mathcal{Z},\mathcal{Y}}\ L(\mathcal{X},\mathcal{Z},\mathcal{Y},\mu)=\min_{1\leqslant i\leqslant N}\|\mathcal{Z}_i\|_*+\langle\mathcal{Y},\mathcal{X}-\mathcal{Z}\rangle+\frac{\mu}{2}\|\mathcal{Z}-\mathcal{X}\|_F^2\\ &\text{s.t.}\quad P_\Omega(\mathcal{X})=P_\Omega(\mathcal{T}),\end{aligned}\tag{3.2}$$

其中 $\mu$ 是正参数.

对上述增广 Lagrange 函数优化模型使用交替方向乘子法可得如下的算法 1.

**算法 1** 求解极小极小优化模型的交替方向乘子算法 (alternating direction method of multipliers for minimizing minimal optimization model, MinADMM)

步 0 给定 $\Omega$, $\mathcal{T}$, $\rho > 1$, $\epsilon > 0$, $\mathcal{Y}^0$, $\mathcal{X}^0$, $\mu^0$, $k := 0$.

步 1 计算

$$\begin{aligned}
\mathcal{X}_{\min}^k &\in \left\{\mathcal{X}_i^k \,\middle|\, \|\mathcal{X}_i^k\|_* = \min_{1\leqslant j\leqslant N} \|\mathcal{X}_j^k\|_*\right\},\\
\mathcal{Z}^{k+1} &= \mathrm{Fold}_{\min}\left(D_{\frac{1}{\mu^k}}\left(\mathcal{X}_{\min}^{k+1} + \frac{1}{\mu^k}\mathcal{Y}_{\min}^k\right)\right),\\
\mathcal{X}^{k+1} &= P_{\widehat{\Omega}}\left(\mathcal{Z}^{k+1} - \frac{1}{\mu^k}\mathcal{Y}^k\right) + P_\Omega(\mathcal{T}),\\
\mathcal{Y}^{k+1} &= \mathcal{Y}^k - \mu^k(\mathcal{Z}^{k+1} - \mathcal{X}^{k+1}).
\end{aligned}$$

步 2

$$\mu^{k+1} = \rho\mu^k.$$

步 3 $\frac{\|\mathcal{X}^{k+1}-\mathcal{X}^k\|}{\|\mathcal{T}\|} \leqslant \epsilon$, 停止; 否则, 返回步 1.

类似地, 将极小极大优化模型 (1.4) 转化为下面等价的可分离优化模型:

$$\begin{aligned}
&\min_{\mathcal{X},\mathcal{Z}} \left\{\max_{1\leqslant i\leqslant N} \|\mathcal{Z}_i\|_*\right\}\\
&\text{s.t. } \mathcal{Z} = \mathcal{X},\\
&\qquad P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{T}),
\end{aligned} \tag{3.3}$$

其带约束的增广 Lagrange 函数优化模型及算法 2 如下:

$$\begin{aligned}
&\min_{\mathcal{X},\mathcal{Z},\mathcal{Y}} L(\mathcal{X},\mathcal{Z},\mathcal{Y},\mu) = \max_{1\leqslant i\leqslant N} \|\mathcal{Z}_i\|_* + \langle \mathcal{Y}, \mathcal{X}-\mathcal{Z}\rangle + \frac{\mu}{2}\|\mathcal{X}-\mathcal{Z}\|_F^2\\
&\text{s.t.} \quad P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{T}).
\end{aligned} \tag{3.4}$$

**算法 2** 求解极小极大优化模型的交替方向乘子算法 (alternating direction method of multipliers for minimizing the minimax optimization model, MaxADMM)

步 0 给定 $\Omega$, $\mathcal{T}$, $\rho > 1$, $\epsilon > 0$, $\mathcal{Y}^0$, $\mathcal{X}^0$, $\mu^0$, $k := 0$.

步 1 计算

$$\begin{aligned}
\mathcal{X}_{\max}^k &\in \left\{\mathcal{X}_i^k \,\middle|\, \|\mathcal{X}_i^k\|_* = \max_{1\leqslant j\leqslant N} \|\mathcal{X}_j^k\|_*\right\},\\
\mathcal{Z}^{k+1} &= \mathrm{Fold}_{\max}\left(D_{\frac{1}{\mu^k}}\left(\mathcal{X}_{\max}^k + \frac{1}{\mu^k}\mathcal{Y}_{\max}^k\right)\right),\\
\mathcal{X}^{k+1} &= P_{\widehat{\Omega}}\left(\mathcal{Z}^{k+1} - \frac{1}{\mu^k}\mathcal{Y}^k\right) + P_\Omega(\mathcal{T}),\\
\mathcal{Y}^{k+1} &= \mathcal{Y}^k - \mu^k(\mathcal{Z}^{k+1} - \mathcal{X}^{k+1}).
\end{aligned}$$

步 2

$$\mu^{k+1} = \rho\mu^k.$$

步 3 $\frac{\|\mathcal{X}^{k+1}-\mathcal{X}^k\|}{\|\mathcal{T}\|} \leqslant \epsilon$, 停止; 否则, 返回步 1.

对于优化模型 (1.5), 引入变量 $\mathcal{Z}_1$ 和 $\mathcal{Z}_2$, 构建如下的可分离优化模型:

$$\begin{aligned} \min_{\mathcal{X},\mathcal{Z}_1,\mathcal{Z}_2} \ & \frac{1}{2}\Big(\min_{1\leqslant i\leqslant N}\|\mathcal{Z}_{1,i}\|_* + \max_{1\leqslant i\leqslant N}\|\mathcal{Z}_{2,i}\|_*\Big) \\ \text{s.t.} \ & P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{T}), \\ & \mathcal{X} = \mathcal{Z}_1, \\ & \mathcal{X} = \mathcal{Z}_2. \end{aligned} \tag{3.5}$$

不失一般性, 这里参数 $\alpha_1 = \alpha_2 = \frac{1}{2}$. 相应的增广 Lagrange 函数优化模型及算法 3 如下:

$$\begin{aligned} \min_{\mathcal{X},\mathcal{Z}_1,\mathcal{Z}_2,\mathcal{Y}_1,\mathcal{Y}_2} \ L(\mathcal{X},\mathcal{Z}_1,\mathcal{Z}_2,\mathcal{Y}_1,\mathcal{Y}_2,\mu) = & \frac{1}{2}\Big(\Big\{\min_{1\leqslant i\leqslant N}\|\mathcal{Z}_{1,i}\|_* + \max_{1\leqslant i\leqslant N}\|\mathcal{Z}_{2,i}\|_*\Big\}\Big) \\ & + \langle \mathcal{Y}_1, \mathcal{X}-\mathcal{Z}_1\rangle + \langle \mathcal{Y}_2, \mathcal{X}-\mathcal{Z}_2\rangle \\ & + \frac{\mu}{2}(\|\mathcal{Z}_1-\mathcal{X}\|_F^2 + \|\mathcal{Z}_2-\mathcal{X}\|_F^2) \\ \text{s.t.} \ P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{T}). & \end{aligned} \tag{3.6}$$

**算法 3** 求解极小极小和极小极大组合优化模型的交替方向乘子算法 (alternating direction method of multipliers for minimizing the minimal and maximal optimization model, MinMaxADMM)

步 0 给定 $\Omega$, $\mathcal{T}$, $\rho > 1$, $\epsilon > 0$, $\mathcal{Y}_1^0$, $\mathcal{Y}_2^0$, $\mathcal{X}^0$, $\mu^0$, $k := 0$.

