Numerical Algorithms
https://doi.org/10.1007/s11075-018-0630-z

ORIGINAL PAPER

# Alternating direction and Taylor expansion minimization algorithms for unconstrained nuclear norm optimization

**Jianxi Zhao[1] · Qian Feng[2] · Lina Zhao[3]**

Received: 3 September 2018 / Accepted: 2 November 2018 / Published online: 12 November 2018


## Abstract

In the past decade, robust principal component analysis (RPCA) and low-rank matrix completion (LRMC), as two very important optimization problems with the view of recovering original low-rank matrix from sparsely and highly corrupted observations or a subset of its entries, have already been successfully adopted in image denoising, video processing, web search, biological information, etc. This paper proposes an efficient and effective algorithm, named the alternating direction and step size minimization (ADSM) algorithm, which employs the alternating direction minimization idea to solve the general relaxed model that can describe small noise (e.g., Gaussian noise). The coupling of sparse noise and small noise makes low-rank matrix recovery more challenging than that of RPCA. We make use of Taylor expansion, singular value decomposition and shrinkage operator as the alternating direction minimization method to deduce iterative direction matrices. A continuous technology is incorporated into ADSM to accelerate convergence. Similarly, the Taylor expansion and step size minimization (TESM) algorithm for LRMC is designed by the above way, but the alternating direction minimization idea needs to be ruled out since there is not a sparse matrix in it. Theoretically, it is proved that the two algorithms globally converge to their respective optimal points based on some conditions. The numerical results are reported, illustrating that ADSM and TESM are quite efficient and effective for recovering large-scale low-rank matrix problems at many cases.

**Keywords** Robust principal component analysis · Alternating direction minimization · Taylor expansion · Low-rank matrix completion

**Mathematics Subject Classification (2010)** 15A29 · 41A29 · 65F10 · 65J20 · 68Q25 · 90C26

✉ Lina Zhao
zhaoln@mail.buct.edu.cn

[1] School of Science, Beijing Information Science and Technology University, Beijing 100192, China

[2] School of Mathematics and Statistics, The University of Melbourne, Melbourne VIC 3010, Australia

[3] School of Science, Beijing University of Chemical Technology, Beijing 100029, China

# 1 Introduction

## 1.1 Problems and models

Suppose that a large data matrix D $\in R^{m\times n}$ is given and we know that it may be decomposed as a low-rank matrix A $\in R^{m\times n}$ and a sparse matrix E $\in R^{m\times n}$, namely, D = A+E, but we do not know the low-dimensional column and row spaces of A, even their dimensions. Moreover, the locations and the number of non-zero entries of E are also unknown. Our main purpose is to recover the low-rank and sparse components both accurately and efficiently. This problem named robust principal component analysis (RPCA) has intensively involved in the fields of face recognition [1], video processing [2], latent semantic indexing [3], ranking and collaborative filtering [4] and so on, whose data have routinely increased to thousands or even billions of dimensions. Let us use video surveillance as an example to specify the optimization problem. If a sequence of surveillance video's frames are given, active components often need to be identified from the background, namely, active and background components would be separated. We can arrange all the video's frames into columns which form the data matrix D. The low-rank component A naturally corresponds to the stationary background and the sparse component E denotes the active objects.

The low-rank matrix A and the sparse matrix E are described by two models: the nuclear norm minimization min $||A||_*$ [5, 6] and the $l_1$-norm minimization min $||E||_1$ [7, 8] respectively, where $||A||_*$ is the sum of the singular values of A and $||E||_1$ is the sum of the absolute values of all entries of E. In 2009, Wright et al. [1] combined min $||A||_*$ and $||E||_1$ with $\gamma = 1/\sqrt{\max(m, n)}$ [2] as a tractable form (1) which was the classical convex optimization model of RPCA.

$$\min_{A,E} ||A||_* + \gamma||E||_1, \; s.t. \; D = A + E. \tag{1}$$

The model (1) enables to correctly recover underlying low-rank structure in the presence of gross errors or outlying observations, or identify underlying sparse structure from the background.

A low-rank matrix A is contaminated by both sparse noise matrix E and small noise matrix N (e.g., Gaussian noise), namely, the observation data matrix D = A + E + N, so it becomes a major concern problem how to recover A from D.

This paper focuses on the following general convex relaxed structured minimization of the classical model (1) in the RPCA field:

$$\min_{A,E} F(A, E) = \min_{A,E} f(A, E) + \mu(||A||_* + \gamma||E||_1), \tag{2}$$

where $f : (R^{m\times n}, R^{m\times n}) \rightarrow R$ is a bivariate bounded continuous differentiable function, and the parameter $\mu > 0$ is used to trade off $f(A, E)$ and $||A||_* + \gamma||E||_1$ for minimization. Given its structure, the term $0.5||D-A-E||_F^2$ for describing small noise can be regarded as a special case of the general smooth function $f(A, E)$.

Recovering a rectangular matrix from a subset of its entries is known as matrix completion [9]. If there are not any additional conditions, the issue is apparently ill due to obtaining infinite solutions from few conditions. In many applications, we hope to recover a low-rank or approximate low-rank matrix with very limited information [9], for instance, the famous Netflix recommended system [4]: users only rate a few items, but one would like to infer their preferences from the incomplete rating matrix. The Netflix data matrix of all user-ratings may be approximately low-rank because it is commonly believed that only a few factors, such as subjects, directors, actors and so forth, contribute to anyone's taste or preference. The similar low-rank recovery problem with incomplete data is named low-rank matrix completion (LRMC) that is also applied to many other practical problems, such as system identification [10], remote sensing [11], video denoising [12], and illumination compensation [13].

In 2010, Recht et al. [5] showed that if a certain restricted isometry property held for the linear transformation defining the constraints, the solution about LRMC could be recovered by solving the model (3), which was the most classical and popular model in the LRMC field.

$$\min \ ||A||_*, \ s.t. \, \mathcal{P}_\Omega(A) = \mathcal{P}_\Omega(D), \tag{3}$$

where $\mathcal{P}_\Omega(\cdot)$ denotes an orthogonal projector onto the span of matrix vanishing outside of $\Omega$, in other words, if $(i, j) \in \Omega$, the $(i, j)$th entry of $\mathcal{P}_\Omega(A)$ is $A_{ij}$, otherwise, it is zero. The model (3) means how we recover the low-rank matrix A when the subset $\mathcal{P}_\Omega(A)$ of A has been known ($\mathcal{P}_\Omega(A) = \mathcal{P}_\Omega(D)$). We try to solve the following general relaxed structured minimization (4) of the model (3):

$$\min F(A) = \min f(A) + \mu||A||_*, \tag{4}$$

where $0.5||\mathcal{P}_\Omega(A) - \mathcal{P}_\Omega(D)||_F^2$ for describing small noise can be seen as a special case of $f(A)$.

### 1.2 Existing algorithms

In the aspect of RPCA, the off-the-shelf interior point methods can be applied to solve the semidefinite program [14] reformulated as the model (1). They show some effectiveness, but only in handling small-scale matrix whose size is $n \times n$ $(n \leq 100)$, due to its high-order complexity $O(n^6)$ where $n$ is the order of a square matrix. In pattern recognition, the sizes of matrices are so huge that the interior point methods have not satisfied the demand of many practical applications due to depending on the second-order information of the objective function essentially. Wright et al. [1] in 2009 presented the iterative thresholding (IT) algorithm, with only $O(n^3)$ complexity, relying on the first-order information for solving the model (1). It can compensate for the drawback that the interior point methods hardly solve large-scale matrix problems, but it converges slowly, resulting in needing much more time. At the same year, Lin et al. [15] developed a complementary method: accelerated proximal gradient (APG). This algorithm still depends on the first-order information for solving the model (2). It is faster and more scalable than IT by combining with a continuation technique, which can remove small noise in theory. Then, Lin et al. proposed that the augmented Lagrange multiplier method could be applied to solve the model (1) that didn't describe small noise. Although some people presented the other models without nuclear norm and the corresponding algorithms, the models were non-convex and the algorithms were not very robust in many cases and not globally convergent.

In the aspect of LRMC, in order to overcome the narrow limitation of interior point methods in terms of matrix size, Cai et al. [9] in 2010 presented the singular value thresholding (SVT) algorithm that shrank the singular values of a sparse matrix at each iteration. It cannot usually recover the matrices that have moderate or high rank efficiently. Considering the problems about both rank and size, Ma et al. in 2011 proposed a linear time approximate singular value decomposition based on fixed point continuation algorithm (FPCA) [16] that made use of an operator splitting technology and synthesized Bregman iterative algorithm, FPC algorithm and linear time approximate SVD [17]. Compared to SVT, FPCA is in the ascendant in terms of time and robustness. Inspired by the fast iterative shrinkage thresholding algorithm for linear inverse problems [18], Toh et al. [19] developed the APGL algorithm with a linesearch-like technology, which had the better iteration complexity than that of SVT and FPCA. The two algorithms APGL and FPCA for solving the model (4) can remove small noise to a certain extent. Further, Lin et al. used the augmented Lagrange multiplier method to solve the model (3) that didn't also describe small noise. Some researchers proposed the other models without nuclear norm, but they were non-convex, so the corresponding algorithms were not robust in many cases and not globally convergent.

Despite such exciting developments in the RPCA and LRMC fields, the current existing algorithms still lose some efficiency and effectiveness for large-scale matrix problems when removing small noise in some situations. Therefore, it is necessary to propose more exciting algorithms for the denoising problem in above fields.

### 1.3 Contributions and organization

The main contributions of this paper are as follows. We propose an algorithm named the alternating direction and step size minimization (ADSM) algorithm that uses the alternating direction minimization idea to solve the general relaxed model (2) in the RPCA field. This paper uses Taylor expansion, SVD, shrinkage operator, and so forth to deduce iterative direction matrices of the low-rank matrix and the sparse matrix. By combining this idea with the direction step size formula, we update the direction matrices and the corresponding step sizes alternately. The Taylor expansion and step size minimization (TESM) algorithm for LRMC is designed by the similar way without the alternating direction minimization idea. We prove their global convergence based on some conditions. Experimentally, compared with the current existing algorithms, the two proposed algorithms are very promising in running time, computational accuracy, robustness, etc. in some situations.

The rest of this paper is organized as below. In Section 2, we deduce the iterative direction matrices of the low-rank matrix and the sparse matrix. The basic steps of

ADSM and TESM are established respectively. In Section 3, it is given that the global convergence theorems in some conditions. In Section 4, we conduct some experiments to show the efficiency and effectiveness of the two algorithms. This paper is concluded briefly in the last section.

## 2 Algorithm analysis

This paper inspired by the non-monotone Barzilai-Borwein gradient algorithm [20] in the compressed sensing field proposes the two algorithms: ADSM for RPCA and TESM for LRMC. By combining the direction step size formulas $A_{k+1} = A_k + \alpha_k M_k$ and $E_{k+1} = E_k + \beta_k N_k$ with the alternating direction minimization idea [21], we deduce the two direction matrices $M_k$ and $N_k$. After the two corresponding step sizes $\alpha_k$ and $\beta_k$ are given, the next iterative matrices $M_k$ and $N_k$ can be computed easily.

