FOUNDATIONS OF COMPUTING AND DECISION SCIENCES
Vol. 49 (2024) No. 1
ISSN 0867-6356
DOI: 10.2478/fcds-2024-0004 e-ISSN 2300-3405

An Improved Alternating Direction Method of Multipliers for Matrix Completion

Xihong Yan, Ning Zhang *, Hao Li †

Abstract. Matrix completion is widely used in information science fields such as machine learning and image processing. The alternating direction method of multipliers (ADMM), due to its ability to utilize the separable structure of the objective function, has become an extremely popular approach for solving this problem. But its subproblems can be computationally demanding. In order to improve computational efficiency, for large scale matrix completion problems, this paper proposes an improved ADMM by using convex combination technique. Under certain assumptions, the global convergence of the new algorithm is proved. Finally, we demonstrate the performance of the proposed algorithms via numerical experiments.

Keywords: matrix completion, alternating direction method of multipliers

1. Introduction

The matrix completion problem mainly discusses how to complete the missing data of the matrix through the observed partial data. Matrix completion has been widely used in information science fields such as machine learning [20, 1], image processing [24] and computer vision [22]. Its mathematical model is as follows:

$$\begin{array}{ll} \min & \operatorname{rank}(X) \\ \text{s.t.} & X_{ij} = M_{ij},\ (i,j) \in \Omega, \end{array} \tag{1}$$

where $X \in R^{m\times n}$ is a matrix to be completed, $M \in R^{m\times n}$ is a sampling matrix, $\operatorname{rank}(X)$ denotes the rank of the matrix X, $\Omega \subset \{1, \cdots, m\}\times\{1, \cdots, n\}$ is the subscript set of known elements, M_{ij} represents the known elements of the matrix M. However,

*Corresponding author.
†College of Mathematics and Statistics, Taiyuan Normal University, Shanxi Key Laboratory for Intelligent Optimization Computing and Blockchain Technology, Jinzhong, Shanxi, P. R. China (030619). E-mail: yanxihong@tynu.edu.cn, {ning10231023, 18401620037}@163.com.

due to the non-convexity of the rank function, model (1) is NP-hard problem, which is difficult to solve directly. Therefore, Candès and Recht [3] relaxed the problem into the following convex optimization problem:

$$\begin{array}{ll} \min & \|X\|_* \\ \text{s.t.} & X_{ij} = M_{ij},\ (i,j) \in \Omega. \end{array} \tag{2}$$

To solve the kernel norm minimization problem (2), many scholars have proposed a variety of fast algorithms. Fazel [7] and Fazel et al. [8] first proposed a semidefinite programming algorithm to solve the optimization problem (2). Toh and Yun [21] proposed an accelerated proximal gradient method (APG). Meanwhile, three accelerated APG [21] are proposed by using linesearch-like technique, continuation technique and truncation technique. Then, Ma et al. [16] used a homotopy approach together with an approximate singular value decomposition procedure to fixed point continuation algorithm (FPC [13]), proposed an fixed point continuation algorithm with approximate SVD(FPCA), and improved the FPC. Cai et al. [2] proposed singular value thresholding algorithm(SVT). Subsequently, Lin et al. [15] proposed an augmented lagrange multiplier method (ALM), and proved through numerical experiments that the ALM algorithm convergence rate is faster than SVT algorithm and APG algorithm.

It is well known that the alternating direction multiplier method (ADMM) [10] has become a very popular method for solving optimization problems, due to its ability to utilize the separable structure of the objective function. Chen et al. [5] reformulated the model (2) as the following linearly constrained convex programming problem:

$$\begin{array}{ll} \min & \|X\|_* \\ \text{s.t.} & X - W = 0 \\ & X \in R^{m\times n},\ W \in \mathbb{U} = \{U \in R^{m\times n} | U_{ij} = M_{ij}\}, \end{array} \tag{3}$$

and first applied the ADMM algorithm to the matrix completion problem. The iterative scheme is as follows:

$$\left\{ \begin{array}{ll} X_n & \in \arg\min_{X\in R^{m\times n}}\{L_\sigma(X, W_{n-1}, Y_{n-1})\}, \\ W_n & \in \arg\min_{W\in \mathbb{U}}\{L_\sigma(X_n, W, Y_{n-1})\}, \\ Y_n & = Y_{n-1} + \sigma(X_n - W_n), \end{array} \right. \tag{4}$$

where (X_n, Y_n, Z_n) is the given triple of iterates, $L_\sigma(X, W, Y) = \|X\|_* + \langle Y,\ X - W\rangle + \frac{\sigma}{2}\|X - W\|_F^2$ is the augmented Lagrange function of (3), $Y \in R^{m\times n}$ is the Lagrange multiplier of the linear constraint, $\sigma > 0$ is the penalty parameter for the violation of the linear constraint. In [11, 12], in order to improve computational efficiency, it was proved that the step for updating the Lagrange multiplier Y^{k+1} can be generalized into

$$Y_n = Y_{n-1} + \gamma\sigma(X_n - W_n) \text{ with } 0 < \gamma < \frac{\sqrt{5}+1}{2}.$$