步 1 计算

$$\begin{aligned} \mathcal{X}_{\min}^k &\in \Big\{\mathcal{X}_i^k \,\Big|\, \|\mathcal{X}_i^k\|_* = \min_{1\leqslant j\leqslant N}\|\mathcal{X}_j^k\|_*\Big\}, \\ \mathcal{X}_{\max}^k &\in \Big\{\mathcal{X}_i^k \,\Big|\, \|\mathcal{X}_i^k\|_* = \max_{1\leqslant j\leqslant N}\|\mathcal{X}_j^k\|_*\Big\}, \\ \mathcal{Z}_1^{k+1} &= \mathrm{Fold}_{\min}\Big(D_{\frac{1}{2\mu^k}}\Big(\mathcal{X}_{\min}^k + \frac{1}{\mu^k}\mathcal{Y}_{1,min}^k\Big)\Big), \\ \mathcal{Z}_2^{k+1} &= \mathrm{Fold}_{\max}\Big(D_{\frac{1}{2\mu^k}}\Big(\mathcal{X}_{\max}^k + \frac{1}{\mu^k}\mathcal{Y}_{2,max}^k\Big)\Big), \\ \mathcal{X}^{k+1} &= \frac{1}{2}P_{\bar{\Omega}}\Big(\mathcal{Z}_1^{k+1} + \mathcal{Z}_2^{k+1} - \frac{1}{\mu^k}\mathcal{Y}_1^k - \frac{1}{\mu^k}\mathcal{Y}_2^k\Big) + P_\Omega(\mathcal{T}), \\ \mathcal{Y}_1^{k+1} &= \mathcal{Y}_1^k - \mu^k(\mathcal{Z}_1^{k+1} - \mathcal{X}^{k+1}), \\ \mathcal{Y}_2^{k+1} &= \mathcal{Y}_2^k - \mu^k(\mathcal{Z}_2^{k+1} - \mathcal{X}^{k+1}). \end{aligned}$$

步 2

$$\mu^{k+1} = \rho\mu^k.$$

步 3 $\frac{\|\mathcal{X}^{k+1}-\mathcal{X}^k\|}{\|\mathcal{T}\|} \leqslant \epsilon$, 停止; 否则, 返回步 1.

**注 3.1** 本文算法 1–3 均采用如下方法计算核范数:

$$\|\mathcal{X}_i^k\|_* = \mathrm{tr}(\mathcal{X}_i^k(\mathcal{X}_i^k)^{\mathrm{T}})^{\frac{1}{2}}.$$

一般地, 由于 $I_i < \prod_{j\neq i} I_j$, 所以该算法减少了 SVD 的计算复杂度, 提高了整个算法的效率.

**注 3.2** 显然, 极小极小模型优化不是凸优化模型, 算法 1 和 3 的收敛性证明与算法 2 不同 (凸优化算法收敛性证明可参见文献 [19]), 又因为算法 1 是算法 3 的一个特殊情形. 因此, 本文仅详细分析算法 3 的收敛性.

## 4 收敛性分析

本节讨论 3 种新模型解的存在性和算法的收敛性质, 给出 3 种新模型的零空间性质和 RIP 条件, 并建立算法 3 的收敛性理论.

首先引入下列引理, 其具体证明详见文献 [14, 定理 7.4.15 和 7.4.51].

**引理 4.1** 设 $X$ 和 $Y$ 是同型矩阵, 则

$$\sum_{i=1}^{m}|\sigma_i(X)-\sigma_i(Y)|\leqslant\|X-Y\|_*,\quad \sum_{i=1}^{m}(\sigma_i(X)-\sigma_i(Y))^2\leqslant\|X-Y\|_F^2.$$

接下来给出 3 个模型解的性质及 RIP 条件.

**定理 4.1** 设张量 $\mathcal{X}^*$ 具有 Tucker 秩 -$(r_1,r_2,\ldots,r_N)$, 且 $\mathcal{X}^*$ 是优化问题 (1.3) 的可行解. 若对任何非零张量 $\mathcal{H}\in\mathrm{Null}(\Omega)$ 都有

$$\sum_{j=r_i+1}^{I_i}\sigma_j(\mathcal{H}_i)\geqslant\sum_{j=1}^{r_i}\sigma_j(\mathcal{H}_i),\quad \forall\, i\in\mathcal{N},\tag{4.1}$$

则 $\mathcal{X}^*$ 是优化问题 (1.3) 的一个最优解.

**证明** 对于任意 $i\in\mathcal{N}$ 和非零张量 $\mathcal{H}\in\mathrm{Null}(\Omega)$, 如下式子成立:

$$\begin{aligned}\|(\mathcal{X}^*+\mathcal{H})_i\|_*&\geqslant\sum_{j=1}^{I_i}|\sigma_j(\mathcal{X}_i^*)-\sigma_j(\mathcal{H}_i)|\\&=\sum_{j=1}^{r_i}|\sigma_j(\mathcal{X}_i^*)-\sigma_j(\mathcal{H}_i)|+\sum_{j=r_i+1}^{I_i}\sigma_j(\mathcal{H}_i)\\&\geqslant\sum_{j=1}^{r_i}(\sigma_j(\mathcal{X}_i^*)-\sigma_j(\mathcal{H}_i))+\sum_{j=r_i+1}^{I_i}\sigma_j(\mathcal{H}_i)\\&=\|\mathcal{X}_i^*\|_*+\sum_{j=r_i+1}^{I_i}\sigma_j(\mathcal{H}_i)-\sum_{j=1}^{r_i}\sigma_j(\mathcal{H}_i)\\&\geqslant\|\mathcal{X}_{\min}^*\|_*+\sum_{j=r_i+1}^{I_i}\sigma_j(\mathcal{H}_i)-\sum_{j=1}^{r_i}\sigma_j(\mathcal{H}_i).\end{aligned}$$

第一个不等式来自引理 4.1, 第一个等式来自 $r(\mathcal{X}_i^*)=r_i$, 最后一个不等式来自 $\mathcal{X}_{\min}^*$ 的定义 (参见算法 1). 因此, 定理 4.1 结论成立. □

同样地, 对于优化模型 (1.4) 和 (1.5), 有下面类似的结论.

**定理 4.2** 设张量 $\mathcal{X}^*$ 具有 Tucker 秩 -$(r_1,r_2,\ldots,r_N)$, 且 $\mathcal{X}^*$ 是优化问题 (1.4) 的可行解. 若对任何非零张量 $\mathcal{H}\in\mathrm{Null}(\Omega)$ 都有

$$\sum_{j=r_i+1}^{I_i}\sigma_j(\mathcal{H}_i)\geqslant\sum_{j=1}^{r_i}\sigma_j(\mathcal{H}_i),\quad \text{存在某个 } i\in\mathcal{N},\tag{4.2}$$

则 $\mathcal{X}^*$ 是优化问题 (1.4) 的一个最优解.

