doi:10.3969/j.issn.1001-4616.2026.03.001

一种求解低秩矩阵最小化问题的非凸分式模型的对称 ADMM 算法研究

张思宇[1],张　欣[2,3],沈博涛[4],葛志利[4]

(1.安徽工业大学微电子与数据科学学院,安徽 马鞍山 243032)
(2.宿迁学院文理学院,江苏 宿迁 223800)
(3.大规模复杂系统数值模拟教育部重点实验室,江苏 南京 210023)
(4.南京特殊教育师范学院数学科学学院,江苏 南京 210038)

[**摘要**]　鉴于低秩矩阵最小化问题的非凸非连续,本文考虑了一种近似秩函数的非凸分式模型. 凭借非凸分式模型的可分性特点,设计了一种对称 ADMM 算法来求解该模型. 区别于传统的 ADMM,该方法在乘子更新项 $\boldsymbol{\Lambda}^{k+\frac{1}{2}}$ 中添加了松弛系数 α,该系数有利于加速算法收敛. 在适当的假设条件下,给出了新算法的子序列收敛性和全局收敛性. 最后,通过数值实验验证该方法的有效性.

[**关键词**]　非凸分式模型,对称 ADMM,收敛性.
[**中图分类号**]O221.4　[**文献标志码**]A　[**文章编号**]1001-4616(2026)03-0001-09

Research on a Symmetric ADMM Algorithm for Solving Nonconvex Fractional Model of Low Rank Matrix Minimization Problems

Zhang Siyu[1],Zhang Xin[2,3],Shen botao[4],Ge Zhili[4]

(1.School of Microelectronics and Data Science,Anhui University of Technology,Ma'anshan 243032,China)
(2.School of Sciences and Arts,Suqian University,Suqian 223800,China)
(3.Key Laboratory of Numerical Simulation for Large Scale Complex Systems,Ministry of Education,Nanjing 210023,China)
(4.School of Mathematical Sciences,Nanjing Normal University of Special Education,Nanjing 210038,China)

Abstract:Due to the nonconvex and discontinuous properties of the low rank matrix minimization problem,this paper considers a nonconvex fractional model that approximates the rank function. By exploiting the separability structure of the nonconvex fractional model,a symmetric ADMM algorithm is developed to solve this model. Compared to traditional ADMM,this method incorporates a relaxation coefficient α into the multiplier update term $\boldsymbol{\Lambda}^{k+\frac{1}{2}}$,which allows to speed up the convergence of the algorithm. Under suitable assumptions,the subsequential and global convergence of the new algorithm are established. Finally,numerical experiments are performed to verify the effectiveness of the proposed method.
Key words:nonconvex fractional model,symmetric ADMM,convergence

近年来,非凸低秩矩阵最小化问题被广泛应用于各个领域,如推荐系统,图像处理和机器学习等[1-6]. 它在数学上可以表示为如下的优化问题:

$$\min_{\boldsymbol{X}} \operatorname{rank}(\boldsymbol{X}) \quad \text{s.t. } A(\boldsymbol{X})=\boldsymbol{b}, \tag{1}$$

其中,A 是一个线性算子,$\boldsymbol{b}$ 是一个向量,$\boldsymbol{X}$ 是需要求解的未知矩阵. 如下的矩阵补全问题可以看成是问题(1)的特例.

$$\min_{\boldsymbol{X}} \operatorname{rank}(\boldsymbol{X}) \quad \text{s.t. } P_{\Omega}(\boldsymbol{X})=P_{\Omega}(\boldsymbol{Y}).$$

收稿日期:2025-08-20.
基金项目:国家自然科学基金资助项目(120081、12471290)、江苏省青蓝工程和大规模复杂系统数值模拟教育部重点实验室开放课题基金资助项目.
通讯作者:葛志利,博士,副教授,研究方向:运筹学与控制论. E-mail:250053@njts.edu.cn

$\boldsymbol{Y}$ 是初始观测到的矩阵,Ω 是观测到的条目的索引集,P 是一个线性算子保证 $\boldsymbol{Y}$ 在 Ω 中的元素是已知的,其他元素未知. 然而,由于秩函数的非凸非连续,问题(1)是 NP-难的. Fazel[7] 提出一种常见的秩函数的凸近似-核范数最小化模型,它的数学表达式如下:

$$\min_{\boldsymbol{X}} \lambda \| \boldsymbol{X} \|_* + \| A(\boldsymbol{X}) - \boldsymbol{b} \|_2^2, \tag{2}$$

其中,λ 是正则化参数,$\| \boldsymbol{X} \|_*$ 是 $\boldsymbol{X}$ 的所有奇异值之和. 式(2)被认为是问题(1)的最紧凸松弛,且有很多有效的方法求解,如,奇异值阈值化(Singular Value Thresholding,SVT)[1]、不动点迭代法[8] 和邻近梯度法[9-10] 等. 但由于核范数往往受较大奇异值主导,往往会导致次优[11]. 因此,很多非凸近似函数应运而生. 它们能够更好地近似秩函数,常见的非凸近似有迭代加权核范数(Iteratively Reweighted Nuclear Norm,IRNN)[12],分段 $p(0<p<1)$ 范数(Schatten p,S_p)[2] 及其他.

最近,有些工作研究了一些新的非凸近似. 如,Phan 等[13] 利用降维技术,即用一个小矩阵的 SVD 分解代替原来的完全的 SVD 分解,提出了一种加速的迭代加权核范数方法(Accelerated Iteratively Reweighted Nuclear Norm,AIRNN). Li 等[5] 合并了 S_p 范数与封顶范数(Capped Norm),提出了一种分段封顶 $p(0<p<1)$ 范数(Schatten Capped p,SC_p). Gao 等[14] 提出了一种新的非凸分式模型,具体形式如下:

$$\min_{\boldsymbol{X}} \lambda \frac{\| \boldsymbol{X} \|_*}{\| \boldsymbol{X} \|_F} + \frac{1}{2} \| A(\boldsymbol{X}) - \boldsymbol{b} \|_2^2. \tag{3}$$

与现有的其他非凸模型相比,问题(3)能够避免参数选择,且能更精确地近似秩函数. 因此,本文将基于问题(3)展开算法研究.