The soft-thresholding operator belongs to the proximity operator whose details can be found in [22]. Theorem 1 about it, which has been introduced in [9], is used to support the derivation process of $M_k$.

**Theorem 1** *Let $A \in R^{m\times n}$ be a low-rank matrix, $\mathcal{D}_\tau(\cdot)$ be the soft-threshold operator and $\mathcal{S}_\tau(\cdot)$ be the shrinkage operator: $\mathcal{S}_\tau(x) = \max\{|x| - \tau, 0\}(x/|x|)$ where $\tau$ is a positive threshold value. The SVD of $Q \in R^{m\times n}$ is $U\Sigma V^T$ where $U \in R^{m\times r}$, $\Sigma \in R^{r\times r}$ and $V \in R^{n\times r}$ are the left orthogonal matrix, the singular matrix and the right orthogonal matrix respectively. Then, the following expression holds.*

$$A = \arg\min_{A} 0.5||A - Q||_F^2 + \tau||A||_* = \mathcal{D}_\tau(Q) = U\mathcal{S}_\tau(\Sigma)V^T.$$

At the $k$th iteration, we see $A_k$ and $E_k$ as the known matrices, and retain the first three terms of Taylor expansion of $f(A_k + M)$ at $A_k$ and the first two terms of Taylor expansion of $||A_k + M||_*$ at $A_k$. The first-order derivative form of $||A_k + M||_*$ at $A_k$ has three new unknown matrices due to the unknown matrix $M$. The difficulty can be overcome by approximated form of derivative definition. Specifically, the expanded form of $F(A_k + M)$ at $A_k$ is as follows:

$$\begin{aligned} & F(A_k + M) \\ & = f(A_k + M) + \mu(||A_k + M||_* + \gamma||E_k||_1) \\ & \approx f(A_k) + \langle \nabla f(A_k), M\rangle + \frac{\lambda_k}{2}||M||_F^2 + \mu\gamma||E_k||_1 + \mu(||A_k||_* \\ & \quad + \frac{||A_k + hM||_* - ||A_k||_*}{h}) \\ & \triangleq P_k(M), \end{aligned} \tag{5}$$

where both $h$ and $\lambda_k$ are small positive numbers and

$$\langle \nabla f(A_k), M\rangle = trace(\nabla f(A_k)^T M).$$

$$
\begin{aligned}
& A_k + hM_k \\
&= \arg \min_{A_k+hM} P_k(M) \\
&= \arg \min_{A_k+hM} \langle \nabla f(A_k), M\rangle + \frac{\lambda_k}{2}||M||_F^2 + \frac{\mu}{h}||A_k + hM||_* \\
&= \arg \min_{A_k+hM} \frac{h^2}{\lambda_k}(\langle \nabla f(A_k), M\rangle + \frac{\lambda_k}{2}||M||_F^2 + \frac{\mu}{h}||A_k + hM||_*) \\
&= \arg \min_{A_k+hM} \frac{1}{2}\langle A_k + hM - (A_k - \frac{h}{\lambda_k}\nabla f(A_k)), A_k + hM - (A_k - \frac{h}{\lambda_k}\nabla f(A_k))\rangle \\
&\quad + \frac{\mu h}{\lambda_k}||A_k + hM||_* \\
&= \arg \min_{A_k+hM} \frac{1}{2}||A_k + hM - (A_k - \frac{h}{\lambda_k}\nabla f(A_k))||_F^2 + \frac{\mu h}{\lambda_k}||A_k + hM||_* \\
&= \mathcal{D}_{\frac{\mu h}{\lambda_k}}(A_k - \frac{h}{\lambda_k}\nabla f(A_k)),
\end{aligned}
$$

where the last equality is supported by Theorem 1, so the formula (6) of $M_k$ can be obtained.

$$
M_k = \frac{1}{h}[\mathcal{D}_{\frac{\mu h}{\lambda_k}}(A_k - \frac{h}{\lambda_k}\nabla f(A_k)) - A_k]. \tag{6}
$$

Lemma 1 shows that the function $L(h)$ approaches the first-order term of Taylor expansion of $||A + M||_*$ at $A$ decreasingly and avoids appearing mutation as $h > 0$ decreases. In addition, Lemma 1 supports the proof of Theorem 2.

**Lemma 1** *For any two matrices* $A, M \in R^{m\times n}$ *, the function* $L(h)$ *increases monotonously as h increases.*

$$
L(h) = \frac{||A + hM||_* - ||A||_*}{h}, \quad h \in (0, +\infty).
$$

*Proof* For $\forall h_1,\ h_2 \in (0, +\infty)$ and $h_1 < h_2$,

$$
\begin{aligned}
& L(h_1) - L(h_2) \\
&= \frac{||A + h_1M||_* - ||A||_*}{h_1} - \frac{||A + h_2M||_* - ||A||_*}{h_2} \\
&= \frac{h_2||A + h_1M||_* - h_2||A||_* - h_1||A + h_2M||_* + h_1||A||_*}{h_1h_2} \\
&= \frac{||h_2A + h_1h_2M||_* - ||h_1A + h_1h_2M||_* + (h_1 - h_2)||A||_*}{h_1h_2} \\
&= \frac{||h_1A + h_1h_2M + h_2A - h_1A||_* - ||h_1A + h_1h_2M||_* + (h_1 - h_2)||A||_*}{h_1h_2} \\
&\le \frac{||h_1A + h_1h_2M||_* + ||h_2A - h_1A||_* - ||h_1A + h_1h_2M||_* + (h_1 - h_2)||A||_*}{h_1h_2} \\
&= 0,
\end{aligned}
$$

namely, $L(h_1) \le L(h_2)$, so $L(h)$ increases monotonously as $h$ increases. □

Theorem 2 shows that the direction matrix $M_k$ defined by the formula (6) is descent if $M_k \neq 0$.

**Theorem 2** *Suppose that* $\lambda_k > 0$ *and the direction matrix* $M_k$ *is defined by the formula (6). Then,* $\exists \theta \in (0, h]$ *such that*

$$F(A_k + \theta M_k) \leq F(A_k) + \theta(\langle \nabla f(A_k), M_k \rangle + \frac{\mu ||A_k + hM_k||_* - \mu ||A_k||_*}{h}) + o(\theta), \tag{7}$$

*and*

$$\langle \nabla f(A_k), M_k \rangle + \frac{\mu ||A_k + hM_k||_* - \mu ||A_k||_*}{h} \leq -\frac{\lambda_k}{2} ||M_k||_F^2. \tag{8}$$

*Proof* By differentiability of the smooth function $f(\cdot)$ and convexity of $||A||_*$, it can be shown that for $\forall \theta \in (0, h]$, namely, $\forall \theta / h \in (0, 1]$,

$$\begin{aligned}
& F(A_k + \theta M_k) - F(A_k) \\
&= f(A_k + \theta M_k) - f(A_k) + \mu ||A_k + \theta M_k||_* - \mu ||A_k||_* + \mu\gamma ||E||_1 - \mu\gamma ||E||_1 \\
&= f(A_k + \theta M_k) - f(A_k) + \mu || \frac{\theta}{h}(A_k + hM_k) + (1 - \frac{\theta}{h})A_k||_* - \mu ||A_k||_* \\
&\leq f(A_k + \theta M_k) - f(A_k) + \frac{\mu\theta}{h} ||A_k + hM_k||_* + \mu(1 - \frac{\theta}{h})||A_k||_* - \mu ||A_k||_* \\
&= f(A_k) + \langle \nabla f(A_k), \theta M_k \rangle + o(\theta) - f(A_k) + \frac{\mu\theta}{h} ||A_k + hM_k||_* + \mu ||A_k||_* \\
&\quad - \frac{\mu\theta}{h} ||A_k||_* - \mu ||A_k||_* \\
&= \theta \langle \nabla f(A_k), M_k \rangle + \frac{\mu\theta}{h} ||A_k + hM_k||_* - \frac{\mu\theta}{h} ||A_k||_* + o(\theta).
\end{aligned}$$

So the inequality (7) is proved.

Note that $M_k$ is the minimizer of the expression (5) and $\theta \in (0, h]$. By (5) and convexity of $||A||_*$,

$$\begin{aligned}
& \langle \nabla f(A_k), M_k \rangle + \frac{\lambda_k}{2} ||M_k||_F^2 + \frac{\mu ||A_k + hM_k||_* - \mu ||A_k||_*}{h} \\
&\leq \theta \langle \nabla f(A_k), M_k \rangle + \frac{\lambda_k \theta^2}{2} ||M_k||_F^2 + \frac{\mu}{h} ||A_k + \theta h M_k||_* - \frac{\mu}{h} ||A_k||_* \\
&= \theta \langle \nabla f(A_k), M_k \rangle + \frac{\lambda_k \theta^2}{2} ||M_k||_F^2 + \frac{\mu}{h} || \frac{\theta}{h} A_k + \theta h M_k + A_k - \frac{\theta}{h} A_k ||_* - \frac{\mu}{h} ||A_k||_* \\
&\leq \theta \langle \nabla f(A_k), M_k \rangle + \frac{\lambda_k \theta^2}{2} ||M_k||_F^2 + \frac{\mu\theta}{h^2} ||A_k + h^2 M_k||_* + \frac{\mu}{h}(1 - \frac{\theta}{h})||A_k||_* - \frac{\mu}{h} ||A_k||_*,
\end{aligned}$$

so

$$\begin{aligned}
& (1 - \theta) \langle \nabla f(A_k), M_k \rangle + \frac{\mu}{h} ||A_k + hM_k||_* - \frac{\mu\theta}{h^2} ||A_k + h^2 M_k||_* \\
& - \frac{\mu}{h}(1 - \frac{\theta}{h})||A_k||_* \leq -\frac{\lambda_k}{2}(1 - \theta^2)||M_k||_F^2.
\end{aligned} \tag{9}$$

The last three terms of the left side of the inequality in (9) can be arranged as

$$\begin{aligned}
&\frac{\mu}{h}||A_k + hM_k||_* - \frac{\mu\theta}{h^2}||A_k + h^2M_k||_* - \frac{\mu}{h}(1 - \frac{\theta}{h})||A_k||_* \\
&= \frac{\mu}{h}(||A_k + hM_k||_* - ||A_k||_* - \theta\frac{||A_k + h^2M_k||_* - ||A_k||_*}{h}) \\
&= \frac{\mu}{h}(||A_k + hM_k||_* - ||A_k||_* - \theta h\frac{||A_k + h^2M_k||_* - ||A_k||_*}{h^2}) \\
&\geq \frac{\mu}{h}(||A_k + hM_k||_* - ||A_k||_* - \theta h\frac{||A_k + hM_k||_* - ||A_k||_*}{h}) \\
&= \frac{\mu}{h}(||A_k + hM_k||_* - ||A_k||_* - \theta||A_k + hM_k||_* + \theta||A_k||_*) \\
&= \frac{\mu}{h}(1-\theta)(||A_k + hM_k||_* - ||A_k||_*)\ ,
\end{aligned} \tag{10}$$

where the inequality is supported by Lemma 1. By combining (9) with (10), we get the inequality

$$(1-\theta)\langle\nabla f(A_k), M_k\rangle + \frac{\mu}{h}(1-\theta)(||A_k + hM_k||_* - ||A_k||_*) \leq -\frac{\lambda_k}{2}(1-\theta^2)||M_k||_F^2,$$

namely,

$$\langle\nabla f(A_k), M_k\rangle + \frac{\mu}{h}(||A_k + hM_k||_* - ||A_k||_*) \leq -\frac{\lambda_k}{2}(1+\theta)||M_k||_F^2 \leq -\frac{\lambda_k}{2}||M_k||_F^2,$$

so the inequality (8) is also proved. □

Theorem 3 about the shrinkage operator, which has been introduced in [15], is used to support the derivation process of $N_k$.