**定理 4.3** 设张量 $\mathcal{X}^*$ 具有 Tucker 秩 -$(r_1, r_2, \ldots, r_N)$, 且 $\mathcal{X}^*$ 是优化问题 (1.5) 的可行解. 若对任何非零张量 $\mathcal{H} \in \mathrm{Null}(\Omega)$ 都有

$$\sum_{j=r_i+1}^{I_i} \sigma_j(\mathcal{H}_i) \geqslant \sum_{j=1}^{r_i} \sigma_j(\mathcal{H}_i), \quad \forall\, i \in \mathcal{N}, \tag{4.3}$$

则 $\mathcal{X}^*$ 是优化问题 (1.5) 的一个最优解.

**定理 4.4** (RIP 条件) 设 $\mathcal{X}^*$ 是一个秩为 $(r_1, r_2, \ldots, r_N)$ 或更低秩的张量. 如果 RIP 条件满足

$$(1-\delta_{(I_1,\ldots,2r_n,\ldots,I_N)})\|\mathcal{X}^*\|_F^2 \leqslant \|P_\Omega(\mathcal{X}^*)\|_F^2 \leqslant (1+\delta_{(I_1,\ldots,2r_n,\ldots,I_N)})\|\mathcal{X}^*\|_F^2 \tag{4.4}$$

及 $\delta_{(I_1,\ldots,2r_n,\ldots,I_N)} < 0.4931$, 则优化问题 (1.3)–(1.5) 能从测量算子 $P_\Omega(\mathcal{T}) = P_\Omega(\mathcal{X}^*)$ 中精确地恢复 $\mathcal{X}^*$.

**注 4.1** 定理 4.4 的证明方法与文献 [25, 定理 8] 类似, 这里省略证明过程. 对于凸优化模型 (1.4), 定理 4.2 成立仅需要一个展开模式矩阵满足条件 (4.2), 其余两个非凸模型则需要所有展开的模式矩阵满足 (4.1) 或 (4.3). 另外, 由于 $\forall\, i \in \{1, 2, \ldots, N\}$, $\|\mathcal{X}_i^*\|_F = \|\mathcal{X}^*\|_F$, 所以 3 个模型的 RIP 条件相同, 正是文献 [25, 定理 8] 给出的 (8) 式.

下面分析算法 3 的收敛性.

**引理 4.2** 设矩阵 $X, Y, Z \in \mathbb{R}^{m\times n}$ $(m \geqslant n)$ 满足 $Y = X + \epsilon Z$, $\epsilon > 0$, $Y$ 的 SVD 为 $Y = U\Sigma V^{\mathrm{T}}$. 令 $X_\tau = X - \tau UV^{\mathrm{T}}$, $0 < \tau \leqslant \sigma_n(X)$, 则存在一个 $\epsilon_0 > 0$, 使得对于任意 $\epsilon \leqslant \epsilon_0$ 都有 $\|X_\tau\|_F \leqslant \|X\|_F$.

**证明** 对于矩阵 $Y$, 由于奇异值是连续的, 所以奇异向量张成的不变子空间也是连续的, 即扰动后的奇异向量不变子空间与原奇异向量不变子空间的夹角是连续的 (参见文献 [11,26,28,29]). 这里就是矩阵 $Y$ 和 $X$ 的奇异向量不变子空间夹角 $\theta(\epsilon) = (\theta_1(\epsilon), \ldots, \theta_d(\epsilon))$ ($d$ 是矩阵 $Y$ 的奇异向量不变子空间的个数) 满足 $\lim_{\epsilon\to 0}\theta(\epsilon) = 0$. 这样在奇异向量不变子空间中一定存在奇异向量组成的矩阵 $U_X$, $V_X$ 和矩阵 $\tilde{U}$, $\tilde{V}$ 满足 $U - U_X = \epsilon_1\tilde{U}$, $V - V_X = \epsilon_2\tilde{V}$ 和 $X = U_X\Sigma_X V_X^{\mathrm{T}}$, 其中, $\|\tilde{U}\|_F = \|\tilde{V}\|_F = 1$, $\epsilon_1, \epsilon_2 > 0$. 从而

$$\begin{aligned}
U_X^{\mathrm{T}} X_\tau V_X &= \Sigma_X - \tau U_X^{\mathrm{T}} U V^{\mathrm{T}} V_X \\
&= \Sigma_X - \tau U_X^{\mathrm{T}}(U_X + \epsilon_1\tilde{U})(V_X + \epsilon_2\tilde{V})^{\mathrm{T}} V_X \\
&= \Sigma_X - \tau(I + \epsilon_1 U_X^{\mathrm{T}}\tilde{U})(I + \epsilon_2\tilde{V}^{\mathrm{T}} V_X) \\
&= \Sigma_X - \tau I - \tau(\epsilon_1 U_X^{\mathrm{T}}\tilde{U} + \epsilon_2\tilde{V}^{\mathrm{T}} V_X + \epsilon_1\epsilon_2 U_X^{\mathrm{T}}\tilde{U}\tilde{V}^{\mathrm{T}} V_X).
\end{aligned}$$

由此推导出

$$\|X_\tau\|_F \leqslant \|\Sigma_X - \tau I\|_F + \tau(\|\epsilon_1 U_X^{\mathrm{T}}\tilde{U}\|_F + \|\epsilon_2\tilde{V}^{\mathrm{T}} V_X\|_F + \|\epsilon_1\epsilon_2 U_X^{\mathrm{T}}\tilde{U}\tilde{V}^{\mathrm{T}} V_X\|_F).$$

经过简单计算, 可得

$$\|\Sigma_X - \tau I\|_F \leqslant \|\Sigma_X\|_F - \frac{\tau\|\Sigma_X\|_*}{\|\Sigma_X\|_F} + \frac{n\tau^2}{2\|\Sigma_X\|_F}.$$

由于 $U_X$, $V_X$, $U$ 和 $V$ 是正交矩阵, 所以

$$\|U_X^{\mathrm{T}}\tilde{U}\|_F \leqslant \sqrt{n}, \quad \|\tilde{V}^{\mathrm{T}} V_X\|_F \leqslant \sqrt{n}, \quad \|U_X^{\mathrm{T}}\tilde{U}\tilde{V}^{\mathrm{T}} V_X\|_F \leqslant n.$$

从而得到

$$\|X_\tau\|_F \leqslant \|X\|_F - \frac{\tau\|X\|_*}{\|X\|_F} + \frac{n\tau^2}{2\|X\|_F} + \tau\max\{\epsilon_1,\epsilon_2\}(2+\max\{\epsilon_1,\epsilon_2\}\sqrt{n})\sqrt{n}.$$

由连续性可知 $\lim_{\epsilon\to 0}\epsilon_1 = \lim_{\epsilon\to 0}\epsilon_2 = 0$. 由此推导出, 存在 $\epsilon_0 > 0$, 当 $\epsilon_1,\epsilon_2 \leqslant \epsilon_0$ 时, 有

$$\frac{n\tau}{2\|X\|_F} + \max\{\epsilon_1,\epsilon_2\}(2+\max\{\epsilon_1,\epsilon_2\}\sqrt{n})\sqrt{n} \leqslant \frac{\|X\|_*}{\|X\|_F}.$$

从而得到 $\|X_\tau\|_F \leqslant \|X\|_F$, 引理得证. □

**假设 4.1** 对于集合 $\Omega$ 以及任意的 $\forall\,\mathcal{X}$ 和 $\mathcal{Y}$, $\|\mathcal{X}\|_F \leqslant \|\mathcal{Y}\|_F$ 隐含 $\|P_{\bar{\Omega}}(\mathcal{X})\|_F \leqslant \|P_{\bar{\Omega}}(\mathcal{Y})\|_F$ 成立.