However, in general the subproblems of the ADMM can be computationally demanding. Therefore, Eckstein [6] proposed a proximal ADMM (PADMM), which introduces Euclidean-norm-squared proximal terms into the ADMM subproblems. Later, Shefi and Teboulle[19] added general weighted-norm proximal terms into the ADMM subproblems. Given $(X_{n-1}, Y_{n-1}, Z_{n-1})$, the iterative scheme of PADMM for solving the problem (3) is as follows:

$$\begin{cases} X_n & \in \arg\min_{X \in R^{m\times n}} \{L_\sigma(X, W_{n-1}, Y_{n-1}) + \frac{1}{2}\|X - X_{n-1}\|_S^2\}, \\ W_n & \in \arg\min_{W \in \mathbb{U}} \{L_\sigma(X_n, W, Y_{n-1}) + \frac{1}{2}\|W - W_{n-1}\|_T^2\}, \\ Y_n & = Y_{n-1} + \sigma(X_n - W_n), \end{cases} \tag{5}$$

where $S,\ T$ are symmetric and positive semidefinite. In particular, when both S and T are zero matrices, PADMM reduces to ADMM, see [4, 9, 23] for more discussions.

Considering that the data in practical applications are often large-scale, it is necessary to further improve the efficiency of ADMM algorithm. There are already some modified ADMM [4, 14, 18]. In fact, convex combination technique is a classical acceleration technique for optimization problems, and has good numerical performance [17]. Therefore, this paper proposes an improved alternating direction method of multipliers(IADMM) for matrix completion by introducing the convex combination technique, and proves the global convergence of the new algorithm.

2. Preliminaries

In this section, we summarize some notations and elementary results to be used for further analysis.

Notation : $R^{m\times n}$ denotes the set of all $m \times n$ real matrices; $tr(X)$ denotes the trace of matrix X; $\|X\|_* := \sum_{i=1}^{r} \varrho_i(X)$ denotes the nuclear norm of matrix X, where $\varrho_i(X)$ denotes the i-largest singular value of matrix X, $\varrho_1 \geq \varrho_2 \geq ... > \varrho_i \geq ... \geq \varrho_r > 0$; $\|X\|_F := \sqrt{tr(X^{\mathrm{T}}X)}$ is the Frobenius norm of matrix X; X^{T} denotes the transpose operation of matrix X; X^{-1} denotes the inverse operation of matrix X; The identity matrix of appropriate sizes will be denoted by I; $\langle A, B\rangle := tr(A^{\mathrm{T}}B)$ denotes the standard trace inner product of matrix $A,\ B$; $diag(\sigma_1, \sigma_2, ..., \sigma_r)$ denotes a diagonal matrix whose diagonal elements are $\sigma_1, \sigma_2, ..., \sigma_r$; Let $\mathbb{N} = \{1, 2, ...\}$ be the sequence of positive integers and $\phi := \frac{\sqrt{5}+1}{2}$ be the golden ratio, which is a key parameter in our proposed algorithm.

There are basic definitions that will be useful in the subsequent sections.

Definition 2.1. The Lagrange function $L(X, W, Y)$ of (3) is

$$L(X, W, Y) = \|X\|_* + \langle Y,\ X - W\rangle,$$

where $Y \in R^{m\times n}$ is the Lagrange multiplier of the linear constraint. There is a triplet $(\bar{X}, \bar{W}, \bar{Y}) \in R^{m\times n} \times R^{m\times n} \times R^{m\times n}$, if

$$L(\bar{X}, \bar{W}, Y) \leq L(\bar{X}, \bar{W}, \bar{Y}) \leq L(X, W, \bar{Y}) \text{ for all } (X, W, Y) \in R^{m\times n} \times R^{m\times n} \times R^{m\times n},$$

then it is called a saddle point of $L(\cdot)$

Definition 2.2. The effective domain of any extended realvalued closed proper convex function H is denoted by $dom(H) := \{X \in R^{m\times n} : H(X) < \infty\}$, if the subdifferential of H at $X \in R^{m\times n}$ exists, it defined as

$$\partial H(X) := \{g \in R^{m\times n} : H(Y) \geq H(X) + \langle g, Y - X\rangle \text{ for all } Y \in R^{m\times n}\}.$$

Definition 2.3. For any matrix $X \in R^{m\times n}$ with rank r, there exist two orthogonal column matrices $U \in R^{m\times r}$ and $V \in R^{n\times r}$ such that

$$X = U\Sigma_r V^{\mathrm{T}},$$

where $\Sigma_r = diag(\sigma_1, \cdots, \sigma_r),\ \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.