文献[14]中,通过引入辅助变量 $\boldsymbol{Y}$,重新将问题(3)写成如下可分离的形式:

$$\begin{aligned} &\min_{\boldsymbol{X},\boldsymbol{Y}} \lambda \frac{\| \boldsymbol{X} \|_*}{\| \boldsymbol{X} \|_F} + \frac{1}{2} \| A(\boldsymbol{Y}) - \boldsymbol{b} \|_2^2 \\ &\text{s.t. } \boldsymbol{X} = \boldsymbol{Y}. \end{aligned} \tag{4}$$

利用经典的交替方向乘子法(Alternating Direction Method of Multipliers,ADMM)求解(4),可以得到

$$\begin{cases} \boldsymbol{X}^{k+1} \in \arg\min_{\boldsymbol{X}} \mathcal{L}_\beta(\boldsymbol{X}, \boldsymbol{Y}^k, \boldsymbol{\Lambda}^k), & (5a) \\ \boldsymbol{Y}^{k+1} = \arg\min_{\boldsymbol{Y}} \mathcal{L}_\beta(\boldsymbol{X}^{k+1}, \boldsymbol{Y}, \boldsymbol{\Lambda}^k), & (5b) \\ \boldsymbol{\Lambda}^{k+1} = \boldsymbol{\Lambda}^k + \beta(\boldsymbol{X}^{k+1} - \boldsymbol{Y}^{k+1}), & (5c) \end{cases}$$

其中,

$$\mathcal{L}_\beta(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{\Lambda}) = \lambda \frac{\| \boldsymbol{X} \|_*}{\| \boldsymbol{X} \|_F} + \frac{1}{2} \| A(\boldsymbol{Y}) - \boldsymbol{b} \|_2^2 + \langle \boldsymbol{\Lambda}, \boldsymbol{X} - \boldsymbol{Y} \rangle + \frac{\beta}{2} \| \boldsymbol{X} - \boldsymbol{Y} \|_F^2, \tag{6}$$

$\boldsymbol{\Lambda}$ 是拉格朗日乘子,β 是罚参数. ADMM 在非凸两分块的向量优化问题中的分析研究有很多相关的工作,见文献[15-19]. Wu 等[18] 提出了一种对称的 ADMM,算法在乘子更新项 $\boldsymbol{\Lambda}^{k+\frac{1}{2}}$ 中附带了一个松弛系数 $\alpha \in (-1,1)$,此系数有利于加速算法收敛,在一些实际问题中通过恰当的选取,可以达到最佳数值效果.

本文受文献[18]的启发,提出一种对称的 ADMM 求解问题(4). 具体形式如下:

$$\begin{cases} \boldsymbol{X}^{k+1} \in \arg\min_{\boldsymbol{X}} \mathcal{L}_\beta(\boldsymbol{X}, \boldsymbol{Y}^k, \boldsymbol{\Lambda}^k), & (7a) \\ \boldsymbol{\Lambda}^{k+\frac{1}{2}} = \boldsymbol{\Lambda}^k + \alpha\beta(\boldsymbol{X}^{k+1} - \boldsymbol{Y}^k), & (7b) \\ \boldsymbol{Y}^{k+1} = \arg\min_{\boldsymbol{Y}} \mathcal{L}_\beta(\boldsymbol{X}^{k+1}, \boldsymbol{Y}, \boldsymbol{\Lambda}^{k+\frac{1}{2}}), & (7c) \\ \boldsymbol{\Lambda}^{k+1} = \boldsymbol{\Lambda}^{k+\frac{1}{2}} + \beta(\boldsymbol{X}^{k+1} - \boldsymbol{Y}^{k+1}). & (7d) \end{cases}$$

区别于传统的 ADMM,对称 ADMM 除增加了乘子的中间更新外,还附带了一个松弛因子 α. α 的融入,使得我们的算法在数值上更有优势. 当 $\alpha=0$ 时,该方法就退化为经典的 ADMM(5). 因此,本文研究的算法是经典 ADMM 的更一般的形式. 在与文献[18]相同的假设条件下,证明了该方法的收敛性. 最后,通过数值实验结果展示了新算法的有效性.

1 预备知识

广义实值函数 g 的定义域为:

$$\mathrm{dom}\ g:=\{\boldsymbol{x}\in\mathbf{R}^n\mid g(\boldsymbol{x})<\infty\}.$$

如果函数 g 是下半连续的,则称函数 g 为闭函数. 如果函数 g 的定义域 $\mathrm{dom}\ g\neq\emptyset$,并且对于定义域内的任意 $\boldsymbol{x}$,都有 $g(\boldsymbol{x})>-\infty$,则称函数 g 为正常函数. 对于任意集合 $S\subseteq\mathbf{R}^n$,任意点 $\boldsymbol{x}\in\mathbf{R}^n$ 到集合 S 的距离定义如下:

$$\mathrm{dist}(\boldsymbol{x},S):=\inf\{\|\boldsymbol{y}-\boldsymbol{x}\|_2\mid\boldsymbol{y}\in S\}.$$

对于一个正常且闭的函数 $g:\mathbf{R}^n\to\mathbf{R}\cup\{\infty\}$,向量 $\boldsymbol{u}\in\partial g(\boldsymbol{x})$ 是 g 在 $\boldsymbol{x}\in\mathrm{dom}\ g$ 上的一个次梯度,其中 ∂g 表示为 g 的次微分,定义为:

$$\partial g(\boldsymbol{x}):=\{\boldsymbol{u}\in\mathbf{R}^n\mid\exists\boldsymbol{x}^k\to\boldsymbol{x},\hat{\partial}g(\boldsymbol{x}^k)\ni\boldsymbol{u}^k\to\boldsymbol{u}\ \text{with}\ g(\boldsymbol{x}^k)\to g(\boldsymbol{x})\},$$

其中,$\hat{\partial}g(\boldsymbol{x})$ 为 g 在 $\boldsymbol{x}$ 上的正则次梯度,定义为:

$$\hat{\partial}g(\boldsymbol{x}):=\{\boldsymbol{u}\in\mathbf{R}^n\mid g(\boldsymbol{y})\geqslant g(\boldsymbol{x})+\langle\boldsymbol{u},\boldsymbol{y}-\boldsymbol{x}\rangle+o(\|\boldsymbol{y}-\boldsymbol{x}\|_2)\quad\forall\boldsymbol{y}\in\mathbf{R}^n\}.$$

若 $f:\mathbf{R}^n\to\mathbf{R}$ 连续可微,且 $g:\mathbf{R}^n\to\mathbf{R}\cup\{\infty\}$ 是正常下半连续函数,有 $\partial(f+g)=\nabla f+\partial g$. 如果点 $\boldsymbol{x}^*$ 满足 $0\in\partial F(\boldsymbol{x}^*)$,则称其为函数 F 的(极限的)稳定点,函数 F 稳定点的集合记为 crit F.