**引理 4.3** 若假设 4.1 成立, 且 $\mu^k$ 充分大, 则算法 3 生成的序列 $\{\mathcal{X}^k, \mathcal{Z}_1^k, \mathcal{Z}_2^k\}$ 是有界的.

**证明** 从算法 3 计算 $\mathcal{Z}_1^{k+1}$ 和 $\mathcal{Z}_2^{k+1}$ 可以得出, 阈值 $\tau = \frac{1}{2\mu^k}$. 由子问题的最优性条件可得下面的等式:

$$\mathcal{Z}_1^{k+1} = \mathcal{X}^k + \frac{1}{\mu^k}\mathcal{Y}_1^k - \frac{1}{2\mu^k}\mathcal{B}^{k+1}, \tag{4.5}$$

$$\mathcal{Z}_2^{k+1} = \mathcal{X}^k + \frac{1}{\mu^k}\mathcal{Y}_2^k - \frac{1}{2\mu^k}\mathcal{C}^{k+1}, \tag{4.6}$$

$$\mathcal{X}^{k+1} = \frac{1}{2}P_{\bar{\Omega}}\left(\mathcal{Z}_1^{k+1} + \mathcal{Z}_2^{k+1} - \frac{1}{\mu^k}(\mathcal{Y}_1^k + \mathcal{Y}_2^k)\right) + P_\Omega(\mathcal{T}), \tag{4.7}$$

其中 $\mathcal{B}^{k+1} = \mathrm{Fold}_{\min}(\partial\|\mathcal{Z}_{1,\min}^{k+1}\|_*)$, $\mathcal{C}^{k+1} = \mathrm{Fold}_{\max}(\partial\|\mathcal{Z}_{2,\max}^{k+1}\|_*)$. 将 (4.5) 和 (4.6) 代入 (4.7), 可推导出

$$\mathcal{X}^{k+1} = P_{\bar{\Omega}}\left(\mathcal{X}^k - \frac{1}{4\mu^k}\mathcal{B}^{k+1} - \frac{1}{4\mu^k}\mathcal{C}^{k+1}\right) + P_\Omega(\mathcal{T}). \tag{4.8}$$

令 $\widetilde{\mathcal{X}}^{k+1} = \mathcal{X}^k - \frac{1}{4\mu^k}\mathcal{B}^{k+1} - \frac{1}{4\mu^k}\mathcal{C}^{k+1}$, 则有

$$\widetilde{\mathcal{X}}^{k+1} = \frac{1}{2}\left(\mathcal{X}^k - \frac{1}{2\mu^k}\mathcal{B}^{k+1} + \mathcal{X}^k - \frac{1}{2\mu^k}\mathcal{C}^{k+1}\right).$$

根据引理 4.2, 当 $\mu^k$ 充分大时, 有 $\|\mathcal{X}^k - \frac{1}{2\mu^k}\mathcal{B}^{k+1}\|_F \leqslant \|\mathcal{X}^k\|_F$ 和 $\|\mathcal{X}^k - \frac{1}{2\mu^k}\mathcal{C}^{k+1}\|_F \leqslant \|\mathcal{X}^k\|_F$ 成立, 从而 $\|\widetilde{\mathcal{X}}^{k+1}\|_F \leqslant \|\mathcal{X}^k\|_F$. 结合假设 4.1, 可得 $\|P_{\bar{\Omega}}(\widetilde{\mathcal{X}}^{k+1})\|_F \leqslant \|P_{\bar{\Omega}}(\mathcal{X}^k)\|_F$. 由于 $P_{\bar{\Omega}}(\mathcal{X}^{k+1}) = P_{\bar{\Omega}}(\widetilde{\mathcal{X}}^{k+1})$ 及 $P_\Omega(\mathcal{X}^{k+1}) = P_\Omega(\mathcal{X}^k)$, 因此 $\|\mathcal{X}^{k+1}\|_F \leqslant \|\mathcal{X}^k\|_F$. 对于 $\mathcal{Y}_1^k$ 和 $\mathcal{Y}_2^k$, 由算法 3 可知

$$\mathcal{Y}_1^k \in \frac{1}{2}\mathrm{Fold}(\partial\|\mathcal{Z}_{1,\min}^k\|_*), \quad \mathcal{Y}_2^k \in \frac{1}{2}\mathrm{Fold}(\partial\|\mathcal{Z}_{2,\max}^k\|_*).$$

从而, 如下不等式成立:

$$\|\mathcal{Z}_1^{k+1}\|_F \leqslant \|\mathcal{X}^k\|_F + \frac{1}{\mu^k}\|\mathcal{Y}_1^k\|_F \leqslant \|\mathcal{X}^0\|_F + \frac{1}{2\mu^k}r(\mathcal{Z}_{1,\min}^k),$$
$$\|\mathcal{Z}_2^{k+1}\|_F \leqslant \|\mathcal{X}^k\|_F + \frac{1}{\mu^k}\|\mathcal{Y}_2^k\|_F \leqslant \|\mathcal{X}^0\|_F + \frac{1}{2\mu^k}r(\mathcal{Z}_{2,\max}^k).$$

这样就推导出, 当 $\mu^k$ 充分大时, $\{\mathcal{X}^k, \mathcal{Z}_1^k, \mathcal{Z}_2^k\}$ 是有界的. □

**定理 4.5** 设假设 4.1 成立. 算法 3 生成的序列 $\{\mathcal{X}^k, \mathcal{Z}_1^k, \mathcal{Z}_2^k\}$ 满足

(i) $\|\mathcal{Z}_1^{k+1} - \mathcal{Z}_1^k\|_F \to 0$, $\|\mathcal{Z}_2^{k+1} - \mathcal{Z}_2^k\|_F \to 0$, $\|\mathcal{X}^{k+1} - \mathcal{X}^k\|_F \to 0$;

(ii) $\lim_{k\to\infty}\mathcal{Z}_1^k = \lim_{k\to\infty}\mathcal{Z}_2^k = \lim_{k\to\infty}\mathcal{X}^k$;

(iii) 当 $k\to\infty$ 时, 存在一个子序列 $\{\mathcal{X}^{k_j}, \mathcal{Z}_1^{k_j}, \mathcal{Z}_2^{k_j}\}$ 收敛到优化问题 (3.5) 的一个 KKT (Karush-Kuhn-Tucker) 点, 也是优化问题 (1.5) 的一个 KKT 点.