Definition 2.4. For any parameter $\tau \geq 0$, matrix $X \in R^{m\times n}$ with rank r, there exists singular value decomposition $X = U\Sigma_r V^T$. The singular value threshold operator D_τ is defined as

$$D_\tau(X) := UD_\tau(\Sigma_r)V^T,$$

where

$$D_\tau(\Sigma_r) = diag(\{\sigma_i - \tau\}_+),$$

$$\{\sigma_i - \tau\}_+ = \begin{cases} \sigma_i - \tau, & if \quad \sigma_i > \tau, \\ 0, & if \quad \sigma_i \leq \tau. \end{cases}$$

The following results will be used in convergence analysis.

Lemma 2.1. For any $A,\ B,\ C,\ D \in R^{m\times n}$, symmetric and positive definite matrix $H \in R^{m\times n}$, $\theta \in R$, there hold

$$2\langle A - B,\ C - D\rangle_H = \|A - D\|_H^2 + \|B - C\|_H^2 - \|A - C\|_H^2 - \|B - D\|_H^2, \tag{6}$$

$$\|\theta A + (1-\theta)B\|_H^2 = \theta\|A\|_H^2 + (1-\theta)\|B\|_H^2 - \theta(1-\theta)\|A - B\|_H^2. \tag{7}$$

Lemma 2.2. Let $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ be two nonnegative real sequences. If there exists an integer $N \in \mathbb{N}$ such that $a_{n+1} \leq a_n - b_n$ for all $n > N$, then $\lim\limits_{n\to\infty} a_n$ exists and $\lim\limits_{n\to\infty} b_n = 0$.

3. Algorithm

In this section, we present IADMM to solve (3), which adopts the convex combination technique in [17] into the ADMM in [11, 12]. The complete algorithm is summarized in Algorithm 1.

Algorithm 1. Improved Alternating Direction Method of Multipliers (IADMM) for Matrix Completion

---

Let T be symmetric and positive semidefinite, Ω be a given subscript set and M be a sampling matrix. $\tau > 0$, $\beta > 0$ and $\psi \in (1, \phi]$. Choose $X_0, Y_0, W_0 \in R^{m\times n}$. Set $Z_0 = X_0$ and $n = 1$.

while the $\frac{\|X_n - X_{n-1}\|_F}{\|X_n\|_F} \le 10^{-6}$ is not satisfied **do**

$$\begin{cases} Z_n = \dfrac{\psi - 1}{\psi} X_{n-1} + \dfrac{1}{\psi} Z_{n-1}, & \text{(8a)} \\ X_n = \arg\min_{X \in R^{m \times n}} \{L(X, W_{n-1}, Y_{n-1}) + \dfrac{1}{2\tau} \|X - Z_n\|_F^2\}, & \text{(8b)} \\ W_n = \arg\min_{W \in \mathbb{U}} \{L_\beta(X_n, W, Y_{n-1}) + \dfrac{1}{2} \|W - W_{n-1}\|_T^2\}, & \text{(8c)} \\ Y_n = Y_{n-1} + \beta(X_n - W_n). & \text{(8d)} \end{cases}$$

end while

Remark 3.1. Note that IADMM requires the stepsizes $\tau > 0$, $\beta > 0$ satisfying the condition $\beta\tau < \psi$, where $\psi \in (1, \phi]$.