定义 1 若点 $(\boldsymbol{X}^*,\boldsymbol{Y}^*,\boldsymbol{\Lambda}^*)$ 满足

$$\begin{cases}\boldsymbol{0}\in\lambda\partial\left(\dfrac{\|\boldsymbol{X}^*\|_*}{\|\boldsymbol{X}^*\|_F}\right)+\boldsymbol{\Lambda}^*,\\ \boldsymbol{0}=A^{\mathrm{T}}(A(\boldsymbol{Y}^*)-\boldsymbol{b})-\boldsymbol{\Lambda}^*,\\ \boldsymbol{0}=\boldsymbol{X}^*-\boldsymbol{Y}^*,\end{cases}$$

则称 $(\boldsymbol{X}^*,\boldsymbol{Y}^*,\boldsymbol{\Lambda}^*)$ 是式(6)定义的增广拉格朗日函数的稳定点. 容易看出,增广拉格朗日函数的稳定点恰好是问题(4)的 KKT 点. 根据文献[14]中精确恢复条件,本文中我们假设至少存在一个点是(4)的 KKT 点.

2 算法

根据式(6)中增广拉格朗日函数的定义和算法 1,下面将说明(4)的 $\boldsymbol{X}$-子问题和 $\boldsymbol{Y}$-子问题是如何求解的.

先看 $\boldsymbol{Y}$-子问题:对于给定的 $\boldsymbol{X}^{k+1}$ 和 $\boldsymbol{\Lambda}^{k+\frac{1}{2}}$,我们通过 $\mathcal{L}_\beta(\boldsymbol{X}^{k+1},\boldsymbol{Y},\boldsymbol{\Lambda}^{k+\frac{1}{2}})$,有

$$\boldsymbol{Y}^{k+1}=\arg\min_{\boldsymbol{Y}}\mathcal{L}_\beta(\boldsymbol{X}^{k+1},\boldsymbol{Y},\boldsymbol{\Lambda}^{k+\frac{1}{2}})=\arg\min_{\boldsymbol{Y}}\frac{1}{2}\|A(\boldsymbol{Y})-\boldsymbol{b}\|_2^2+\frac{\beta}{2}\left\|\boldsymbol{X}^{k+1}-\boldsymbol{Y}+\frac{\boldsymbol{\Lambda}^{k+\frac{1}{2}}}{\beta}\right\|_F^2=$$

$$(A^{\mathrm{T}}A+\beta\boldsymbol{I})^{-1}(A^{\mathrm{T}}\boldsymbol{b}+\beta\boldsymbol{X}^{k+1}+\boldsymbol{\Lambda}^{k+\frac{1}{2}}).\tag{8}$$

下面再看 $\boldsymbol{X}$-子问题,注意到 $\boldsymbol{X}$-子问题无法直接得到显式解. 因此,通过引入变量 $\boldsymbol{M}$,重新将 $\boldsymbol{X}$-子问题写成如下形式:

$$\begin{aligned}&\min_{\boldsymbol{X},\boldsymbol{M}}\lambda\frac{\|\boldsymbol{X}\|_*}{\|\boldsymbol{M}\|_F}+\frac{\beta}{2}\|\boldsymbol{X}-\boldsymbol{S}^k\|_F^2\\&\text{s.t. }\boldsymbol{X}=\boldsymbol{M},\end{aligned}\tag{9}$$

其中,$\boldsymbol{S}^k=\boldsymbol{Y}^k-\dfrac{\boldsymbol{\Lambda}^k}{\beta}$. 式(9)的增广拉格朗日函数为

$$\mathcal{L}_\mu^k(\boldsymbol{X},\boldsymbol{M},\boldsymbol{\xi})=\lambda\frac{\|\boldsymbol{X}\|_*}{\|\boldsymbol{M}\|_F}+\frac{\beta}{2}\|\boldsymbol{X}-\boldsymbol{S}^k\|_F^2+\langle\boldsymbol{\xi},\boldsymbol{X}-\boldsymbol{M}\rangle+\frac{\mu}{2}\|\boldsymbol{X}-\boldsymbol{M}\|_F^2,$$

其中,$\boldsymbol{\xi}$ 是拉格朗日乘子,μ 是罚因子. 我们用 ADMM 求解(10),迭代格式如下:

$$\begin{cases}\boldsymbol{X}^{t+1}=\arg\min_{\boldsymbol{X}}\mathcal{L}_\mu^k(\boldsymbol{X},\boldsymbol{M}^t,\boldsymbol{\xi}^t),&(10a)\\ \boldsymbol{M}^{t+1}=\arg\min_{\boldsymbol{M}}\mathcal{L}_\mu^k(\boldsymbol{X}^{t+1},\boldsymbol{M},\boldsymbol{\xi}^t),&(10b)\\ \boldsymbol{\xi}^{t+1}=\boldsymbol{\xi}^t+\mu(\boldsymbol{X}^{t+1}-\boldsymbol{M}^{t+1}),&(10c)\end{cases}$$

t 是内迭代次数. 因此,$\boldsymbol{X}$-子问题最优解为

$$\boldsymbol{X}^{t+1}=\arg\min_{\boldsymbol{X}}\mathcal{L}_{\mu}^{k}(\boldsymbol{X},\boldsymbol{M}^{t},\boldsymbol{\xi}^{t})=\arg\min_{\boldsymbol{X}}\frac{\lambda}{\|\boldsymbol{M}^{t}\|_{F}}\|\boldsymbol{X}\|_{*}+\frac{\beta+\mu}{2}\left\|\boldsymbol{X}-\frac{\beta\boldsymbol{S}^{k}+\mu\boldsymbol{C}^{t}}{\beta+\mu}\right\|_{F}^{2}=\boldsymbol{U}\mathcal{D}(\boldsymbol{\Sigma})\boldsymbol{V}^{\mathrm{T}},\tag{11}$$

其中,$\boldsymbol{C}^{t}=\boldsymbol{M}^{t}-\frac{\boldsymbol{\xi}^{t}}{\mu}$,$\mathcal{D}(\boldsymbol{\Sigma})=\mathrm{diag}\left(\max\left\{\boldsymbol{\Sigma}_{ii}-\frac{\lambda}{(\beta+\mu)\|\boldsymbol{M}^{t}\|_{F}},0\right\}\right)$且要满足

$$\beta\boldsymbol{S}^{k}+\mu\left(\boldsymbol{M}^{t}-\frac{\boldsymbol{\xi}^{t}}{\mu}\right)=(\beta+\mu)\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}.$$