**证明** (i) 由 (4.8) 以及 $\mathcal{B}^{k+1}$ 和 $\mathcal{C}^{k+1}$ 的定义可得 $\mathcal{X}^{k+1}-\mathcal{X}^k=\frac{1}{4\mu^k}P_{\bar{\Omega}}(\mathcal{B}^{k+1}+\mathcal{C}^{k+1})$ 且 $\|\mathcal{B}^{k+1}\|_2=\|\mathcal{C}^{k+1}\|_2=1$. 由算法可知, 当 $k\to\infty$ 时, $\mu^k\to\infty$, 从而 $\|\mathcal{X}^{k+1}-\mathcal{X}^k\|_F\leqslant\frac{1}{4\mu^k}\|\mathcal{B}^{k+1}+\mathcal{C}^{k+1}\|_F\to 0$. 另外, 依据 (4.5) 和 (4.8), 有

$$\begin{aligned}\|\mathcal{Z}_1^{k+1}-\mathcal{Z}_1^k\|_F&=\|\mathcal{Z}_1^{k+1}-\mathcal{X}^k+\mathcal{X}^k-\mathcal{X}^{k-1}+\mathcal{X}^{k-1}-\mathcal{Z}_1^k\|_F\\&\leqslant\|\mathcal{Z}_1^{k+1}-\mathcal{X}^k\|_F+\|\mathcal{X}^k-\mathcal{X}^{k-1}\|_F+\|\mathcal{X}^{k-1}-\mathcal{Z}_1^k\|_F\\&\leqslant\frac{1}{2\mu^k}\|2\mathcal{Y}_1^k-\mathcal{B}^{k+1}\|_F+\frac{1}{4\mu^{k-1}}\|\mathcal{B}^k+\mathcal{C}^k\|_F+\frac{1}{2\mu^{k-1}}\|\mathcal{B}^k-2\mathcal{Y}_1^{k-1}\|_F.\end{aligned}$$

由于 $\mathcal{B}^{k+1}$, $\mathcal{Y}_1^k$, $\mathcal{B}^k$, $\mathcal{C}^k$ 和 $\mathcal{Y}_1^{k-1}$ 有界, 所以当 $k\to\infty$ 时, $\|\mathcal{Z}_1^{k+1}-\mathcal{Z}_1^k\|_F\to 0$. 同理可得, 当 $k\to\infty$ 时, $\|\mathcal{Z}_2^{k+1}-\mathcal{Z}_2^k\|_F\to 0$.

(ii) 由 (4.5) 和 (4.6) 可推导出不等式 $\|\mathcal{Z}_1^{k+1}-\mathcal{X}^k\|_F\leqslant\frac{1}{2\mu^k}\|2\mathcal{Y}_1^k-\mathcal{B}^{k+1}\|_F$ 和 $\|\mathcal{Z}_2^{k+1}-\mathcal{X}^k\|_F\leqslant\frac{1}{2\mu^k}\|2\mathcal{Y}_2^k-\mathcal{C}^{k+1}\|_F$, 从而可得

$$\lim_{k\to\infty}\mathcal{Z}_1^k=\lim_{k\to\infty}\mathcal{Z}_2^k=\lim_{k\to\infty}\mathcal{X}^k.$$

(iii) 由引理 4.3 可知, 若 $k$ 足够大, 则 $\{\mathcal{X}^k,\mathcal{Z}_1^k,\mathcal{Z}_2^k\}$ 是有界的. 因此存在一个子序列 $\{\mathcal{X}^{k_j},\mathcal{Z}_1^{k_j},\mathcal{Z}_2^{k_j}\}$ 收敛于点 $\{\mathcal{X}^*,\mathcal{Z}_1^*,\mathcal{Z}_2^*\}$, 而且次梯度 [28] 满足

$$\begin{aligned}&0\in\frac{\partial L(\mathcal{X}^{k_j},\mathcal{Z}_1^{k_j+1},\mathcal{Z}_2^{k_j},\mathcal{Y}_1^{k_j},\mathcal{Y}_2^{k_j},\mu^{k_j})}{\partial\mathcal{Z}_1},\\&0\in\frac{\partial L(\mathcal{X}^{k_j},\mathcal{Z}_1^{k_j+1},\mathcal{Z}_2^{k_j+1},\mathcal{Y}_1^{k_j},\mathcal{Y}_2^{k_j},\mu^{k_j})}{\partial\mathcal{Z}_2},\\&0\in P_{\bar{\Omega}}\left(\frac{\partial L(\mathcal{X}^{k_{j+1}},\mathcal{Z}_1^{k_j+1},\mathcal{Z}_2^{k_j+1},\mathcal{Y}_1^{k_j},\mathcal{Y}_2^{k_j},\mu^{k_j})}{\partial\mathcal{X}}\right).\end{aligned}$$

由 (ii) 可知

$$\lim_{k_j\to\infty}\mathcal{Z}_1^{k_j}=\lim_{k_j\to\infty}\mathcal{Z}_2^{k_j}=\lim_{k_j\to\infty}\mathcal{X}^{k_j}=\mathcal{X}^*=\mathcal{Z}_1^*=\mathcal{Z}_2^*.$$

依据算法 3 中 $\mathcal{Y}_1^k$ 和 $\mathcal{Y}_2^k$ 的计算可知

$$\lim_{k_j\to\infty}\mathcal{Y}_1^{k_j}=\mathcal{Y}_1^*,\quad\lim_{k_j\to\infty}\mathcal{Y}_2^{k_j}=\mathcal{Y}_2^*.$$

另外, 对于足够大的 $k_j$, 如下次梯度包含关系成立 (参见文献 [28]):

$$\begin{aligned}&\frac{\partial L(\mathcal{X}^{k_j},\mathcal{Z}_1^{k_j+1},\mathcal{Z}_2^{k_j},\mathcal{Y}_1^{k_j},\mathcal{Y}_2^{k_j},\mu^{k_j})}{\partial\mathcal{Z}_1}\subseteq\frac{L(\mathcal{X}^*,\mathcal{Z}_1^*,\mathcal{Z}_2^*,\mathcal{Y}_1^*,\mathcal{Y}_2^*)}{\partial\mathcal{Z}_1},\\&\frac{\partial L(\mathcal{X}^{k_j},\mathcal{Z}_1^{k_j+1},\mathcal{Z}_2^{k_j},\mathcal{Y}_1^{k_j},\mathcal{Y}_2^{k_j},\mu^{k_j})}{\partial\mathcal{Z}_2}\subseteq\frac{L(\mathcal{X}^*,\mathcal{Z}_1^*,\mathcal{Z}_2^*,\mathcal{Y}_1^*,\mathcal{Y}_2^*)}{\partial\mathcal{Z}_2},\\&\frac{\partial L(\mathcal{X}^{k_j},\mathcal{Z}_1^{k_j+1},\mathcal{Z}_2^{k_j},\mathcal{Y}_1^{k_j},\mathcal{Y}_2^{k_j},\mu^{k_j})}{\partial\mathcal{X}}\subseteq\frac{L(\mathcal{X}^*,\mathcal{Z}_1^*,\mathcal{Z}_2^*,\mathcal{Y}_1^*,\mathcal{Y}_2^*)}{\partial\mathcal{X}}.\end{aligned}$$