**Theorem 3** *Let $\tau$ be a positive shrinkage thresholding, $E \in R^{m\times n}$ be a sparse matrix, $\mathcal{S}_\tau(\cdot)$ be the shrinkage operator and $Q \in R^{m\times n}$. Then,*

$$E = \arg\min_E 0.5||E - Q||_F^2 + \tau||E||_1 = \mathcal{S}_\tau(Q).$$

At the $k$th iteration, we see $A_{k+1}$ and $E_k$ as the known matrices and retain the first three terms of Taylor expansion of $f(E_k + N)$ and the first two terms of Taylor expansion of $||E_k + N||_1$ at $E_k$. The first-order derivative form of $||E_k + N||_1$ can be replaced by an approximated form of derivative definition due to non-differentiability of $||\cdot||_1$. The expanded form of $F(E_k + N)$ at $E_k$ is as follows:

$$\begin{aligned}
&F(E_k + N) \\
&= f(E_k + N) + \mu(||A_{k+1}||_* + \gamma||E_k + N||_1) \\
&\approx f(E_k) + \langle\nabla f(E_k), N\rangle + \frac{\rho_k}{2}||N||_F^2 + \mu||A_{k+1}||_* + \mu\gamma(||E_k||_1 \\
&\quad + \frac{||E_k + gN||_1 - ||E_k||_1}{g}) \\
&\triangleq Q_k(N),
\end{aligned} \tag{11}$$

where both $g$ and $\rho_k$ are small positive numbers.

$$
\begin{aligned}
& E_k + gN_k \\
= & \arg \min_{E_k+gN} Q_k(N) \\
= & \arg \min_{E_k+gN} \langle \nabla f(E_k), N \rangle + \frac{\rho_k}{2}||N||_F^2 + \frac{\mu\gamma}{g}||E_k + gN||_1 \\
= & \arg \min_{E_k+gN} \frac{g^2}{\rho_k}(\langle \nabla f(E_k), N \rangle + \frac{\rho_k}{2}||N||_F^2 + \frac{\mu\gamma}{g}||E_k + gN||_1) \\
= & \arg \min_{E_k+gN} \frac{1}{2}\langle E_k + gN - (E_k - \frac{g}{\rho_k}\nabla f(E_k)), E_k + gN - (E_k - \frac{g}{\rho_k}\nabla f(E_k))\rangle \\
& + \frac{\gamma\mu g}{\rho_k}||E_k + gN||_1 \\
= & \arg \min_{E_k+gN} \frac{1}{2}||E_k + gN - (E_k - \frac{g}{\rho_k}\nabla f(E_k))||_F^2 + \frac{\gamma\mu g}{\rho_k}||E_k + gN||_1 \\
= & \mathcal{S}_{\frac{\gamma\mu g}{\rho_k}}(E_k - \frac{g}{\rho_k}\nabla f(E_k)),
\end{aligned}
$$

where the last equality is supported by Theorem 3, so the formula (12) of $N_k$ can be obtained.

$$
N_k = \frac{1}{g}[\mathcal{S}_{\frac{\gamma\mu g}{\rho_k}}(E_k - \frac{g}{\rho_k}\nabla f(E_k)) - E_k]. \tag{12}
$$

Lemma 2 shows that the function $L(g)$ decreasingly approaches the first-order term of Taylor expansion of $||E_k + N||_1$ at $E_k$ and avoids appearing mutation as $g$ decreases. Furthermore, Lemma 2 supports the proof of Theorem 4.

**Lemma 2** *For any two matrices $E, N \in R^{m\times n}$ , the function $L(g)$ increases monotonously as g increases.*

$$
L(g) = \frac{||E + gN||_1 - ||E||_1}{g}, \quad g \in (0, +\infty).
$$

*Proof* For $\forall g_1,\ g_2 \in (0, +\infty)$ and making $g_1 < g_2$,

$$
\begin{aligned}
& L(g_1) - L(g_2) \\
= & \frac{||E + g_1N||_1 - ||E||_1}{g_1} - \frac{||E + g_2N||_1 - ||E||_1}{g_2} \\
= & \frac{g_2||E + g_1N||_1 - g_2||E||_1 - g_1||E + g_2N||_1 + g_1||E||_1}{g_1g_2} \\
= & \frac{||g_2E + g_1g_2N||_1 - ||g_1E + g_1g_2N||_1 + (g_1 - g_2)||E||_1}{g_1g_2} \\
= & \frac{||g_1E + g_1g_2N + g_2E - g_1E||_1 - ||g_1E + g_1g_2N||_1 + (g_1 - g_2)||E||_1}{g_1g_2} \\
\le & \frac{||g_1E + g_1g_2N||_1 + ||g_2E - g_1E||_1 - ||g_1E + g_1g_2N||_1 + (g_1 - g_2)||E||_1}{g_1g_2} \\
= & 0,
\end{aligned}
$$

namely, $L(g_1) \le L(g_2)$, so $L(g)$ increases as $g$ increases. □

Theorem 4 shows that the direction matrix $N_k$ defined by the formula (12) is descent if $N_k \neq 0$.

**Theorem 4** *Suppose that* $\rho_k > 0$ *and the direction matrix* $N_k$ *is defined by the formula (12). Then,* $\exists \eta \in (0, g]$ *such that*

$$\begin{aligned} &F(E_k + \eta N_k) \\ &\leq F(E_k) + \eta(\langle \nabla f(E_k), N_k\rangle + \frac{\mu\gamma||E_k + gN_k||_1 - \mu\gamma||E_k||_1}{g}) + o(\eta), \end{aligned} \tag{13}$$

*and*

$$\langle \nabla f(E_k), N_k\rangle + \frac{\mu\gamma||E_k + gN_k||_1 - \mu\gamma||E_k||_1}{g} \leq -\frac{\rho_k}{2}||N_k||_F^2. \tag{14}$$

*Proof* By differentiability of $f(\cdot)$ and convexity of $||E||_1$, it can be shown that for $\forall \eta \in (0, g]$, namely, $\forall \eta/g \in (0, 1]$,

$$\begin{aligned} &F(E_k + \eta N_k) - F(E_k) \\ &= f(E_k + \eta N_k) - f(E_k) + \mu||A_{k+1}||_* - \mu||A_{k+1}||_* + \mu\gamma||E_k + \eta N_k||_1 - \mu\gamma||E_k||_1 \\ &= f(E_k + \eta N_k) - f(E_k) + \mu\gamma||\frac{\eta}{g}(E_k + gN_k) + (1 - \frac{\eta}{g})E_k||_1 - \mu\gamma||E_k||_1 \\ &\leq f(E_k) + \langle \nabla f(E_k), \eta N_k\rangle + o(\eta) - f(E_k) + \frac{\mu\gamma\eta}{g}||E_k + \eta N_k||_1 + \mu\gamma||E_k||_1 \\ &\quad - \frac{\mu\gamma\eta}{g}||E_k||_1 - \mu\gamma||E_k||_1 \\ &= \eta\langle \nabla f(E_k), N_k\rangle + o(\eta) + \frac{\mu\gamma\eta}{g}||E_k + \eta N_k||_1 - \frac{\mu\gamma\eta}{g}||E_k||_1, \end{aligned}$$

so the inequality (13) is proved.

Note that $N_k$ is the minimizer of the form (11) and $\eta \in (0, g]$. By (11) and convexity of $||E||_1$,

$$\begin{aligned} &\langle \nabla f(E_k), N_k\rangle + \frac{\rho_k}{2}||N_k||_F^2 + \frac{\mu\gamma||E_k + gN_k||_1 - \mu\gamma||E_k||_1}{g} \\ &\leq \eta\langle \nabla f(E_k), N_k\rangle + \frac{\rho_k\eta^2}{2}||N_k||_F^2 + \frac{\mu\gamma}{g}||E_k + \eta gN_k||_1 - \frac{\mu\gamma}{g}||E_k||_1 \\ &= \eta\langle \nabla f(E_k), N_k\rangle + \frac{\rho_k\eta^2}{2}||N_k||_F^2 + \frac{\mu\gamma}{g}||\frac{\eta}{g}E_k + \eta gN_k + E_k - \frac{\eta}{g}E_k||_1 - \frac{\mu\gamma}{g}||E_k||_1 \\ &\leq \eta\langle \nabla f(E_k), N_k\rangle + \frac{\rho_k\eta^2}{2}||N_k||_F^2 + \frac{\mu\gamma}{g}||\frac{\eta}{g}E_k + \eta gN_k||_1 + \frac{\mu\gamma}{g}(1 - \frac{\eta}{g})||E_k||_1 - \frac{\mu\gamma}{g}||E_k||_1 \\ &= \eta\langle \nabla f(E_k), N_k\rangle + \frac{\rho_k\eta^2}{2}||N_k||_F^2 + \frac{\mu\gamma}{g}||\frac{\eta}{g}E_k + \eta gN_k||_1 - \frac{\mu\gamma\eta}{g^2}||E_k||_1, \end{aligned}$$

so

$$\begin{aligned} &(1 - \eta)\langle \nabla f(E_k), N_k\rangle + \frac{\mu\gamma}{g}||E_k + gN_k||_1 - \frac{\mu\gamma\eta}{g^2}||E_k + g^2N_k||_1 \\ &- \frac{\mu\gamma}{g}(1 - \frac{\eta}{g})||E_k||_1 \leq -\frac{\rho_k}{2}(1 - \eta^2)||N_k||_F^2. \end{aligned} \tag{15}$$