类似于文献[20]中的讨论,$\boldsymbol{M}$-子问题的求解格式如下:

$$\boldsymbol{M}^{t+1}=\arg\min_{\boldsymbol{M}}\mathcal{L}_{\mu}^{k}(\boldsymbol{X}^{t+1},\boldsymbol{M},\boldsymbol{\xi}^{t})=\arg\min_{\boldsymbol{M}}\frac{\lambda\|\boldsymbol{X}^{t+1}\|_{*}}{\|\boldsymbol{M}\|_{F}}+\frac{\mu}{2}\left\|\boldsymbol{X}^{t+1}-\boldsymbol{M}+\frac{\boldsymbol{\xi}^{t}}{\mu}\right\|_{F}^{2}=\begin{cases}\gamma^{t}\left(\boldsymbol{X}^{t+1}+\frac{\boldsymbol{\xi}^{t}}{\mu}\right), & \boldsymbol{X}^{t+1}+\frac{\boldsymbol{\xi}^{t}}{\mu}\neq\boldsymbol{0},\\ \boldsymbol{P}^{t}, & \boldsymbol{X}^{t+1}+\frac{\boldsymbol{\xi}^{t}}{\mu}=\boldsymbol{0}.\end{cases}\tag{12}$$

其中,$\boldsymbol{P}^{t}$ 是一个随机矩阵满足 $\|\boldsymbol{P}^{t}\|_{F}^{3}=\frac{d^{t}}{\mu}$ 且 $d^{t}=\lambda\|\boldsymbol{X}^{t+1}\|_{*}$,$\gamma^{t}=\frac{1}{3}+\frac{1}{3}\left(a^{t}+\frac{1}{a^{t}}\right)$,其中 $a^{t}=\sqrt[3]{\frac{27q^{t}+2+\sqrt{(27q^{t}+2)^{2}-4}}{2}}$,$q^{t}=\frac{d^{t}}{\mu\left\|\boldsymbol{X}^{t+1}+\frac{\boldsymbol{\xi}^{t}}{\mu}\right\|_{F}^{3}}$. 因此,得到求解问题(4)的对称 ADMM 算法如下:

算法 1　对称交替方向乘子法(Symmetric Alternating Direction Method of Multipliers,SADMM)

输入:$A,\boldsymbol{b},\lambda,\beta,\mu,k_{\max},t_{\max}$
初始化:$\boldsymbol{X}^{0},\boldsymbol{Y}^{0},\boldsymbol{M}^{0},\boldsymbol{\Lambda}^{0},\boldsymbol{\xi}^{0}$ 且令 $k,t=0$
　while $\frac{\|\boldsymbol{X}^{k+1}-\boldsymbol{X}^{k}\|_{F}}{\|\boldsymbol{X}\|_{F}^{k}}>10^{-5}$ 且 $k\leqslant k_{\max}$ do
　　while $\frac{\|\boldsymbol{X}^{t+1}-\boldsymbol{X}^{t}\|_{F}}{\|\boldsymbol{X}\|_{F}^{t}}>10^{-5}$ 且 $t\leqslant t_{\max}$ do
　　通过式(11)更新 $\boldsymbol{X}^{t+1}$
　　通过式(12)更新 $\boldsymbol{M}^{t+1}$
　　通过式(10c)更新 $\boldsymbol{\xi}^{t+1}$
　　令 $t=t+1$
　end while
　return $\boldsymbol{X}^{k+1}=\boldsymbol{X}^{t}$
　通过式(8)更新 $\boldsymbol{Y}^{k+1}$
　通过式(7b)更新 $\boldsymbol{\Lambda}^{k+\frac{1}{2}}$
　通过式(7d)更新 $\boldsymbol{\Lambda}^{k+1}$
　令 $k=k+1$ 和 $t=0$
end while
return $\boldsymbol{X}^{*}=\boldsymbol{X}^{k}$

3　收敛性

本节将分析算法 1 的收敛性. 回顾迭代格式(7a)-(7d),首先可以给出算法 1 中子问题的一阶最优性条件.

引理 1　设 $\boldsymbol{W}^{k}:=(\boldsymbol{X}^{k},\boldsymbol{Y}^{k},\boldsymbol{\Lambda}^{k})$,则有

$$\begin{cases}\boldsymbol{0}\in\lambda\partial\left(\frac{\|\boldsymbol{X}^{k+1}\|_{*}}{\|\boldsymbol{X}^{k+1}\|_{F}}\right)+\boldsymbol{\Lambda}^{k+1}+\frac{\beta}{\alpha+1}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k})-\frac{\alpha}{\alpha+1}(\boldsymbol{\Lambda}^{k+1}-\boldsymbol{\Lambda}^{k}),\\ \boldsymbol{0}=A^{\mathrm{T}}(A(\boldsymbol{Y}^{k+1})-\boldsymbol{b})-\boldsymbol{\Lambda}^{k+1},\\ \boldsymbol{\Lambda}^{k+1}=\boldsymbol{\Lambda}^{k}+\beta[(\alpha+1)(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k})-(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k})].\end{cases}$$

证明　将式(7b)和式(7d)相加,可以得到

$$\boldsymbol{\Lambda}^{k+1}-\boldsymbol{\Lambda}^{k}=\alpha\beta(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k})+\beta(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k+1})=(\alpha+1)\beta(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k})-\beta(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k})+\beta(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k+1})=$$
$$(\alpha+1)\beta(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k})-\beta(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}).$$

因此

$$\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k}=\frac{1}{(\alpha+1)\beta}(\boldsymbol{\Lambda}^{k+1}-\boldsymbol{\Lambda}^{k})+\frac{1}{\alpha+1}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}). \tag{13}$$

由(7a)的最优性条件,可得

$$\boldsymbol{0}\in\lambda\partial\left(\frac{\|\boldsymbol{X}^{k+1}\|_{*}}{\|\boldsymbol{X}^{k+1}\|_{F}}\right)+\boldsymbol{\Lambda}^{k}+\beta(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k})=\lambda\partial\left(\frac{\|\boldsymbol{X}^{k+1}\|_{*}}{\|\boldsymbol{X}^{k+1}\|_{F}}\right)+\boldsymbol{\Lambda}^{k+1}-(\boldsymbol{\Lambda}^{k+1}-\boldsymbol{\Lambda}^{k})+\beta(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k}).$$

将(13)代入上式中,可以得到

$$\boldsymbol{0}\in\lambda\partial\left(\frac{\|\boldsymbol{X}^{k+1}\|_{*}}{\|\boldsymbol{X}^{k+1}\|_{F}}\right)+\boldsymbol{\Lambda}^{k+1}+\frac{\beta}{\alpha+1}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k})-\frac{\alpha}{\alpha+1}(\boldsymbol{\Lambda}^{k+1}-\boldsymbol{\Lambda}^{k}).$$

同理,由(7c)的最优性条件,可得 $A^{\mathrm{T}}(A(\boldsymbol{Y}^{k+1})-\boldsymbol{b})-\boldsymbol{\Lambda}^{k+\frac{1}{2}}-\beta(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k+1})=\boldsymbol{0}$. 再由(7d)可得 $A^{\mathrm{T}}(A(\boldsymbol{Y}^{k+1})-\boldsymbol{b})-\boldsymbol{\Lambda}^{k+1}=\boldsymbol{0}$.