从而有

$$0\in\frac{\partial L(\mathcal{X}^*,\mathcal{Z}_1^*,\mathcal{Z}_2^*,\mathcal{Y}_1^*,\mathcal{Y}_2^*)}{\partial\mathcal{Z}_1},\quad 0\in\frac{\partial L(\mathcal{X}^*,\mathcal{Z}_1^*,\mathcal{Z}_2^*,\mathcal{Y}_1^*,\mathcal{Y}_2^*)}{\partial\mathcal{Z}_2},\quad 0\in P_{\bar{\Omega}}\left(\frac{\partial L(\mathcal{X}^*,\mathcal{Z}_1^*,\mathcal{Z}_2^*,\mathcal{Y}_1^*,\mathcal{Y}_2^*)}{\partial\mathcal{X}}\right).$$

因为 $\mathcal{X}^* = \mathcal{Z}_1^* = \mathcal{Z}_2^*$ 和 $P_\Omega(\mathcal{X}^*) = P_\Omega(\mathcal{T})$, 所以 $\{\mathcal{X}^*, \mathcal{Z}_1^*, \mathcal{Z}_2^*\}$ 是优化问题 (3.5) 的一个 KKT 点. 由算法 3 推导出

$$\mathcal{Y}_1^k \in \frac{1}{2}\mathrm{Fold}(\partial\|\mathcal{Z}_{1,\min}^k\|_*), \quad \mathcal{Y}_2^k \in \frac{1}{2}\mathrm{Fold}(\partial\|\mathcal{Z}_{2,\max}^k\|_*), \quad P_{\bar{\Omega}}(\mathcal{Y}_1^k + \mathcal{Y}_2^k) = 0, \quad k = 1, 2, \ldots$$

所以, 由 $\mathcal{X}^* = \mathcal{Z}_1^* = \mathcal{Z}_2^*$ 可得

$$\mathcal{Y}_1^* \in \frac{1}{2}\mathrm{Fold}(\partial\|\mathcal{X}_{\min}^*\|_*), \quad \mathcal{Y}_2^* \in \frac{1}{2}\mathrm{Fold}(\partial\|\mathcal{X}_{\max}^*\|_*), \quad P_{\bar{\Omega}}(\mathcal{Y}_1^* + \mathcal{Y}_2^*) = 0.$$

从而 $\mathcal{X}^*$ 是优化问题 (1.5) 的一个 KKT 点. □

## 5 数值实验

本节通过随机生成的张量补全例子展示 3 个新优化模型和传统模型的计算效果, 其中传统模型 (1.2) 求解采用文献 [20] 中的 HaLRTC 方法. 随机张量是通过张量的 Tucker 分解生成的:

$$\mathcal{T} = \mathcal{G} \times_1 U_1 \times_2 U_2 \times_3 \cdots \times_N U_N,$$

其中 $\mathcal{G} \in \mathbb{R}^{r_1 \times r_2 \times \cdots \times r_N}$ 是核张量, $U_n \in \mathbb{R}^{I_n \times r_n}$, $n = 1, 2, \ldots, N$ 是矩阵. 在实验中, 采样率 $sr = m/\prod_{i=1}^N I_i$, 其中 $m$ 表示已知样本的数量.

参数设置: $\epsilon = 10^{-7}$, $\mu^0 = 10^{-7}$, $\rho = 1.1$, $\mathcal{X}^0 = P_\Omega(\mathcal{T})$, $\mathcal{Y}_1^0 = \mathcal{Y}_2^0 = 0$ (在算法 1 和 2 中, $\mathcal{Y}^0 = 0$). 随机张量补全的误差由以下公式计算: $\mathrm{RSE} = \frac{\|\mathcal{X}-\mathcal{T}\|_F}{\|\mathcal{T}\|_F}$.

从上述随机数值实验结果表 1–3 和图 1 可知, 极小极小和极小极大组合优化模型 (1.5) 的计算结果在 CPU 时间优于传统优化模型 (1.2). 在相同精度的条件下, 极小极小和极小极大组合优化模型

**表 1 当 $sr = 0.3$ 时, 新算法与 HaLRTC 算法的比较**

| $I_1 \times I_2 \times I_3 \times I_4$ | 秩 | 算法 | 迭代次数 | CPU 时间 (s) | RSE |
|---|---|---|---|---|---|
| $(50 \times 50 \times 50 \times 50)$ | $(2, 2, 2, 2)$ | MinADMM | 164 | 34.6520 | 7.1755E−07 |
| | | MaxADMM | 118 | 26.3040 | 2.0709E−04 |
| | | MinMaxADMM | 76 | 22.4166 | 1.1782E−07 |
| | | HaLRTC | 73 | 29.3236 | 1.7948E−07 |
| $(50 \times 50 \times 50 \times 50)$ | $(2, 2, 3, 3)$ | MinADMM | 165 | 32.7725 | 1.2961E−05 |
| | | MaxADMM | 114 | 24.5071 | 2.9533E−06 |
| | | MinMaxADMM | 77 | 22.4824 | 1.3499E−07 |
| | | HaLRTC | 74 | 30.9102 | 9.3507E−08 |
| $(50 \times 60 \times 70 \times 80)$ | $(3, 3, 3, 3)$ | MinADMM | 153 | 95.1939 | 1.5953E−03 |
| | | MaxADMM | 115 | 76.9551 | 1.9631E−06 |
| | | MinMaxADMM | 87 | 68.8280 | 1.6059E−07 |
| | | HaLRTC | 72 | 81.9131 | 1.5718E−07 |
| $(50 \times 60 \times 70 \times 80)$ | $(2, 3, 4, 5)$ | MinADMM | 156 | 93.4482 | 1.2907E−03 |
| | | MaxADMM | 114 | 67.2797 | 1.6653E−05 |
| | | MinMaxADMM | 77 | 58.7081 | 1.6449E−07 |
| | | HaLRTC | 72 | 82.7735 | 1.8411E−07 |