Remark 3.2. The solution of (8b) can be formulate by the following expression

$$X_n = D_\tau(Z_n - \tau Y_{n-1}).$$

The solution of (8c) can be formulate by the following expression

$$W_n = \begin{cases} M, \ (i,j) \in \Omega, \\ (\beta I + T)^{-1}(Y_{n-1} + \beta X_n + T W_{n-1}), \ \textit{otherwise}. \end{cases}$$

Proof. First, it holds that

$$\begin{aligned} X_n &= \arg\min_{X \in R^{m \times n}} \{L(X, W_{n-1}, Y_{n-1}) + \tfrac{1}{2\tau} \|X - Z_n\|_F^2\} \\ &= \arg\min_{X \in R^{m \times n}} \{\|X\|_* + \langle Y_{n-1}, X - W_{n-1}\rangle + \tfrac{1}{2\tau} \|X - Z_n\|_F^2\} \\ &= \arg\min_{X \in R^{m \times n}} \{\|X\|_* + \tfrac{1}{2\tau} \|X - Z_n + \tau Y_{n-1}\|_F^2\} \\ &= D_\tau(Z_n - \tau Y_{n-1}). \end{aligned}$$

Next, we consider the solution of (8c). When $(i,j) \notin \Omega$, we have

$$\begin{aligned} W_n &= \arg\min_{W \in \mathbb{U}} \{L_\beta(X_n, W, Y_{n-1}) + \tfrac{1}{2} \|W - W_{n-1}\|_T^2\} \\ &= \arg\min_{W \in \mathbb{U}} \{\|X_n\|_* + \langle Y_{n-1}, X_n - W\rangle + \tfrac{\beta}{2} \|X_n - W\|_F^2 + \tfrac{1}{2} \|W - W_{n-1}\|_T^2\} \\ &= \arg\min_{W \in \mathbb{U}} \{\langle Y_{n-1}, X_n - W\rangle + \tfrac{\beta}{2} \|X_n - W\|_F^2 + \tfrac{1}{2} \|W - W_{n-1}\|_T^2\} \\ &= (\beta I + T)^{-1}(Y_{n-1} + \beta X_n + T W_{n-1}). \end{aligned}$$

When $(i,j) \in \Omega$, we have $W_n = M$.

4. Convergence analysis

In this section, we give the convergence analysis of IADMM. Firstly, we give an important lemma, which plays a key role in our subsequent analysis.

Lemma 4.1. Let $\xi := \sqrt{\frac{\beta\tau}{\psi}}$ and $\{(Z_n, X_n, W_n, Y_n)\}$ be the sequence generated by Algorithm 1. Then, for any solution $(\bar{X}, \bar{W})$ of (3), $X, \ W, \ Y \in R^{m \times n}$ and $n > 1$, we have

$$\begin{array}{ll} L(X_n, W_n, Y) \leq & \frac{\psi}{2\tau(\psi-1)}\|Z_{n+1}-\bar{X}\|_F^2 + \frac{1}{2}\|W_{n-1}-\bar{W}\|_T^2 + \frac{1}{2\beta}\|Y_{n-1}-Y\|_F^2 \\ -\|\bar{X}\|_* & -(\frac{\psi}{2\tau(\psi-1)}\|Z_{n+2}-\bar{X}\|_F^2 + \frac{1}{2}\|W_n-\bar{W}\|_T^2 + \frac{1}{2\beta}\|Y_n-Y\|_F^2) \\ & -(\frac{(1-\xi)\psi}{2\tau}\|X_{n+1}-X_n\|_F^2 + \frac{1}{2}\|W_n-W_{n-1}\|_T^2 + \frac{1-\xi}{2\beta}\|Y_n-Y_{n-1}\|_F^2) \\ & -\frac{\psi}{2\tau}\|Z_{n+1}-X_n\|_F^2 - \frac{1}{2\tau}(1+\frac{1}{\psi}-\psi)\|X_{n+1}-Z_{n+1}\|_F^2. \end{array} \tag{9}$$

Proof. From (8b) and (8c), and utilizing (8d), we obtain

$$\begin{array}{l} 0 \in \partial\|X_n\|_* + Y_{n-1} + \frac{1}{\tau}(X_n - Z_n), \\ 0 = -Y_n + T(W_n - W_{n-1}). \end{array}$$

By the convexity of $\|\cdot\|_*$ and the subgradient inequality, it holds that

$$\|X_n\|_* - \|X\|_* \leq \langle Y_{n-1}, X - X_n\rangle + \frac{1}{\tau}\langle X_n - Z_n, X - X_n\rangle, \tag{10}$$

$$0 = -\langle Y_n, W - W_n\rangle + \langle T(W_n - W_{n-1}), W - W_n\rangle. \tag{11}$$

Similar to (10), we have

$$\|X_{n+1}\|_* - \|X\|_* \leq \langle Y_n, X - X_{n+1}\rangle + \frac{1}{\tau}\langle X_{n+1} - Z_{n+1}, X - X_{n+1}\rangle. \tag{12}$$