The last three terms of the left side of the inequality in (15) can be arranged as

$$
\begin{aligned}
&\frac{\mu\gamma}{g}[||E_k + gN_k||_1 - \frac{\eta}{g}||E_k + g^2 N_k||_1 - (1 - \frac{\eta}{g})||E_k||_1] \\
&= \frac{\mu\gamma}{g}[||E_k + gN_k||_1 - ||E_k||_1 - \eta g\frac{||E_k + g^2 N_k||_1 - ||E_k||_1}{g^2}] \\
&\geq \frac{\mu\gamma}{g}[||E_k + gN_k||_1 - ||E_k||_1 - \eta g\frac{||E_k + gN_k||_1 - ||E_k||_1}{g}] \\
&= \frac{\mu\gamma}{g}[||E_k + gN_k||_1 - ||E_k||_1 - \eta||E_k + gN_k||_1 + \eta||E_k||_1] \\
&= \frac{\mu\gamma}{g}[(1-\eta)||E_k + gN_k||_1 - (1-\eta)||E_k||_1] \\
&= \frac{\mu\gamma}{g}(1-\eta)(||E_k + gN_k||_1 - ||E_k||_1),
\end{aligned}
\tag{16}
$$

where the inequality is supported by Lemma 2. By combining (15) with (16), we get the inequality

$$
(1-\eta)\langle\nabla f(E_k), N_k\rangle + \frac{\mu\gamma}{g}(1-\eta)(||E_k + gN_k||_1 - ||E_k||_1) \leq -\frac{\rho_k}{2}(1-\eta^2)||N_k||_F^2,
$$

namely,

$$
\langle\nabla f(E_k), N_k\rangle + \frac{\mu\gamma}{g}(||E_k + gN_k||_1 - ||E_k||_1) \leq -\frac{\rho_k}{2}(1+\eta)||N_k||_F^2 \leq -\frac{\rho_k}{2}||N_k||_F^2,
$$

so the inequality (14) is also proved. □

The large values $\mu h/\lambda_k$ in (6) and $\gamma\mu g/\rho_k$ in (12) are desired to shrink singular values and all entries of matrix respectively. Hale et al. [23] use a continuation technology (17) to dynamically adjust $\mu$ for accelerating convergence. They define a decreasing sequence $\{\mu_k\}$, as opposed to fixing the two terms $\bar{\mu}h/\lambda_k$ and $\gamma\bar{\mu}g/\rho_k$. When the model (2) associated with the next $\mu_{k+1}$ is to be solved, the approximate solution $(A(\mu_k), E(\mu_k))$ about the current $\mu_k$ is used as the starting point in the iteration. In fact, this framework approximately follows the path $(A(\mu), E(\mu))$ in the interval $[\bar{\mu},\ \mu_0]$.

$$
\mu_{k+1} = \max\{\tau\mu_k, \bar{\mu}\}, \quad k = 0, 1, 2, \ldots, L-1. \tag{17}
$$

According to all the above derivations, the basic steps of ADSM are designed as Algorithm 1 based on the alternating direction minimization idea. Similarly, TESM is also proposed by the way, but the idea is not adopted due to non-existing sparse matrix and the formula (6) of the direction matrix is deduced from the model (4) instead of (2). Its basic steps are designed as Algorithm 2.

**Algorithm 1** Alternating Direction and Step Size Minimization

**Step 0: Initialization.** The initial point $(A_0, E_0)$, the initial parameter $\mu_0 > 0$, the constant $\tau > 1$ and the stopping threshold $tol > 0$ are given. Set the iteration $k = 0$.
**Step 1: Stopping criterion.** If $||D - A_{k+1} - E_{k+1}||_F/||D||_F < tol$, stop; otherwise, continue.
**Step 2:** Compute the direction matrix $M_k$ of $A_k$ by (6).
**Step 3:** Select the optimal step size $\alpha_k$ of $A_k$.
**Step 4:** Update the low-rank matrix $A_{k+1} = A_k + \alpha_k M_k$.
**Step 5:** Compute the direction matrix $N_k$ of $E_k$ by (12).
**Step 6:** Select the optimal step size $\beta_k$ of $E_k$.
**Step 7:** Update the sparse matrix $E_{k+1} = E_k + \beta_k N_k$.
**Step 8:** Update the penalty parameter $\mu_{k+1}$ by (17).
**Step 9: Loop.** Let $k = k + 1$. Go to Step 1.
**Step 10: Output.** $(\hat{A}, \hat{E}) = (A_{k+1}, E_{k+1})$.

**Algorithm 2** Taylor Expansion and Step Size Minimization

**Step 0: Initialization.** The initial point $A_0$, the initial parameter $\mu_0 > 0$, the constant $\tau > 1$ and the stopping threshold $tol > 0$ are given. Set the iteration $k = 0$.
**Step 1: Stopping criterion.** If $||A_{k+1} - A_k||_F/||A_k||_F < tol$, stop; otherwise, continue.
**Step 2:** Compute the direction matrix $M_k$ of $A_k$ by (6) .
**Step 3:** Select the optimal step size $\alpha_k$ of $A_k$.
**Step 4:** Update the low-rank matrix $A_{k+1} = A_k + \alpha_k M_k$.
**Step 5:** Update the penalty parameter $\mu_{k+1}$ by (17).
**Step 6: Loop.** Let $k = k + 1$. Go to Step 1.
**Step 7: Output.** $\hat{A} = A_{k+1}$.

## 3 Convergence analysis

If the step sizes $\alpha_k$ of $A_k$ and $\beta_k$ of $E_k$ are determined by the non-monotone line search method [24], the convergence of ADSM and TESM can be analyzed on the basis of Assumption 1 and Assumption 2 respectively, which is inspired by the literature [21].

**Assumption 1** *The level set* $\Omega = \{(A, E) : F(A, E) \leq F(A_0, E_0)\}$ *is bounded.*

**Lemma 3** *Suppose that the direction matrices* $M_k$ *and* $N_k$ *are defined by (6) and (12) respectively, where* $\lambda_k, \rho_k > 0$, $h, g \in (0, 1]$, *and the step sizes* $\alpha_k, \beta_k > 0$. *Then,* $F(A_k, E_k) \geq F(A_{k+1}, E_{k+1})$.

*Proof* On the one hand, according to the inequality (8), we get

$$\alpha_k\langle\nabla f(A_k), M_k\rangle + \frac{\lambda_k\alpha_k^2}{2}||M_k||_F^2 + \frac{\mu_k}{h}||A_k + \alpha_k h M_k||_* \leq \frac{\mu_k}{h}||A_k||_*. \tag{18}$$

Combining Taylor expansion of $f(A_k + \alpha_k M_k)$ at $A_k$.

$$f(A_k + \alpha_k M_k) - f(A_k) = \alpha_k\langle\nabla f(A_k), M_k\rangle + o(\alpha_k)$$

with (18),

$$\begin{aligned}
&\frac{\partial F(A_k; M_k)}{\partial A_k}\\
&= \lim_{\alpha_k\to 0}\frac{F(A_k + \alpha_k M_k) - F(A_k)}{\alpha_k}\\
&= \lim_{\alpha_k\to 0}\frac{1}{\alpha_k}[f(A_k + \alpha_k M_k) + \mu_k(||A_k + \alpha_k M_k||_* + \gamma||E||_1) - f(A_k)\\
&\quad - \mu_k(||A_k||_* + \gamma||E_k||_1)]\\
&= \lim_{\alpha_k\to 0}\frac{1}{\alpha_k}[\alpha_k\langle\nabla f(A_k), M_k\rangle + o(\alpha_k) + \mu_k||A_k + \alpha_k M_k||_* - \mu_k||A_k||_*]\\
&\leq \lim_{\alpha_k\to 0}\frac{1}{\alpha_k}[-\frac{\lambda_k\alpha_k^2}{2}||M_k||_F^2 + o(\alpha_k) + (\mu_k||A_k + \alpha_k M_k||_* - \mu_k||A_k||_*)\\
&\quad + (\frac{\mu_k}{h}||A_k||_* - \frac{\mu_k}{h}||A_k + \alpha_k h M_k||_*)]\\
&= \lim_{\alpha_k\to 0}\frac{-\frac{\lambda_k\alpha_k^2}{2}||M_k||_F^2 + o(\alpha_k)}{\alpha_k} + 0\\
&= 0.
\end{aligned}$$

So $F(A_k; M_k)$ decreases monotonously as $k \to +\infty$, namely,

$$F(A_k, E_k) \geq F(A_{k+1}, E_k).$$

On the other hand, according to the inequality (14), we get

$$\beta_k\langle\nabla f(E_k), N_k\rangle + \frac{\rho_k\beta_k^2}{2}||N_k||_F^2 + \frac{\mu_k\gamma}{g}||E_k + \mathrm{g}\beta_k N_k||_1 \leq \frac{\mu_k\gamma}{g}||E_k||_1. \tag{19}$$

Combining Taylor expansion of $f(E_k + \beta_k N_k)$ at $E_k$

$$f(E_k + \beta_k N_k) - f(E_k) = \beta_k\langle\nabla f(E_k), N_k\rangle + o(\beta_k)$$

with (19),

$$
\begin{aligned}
&\frac{\partial F(E_k; N_k)}{\partial E_k} \\
&= \lim_{\beta_k \to 0} \frac{F(E_k + \beta_k N_k) - F(E_k)}{\beta_k} \\
&= \lim_{\beta_k \to 0} \frac{1}{\beta_k}[f(E_k + \beta_k N_k) + \mu_k(||A_k||_* + \gamma||E_k + \beta_k N_k||_1) - f(E_k) \\
&\quad - \mu_k(||A_k||_* + \gamma||E_k||_1)] \\
&= \lim_{\beta_k \to 0} \frac{1}{\beta_k}[\beta_k\langle \nabla f(E_k), N_k\rangle + o(\beta_k) + \mu_k\gamma||E_k + \beta_k N_k||_1 - \mu_k\gamma||E_k||_1] \\
&\leq \lim_{\beta_k \to 0} \frac{1}{\beta_k}[-\frac{\rho_k\beta_k^2}{2}||N_k||_F^2 + o(\beta_k) + (\mu_k\gamma||E_k + \beta_k N_k||_1 - \mu_k\gamma||E_k||_1) \\
&\quad + (\frac{\mu_k\gamma}{g}||E_k||_1 - \frac{\mu_k\gamma}{g}||E_k + g\beta_k N_k||_1)] \\
&= \lim_{\beta_k \to 0} \frac{-\frac{\rho_k\beta_k^2}{2}||N_k||_F^2 + o(\beta_k)}{\beta_k} + 0 \\
&= 0.
\end{aligned}
$$

So $F(E_k; N_k)$ decreases monotonously as $k \to +\infty$, namely,

$$F(A_{k+1}, E_k) \geq F(A_{k+1}, E_{k+1}).$$

By combining $F(A_k, E_k) \geq F(A_{k+1}, E_k)$ and $F(A_{k+1}, E_k) \geq F(A_{k+1}, E_{k+1})$, we have the inequality $F(A_k, E_k) \geq F(A_{k+1}, E_{k+1})$. □

**Lemma 4** *Let $l_1(k)$ be an integer such that*

$$k - m_1(k) \leq l_1(k) \leq k \text{ and } F(A_{l_1(k)}) = \max_{0 \leq j \leq m_1(k)} F(A_{k-j})$$

*and let $l_2(k)$ be an integer such that*

$$k - m_2(k) \leq l_2(k) \leq k \text{ and } F(E_{l_2(k)}) = \max_{0 \leq j \leq m_2(k)} F(E_{k-j}),$$

*where*

$$\begin{cases} m_1(0) = 0, \ 0 \leq m_1(k) \leq \min\{m_1(k-1) + 1, \ \widetilde{m_1}\}, \\ m_2(0) = 0, \ 0 \leq m_2(k) \leq \min\{m_2(k-1) + 1, \ \widetilde{m_2}\}. \end{cases}$$

*If the step sizes $\alpha_k$ and $\beta_k$ are determined by the non-monotone line search method* [24]*, the direction matrices $M_k$ and $N_k$ satisfy*

$$\lim_{k \to \infty} \alpha_k||M_k||_F = 0 \text{ and } \lim_{k \to \infty} \beta_k||N_k||_F = 0$$

*respectively.*

*Proof* The non-monotone line search method is used to determine the step sizes $\alpha_k$ and $\beta_k$, please see the formulas (20)–(23).