**表 2 当 $sr = 0.2$ 时, 新算法与 HaLRTC 算法的比较**

| $I_1 \times I_2 \times I_3 \times I_4$ | 秩 | 算法 | 迭代次数 | CPU 时间 (s) | RSE |
|---|---|---|---|---|---|
| $(50 \times 50 \times 50 \times 50)$ | $(2,2,2,2)$ | MinADMM | 161 | 30.1703 | 1.7336E−02 |
| | | MaxADMM | 113 | 23.6126 | 9.5989E−04 |
| | | MinMaxADMM | 113 | 32.1671 | 2.1584E−07 |
| | | HaLRTC | 101 | 40.2519 | 1.8859E−07 |
| $(50 \times 50 \times 50 \times 50)$ | $(2,2,3,3)$ | MinADMM | 160 | 29.7121 | 1.9606E−02 |
| | | MaxADMM | 111 | 22.1707 | 3.2379E−03 |
| | | MinMaxADMM | 108 | 29.9584 | 2.1371E−07 |
| | | HaLRTC | 102 | 41.3960 | 1.8469E−07 |
| $(50 \times 60 \times 70 \times 80)$ | $(3,3,3,3)$ | MinADMM | 148 | 73.1436 | 2.7137E−02 |
| | | MaxADMM | 107 | 54.0089 | 6.5913E−04 |
| | | MinMaxADMM | 117 | 89.7480 | 1.8335E−07 |
| | | HaLRTC | 99 | 118.2255 | 1.8471E−07 |
| $(50 \times 60 \times 70 \times 80)$ | $(2,3,4,5)$ | MinADMM | 150 | 75.2818 | 1.8931E−02 |
| | | MaxADMM | 108 | 57.4200 | 3.9843E−03 |
| | | MinMaxADMM | 107 | 75.6992 | 2.0679E−07 |
| | | HaLRTC | 100 | 110.6125 | 1.8124E−07 |

**表 3 当 $sr = 0.1$ 时, 新算法与 HaLRTC 算法的比较**

| $I_1 \times I_2 \times I_3 \times I_4$ | 秩 | 算法 | 迭代次数 | CPU 时间 (s) | RSE |
|---|---|---|---|---|---|
| $(50 \times 50 \times 50 \times 50)$ | $(2,2,2,2)$ | MinADMM | 152 | 21.2821 | 1.4358E−01 |
| | | MaxADMM | 121 | 19.0087 | 4.2304E−02 |
| | | MinMaxADMM | 158 | 39.5576 | 3.7741E−07 |
| | | HaLRTC | 149 | 57.1563 | 3.2804E−07 |
| $(50 \times 50 \times 50 \times 50)$ | $(2,2,3,3)$ | MinADMM | 151 | 25.2896 | 2.0567E−01 |
| | | MaxADMM | 125 | 22.9616 | 1.0666E−01 |
| | | MinMaxADMM | 164 | 42.4373 | 5.7902E−07 |
| | | HaLRTC | 149 | 58.8018 | 2.7878E−07 |
| $(50 \times 60 \times 70 \times 80)$ | $(3,3,3,3)$ | MinADMM | 135 | 63.4821 | 2.6014E−01 |
| | | MaxADMM | 119 | 59.7458 | 6.5913E−03 |
| | | MinMaxADMM | 150 | 111.3659 | 3.4199E−07 |
| | | HaLRTC | 144 | 159.7726 | 2.2051E−07 |
| $(50 \times 60 \times 70 \times 80)$ | $(2,3,4,5)$ | MinADMM | 141 | 65.4778 | 2.10211E−01 |
| | | MaxADMM | 116 | 56.5350 | 6.6970E−02 |
| | | MinMaxADMM | 152 | 111.6332 | 3.4481E−07 |
| | | HaLRTC | 145 | 161.8855 | 7.0265E−03 |

(1.5) 的计算时间通常约为传统优化模型 (1.2) 的 0.7 倍. 另外, 极小极小优化模型和极小极大优化模型计算结果的精度较低, 极小极大凸优化模型没有计算优势. 在这几种优化模型对比中, 极小极小优化模型的计算效果最差, 极小极小与极小极大组合模型计算效果最佳.

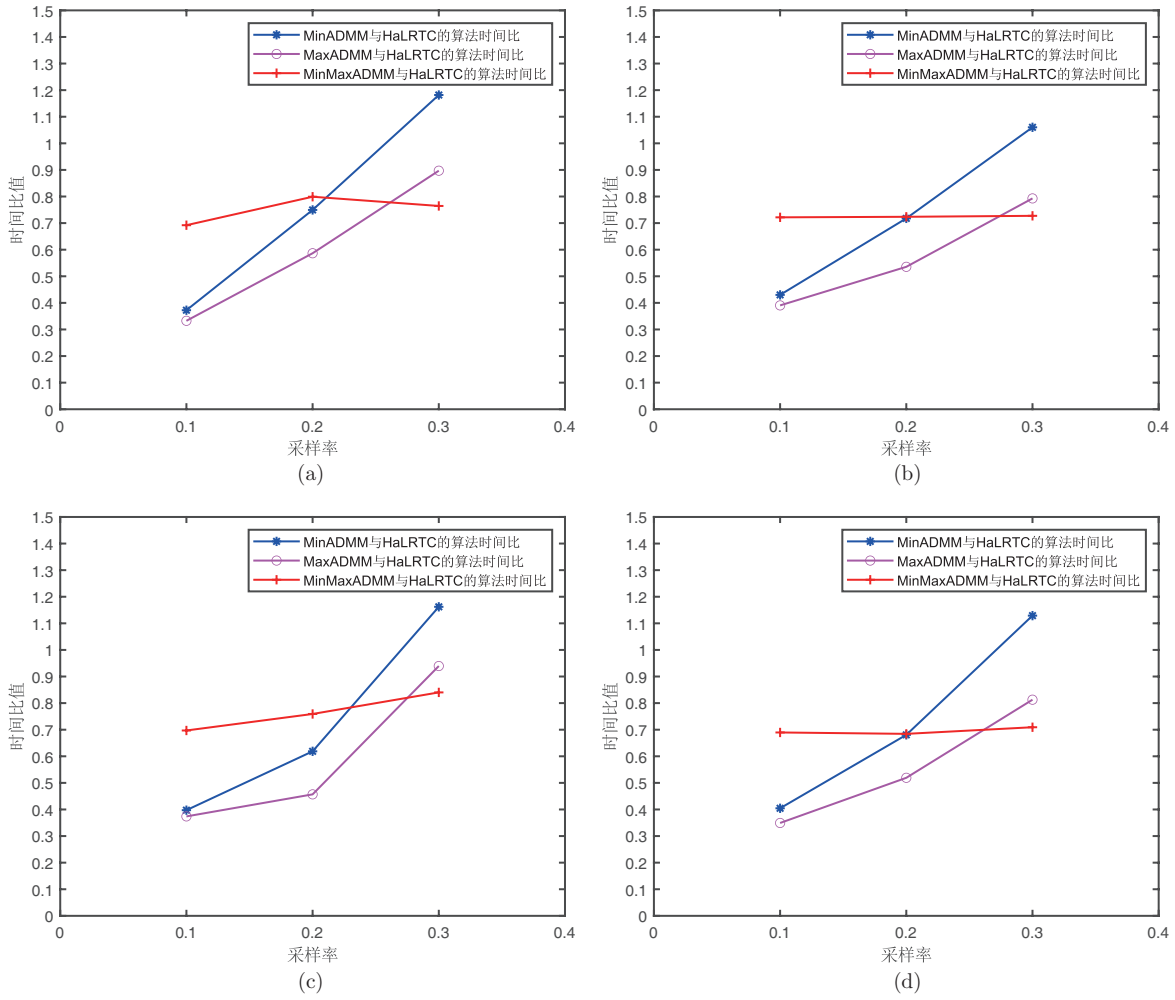


**图 1 (网络版彩图) 3 种算法与 HaLRTC 算法的时间比示意图. (a) $50 \times 50 \times 50 \times 50$, $r = (2, 2, 2, 2)$; (b) $50 \times 50 \times 50 \times 50$, $r = (2, 2, 3, 3)$; (c) $50 \times 60 \times 70 \times 80$, $r = (3, 3, 3, 3)$; (d) $50 \times 60 \times 70 \times 80$, $r = (2, 3, 4, 5)$**