Setting $X = X_{n+1}, W = \bar{W}$ and $X = \bar{X}$ in (10), (11) and (12), respectively, adding them together, utilizing $\bar{X} - \bar{W} = 0$, and rearranging terms, it holds that

$$\begin{array}{ll} \|X_n\|_* - \|\bar{X}\|_* \leq & \frac{1}{\tau}(\langle X_n - Z_n, X_{n+1} - X_n\rangle + \langle X_{n+1} - Z_{n+1}, \bar{X} - X_{n+1}\rangle) \\ & +\langle T(W_n - W_{n-1}), \bar{W} - W_n\rangle - \langle Y_n, X_n - W_n\rangle \\ & +\langle Y_{n-1} - Y_n, X_{n+1} - X_n\rangle. \end{array} \tag{13}$$

From (8a), we have

$$X_n - Z_n = \psi(X_n - Z_{n+1}).$$

Adding $\langle Y, X_n - W_n\rangle$ to both sides of (13), using the above equation and (8d), we obtain

$$\begin{array}{ll} L(X_n, W_n, Y) \leq & \frac{1}{\tau}(\psi\langle X_n - Z_{n+1}, X_{n+1} - X_n\rangle + \langle X_{n+1} - Z_{n+1}, \bar{X} - X_{n+1}\rangle) \\ -\|\bar{X}\|_* & +\langle T(W_n - W_{n-1}), \bar{W} - W_n\rangle + \frac{1}{\beta}\langle Y - Y_n, Y_n - Y_{n-1}\rangle \\ & +\langle Y_{n-1} - Y_n, X_{n+1} - X_n\rangle. \end{array} \tag{14}$$

From (8a), we have $Z_{n+2} - Z_{n+1} = \frac{\psi-1}{\psi}(X_{n+1} - Z_{n+1})$.

By applying identity (7) to $X_{n+1} - \bar{X} = \frac{\psi}{\psi-1}(Z_{n+2} - \bar{X}) - \frac{1}{\psi-1}(Z_{n+1} - \bar{X})$, it holds that

$$\|X_{n+1} - \bar{X}\|_F^2 = \frac{\psi}{\psi-1}\|Z_{n+2} - \bar{X}\|_F^2 - \frac{1}{\psi-1}\|Z_{n+1} - \bar{X}\|_F^2 + \frac{1}{\psi}\|Z_{n+1} - X_{n+1}\|_F^2. \tag{15}$$

Next, we treat the terms on the right hand side of (14). First, by using (6) and (15), we have

$$\begin{array}{rl} & \psi\langle X_n - Z_{n-1}, X_{n+1} - X_n\rangle + \langle X_{n+1} - Z_{n+1}, \bar{X} - X_{n+1}\rangle \\ = & \frac{\psi}{2}(\|Z_{n+1} - X_{n+1}\|_F^2 - \|X_{n+1} - X_n\|_F^2 - \|Z_{n+1} - X_n\|_F^2) \\ & +\frac{1}{2}(\|Z_{n+1} - \bar{X}\|_F^2 - \|X_{n+1} - \bar{X}\|_F^2 - \|Z_{n+1} - X_{n+1}\|_F^2) \\ = & \frac{\psi}{2(\psi-1)}(\|Z_{n+1} - \bar{X}\|_F^2 - \|Z_{n+2} - \bar{X}\|_F^2) - \frac{1}{2}(1 + \frac{1}{\psi} - \psi)\|Z_{n+1} - X_{n+1}\|_F^2 \\ & -\frac{\psi}{2}(\|X_{n+1} - X_n\|_F^2 + \|Z_{n+1} - X_n\|_F^2). \end{array} \tag{16}$$

Second, using the identity (6), we have

$$\langle T(W_n - W_{n-1}), \bar{W} - W_n\rangle = \frac{1}{2}(\|W_{n-1} - \bar{W}\|_T^2 - \|W_n - \bar{W}\|_T^2 - \|W_n - W_{n-1}\|_T^2), \tag{17}$$

$$\langle Y - Y_n, Y_n - Y_{n-1}\rangle = \frac{1}{2}(\|Y_{n-1} - Y\|_F^2 - \|Y_n - Y\|_F^2 - \|Y_n - Y_{n-1}\|_F^2). \tag{18}$$

Third, according to $\xi = \sqrt{\frac{\beta\tau}{\psi}}$, the Cauchy-Schwarz inequalities and Young's inequalities, we obtain

$$\langle Y_n - Y_{n-1}, X_n - X_{n+1}\rangle \leq \frac{\xi}{2}(\frac{1}{\beta}\|Y_n - Y_{n-1}\|_F^2 + \frac{\psi}{\tau}\|X_{n+1} - X_n\|_F^2). \tag{19}$$

Finally, substituting (16)-(19) into (14) and rearranging the terms, (9) holds.