Let $\delta_1 \in (0, 1)$, $\rho_1 \in (0, 1)$, and $\tilde{m}_1$ be a positive integer. The method is to choose the smallest non-negative integer $j_k^1$ so that the step size $\alpha_k = \tilde{\alpha}\rho_1^{j_k^1}$ satisfies

$$F(A_k + \alpha_k M_k) \leq \max_{0 \leq j_k^1 \leq m_1(k)} F(A_{k-j_k^1}) + \delta_1 \alpha_k \Delta_k^1, \tag{20}$$

where

$$\Delta_k^1 = \langle \nabla f(A_k), M_k \rangle + \frac{\mu_k ||A_k + hM_k||_* - \mu_k ||A_k||_*}{h}, \tag{21}$$

$$m_1(0) = 0,\ 0 \leq m_1(k) \leq \min\{m_1(k-1) + 1, \tilde{m}_1\}.$$

The inequality (8) apparently shows if $M_k \neq 0$,

$$\Delta_k^1 \quad \leq -\frac{\lambda_k}{2} ||M_k||_F^2 < 0$$

Let $\delta_2 \in (0, 1)$, $\rho_2 \in (0, 1)$, and $\tilde{m}_2$ be a positive integer. It is to choose the smallest non-negative integer $j_k^2$ so that the step size $\beta_k = \tilde{\beta}\rho_2^{j_k^2}$ satisfies

$$F(E_k + \beta_k N_k) \leq \max_{0 \leq j_k^2 \leq m_2(k)} F(E_{k-j_k^2}) + \delta_2 \beta_k \Delta_k^2, \tag{22}$$

where

$$\Delta_k^2 = \langle \nabla f(E_k), N_k \rangle + \frac{\gamma \mu_k ||E_k + gN_k||_1 - \gamma \mu_k ||E_k||_1}{g}, \tag{23}$$

$$m_2(0) = 0,\ 0 \leq m_2(k) \leq \min\{m_2(k-1) + 1, \tilde{m}_2\}.$$

The inequality (14) apparently shows that if $N_k \neq 0$,

$$\Delta_k^2 \quad \leq -\frac{\rho_k}{2} ||N_k||_F^2 < 0.$$

By the inequalities (20) and (22), for $\forall k > \max\{\widetilde{m_1}, \widetilde{m_2}\}$,

$$\begin{aligned}
&F(A_{l_1(k)}, E_{l_2(k)}) \\
&= F(A_{l_1(k)-1} + \alpha_{l_1(k)-1}\Delta^1_{l_1(k)-1}, E_{l_2(k)-1} + \beta_{l_2(k)-1}\Delta^2_{l_2(k)-1}) \\
&\leq \max_{0 \leq j_1 \leq m_1(l_1(k)-1)} F(A_{l_1(k)-1-j_1}, E_{l_2(k)-1} + \beta_{l_2(k)-1}\Delta^2_{l_2(k)-1}) + \delta_1 \alpha_{l_1(k)-1}\Delta^1_{l_1(k)-1} \\
&\leq \max_{0 \leq j_1 \leq m_1(l_1(k)-1)} [\max_{0 \leq j_2 \leq m_2(l_2(k)-1)} F(A_{l_1(k)-1-j_1}, E_{l_2(k)-1-j_2}) + \delta_2 \beta_{l_2(k)-1}\Delta^2_{l_2(k)-1}] \\
&\quad + \delta_1 \alpha_{l_1(k)-1}\Delta^1_{l_1(k)-1} \\
&= \max_{0 \leq j_1 \leq m_1(l_1(k)-1)} \max_{0 \leq j_2 \leq m_2(l_2(k)-1)} F(A_{l_1(k)-1-j_1}, E_{l_2(k)-1-j_2}) + \delta_1 \alpha_{l_1(k)-1}\Delta^1_{l_1(k)-1} \\
&\quad + \delta_2 \beta_{l_2(k)-1}\Delta^2_{l_2(k)-1} \\
&= F(A_{l_1(l_1(k)-1)}, E_{l_2(l_2(k)-1)}) + \delta_1 \alpha_{l_1(k)-1}\Delta^1_{l_1(k)-1} + \delta_2 \beta_{l_2(k)-1}\Delta^2_{l_2(k)-1}\ .
\end{aligned}$$

According to Assumption 1, the sequence $\{F(A_{l_1(k)}, E_{l_2(k)})\}$ converges to a limit as $k \to \infty$.

From $\alpha_k = \widetilde{\alpha}\rho_1^{j_k^1} > 0$, $\beta_k = \widetilde{\beta}\rho_2^{j_k^2} > 0$, $\Delta^1_{l_1(k)-1} \leq 0$, $\Delta^2_{l_2(k)-1} \leq 0$, we get

$$0 \leq \lim_{k \to \infty} \delta_1 \alpha_{l_1(k)-1}\Delta^1_{l_1(k)-1} + \lim_{k \to \infty} \delta_2 \beta_{l_2(k)-1}\Delta^2_{l_2(k)-1} \leq 0.$$

Thus,

$$\lim_{k\to\infty} \delta_1 \alpha_{l_1(k)-1} \Delta^1_{l_1(k)-1} + \lim_{k\to\infty} \delta_2 \alpha_{l_2(k)-1} \Delta^2_{l_2(k)-1} = 0.$$

According to $\lim_{k\to\infty} \delta_1 \alpha_{l_1(k)-1} \Delta^1_{l_1(k)-1} \le 0$ and $\lim_{k\to\infty} \delta_2 \beta_{l_2(k)-1} \Delta^2_{l_2(k)-1} \le 0$, we get

$$\lim_{k\to\infty} \delta_1 \alpha_{l_1(k)-1} \Delta^1_{l_1(k)-1} = 0, \tag{24}$$

$$\lim_{k\to\infty} \delta_2 \beta_{l_2(k)-1} \Delta^2_{l_2(k)-1} = 0. \tag{25}$$

From (24), (25) and the analysis of Grippo et al. in [24], we can deduce the following conclusions: the direction matrices $M_k$ and $N_k$ satisfy

$$\lim_{k\to\infty} \alpha_k ||M_k||_F = 0 \text{ and } \lim_{k\to\infty} \beta_k ||N_k||_F = 0.$$

respectively. □

Based on Assumption 1, Lemma 3 and Lemma 4, the global convergence of ADSM that is shown in Theorem 5 is proved.

**Theorem 5** *Let sequences $\{(A_k, E_k)\}$, $\{M_k\}$, and $\{N_k\}$ be generalized by Algorithm 1. Then, the sequence $\{(A_k, E_k)\}$ converges to the globally optimal solution $(A^*, E^*)$ of the general model (2).*

*Proof* According to $\lim_{k\to\infty} \alpha_k ||M_k||_F = 0$ in Lemma 4, we get

$$\lim_{k\to\infty} ||M_k||_F = 0$$

or

$$\lim_{k\to\infty} ||M_k||_F \ne 0 \text{ and } \lim_{k\to\infty} \alpha_k = 0.$$

When $\lim_{k\to\infty} ||M_k||_F = 0$, the model $A_{k+1} = A_k + \alpha_k M_k$ yields that

$\forall \varepsilon > 0$, $\exists N \in N^*$, $\forall k > N$, $||A_{k+1} - A_k||_F < \varepsilon$. By Cauchy's test for convergence, $\lim_{k\to\infty} A_k = A^*$.

When $\lim_{k\to\infty} ||M_k||_F \ne 0$ and $\lim_{k\to\infty} \alpha_k = 0$, because $\alpha_k$ is the first value satisfying (20), $\exists \bar{k}_1 \in N^*$ such that for $\forall k \ge \bar{k}_1$, we get (Lemma 3)

$$F(A_k + \frac{\alpha_k}{\rho_1} M_k) > \max_{0\le j\le m_1(k)} F(A_{k-j}) + \delta_1 \frac{\alpha_k}{\rho_1} \Delta^1_k \ge F(A_k) + \delta_1 \frac{\alpha_k}{\rho_1} \Delta^1_k. \tag{26}$$

Because of the smoothness of $f(\cdot)$, we can Taylor expand it at $A_k$ such that

$$f(A_k + \frac{\alpha_k}{\rho_1} M_k) - f(A_k) = \frac{\alpha_k}{\rho_1} \langle \nabla f(A_k + \theta_k \frac{\alpha_k}{\rho_1} M_k), M_k \rangle, \tag{27}$$

where $\theta_k \in (0, 1)$ is a constant.

Combining (26) and (27),

$$\frac{\alpha_k}{\rho_1} \langle \nabla f(A_k + \theta_k \frac{\alpha_k}{\rho_1} M_k), M_k \rangle + \mu_k ||A_k + \frac{\alpha_k}{\rho_1} M_k||_* - \mu_k ||A_k||_* > \delta_1 \Delta^1_k \frac{\alpha_k}{\rho_1},$$

namely,

$$\langle \nabla f(A_k + \theta_k \frac{\alpha_k}{\rho_1} M_k), M_k \rangle + \frac{\mu_k ||A_k + \frac{\alpha_k}{\rho_1} M_k||_* - \mu_k ||A_k||_*}{\frac{\alpha_k}{\rho_1}} > \delta_1 \Delta_k^1.$$

Let $\widetilde{\alpha} = h$. Because $\alpha_k \to 0$ as $k \to \infty$, we obtain $\alpha_k < \rho_1 h$ as $k \to \infty$. Based on Lemma 1, the following inequality can be obtained.

$$\frac{\mu_k ||A_k + \frac{\alpha_k}{\rho_1} M_k||_* - \mu_k ||A_k||_*}{\frac{\alpha_k}{\rho_1}} - \frac{\mu_k ||A_k + h M_k||_* - \mu_k ||A_k||_*}{h} \leq 0.$$

$$\begin{aligned} &\langle \nabla f(A_k + \theta_k \frac{\alpha_k}{\rho_1} M_k), M_k \rangle - \langle \nabla f(A_k), M_k \rangle \\ &\geq \langle \nabla f(A_k + \theta_k \frac{\alpha_k}{\rho_1} M_k), M_k \rangle - \langle \nabla f(A_k), M_k \rangle \\ &\quad + \frac{\mu_k ||A_k + \frac{\alpha_k}{\rho_1} M_k||_* - \mu_k ||A_k||_*}{\frac{\alpha_k}{\rho_1}} - \frac{\mu_k ||A_k + h M_k||_* - \mu_k ||A_k||_*}{h} \\ &> \delta_1 \Delta_k^1 - \Delta_k^1 \\ &= -(1 - \delta_1) \Delta_k^1 \\ &\geq (1 - \delta_1) \frac{\lambda_{(\min)}}{2} ||M_k||_F^2. \end{aligned} \tag{28}$$

By taking the limits in both sides of the inequality (28), we get

$$0 > (1 - \delta_1) \frac{\lambda_{(\min)}}{2} \lim_{k \to \infty} ||M_k||_F^2 > 0,$$

namely, $\lim_{k \to \infty} ||M_k||_F = 0$. So $\lim_{k \to \infty} A_k = A^*$.