为后面的分析方便,我们令 $L=\sigma_{\max}(A^{\mathrm{T}}A)$. 引理 2 给出了序列 $\{\boldsymbol{W}^{k}\}$ 的单调非增性.

引理 2 设 $\{\boldsymbol{W}^{k}\}$ 为算法 1 生成的序列. 假设 $\alpha\in(-1,1)$,$\beta>\frac{2}{1-\alpha}L$,则有序列 $\{\mathcal{L}_{\beta}(\boldsymbol{W}^{k})\}$ 非增,且

$$\mathcal{L}_{\beta}(\boldsymbol{W}^{k})-\mathcal{L}_{\beta}(\boldsymbol{W}^{k+1})\geqslant c\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2},$$

式中,$c=\left(\frac{1}{\alpha+1}-\frac{1}{2}\right)\beta-\frac{L}{2}-\frac{L^{2}}{(\alpha+1)\beta}>0$.

证明 由(7a)可得

$$\mathcal{L}_{\beta}(\boldsymbol{X}^{k},\boldsymbol{Y}^{k},\boldsymbol{\Lambda}^{k})-\mathcal{L}_{\beta}(\boldsymbol{X}^{k+1},\boldsymbol{Y}^{k},\boldsymbol{\Lambda}^{k})\geqslant 0 \tag{14}$$

根据式(6)中定义的增广拉格朗日函数 $\mathcal{L}_{\beta}(\cdot)$,可得

$$\mathcal{L}_{\beta}(\boldsymbol{X}^{k+1},\boldsymbol{Y}^{k},\boldsymbol{\Lambda}^{k+\frac{1}{2}})-\mathcal{L}_{\beta}(\boldsymbol{X}^{k+1},\boldsymbol{Y}^{k+1},\boldsymbol{\Lambda}^{k+\frac{1}{2}})=\lambda\frac{\|\boldsymbol{X}^{k+1}\|_{*}}{\|\boldsymbol{X}^{k+1}\|_{F}}+\frac{1}{2}\|A(\boldsymbol{Y}^{k})-\boldsymbol{b}\|_{2}^{2}+\langle\boldsymbol{\Lambda}^{k+\frac{1}{2}},\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k}\rangle+\frac{\beta}{2}\|\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}-$$
$$\lambda\frac{\|\boldsymbol{X}^{k+1}\|_{*}}{\|\boldsymbol{X}^{k+1}\|_{F}}-\frac{1}{2}\|A(\boldsymbol{Y}^{k+1})-\boldsymbol{b}\|_{2}^{2}-\langle\boldsymbol{\Lambda}^{k+\frac{1}{2}},\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k+1}\rangle-\frac{\beta}{2}\|\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k+1}\|_{F}^{2}=$$
$$\frac{1}{2}[\|A(\boldsymbol{Y}^{k})-\boldsymbol{b}\|_{2}^{2}-\|A(\boldsymbol{Y}^{k+1})-\boldsymbol{b}\|_{2}^{2}]+\langle\boldsymbol{\Lambda}^{k+\frac{1}{2}},\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\rangle+\frac{\beta}{2}\|\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}-\frac{\beta}{2}\|\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k+1}\|_{F}^{2}=$$
$$\frac{1}{2}[\|A(\boldsymbol{Y}^{k})-\boldsymbol{b}\|_{2}^{2}-\|A(\boldsymbol{Y}^{k+1})-\boldsymbol{b}\|_{2}^{2}]+[\boldsymbol{\Lambda}^{k}+(\alpha+1)\beta(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k})]^{\mathrm{T}}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k})-\beta(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k})^{\mathrm{T}}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k})+$$
$$\frac{\beta}{2}\|\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}-\frac{\beta}{2}\|\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k+1}\|_{F}^{2}+\frac{\beta}{2}\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}-\frac{\beta}{2}\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}=$$
$$\frac{1}{2}[\|A(\boldsymbol{Y}^{k})-\boldsymbol{b}\|_{2}^{2}-\|A(\boldsymbol{Y}^{k+1})-\boldsymbol{b}\|_{2}^{2}]+[\boldsymbol{\Lambda}^{k}+(\alpha+1)\beta(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k})]^{\mathrm{T}}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k})-\frac{\beta}{2}\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}=$$
$$\frac{1}{2}[\|A(\boldsymbol{Y}^{k})-\boldsymbol{b}\|_{2}^{2}-\|A(\boldsymbol{Y}^{k+1})-\boldsymbol{b}\|_{2}^{2}]+\langle\boldsymbol{\Lambda}^{k+1},\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\rangle+\frac{\beta}{2}\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}\geqslant$$
$$-\frac{L}{2}\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}+\frac{\beta}{2}\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}\geqslant\frac{\beta-L}{2}\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}, \tag{15}$$

其中,倒数第二个不等式根据文献[21]的引理 5.7 得到.

接下来,通过增广拉格朗日函数的定义和(7b),可得

$$\mathcal{L}_{\beta}(\boldsymbol{X}^{k+1},\boldsymbol{Y}^{k},\boldsymbol{\Lambda}^{k})-\mathcal{L}_{\beta}(\boldsymbol{X}^{k+1},\boldsymbol{Y}^{k},\boldsymbol{\Lambda}^{k+\frac{1}{2}})+\mathcal{L}_{\beta}(\boldsymbol{X}^{k+1},\boldsymbol{Y}^{k+1},\boldsymbol{\Lambda}^{k+\frac{1}{2}})-\mathcal{L}_{\beta}(\boldsymbol{X}^{k+1},\boldsymbol{Y}^{k+1},\boldsymbol{\Lambda}^{k+1})=$$
$$(\boldsymbol{\Lambda}^{k}-\boldsymbol{\Lambda}^{k+\frac{1}{2}})^{\mathrm{T}}(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k})+(\boldsymbol{\Lambda}^{k+\frac{1}{2}}-\boldsymbol{\Lambda}^{k+1})^{\mathrm{T}}(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k+1})=(\boldsymbol{\Lambda}^{k}-\boldsymbol{\Lambda}^{k+1})^{\mathrm{T}}(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k+1})+(\boldsymbol{\Lambda}^{k}-\boldsymbol{\Lambda}^{k+\frac{1}{2}})^{\mathrm{T}}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k})=$$
$$(\boldsymbol{\Lambda}^{k}-\boldsymbol{\Lambda}^{k+1})^{\mathrm{T}}(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k+1})-\alpha\beta(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k})^{\mathrm{T}}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}). \tag{16}$$