Now, we give the global convergence theorem of Algorithm 1.

Theorem 4.1. The sequence $\{(X_n, W_n, Y_n)\}$ generated by Algorithm 1 converges to a saddle point of $L(\cdot)$.

Proof. Let $(\bar{X}, \bar{W}, \bar{Y})$ be a saddle point of $L(\cdot)$, and $\bar{X} - \bar{W} = 0$, we have

$$\|\bar{X}\|_* = L(\bar{X}, \bar{W}, \bar{Y}).$$

Since $\psi \in (1, \phi]$, the term $(1 + \frac{1}{\psi} - \psi)\|X_{n+1} - Z_{n+1}\|_F^2$ is nonnegative.

Next, by setting $Y = \bar{Y}$ in (9), we obtain

$$0 \leq L(X_n, W_n, \bar{Y}) - L(\bar{X}, \bar{W}, \bar{Y}) \leq \frac{1}{2}(a_n - a_{n+1} - b_n),$$

where

$$a_n = \frac{\psi}{\tau(\psi - 1)}\|Z_{n+1} - \bar{X}\|_F^2 + \|W_{n-1} - \bar{W}\|_T^2 + \frac{1}{\beta}\|Y_{n-1} - \bar{Y}\|_F^2, \tag{20}$$

$$\begin{array}{rl} b_n = & \frac{(1-\xi)\psi}{\tau}\|X_{n+1} - X_n\|_F^2 + \|W_n - W_{n-1}\|_T^2 \\ & +\frac{1-\xi}{\beta}\|Y_n - Y_{n-1}\|_F^2 + \frac{\psi}{\tau}\|Z_{n+1} - X_n\|_F^2. \end{array} \tag{21}$$

Because $(\bar{X}, \bar{W}, \bar{Y})$ is a saddle point of $L(\cdot)$, the first inequality holds.

According to $\xi = \sqrt{\frac{\beta\tau}{\psi}}$ and $\tau < \frac{\psi}{\beta}$, we have $\xi \in (0, 1)$. Therefore, both a_n and b_n are nonnegative.

Then, according Lemma 2.2, we have $\lim_{n\to\infty} a_n$ exists and $\lim_{n\to\infty} b_n = 0$. Due to the positive definiteness of T, we infer from (21) that

$$\begin{aligned}\lim_{n\to\infty}\|X_{n+1}-X_n\|_F = \lim_{n\to\infty}\|W_n - W_{n-1}\|_F = \\ \lim_{n\to\infty}\|Y_n - Y_{n-1}\|_F = \lim_{n\to\infty}\|Z_{n+1}-X_n\|_F = 0.\end{aligned} \tag{22}$$

Form (8a), we obtain $\lim_{n\to\infty}\|Z_n - X_n\|_F = \lim_{n\to\infty}\psi\|Z_{n+1}-X_n\|_F = 0$.
On the other hand, it follows from the definition of a_n and the existence of $\lim_{n\to\infty} a_n$ that the sequences $\{Z_n\}$, $\{W_n\}$ and $\{Y_n\}$ are bounded and $\{X_n\}$ is also bounded due to $\lim_{n\to\infty}\|Z_n - X_n\| = 0$.
Let $\{X_{n_k+1}\}$, $\{W_{n_k}\}$, and $\{Y_{n_k}\}$ be subsequences of $\{X_{n+1}\}$, $\{W_n\}$ and $\{Y_n\}$, respectively, such that $\lim_{k\to\infty}(X_{n_k+1}, W_{n_k}, Y_{n_k}) = (X^*, W^*, Y^*)$.
Combining (11) and (12), for any (X, W), it holds that

$$\left\{\begin{array}{rcl}\|X_{n_k+1}\|_* - \|X\|_* & \le & \langle Y_{n_k}, X - X_{n_k+1}\rangle + \frac{1}{\tau}\langle X_{n_k+1} - Z_{n_k+1}, X - X_{n_k+1}\rangle,\\ 0 & = & -\langle Y_{n_k}, W - W_{n_k}\rangle + \langle T(W_{n_k} - W_{n_k-1}), W - W_{n_k}\rangle.\end{array}\right. \tag{23}$$

Taking the limits over both sides of (23) and using the lower semi-continuity of $\|\cdot\|_*$, we have

$$\left\{\begin{array}{rcl}\|X^*\|_* - \|X\|_* & \le & \langle Y^*, X - X^*\rangle,\ \forall X \in R^{m\times n},\\ 0 & = & -\langle Y^*, W - W^*\rangle,\ \forall W \in R^{m\times n}.\end{array}\right. \tag{24}$$