According to $\lim_{k \to \infty} \beta_k ||N_k||_F = 0$ in Lemma 4, we get

$$\lim_{k \to \infty} ||N_k||_F = 0$$

or

$$\lim_{k \to \infty} ||N_k||_F \neq 0 \text{ and } \lim_{k \to \infty} \beta_k = 0.$$

When $\lim_{k \to \infty} ||N_k||_F = 0$, the model $E_{k+1} = E_k + \beta_k N_k$ yields that

$\forall \varepsilon' > 0$, $\exists N' \in N^*$, $\forall k > N'$, $||E_{k+1} - E_k||_F < \varepsilon'$. By Cauchy's test for convergence, $\lim_{k \to \infty} E_k = E^*$.

When $\lim_{k \to \infty} ||N_k||_F \neq 0$ and $\lim_{k \to \infty} \beta_k = 0$, because $\beta_k$ is the first value satisfying (22), $\exists \overline{k}_2 \in N^*$ such that for $\forall k \geq \overline{k}_2$, we get (Lemma 3)

$$F(E_k + \frac{\beta_k}{\rho_2} N_k) > \max_{0 \leq j \leq m_2(k)} F(E_{k-j}) + \delta_2 \frac{\beta_k}{\rho_2} \Delta_k^2 \geq F(E_k) + \delta_2 \frac{\beta_k}{\rho_2} \Delta_k^2. \tag{29}$$

Because of the smoothness of $f(\cdot)$, we can Taylor expand it at $E_k$ such that

$$f(E_k + \frac{\beta_k}{\rho_2} N_k) - f(E_k) = \frac{\beta_k}{\rho_2} \langle \nabla f(E_k + \theta_k' \frac{\beta_k}{\rho_2} N_k), N_k \rangle, \tag{30}$$

where $\theta_k' \in (0, 1)$ is a constant.

Combining (29) and (30),

$$\frac{\beta_k}{\rho_2}\langle \nabla f(E_k+\theta_k'\frac{\beta_k}{\rho_2}N_k), N_k\rangle + \mu_k\gamma||E_k+\frac{\beta_k}{\rho_2}N_k||_1 - \mu_k\gamma||E_k||_1 > \delta_2\Delta_k^2\frac{\beta_k}{\rho_2},$$

namely,

$$\langle \nabla f(E_k+\theta_k'\frac{\beta_k}{\rho_2}N_k), N_k\rangle + \frac{\mu_k\gamma||E_k+\frac{\beta_k}{\rho_2}N_k||_1 - \mu_k\gamma||E_k||_1}{\frac{\beta_k}{\rho_2}} > \delta_2\Delta_k^2.$$

Let $\widetilde{\beta}=g$. Because $\beta_k\to 0$ as $k\to\infty$, we obtain $\beta_k<\rho_2 g$ as $k\to\infty$. Based on Lemma 2, the following inequality can be obtained.

$$\frac{\mu_k\gamma||E_k+\frac{\beta_k}{\rho_2}N_k||_1 - \mu_k\gamma||E_k||_1}{\frac{\beta_k}{\rho_2}} - \frac{\mu_k\gamma||E_k+gN_k||_1 - \mu_k\gamma||E_k||_1}{g} \le 0.$$

$$\begin{aligned}
&\langle \nabla f(E_k+\theta_k'\frac{\beta_k}{\rho_2}N_k), N_k\rangle - \langle \nabla f(E_k), N_k\rangle \\
&\ge \langle \nabla f(E_k+\theta_k'\frac{\beta_k}{\rho_2}N_k), N_k\rangle - \langle \nabla f(E_k), N_k\rangle \\
&\quad + \frac{\mu_k\gamma||E_k+\frac{\beta_k}{\rho_2}N_k||_1 - \mu_k\gamma||E_k||_1}{\frac{\beta_k}{\rho_2}} - \frac{\mu_k\gamma||E_k+gN_k||_1 - \mu_k\gamma||E_k||_1}{g} \\
&> \delta_2\Delta_k^2 - \Delta_k^2 \\
&= -(1-\delta_2)\Delta_k^2 \\
&\ge (1-\delta_2)\frac{\rho_{(\min)}}{2}||N_k||_F^2. \qquad (31)
\end{aligned}$$

By taking the limits in both sides of the inequality (31), we get

$$0 > (1-\delta_2)\frac{\rho_{(\min)}}{2}\lim_{k\to\infty}||N_k||_F^2 > 0,$$

namely, $\lim_{k\to\infty}||N_k||_F=0$. So $\lim_{k\to\infty}E_k=E^*$.

In addition, we note that $A_{k+1}$ depends on $A_k$ and $E_k$, $E_{k+1}$ depends on $A_{k+1}$ and $E_k$, both $A_k$ and $E_k$ simultaneously converge to their own limits as $k\to\infty$, so we obtain $\lim_{k\to\infty}(A_k, E_k)=(A^*, E^*)$. □

**Assumption 2** *The level set* $\Omega=\{A : F(A)\le F(A_0)\}$ *is bounded.*

**Lemma 5** *Suppose that the direction matrix* $M_k$ *is defined by (6), where* $\lambda_k>0$, $h\in(0,1]$, *and the step size* $\alpha_k>0$. *Then,* $F(A_k)\ge F(A_{k+1})$.

**Lemma 6** *Let $l(k)$ be an integer such that*

$$k - m(k) \leq l(k) \leq k \text{ and } F(A_{l(k)}) = \max_{0 \leq j \leq m(k)} F(A_{k-j}),$$

*where*

$$m(0) = 0,\ 0 \leq m(k) \leq \min\{m(k-1) + 1, \widetilde{m}\}.$$

*If the step size $\alpha_k$ is determined by the non-monotone line search method, the direction matrix $M_k$ satisfies*

$$\lim_{k \to \infty} \alpha_k ||M_k||_F = 0.$$

**Theorem 6** *Let sequences $\{A_k\}$ and $\{M_k\}$ be generalized by Algorithm 2. Then, the sequence $\{A_k\}$ globally converges to the optimal solution $A^*$ of the general model* (4).

Theorem 6 that is the global convergence theorem of LRMC is based on Assumption 2, Lemma 5, and Lemma 6. In fact, Lemma 5, Lemma 6, and Theorem 6 can be proved easily by referring to the proof procedure of Lemma 3, Lemma 4, and Theorem 5 respectively.

At each iteration of ADSM and TESM, the intervals of the proper step sizes $\alpha_k$ and $\beta_k$ determined by the non-monotone line search method are about (0.09, 0.11). Thus, in order to be in pursuit of fast convergence, this paper takes a fixing value from the interval instead of the linear search methods at all iterations. We think that the global convergence of the two algorithms can also be guaranteed in some sense.

## 4 Numerical experiments

All numerical experiments about RPCA and LRMC are implemented with a Matlab R2010b mathematical computing software on a Windows 7 system installed on a HP desktop computer with an Intel(R) Core(TM), which has a 2.27 GHz i5 CPU, a dual-core processer, and 3.87 GB of RAM.

### 4.1 Experiments about RPCA

We make the step sizes $\alpha_k$ and $\beta_k$ be 0.10 at all iterations. Without loss of generality, let the observation data matrix D be an $n$-order square matrix. The true solution is denoted by an ordered pair $(A^*, E^*) \in (R^{n\times n}, R^{n\times n})$. The low-rank matrix $A^*$ with rank $r$ for simulation is generated by the product of two independently and randomly generated matrices $A_L^{n\times r}$ and $A_R^{n\times r}$ whose entries obey independently standard normal distribution, namely, $A^* = A_L A_R^T$. All the entries in the support set of the sparse matrix $E^*$ is chosen independently and uniformly at random in the interval $[-500, 500]$. In reality, small noise refers to Gaussian noise generally, so we suppose that the small noise matrix $N^* \in R^{n\times n}$ is a Gaussian noise matrix whose entries independently subject to standard normal distribution. Let $D = A^* + E^* + \sigma N^*$ where $\sigma \in [0, 1]$ is the level of Gaussian noise. In the experiments, $(\hat{A}, \hat{E})$ denotes the ordered output matrix pair of Algorithm 1.

Considering that the observation data matrix is contaminated by Gaussian noise ($\sigma \neq 0$), we compare ADSM with APGL. The stopping criterions of ADSM and APGL are forcedly set to be the same generalized formula (32).

$$\frac{||D - A_{k+1} - E_{k+1}||_F}{||D||_F} < tol, \tag{32}$$

where $tol$ is a proper small positive number. The relative error of $\hat{A}$ is defined as (33).

$$relerr = \frac{||A^* - \hat{A}||_F}{||A^*||_F}. \tag{33}$$

If someone wants to solve the models (2) and (4) by the augmented Lagrange multiplier method, the two papers [25] written by Yunhai Xiao et al. and [21] written by Caihua Chen et al. can be referred to and we may be inspired by their related research.

In order to study how different orders of magnitude of the level $\sigma$ of Gaussian noise have effect on low-rank matrix recovery, we set three cases: $\sigma = 10^{-3}$, $\sigma = 10^{-2}$ and $\sigma = 10^{-1}$ as follows. The recovery effects of the low-rank matrix $\hat{A}$ are shown in Fig. 1 and Table 1 intuitively. In Fig. 1, the figures in the first row are corresponding to $\sigma = 10^{-3}$, those in the second row are corresponding to $\sigma = 10^{-2}$ and those in the third row are corresponding to $\sigma = 10^{-1}$. In Table 1, relerr is the relative error of the low-rank matrix $\hat{A}$; t(s) is the running time and #iters is iteration counts.

Let the rank $r$ of $A^*$ be 5, the threshold $tol$ be $10^{-4}$, the level $\sigma$ be $10^{-3}$, the ratio of non-zero entries of $E^*$ be 0.05 and other parameters values be default. It is shown that the relative error of ADSM is a little less than that of APGL and the running time of ADSM is apparently superior to that of APGL as the order $n$ grows from 200 to 2600.

Let $tol$ be $10^{-3}$, the level $\sigma$ be $10^{-2}$ and other parameters values be in accordance with above. It can be seen that the relative error of ADSM is a little less than that of APGL as the matrix order grows from 200 to 300, and the running time of ADSM is apparently superior to that of APGL as the order $n$ grows from 100 to 2600 because the number of SVD evaluations of ADSM is much less than that of APGL. The number of SVD evaluations of the two algorithms ADSM and APGL are equal to their own iteration counts.

Let $tol$ be $10^{-2}$, the level $\sigma$ be $10^{-1}$, and other parameter values be in accordance with above. It can be seen that the relative error of ADSM is a little less than that of APGL as the order $n$ grows from 800 to 2600 and the running time of ADSM is apparently superior to that of APGL as the order $n$ grows from 100 to 2600. The reason for the phenomenon is as same as above.