由式(13),可以得到

$$\boldsymbol{X}^{k+1}-\boldsymbol{Y}^{k+1}=\frac{1}{(\alpha+1)\beta}(\boldsymbol{\Lambda}^{k+1}-\boldsymbol{\Lambda}^{k})-\frac{\alpha}{\alpha+1}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}). \tag{17}$$

将式(13)和式(17)代入到式(16)中,可以得到

$$\mathcal{L}_{\beta}(\boldsymbol{X}^{k+1},\boldsymbol{Y}^{k},\boldsymbol{\Lambda}^{k})-\mathcal{L}_{\beta}(\boldsymbol{X}^{k+1},\boldsymbol{Y}^{k},\boldsymbol{\Lambda}^{k+\frac{1}{2}})+\mathcal{L}_{\beta}(\boldsymbol{X}^{k+1},\boldsymbol{Y}^{k+1},\boldsymbol{\Lambda}^{k+\frac{1}{2}})-\mathcal{L}_{\beta}(\boldsymbol{X}^{k+1},\boldsymbol{Y}^{k+1},\boldsymbol{\Lambda}^{k+1})=$$

$$(\boldsymbol{\Lambda}^{k}-\boldsymbol{\Lambda}^{k+1})^{\mathrm{T}}\left[\frac{1}{(\alpha+1)\beta}(\boldsymbol{\Lambda}^{k+1}-\boldsymbol{\Lambda}^{k})-\frac{\alpha}{\alpha+1}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k})\right]-\alpha\beta\left[\frac{1}{(\alpha+1)\beta}(\boldsymbol{\Lambda}^{k+1}-\boldsymbol{\Lambda}^{k})+\frac{1}{\alpha+1}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k})\right]^{\mathrm{T}}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k})=$$

$$-\frac{1}{(\alpha+1)\beta}\|\boldsymbol{\Lambda}^{k+1}-\boldsymbol{\Lambda}^{k}\|_{2}^{2}-\frac{\alpha\beta}{\alpha+1}\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}\geqslant-\left(\frac{L^{2}}{(\alpha+1)\beta}+\frac{\alpha\beta}{\alpha+1}\right)\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}$$

最后一个不等式是根据 $\boldsymbol{\Lambda}^{k+1}=A^{\mathrm{T}}(A(\boldsymbol{Y}^{k+1})-\boldsymbol{b})$ 和 $\boldsymbol{\Lambda}^{k}=A^{\mathrm{T}}(A(\boldsymbol{Y}^{k})-\boldsymbol{b})$,得到

$$\|\boldsymbol{\Lambda}^{k+1}-\boldsymbol{\Lambda}^{k}\|_{2}^{2}=\|A^{\mathrm{T}}A(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k})\|_{2}^{2}\leqslant L^{2}\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}.$$

将式(14)、式(15)和式(18)相加,可以得到

$$\mathcal{L}_{\beta}(\boldsymbol{W}^{k})-\mathcal{L}_{\beta}(\boldsymbol{W}^{k+1})\geqslant 0+\frac{\beta-L}{2}\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}-\left[\frac{L^{2}}{(\alpha+1)\beta}+\frac{\alpha\beta}{\alpha+1}\right]\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}=$$

$$\left(\frac{\beta}{2}-\frac{L}{2}-\frac{L^{2}}{(\alpha+1)\beta}-\frac{\alpha\beta}{\alpha+1}\right)\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}=\left[\left(\frac{1}{\alpha+1}-\frac{1}{2}\right)\beta-\frac{L}{2}-\frac{L^{2}}{(\alpha+1)\beta}\right]\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}.$$

引理 3　设 $\{\boldsymbol{W}^{k}\}$ 是由算法 1 生成的序列,$\alpha\in(-1,1)$,$\beta>\frac{2}{1-\alpha}L$. 若 $\{\boldsymbol{X}^{k}\}$ 有界,则有 $\{\boldsymbol{W}^{k}\}$ 有界且 $\lim\limits_{k\to\infty}\|\boldsymbol{X}^{k+1}-\boldsymbol{X}^{k}\|_{F}^{2}=0$,$\lim\limits_{k\to\infty}\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}=0$ 和 $\lim\limits_{k\to\infty}\|\boldsymbol{\Lambda}^{k+1}-\boldsymbol{\Lambda}^{k}\|_{2}^{2}=0$.

注 1　在标准的有界性分析中,通常使用目标函数的强制来保证. 不幸的是,$\frac{\|\cdot\|_{*}}{\|\cdot\|_{F}}$ 不是强制的,所以需要假设序列 $\{\boldsymbol{X}^{k}\}$ 有界,详细的证明过程见[14]的定理 4.1(i)或[22]的引理 3,本文略.

引理 4　设 $\{\boldsymbol{X}^{k}\}$ 是由算法 1 生成的序列. 定义

$$\begin{cases}\boldsymbol{w}_{X}^{k+1}=(\boldsymbol{\Lambda}^{k+1}-\boldsymbol{\Lambda}^{k})-\beta(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}),\\ \boldsymbol{w}_{Y}^{k+1}=\frac{1}{\alpha+1}(\boldsymbol{\Lambda}^{k}-\boldsymbol{\Lambda}^{k+1})+\frac{\alpha\beta}{\alpha+1}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}),\\ \boldsymbol{w}_{\Lambda}^{k+1}=\frac{1}{(\alpha+1)\beta}(\boldsymbol{\Lambda}^{k+1}-\boldsymbol{\Lambda}^{k})-\frac{\alpha}{\alpha+1}(\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}),\end{cases} \tag{19}$$

且有 $(\boldsymbol{w}_{X}^{k+1},\boldsymbol{w}_{Y}^{k+1},\boldsymbol{w}_{\Lambda}^{k+1})\in\partial\mathcal{L}_{\beta}(\boldsymbol{W}^{k+1})$. 若 $\{\boldsymbol{X}^{k}\}$ 有界,$\alpha\in(-1,1)$,$\beta>\frac{2}{1-\alpha}L$,则存在一个正常数 m 使得

$$\|\operatorname{dist}(\boldsymbol{0},\partial\mathcal{L}_{\beta}(\boldsymbol{W}^{k+1}))\|_{F}^{2}\leqslant\|(\boldsymbol{w}_{X}^{k+1},\boldsymbol{w}_{Y}^{k+1},\boldsymbol{w}_{\Lambda}^{k+1})\|_{F}^{2}\leqslant m(\|\boldsymbol{Y}^{k+1}-\boldsymbol{Y}^{k}\|_{F}^{2}).$$

注 2　引理 4 的证明,直接由增广拉格朗日函数的定义和(19)可得,本文略.