Furthermore, by passing to the limit over both sides of $Y_{n_k+1} - Y_{n_k} = \beta(X_{n_k+1} - W_{n_k+1})$, we have $X^* - W^* = 0$. Therefore, according the definition of $L(\cdot)$, it obtains

$$L(X^*, W^*, Y) = L(X^*, W^*, Y^*),\ \forall Y \in R^{m\times n}.$$

Then, adding the two inequalities in (24) and using $X^* - W^* = 0$, we have

$$\|X^*\|_* - \|X\|_* \le \langle Y^*, X - W\rangle.$$

This implies that

$$L(X^*, W^*, Y^*) \le L(X, W, Y^*),\ \forall X,\ W \in R^{m\times n}.$$

Therefore, (X^*, W^*, Y^*) is a saddle point of $L(\cdot)$.
Next, we show that (X^*, W^*, Y^*) is the only limit point of $\{(X_n, W_n, Y_n)\}$.
First, since $\lim_{k\to\infty}(X_{n_k+1}, W_{n_k}, Y_{n_k}) = (X^*, W^*, Y^*)$ and (22), we have

$$\lim_{k\to\infty}\|Z_{n_k+1} - X^*\|_F = \lim_{k\to\infty}\|W_{n_k-1} - W^*\|_F = \lim_{k\to\infty}\|Y_{n_k-1} - Y^*\|_F = 0.$$

Because (X^*, W^*, Y^*) is a saddle point of $L(\cdot)$, we can replace $(\bar{X}, \bar{W}, \bar{Y})$ by (X^*, W^*, Y^*) in the definition of a_n.
Then, it follows that $\lim_{k\to\infty} a_{n_k} = 0$. Since $\lim_{n\to\infty} a_n$ exists, there must hold $\lim_{n\to\infty} a_n = \lim_{k\to\infty} a_{n_k} = 0$, which implies

$$\lim_{n\to\infty}\|Z_{n+1} - X^*\|_F = \lim_{n\to\infty}\|W_{n-1} - W^*\|_F = \lim_{n\to\infty}\|Y_{n-1} - Y^*\|_F = 0.$$

Noting (22), the convergence of Algorithm 1 is proved.

5. Numerical experiments

In this section, we test the performance of the proposed Algorithm 1 by numerical experiments of random matrix completion. We shall compare IADMM and ADMM[11, 12]. All the experiments were implemented in MATLAB (R2022b) running on a PC with an 13th Gen Intel(R) Core(TM) i9-13900HX 2.20 GHz and 16.0 GB.

Specifically, we create random matrices $M \in R^{n\times n}$ with rank r by the following procedure. We first generate random matrices $M_1 \in R^{n\times r}$ and $M_2 \in R^{n\times r}$ with i.i.d. Gaussian entries and then set $M = M_1 M_2^{\mathrm{T}}$. Then, we sample a subset Ω of m entries uniformly at random. We use $p = m/n^2$, i.e., the number of measurements divided by the number of entries of the matrix, to denote the sampling ratio. We use iter and time to denote the the number of iterations of the algorithm and CPU time of the algorithm, respectively. The stopping criterion is

$$\frac{\|X_n - X_{n-1}\|_F}{\|X_n\|_F} \le 10^{-6}.$$

We define two kinds of the errors of the experiment by

$$\frac{\|X_n - X_{n-1}\|_F}{\|X_n\|_F},\ \frac{\|X_n - M\|_F}{\|M\|_F}.$$

We choose $T = 0$, $n \in \{1000,\ 2000,\ 3000,\ 4000,\ 5000,\ 10000\}$, $r \in \{5,\ 10\}$, $p \in \{0.3,\ 0.4,\ 0.5\}$. Other parameters in the algorithm are set as follows:

$$\psi = 1.618,\ \beta = 0.008,\ \tau = \psi/\beta.$$

Moreover, we list the numerical results of ADMM and IADMM in terms of iteration times, CPU time and two kinds of the algorithm errors for the randomly generated low-rank matrix completion problems, as shown in Table $1-3$. It can be seen from the numerical results that compared with the algorithm ADMM, the new algorithm has shorter CPU time and smaller errors, which also verifies the advantages of convex combination technique for solving large-scale problems.