## 6 结论

本文针对张量补全提出了极小极小、极小极大、极小极小与极小极大组合的 3 种优化模型, 采用增广 Lagrange 乘子算法框架提出了相应的算法. 通过本文提出的算法与求解传统优化模型 (1.2) 的 HaLRTC 算法进行数值实验比较, 我们发现极小极小和极小极大凸组合优化模型 (1.5) 计算效果较为稳定. 在精度基本相同的条件下, 计算时间优于传统的优化模型 (1.2), 大约为 0.7 倍.

## 参考文献

1 Bertalmio M, Bertozzi A L, Sapiro G. Navier-Stokes, fluid dynamics, and image and video inpainting. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE, 2001, 355–362

2 Bertalmio M, Sapiro G, Caselles V, et al. Image inpainting. In: SIGGRAPH Conference. New York: ACM Press/Addison-Wesley, 2000, 417–424

3 Cai J F, Candès E J, Shen Z. A singular value thresholding algorithm for matrix completion. SIAM J Optim, 2010,

20: 1956–1982
4 Candès E J, Recht B. Exact matrix completion via convex optimization. Found Comput Math, 2009, 9: 717–772
5 Chen L Q, Li H X, Xie S L. Modified high-order SVD for spatiotemporal modeling of distributed parameter systems. IEEE Trans Ind Electron, 2022, 69: 4296–4304
6 de Lucas S, Vega J M, Velazquez A. Aeronautic conceptual design optimization method based on high-order singular value decomposition. AIAA J, 2015, 49: 2713–2725
7 Gandy S, Recht B, Yamada I. Tensor completion and low-n-rank tensor recovery via convex optimization. Inverse Problems, 2011, 27: 025010
8 Golub G H, Van Loan C F. Matrix Computations, 4th ed. Baltimore: Johns Hopkins Univ Press, 2013
9 Grasedyck L, Kluge M, Krämer S. Variants of alternating least squares tensor completion in the tensor train format. SIAM J Sci Comput, 2015, 37: A2424–A2450
10 Grasedyck L, Kressner D, Tobler C. A literature survey of low-rank tensor approximation techniques. GAMM-Mitt, 2013, 36: 53–78
11 Greenbaum A, Li R C, Overton M L. First-order perturbation theory for eigenvalues and eigenvectors. SIAM Rev, 2020, 62: 463–482
12 Han Z F, Leung C S, Huang L T, et al. Sparse and truncated nuclear norm based tensor completion. Neural Process Lett, 2017, 45: 729–743
13 Hitchcock F L. The expression of a tensor or a polyadic as a sum of products. J Math Phys, 1927, 6: 164–189
14 Horn R A, Johnson C R. Matrix Analysis. Cambridge: Cambridge Univ Press, 1990
15 Huang L T, So H C, Chen Y, et al. Truncated nuclear norm minimization for tensor completion. In: Proceedings of the IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM). New York: IEEE, 2014, 417–420
16 Komodakis N. Image completion using global optimization. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2006, 442–452
17 Kressner D, Steinlechner M, Vandereycken B. Low-rank tensor completion by Riemannian optimization. BIT, 2013, 23: 1–22
18 Li X, Shen H, Li H F, et al. Patch matching-based multitemporal group sparse representation for the missing information reconstruction of remote-sensing images. IEEE J Sel Top Appl Earth Observat Remote Sens, 2016, 9: 3629–3641
19 Lin Z C, Chen M N, Ma Y. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices. arXiv:1009.5055, 2013
20 Liu J, Musialski P, Wonka P, et al. Tensor completion for estimating missing values in visual data. IEEE Trans Pattern Anal Mach Intell, 2013, 35: 208–220
21 Ng M K P, Yuan Q, Yan L, et al. An Adaptive Weighted tensor completion method for the recovery of remote sensing images with missing data. IEEE Trans Geosci Remote Sens, 2017, 55: 3367–3381
22 Oseledets I V. Tensor-train decomposition. SIAM J Sci Comput, 2011, 33: 2295–2317
23 Pauly J M, Mitra N J, Giesen J, et al. Example-based 3D scan completion. In: Proceedings of the Third Eurographics Symposium on Geometry Processing. Goslar: Eurographics Association, 2005, 23–32
24 Rauhut H, Schneider R, Stojanac Ž. Low rank tensor recovery via iterative hard thresholding. Linear Algebra Appl, 2017, 523: 220–262
25 Shi Z Q, Han J Q, Zheng T R, et al. Guarantees of augmented trace norm models in tensor recovery. arXiv:1207.5326, 2012
26 Stewart G W. Perturbation theory for the singular value decomposition. In: SVD and Signal Processing, II: Algorithm, Analysis and Applications. Amsterdam: Elsevier, 1991, 99–109
27 Tucker L R. Some mathematical notes on three-mode factor analysis. Psychometrika, 1966, 31: 279–311
28 Watson G A. Characterization of the subdifferential of some matrix norms. Linear Algebra Appl, 1992, 170: 33–45
29 Wedin P Å. Perturbation bounds in connection with singular value decomposition. BIT, 1972, 12: 99–111
30 Xue S, Qiu W, Liu F, et al. Truncated nuclear norm regularization for low-rank tensor completion. arXiv:1901.01997, 2019
31 Zhao Q B, Zhou G X, Cichocki A. Tensor ring decomposition. arXiv:1606.05535, 2016

# Bi-level models and algorithms based on Tucker rank for tensor completion

Chuanlong Wang, Linjiang Zhong & Xihong Yan

**Abstract** In this paper, the minimal and maximal nuclear norm of $N$-mode matrices as well as their combination for tensor completion are proposed based on Tucker rank. For the three optimization models, we study the null space property and the RIP (restricted isometry property) condition. Based on the framework of the augmented Lagrange multiplier method, we design three modified augmented Lagrange multiplier methods for solving tensor completion. Convergence of the algorithm for solving the non-convex minimal and maximal combination model is established. Finally, we compare the proposed three algorithms with the high accuracy low-rank tensor completion algorithm, and the experimental results of the randomly generated tensor completion show that the new non-convex minimal and maximal combination optimization model outperforms less CPU (central processing unit) time than the traditional nuclear combination optimization model under the same precision.

**Keywords** **tensor completion, bilayer models, convergence, null space property, RIP condition**

**MSC(2020)** **90C26, 90C47, 15A69**

**doi: 10.1360/SSM-2023-0095**