In order to study how different orders $n$ have effect on low-rank matrix recovery, we set three cases again: $n = 500$, $n = 1000$, and $n = 2000$. The recovery effects of

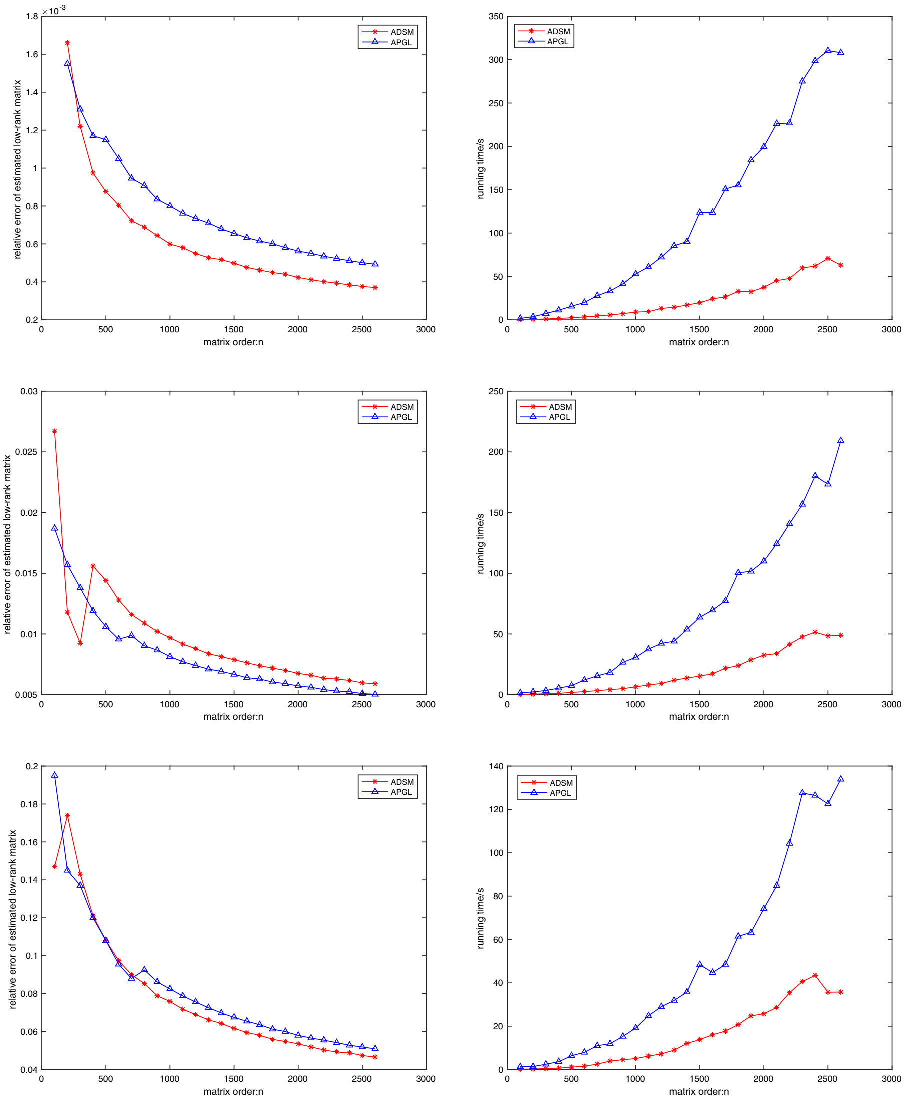


**Fig. 1** Recovery effects with different levels $\sigma$

the low-rank matrix $\hat{A}$ are shown in Table 2 intuitively. Let the rank $r$ grow from 10 to 50 by 10, the level $\sigma$ be $10^{-2}$, the threshold value $tol$ be $10^{-3}$. It can be seen from Table 2 that the relative error of ADSM is close to that of APGL, but the running time and #iters of ADSM are nearly half of those of APGL. The advantages of ADSM are obvious, especially in the aspect of the running time.

We select a real data example about face image denoising whose pictures are shown in Fig. 2 that are taken from the Extended Yale Face B database. Their sizes

**Table 1** Recovery effects with different levels $\sigma$

| Parameter setting | | ADSM | | | APGL | | |
|---|---|---|---|---|---|---|---|
| $\sigma$ | $n$ | relerr | t(s) | #iters | relerr | t(s) | #iters |
| 1.00E−03 | 500 | 8.76E−04 | 2.35 | 14 | 1.15E−03 | 15.62 | 83 |
| 1.00E−03 | 1000 | 5.99E−04 | 8.93 | 14 | 8.00E−04 | 52.84 | 83 |
| 1.00E−03 | 1500 | 4.98E−04 | 19.73 | 14 | 6.55E−04 | 123.79 | 83 |
| 1.00E−03 | 2000 | 4.23E−04 | 37.38 | 14 | 5.62E−04 | 199.54 | 83 |
| 1.00E−03 | 2500 | 3.76E−04 | 70.71 | 14 | 5.01E−04 | 310.36 | 83 |
| 1.00E−02 | 500 | 1.44E−02 | 1.85 | 10 | 1.06E−02 | 7.40 | 62 |
| 1.00E−02 | 1000 | 9.69E−03 | 6.43 | 10 | 8.14E−03 | 30.76 | 61 |
| 1.00E−02 | 1500 | 7.88E−03 | 15.29 | 10 | 6.67E−03 | 63.78 | 61 |
| 1.00E−02 | 2000 | 6.75E−03 | 32.55 | 10 | 5.72E−03 | 109.96 | 61 |
| 1.00E−02 | 2500 | 5.97E−03 | 48.38 | 10 | 5.09E−03 | 173.34 | 61 |
| 1.00E−01 | 500 | 1.08E−01 | 1.14 | 7 | 1.08E−01 | 6.38 | 40 |
| 1.00E−01 | 1000 | 7.59E−02 | 5.08 | 7 | 8.25E−02 | 19.16 | 39 |
| 1.00E−01 | 1500 | 6.17E−02 | 13.83 | 7 | 6.76E−02 | 48.36 | 39 |
| 1.00E−01 | 2000 | 5.36E−02 | 25.74 | 7 | 5.80E−02 | 74.20 | 39 |
| 1.00E−01 | 2500 | 4.74E−02 | 35.69 | 7 | 5.19E−02 | 122.58 | 39 |

are all 192 $\times$ 168, namely, each picture has 32,256 pixel points. The face of a participant is irradiated by a continuous change light source, so the shadows, reflectors, and so on appear on the facial pictures (see the first column), resulting in poor display effects. If the face is not irradiated by any change light source, the similarity of all his facial pictures data should be very high. In other words, if the values of all pixel points of the $i$th picture of the participant can be arranged into a column vector whose dimension is 32,256, the observation data matrix D formed by all his facial pictures is low-rank. Once illumination is considered, D is contaminated by the shadows, the reflectors, and so on so that it is not a low-rank matrix until they are eliminated thoroughly. In fact, the shadows, the reflectors, and so on occupy a small proportion in the pictures and the positions of the non-zero entries of the arranged matrix are scattered, so they can form a sparse matrix E. In the real world, pictures are often contaminated by small noise (e.g., Gaussian noise), so we add Gaussian noise whose level is $10^{-3}$ into the pictures (see the second column). We stack the data of all the pictures as $D \in R^{32,256\times 58}$. The pictures in the third and fourth columns are recovered by ADSM and APGL respectively. It is seen that the recovered pictures by ADSM is a little clearer than that by APGL, especially in the aspect of removing Gaussian noise. In addition, the running time of ADSM and APGL is 57.32 s and 149.61 s respectively. Obviously, the former is much less than the latter.

**Table 2** Recovery effects with different orders $n$

| Parameter setting | | ADSM | | | APGL | | |
|---|---|---|---|---|---|---|---|
| $n$ | $r$ | relerr | t(s) | #iters | relerr | t(s) | #iters |
| 500 | 10 | 1.20E−02 | 5.02 | 29 | 1.06E−02 | 10.83 | 62 |
| 500 | 20 | 9.81E−03 | 5.52 | 30 | 8.65E−03 | 11.91 | 64 |
| 500 | 30 | 7.95E−03 | 6.00 | 31 | 7.84E−03 | 13.68 | 65 |
| 500 | 40 | 8.01E−03 | 6.47 | 31 | 7.23E−03 | 13.43 | 66 |
| 500 | 50 | 6.94E−03 | 7.60 | 32 | 7.27E−03 | 16.12 | 66 |
| 1000 | 10 | 8.31E−03 | 15.90 | 29 | 7.34E−03 | 32.70 | 62 |
| 1000 | 20 | 8.38E−03 | 18.20 | 29 | 7.36E−03 | 38.14 | 62 |
| 1000 | 30 | 6.77E−03 | 18.32 | 30 | 6.71E−03 | 41.89 | 63 |
| 1000 | 40 | 6.81E−03 | 20.21 | 30 | 6.03E−03 | 43.58 | 64 |
| 1000 | 50 | 6.93E−03 | 22.34 | 30 | 6.10E−03 | 46.83 | 64 |
| 2000 | 10 | 5.83E−03 | 76.65 | 29 | 5.72E−03 | 132.85 | 61 |
| 2000 | 20 | 5.87E−03 | 68.36 | 29 | 5.19E−03 | 176.02 | 62 |
| 2000 | 30 | 5.90E−03 | 76.27 | 29 | 5.18E−03 | 154.46 | 62 |
| 2000 | 40 | 5.93E−03 | 95.30 | 29 | 5.22E−03 | 187.53 | 62 |
| 2000 | 50 | 4.78E−03 | 90.07 | 30 | 4.72E−03 | 179.85 | 63 |

**Fig. 2** Recovery effects of face image denoising

### 4.2 Experiments about LRMC

We also make the step size $\alpha_k$ be 0.10 at all iterations and the observation data matrix $\mathcal{P}_\Omega(D)$ be an $n$-order square matrix. Let $\mathcal{P}_\Omega(D) = \mathcal{P}_\Omega(A^*) + \mathcal{P}_\Omega(\sigma N^*)$. The stopping criterion of TESM is set to be the expression (34).

$$\frac{||A_{k+1} - A_k||_F}{||A_k||_F} < tol. \tag{34}$$

The relative error formula of the recovered low-rank matrix $\hat{A}$ is defined as the expression (35).

$$relerr = \frac{||\hat{A} - D||_F}{||D||_F}. \tag{35}$$

Other parameters are set to be as same as Section 4.1.