在下面的定理中,将提出子序列的收敛性,这意味着生成的序列的任何聚点都是问题(4)的稳定点.

定理 1　设 $\{\boldsymbol{W}^{k}\}$ 是算法 1 生成的序列. 若 $\{\boldsymbol{X}^{k}\}$ 有界,$\alpha\in(-1,1)$,$\beta>\frac{2}{1-\alpha}L$,则序列 $\{\boldsymbol{W}^{k}\}$ 的任意聚点是式(6)定义的增广拉格朗日函数 $\mathcal{L}_{\beta}(\cdot)$ 的稳定点.

注 3　上述定理的证明类似于文献[14]中定理 4.1(iii),本文略. 此外,可以借助 Kurdyka-Lojasiewicz 性质和上面得到的结论,保证整个生成序列全局收敛到式(6)中增广拉格朗日函数 $\mathcal{L}_{\beta}(\cdot)$ 的稳定点. 在下面的定理中可以得出结论,详细的证明过程见[23]中的定理 8 及[16]的定理 3.1 的证明. 由于篇幅有限,本文省略了证明.

定理 2　设 $\{\boldsymbol{W}^{k}\}$ 为由算法 1 生成的序列. 若 $\{\boldsymbol{X}^{k}\}$ 有界,$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}>\boldsymbol{0}$,$\alpha\in(-1,1)$,$\beta>\frac{2}{1-\alpha}L$,则 $\{\boldsymbol{W}^{k}\}$ 有限. 即 $\sum\limits_{k=1}^{\infty}\|\boldsymbol{W}^{k+1}-\boldsymbol{W}^{k}\|_{F}<\infty$. 因此,$\{\boldsymbol{W}^{k}\}$ 全局收敛于式(6)定义的增广拉格朗日函数 $\mathcal{L}_{\beta}(\cdot)$ 的稳定点.

4　数值实验

本节中,我们分别用人造数据和真实数据进行了数值实验,用来评估所提出 SADMM 的性能. 实验解

决的是如下的矩阵补全问题:

$$\min_{X} \lambda \frac{\| X \|_*}{\| X \|_F} + \frac{1}{2} \| P_\Omega(X-Y) \|_F^2,$$

其中,$\lambda>0$ 是正则化参数. 所有代码均使用 MATLAB 2021a 在配备 Intel(R) Core(TM) i7-1165G7、2.80 GHz 和 16 GB RAM 的 Windows 10 笔记本电脑上编写和运行.

4.1 人造数据实验

在本小节的数值实验中,目的是比较经典的 ADMM 和本文所提出的 SADMM 在相同条件下的实验结果. 在实验中,将测试 $\alpha \in (-1,1)$ 取不同值对算法结果的影响.

首先,生成一个秩为 r 的矩阵 $\boldsymbol{X}$ 为 $\boldsymbol{M}_L\boldsymbol{M}_R$,其中 $\boldsymbol{M}_L \in \mathbf{R}^{m\times r}$ 和 $\boldsymbol{M}_R \in \mathbf{R}^{r\times n}$ 是由 MATLAB 命令 rand 随机生成,且让 $\boldsymbol{X}$ 随机缺失 50%. 实验参数设置 $m=n=200$,$r=20$,$\lambda=300$,$\beta=0.3$,$\mu=0.03$,最大外部迭代 $k_{\max}$ 和内部迭代 $t_{\max}$ 设置为 500 和 5.

表 1 中列出了 α 取不同值的数值结果,$\alpha=0$ 即为传统的 ADMM,其中 Time(s) 表示以秒为单位的计算时间,Iter.表示迭代次数. 由表 1,可以得到合适的 α 使得 SADMM 快速收敛. $\alpha=0.7$ 时的结果表现最优.

表 1 不同的 α 值对应的迭代时间和迭代次数

Table 1 Iteration time and number of iterations for different α values

α	−0.5	−0.3	−0.1	0	0.1	0.3	0.5	**0.7**	0.9
Time/s	20.31	12.77	9.30	8.16	7.31	6.15	5.52	**5.27**	5.39
Iter.	398	248	177	155	139	118	104	**99**	100

4.2 真实数据实验

本小节将通过对真实图像数据进行实验来评估本文所提算法的有效性,实验中参数设置为 $\lambda=50$,$\beta=0.5$,$\mu=0.05$. 虽然真实的图像处理中,相应的矩阵往往不是低秩的,但它们的主要信息是由较大的奇异值决定. 因此,可以用低秩矩阵来近似它们. 此外,待恢复的图像是由于部分被损坏或者遮挡,所以可以使用矩阵补全算法来恢复这些图像. 将上一节得到的 $\alpha=0.7$ 时的 SADMM 与文献[1]的 SVT,[5]中的 SC_p 以及[13]中的 AIRNN 进行比较.

图 1 分别给出了三张不同的黑白图像:Boat(512×512)、Spillikins(512×512)、Peppers(512×512)和三张不同的彩色图像:Tree(300×300)、House(256×256)、Sky(512×512),并且通过随机丢失 50%的真实数据来模拟图像被部分损坏或遮挡情况. 同时,在图 1 中,给出了使用 SADMM 的恢复图像.



图 1 从左到右:Boat、Spillikins、Peppers、Tree、House、Sky. 第一行:原始的图像,第二行:损坏后的图像,第三行:SADMM 恢复出的图像.

Fig. 1 Left to right: Boat, Spillikins, Peppers, Tree, House, Sky. First row: original image, second row: damaged image, third row: image restored by SADMM.