We compare TESM with APGL, FPCA in the LRMC field. When $\sigma = 10^{-3}$, the numerical results are shown in Table 3 where $m$ is the number of known samples, $df = r(2n - r)$ is the degree of freedom, $SR{=}m/n^2$ is the sample ratio, rank is the rank of the recovered low-rank matrix, and error is input error, in other word, FPCA cannot work at the cases. It can be seen that FPCA does not work well in the two aspects. On the one hand, it estimates the low-rank matrix inaccurately in the conditions: $n$=500 and $r/n > 0.1$. On the other hand, it cannot work and outputs input error. APGL performs better than FPCA, but APGL does not have excellent

**Table 3** Numerical results of TESM, APGL, and FPCA at the level $\sigma = 10^{-3}$

| Parameter setting | | | | Running time (s) | | | Relative error | | |
|---|---|---|---|---|---|---|---|---|---|
| $n$ | $r$ | m/df | SR | TESM | APGL | FPCA | TESM | APGL | FPCA |
| 500 | 30 | 6 | 0.70 | 4.97 | 5.80 | 42.50 | 3.25E−03 | 5.97E−04 | 1.05E−02 |
| 500 | 50 | 5 | 0.95 | 6.80 | 7.92 | 45.28 | 3.33E−03 | 4.96E−04 | 8.32E−01 |
| 500 | 50 | 4 | 0.76 | 6.87 | 9.06 | 44.74 | 3.81E−03 | 9.81E−04 | 1.90E−01 |
| 500 | 80 | 3 | 0.88 | 11.23 | 12.59 | 46.12 | 3.53E−03 | 7.89E−04 | 8.85E−01 |
| 1000 | 50 | 6 | 0.59 | 18.55 | 19.56 | 39.76 | 3.30E−03 | 8.69E−04 | 3.14E−04 |
| 1000 | 80 | 6 | 0.92 | 29.09 | 32.08 | Error | 3.27E−03 | 5.31E−04 | Error |
| 1000 | 80 | 5 | 0.77 | 29.49 | 33.58 | Error | 3.26E−03 | 1.40E−03 | Error |
| 1000 | 100 | 5 | 0.95 | 36.77 | 46.70 | Error | 3.30E−03 | 4.89E−04 | Error |
| 1000 | 100 | 4 | 0.76 | 36.38 | 47.12 | Error | 3.74E−03 | 7.07E−04 | Error |
| 2000 | 100 | 6 | 0.59 | 110.84 | 115.53 | 168.94 | 3.30E−03 | 1.38E−03 | 7.04E−05 |
| 2000 | 200 | 5 | 0.95 | 224.49 | 267.99 | Error | 3.36E−03 | 3.53E−01 | Error |
| 2000 | 200 | 4 | 0.76 | 225.08 | 222.94 | Error | 3.80E−03 | 3.63E−01 | Error |
| 2000 | 300 | 3 | 0.83 | 374.86 | 244.74 | Error | 4.02E−03 | 5.38E−01 | Error |
| 3000 | 200 | 6 | 0.77 | 402.92 | 499.78 | error | 3.79E−03 | 3.86E−01 | Error |
| 3000 | 200 | 5 | 0.64 | 429.65 | 427.04 | 710.54 | 3.29E−03 | 3.93E−01 | 1.14E−04 |
| 3000 | 300 | 5 | 0.95 | 694.20 | 579.75 | Error | 3.35E−03 | 5.62E−01 | Error |
| 3000 | 300 | 4 | 0.76 | 701.53 | 507.59 | Error | 3.76E−03 | 5.73E−01 | Error |

**Table 4** Numerical results of TESM, APGL, and FPCA at the level $\sigma = 10^{-2}$

| Parameter setting | | | | Running time (s) | | | Relative error | | |
|---|---|---|---|---|---|---|---|---|---|
| $n$ | $r$ | m/df | SR | TESM | APGL | FPCA | TESM | APGL | FPCA |
| 500 | 30 | 6 | 0.70 | 6.11 | 6.43 | 47.18 | 3.33E−03 | 4.22E−03 | 1.53E−01 |
| 500 | 50 | 5 | 0.95 | 7.43 | 8.45 | 49.63 | 3.39E−03 | 4.50E−03 | 8.36E−01 |
| 500 | 50 | 4 | 0.76 | 7.70 | 9.61 | 50.15 | 3.88E−03 | 5.27E−03 | 3.31E−01 |
| 500 | 80 | 3 | 0.88 | 12.66 | 13.43 | 51.34 | 3.59E−03 | 5.96E−03 | 8.05E−01 |
| 1000 | 50 | 6 | 0.59 | 20.28 | 22.16 | 248.80 | 3.35E−03 | 4.33E−03 | 1.53E−03 |
| 1000 | 80 | 6 | 0.92 | 31.67 | 33.09 | Error | 3.30E−03 | 4.13E−03 | Error |
| 1000 | 80 | 5 | 0.77 | 31.51 | 35.34 | Error | 3.30E−03 | 4.79E−03 | Error |
| 1000 | 100 | 5 | 0.95 | 42.81 | 48.49 | Error | 3.33E−03 | 4.50E−03 | Error |
| 1000 | 100 | 4 | 0.76 | 42.92 | 45.33 | Error | 3.77E−03 | 5.26E−03 | Error |
| 2000 | 100 | 6 | 0.59 | 124.10 | 116.57 | 187.61 | 3.32E−03 | 4.44E−03 | 6.93E−04 |
| 2000 | 200 | 5 | 0.95 | 233.80 | 237.45 | Error | 3.37E−03 | 3.53E−01 | Error |
| 2000 | 200 | 4 | 0.76 | 230.85 | 205.15 | Error | 3.82E−03 | 3.63E−01 | Error |
| 2000 | 300 | 3 | 0.83 | 389.67 | 216.94 | Error | 4.03E−03 | 5.38E−01 | Error |
| 3000 | 200 | 6 | 0.77 | 424.19 | 450.75 | Error | 3.80E−03 | 3.86E−01 | Error |
| 3000 | 200 | 5 | 0.64 | 455.99 | 379.19 | 669.99 | 3.31E−03 | 3.93E−01 | 6.10E−04 |
| 3000 | 300 | 5 | 0.95 | 723.91 | 553.78 | Error | 3.36E−03 | 5.63E−01 | Error |
| 3000 | 300 | 4 | 0.76 | 740.81 | 478.13 | Error | 3.77E−03 | 5.73E−01 | Error |

performance at the cases in which $r$ exceeds 150. TESM can cover the shortages of FPCA and APGL, namely, not only TESM can accurately estimate the low-rank matrix, but also its running time is the shortest among the three algorithms. When $\sigma = 10^{-2}$, the numerical results are shown in Table 4, from which we can get the similar conclusions.

## 5 Conclusions

In recent years, the popular RPCA and LRMC problems for recovering a low-rank component have extensive applications in pattern recognition. We propose, analyze, and test the new practical algorithms ADSM and TESM, for solving the general relaxed models (2) and (4) respectively. This paper utilizes Taylor expansion, SVD, shrinkage operator, and so on to deduce the iterative direction matrices. We combine the direction step size formula with the alternating direction minimization idea to design the structure of ADSM. This paper presents the global convergence theorems of the two algorithms by supposing that the objective function $F(\cdot)$ is bounded. The experimental results illustrate that they are effective instruments to recover low-rank component. The performance comparisons with the efficient solver APGL verify the advantages of ADSM in terms of running time and relative error. TESM makes up for the defects that FPCA inaccurately estimates or cannot estimate the solutions of the model (4) at many cases and APGL does not have good performance in some scenarios.

**Acknowledgements** All the authors would like to thank the reviewers for their valuable suggestions.

**Funding information** This research is supported in part by the Natural Science Foundation of China (Grant NSFC-11301021).

## References

1. Wright, J., Ganesh, A., Rao, S., Peng, Y., Ma, Y.: Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. Adv. Neural Inf. Process. Syst., 2080–2088 (2009)
2. Candès, E.J., Li, X., Ma, Y., Wright, J.: Robust principal component analysis. J. ACM. **58**, 11 (2011)
3. Min, K., Zhang, Z., Wright, J., Ma, Y.: Decomposing background topics from keywords by principal component pursuit. In: Proceedings of the 19th ACM International Conference, pp. 269–278 (2010)
4. Bennett, J., Lanning, S.: The netflix prize. In: Proceedings of KDD Cup and Workshop, p. 35 (2007)
5. Recht, B., Fazel, M., Parril, P.A.: Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. SIAM Rev. **52**, 471–501 (2010)
6. Candès, E.J., Recht, B.: Exact matrix completion via convex optimization. Found. Comput. Math. **9**, 717–772 (2009)
7. Candès, E.J., Tao, T.: Decoding by linear programming. IEEE. T. Inform. Theory **51**, 4203–4215 (2005)
8. Donoho, D.L.: For most large underdetermined systems of linear equations the minimal $l_1$-norm solution is also the sparsest solution. Commun. Pur. Appl. Math. **59**, 797–829 (2006)
9. Cai, J., Candès, E.J., Shen, Z.: A singular value thresholding algorithm for matrix completion. SIAM J. Optim. **20**, 1956–1982 (2010)
10. Liu, Z., Vandenberghe, L.: Interior-point method for nuclear norm approximation with application to system identification. SIAM J. Matrix. Anal. A **31**, 1235–1256 (2009)
11. Schmidt, R.: Multiple emitter location and signal parameter estimation. IEEE. T. Antenn. Propag. **34**, 276–280 (1986)
12. Ji, H., Liu, C., Shen, Z., Xu, Y.: Robust video denoising using low-rank matrix completion. In: Proceedings of IEEE Conference (2010)
13. Zhang, X., Xiong, H.: Illumination compensation via low rank matrix completion for multiview video coding. In: 2013 IEEE International Conference, pp. 1865–1869 (2013)
14. Chandrasekaran, V., Sanghavi, S., Parrilo, P.A., Willsky, A.S.: Rank-sparsity incoherence for matrix decomposition. SIAM J. Optimiz. **21**, 572–596 (2011)
15. Lin, Z., Ganesh, A., Wright, J., Wu, L., Chen, M., Ma, Y.: Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix. Comput. Adv. Multi-Sensor Adapt. Process., 61 (2009)
16. Ma, S., Goldfarb, D., Chen, L.: Fixed point and Bregman iterative methods for matrix rank minimization. Math. Program. **128**, 321–353 (2011)
17. Drineas, P., Kannan, R., Mahoney, M.W.: Fast Monte Carlo algorithms for matrices II: computing a low-rank approximation to a matrix. SIAM J. Comput. **36**, 158–183 (2006)
18. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging. Sci. **2**, 183–202 (2009)
19. Toh, K.C., Yun, S.: An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems. Pac. J. Optim. **6**, 15 (2010)
20. Xiao, Y.H., Wu, S.Y., Qi, L.: Nonmonotone Barzilai-Borwein gradient algorithm for $l_1$-regularized nonsmooth minimization in compressive sensing. J. Sci. Comput. **61**, 17–41 (2014)
21. Chen, C., He, B., Yuan, X.: Matrix completion via an alternating direction method. IMA J. Numer. Anal. **32**, 227–245 (2012)
22. Hiriart-Urruty, J.B., Lemarechal, C.: Convex Analysis and Minimization Algorithms I: Fundamentals. Springer Science & Business Media (2013)
23. Hale, E.T., Yin, W., Zhang, Y.: A Fixed-point Continuation Method for $l_1$-regularized Minimization with Applications to Compressed Sensing. Rice University, pp. 43–44 (2007)
24. Grippo, L., Lampariello, F., Lucidi, S.: A nonmonotone line search technique for Newton's method. SIAM J. Numer. Anal. **23**, 707–716 (1986)
25. Xiao, Y.H., Jin, Z.: An alternating direction method for linear-constrained matrix nuclear norm minimization. Numer. Linear. Algebr. **19**, 541–554 (2012)