为评估本文所提出 SADMM,表 2 中分别给出了 AIRNN、SVT、SC_p 和本文的 SADMM 这四种方法的原始图像 $\boldsymbol{X}$ 与恢复图像 $\boldsymbol{X}^*$ 的信噪比(Signal-to-Noise-Ratio,SNR)和相对误差(Relative Error,RE)的数值结果. 其中 $\mathrm{SNR}:=10\log_{10}\frac{\|\boldsymbol{X}-\bar{\boldsymbol{X}}\|_F^2}{\|\boldsymbol{X}-\boldsymbol{X}^*\|_F^2}$,$\bar{\boldsymbol{X}}$ 是原始数据的均值,相对误差 $\mathrm{RE}:=\frac{\|\boldsymbol{X}^*-\boldsymbol{X}\|_F}{\|\boldsymbol{X}\|_F}$. 容易看出,SNR 值越大越好,RE 值越小越好. 这里限制了最大外部迭代 $k_{\max}$ 和内部迭代 $t_{\max}$ 设置为 120 和 5. 由表 2 可知,相较于其他三种算法,本文所提出的 SADMM 算法可实现更优的恢复效果.

表 2　不同算法在不同图像上的数值结果

Table 2　Numerical results of different algorithms on different images

图像名称	AIRNN		SVT		SC_p		SADMM	
	SNR	RE	SNR	RE	SNR	RE	SNR	RE
Boat	23.70	0.097 1	17.39	0.136 9	24.20	0.077 6	**26.43**	**0.073 0**
Spillikins	21.45	0.088 4	18.70	0.110 6	22.65	0.063 9	**29.14**	**0.052 2**
Peppers	21.12	0.073 3	16.69	0.104 4	18.17	0.073 2	**26.24**	**0.052 8**
Tree	22.48	0.159 9	20.28	0.175 6	24.86	0.130 7	**26.23**	**0.123 5**
House	20.20	0.078 6	18.45	0.097 1	22.65	0.065 8	**24.86**	**0.061 9**
Sky	26.96	0.107 2	23.74	0.120 1	29.34	0.082 9	**30.79**	**0.076 4**
Average	22.65	0.100 7	19.21	0.124 1	23.65	0.082 3	**27.28**	**0.073 3**

5　结论

本文主要研究了利用对称 ADMM 求解低秩矩阵最小化问题的非凸分式模型. 该算法是在经典的 ADMM 求解非凸分式模型的基础上,添加了对乘子更新迭代项 $\boldsymbol{\Lambda}^{k+\frac{1}{2}}$,且附带了一个松弛系数 α,$\alpha\in(-1,1)$在适当的条件下,分析了算法的收敛性. 最后,通过人造和真实数据实验展示了算法的高效性和稳定性.

[参考文献]

[1] Cai J F, Candès E J, Shen Z W. A singular value thresholding algorithm for matrix completion[J]. SIAM Journal on Optimization, 2010, 20(4): 1956-1982.

[2] Nie F P, Huang H, Ding C. Low-rank matrix recovery via efficient schatten p-norm minimization[C]//Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence. Toronto, Ontario, Canada, 2012, 26(1): 655-661.

[3] Malek-Mohammadi M, Babaie-Zadeh M, Skoglund M. Iterative concave rank approximation for recovering low-rank matrices [J]. IEEE Transactions on Signal Processing, 2014, 62(20): 5213-5226.

[4] Lu C Y, Tang J H, Yan S C, et al. Nonconvex nonsmooth low rank minimization via iteratively reweighted nuclear norm[J]. IEEE Transactions on Image Processing, 2016, 25(2): 829-839.

[5] Li G R, Guo G, Peng S C, et al. Matrix completion via schatten capped p norm[J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 34(1): 394-404.

[6] Xiong Z K, Wei Y M, Xu R J, et al. Low-rank traffic matrix completion with marginal information[J]. Journal of Computational and Applied Mathematics, 2022, 410: 114219.

[7] Fazel M. Matrix rank minimization with applications[D]. Stanford: Stanford University, 2002.

[8] Ma S Q, Goldfarb D, Chen L F. Fixed point and Bregman iterative methods for matrix rank minimization[J]. Mathematical Programming, 2011, 128: 321-353.

[9] Toh K C, Yun S. An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems[J]. Pacific Journal of Optimization, 2010, 6(3): 615-640.

[10] Kuang S F, Chao H Y, Li Q. Majorized proximal alternating imputation for regularized rank constrained matrix completion[J]. Journal of Computational and Applied Mathematics, 2020, 371: 112679.

[11] Candès E J, Recht B. Exact matrix completion via convex optimization[J]. Foundations of Computational Mathematics, 2009, 9(6): 717-772.

[12] Li Y F, Zhang Y J, Huang Z H. A reweighted nuclear norm minimization algorithm for low rank matrix recovery[J]. Journal of Computational and Applied Mathematics, 2014, 263: 338-350.

[13] Phan D N,Nguyen T N. An accelerated IRNN-Iteratively Reweighted Nuclear Norm algorithm for nonconvex nonsmooth low-rank minimization problems[J]. Journal of Computational and Applied Mathematics,2021,396:113602.

[14] Gao K X,Huang Z H,Guo L L. Low-rank matrix recovery problem minimizing a new ratio of two norms approximating the rank function then using an ADMM-type solver with applications[J]. Journal of Computational and Applied Mathematics,2024,438:115564.

[15] 彭建文,雷宏旺. 非凸两分块优化问题的一类惯性对称正则化交替方向乘子法[J]. 运筹学学报,2023,27(3):37-52.

[16] Guo K,Han D R,Wu T T. Convergence of alternating direction method for minimizing sum of two nonconvex functions with linear constraints[J]. International Journal of Computational Mathematics,2017,94(8):1653-1669.

[17] Bot R I,Nguyen D K. The proximal alternating direction method of multipliers in the non-convex setting:convergence analysis and rates[J]. Mathematics of Operations Research,2020,45(2):682-712.

[18] Wu Z M,Li M,Wang D Z W,et al. A symmetric alternating direction method of multipliers for separable nonconvex minimization problems[J]. Asia-Pacific Journal of Operational Research,2017,34(6):1750030.

[19] Wang Y,Yin W T,Zeng J S. Global convergence of ADMM in nonconvex nonsmooth optimization[J]. Journal of Scientific Computing,2019,78(1):29-63.

[20] Rahimi Y,Wang C,Dong H B,et al. A scale-invariant approach for sparse signal recovery[J]. SIAM Journal on Scientific Computing,2019,41(6):A3649-A3672.

[21] Beck A. First-Order methods in optimization[M]. Philadelphia:SIAM Society for Industrial and Applied Mathematics,2017.

[22] Wu Z M,Sun K X,Ge Z L,et al. Sparse portfolio optimization l_1 via l_2 over regularization[J]. European Journal of Operational Research,2024,319(3):820-833.

[23] Tao M,Zhang X P. Study on over minimization for nonnegative signal recovery[J]. Journal of Scientific Computing,2023,95:94-118.

[责任编辑:陆炳新]