Figure 1-2 shows the decline curves of the two algorithms on the behavior of the function value residual $\frac{\left|\|X^{k+1}\|_* - \|X^*\|_*\right|}{\|X^*\|_*}$ versus both the number of iterations and the CPU time. We set $n = 5000,\ 10000$, $r = 10$, $p = 0.3,\ 0.4,\ 0.5$. The new algorithm performs better than the algorithm ADMM in both the number of iterations and the CPU time, and can more quickly converge to the optimal solution.

6. Conclusion

In this paper, we have proposed an Improved ADMM (IADMM), which incorporates the convex combination technique into the proximal ADMM framework, for solving the matrix completion problem. Under relatively loose assumptions, its global

Table 1. Comparison of the results of ADMM and IADMM(p=0.3)

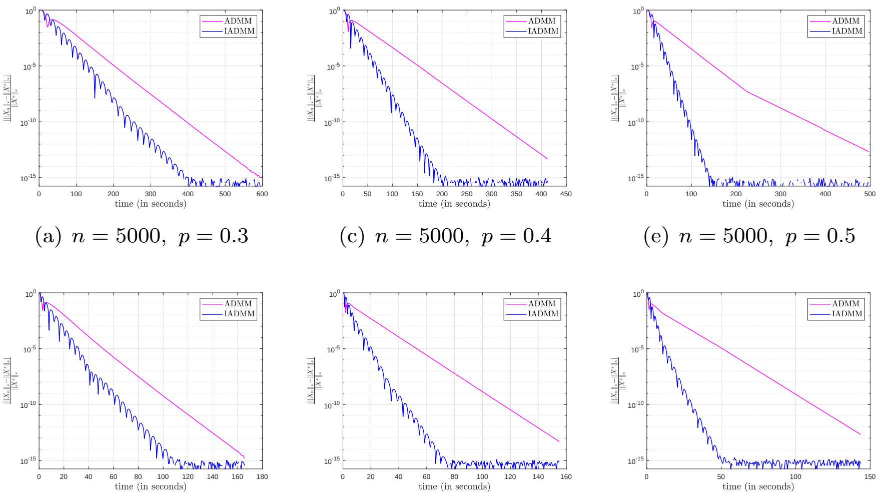
$n \times n$	r	ADMM				IADMM			
		iter	time/s	$\frac{\|X_n-X_{n-1}\|_F}{\|X_n\|_F}$	$\frac{\|X_n-M\|_F}{\|M\|_F}$	iter	time/s	$\frac{\|X_n-X_{n-1}\|_F}{\|X_n\|_F}$	$\frac{\|X_n-M\|_F}{\|M\|_F}$
1000	5	101	1.4119	9.5678e-07	8.8109e-06	72	0.9934	8.7314e-07	2.5673e-06
1000	10	106	1.2994	9.2361e-07	7.7516e-06	70	0.9706	8.9339e-07	2.3134e-06
2000	5	100	4.7269	9.0913e-07	7.9889e-06	73	3.7799	8.6120e-07	2.0823e-06
2000	10	101	5.4022	9.5376e-07	8.4529e-06	74	4.3605	8.4989e-07	1.9393e-06
3000	5	99	10.8490	9.8429e-07	8.3981e-06	74	9.0119	9.1952e-07	1.8489e-06
3000	10	100	13.0014	9.7021e-07	8.5037e-06	75	11.0213	9.6685e-07	1.9930e-06
4000	5	99	20.2546	9.6283e-07	8.0473e-06	74	16.8570	9.8543e-07	2.0918e-06
4000	10	100	22.7662	9.2118e-07	7.9402e-06	75	18.9113	9.6964e-07	2.3568e-06
5000	5	99	30.8722	9.5251e-07	7.8608e-06	75	25.2198	9.0432e-07	1.6424e-06
5000	10	100	36.0881	8.9919e-07	7.6525e-06	76	30.1085	9.2066e-07	2.0620e-06
10000	5	99	119.4883	9.3244e-07	7.4691e-06	75	110.6117	9.0252e-07	1.8136e-06
10000	10	99	146.7775	9.5782e-07	7.8425e-06	76	124.7390	8.1456e-07	1.9646e-06

Table 2. Comparison of the results of ADMM and IADMM(p=0.4)

$n \times n$	r	ADMM				IADMM			
		iter	time/s	$\frac{\|X_n-X_{n-1}\|_F}{\|X_n\|_F}$	$\frac{\|X_n-M\|_F}{\|M\|_F}$	iter	time/s	$\frac{\|X_n-X_{n-1}\|_F}{\|X_n\|_F}$	$\frac{\|X_n-M\|_F}{\|M\|_F}$
1000	5	105	1.1702	9.6564e-07	9.2651e-06	54	0.6534	8.5622e-07	1.5170e-06
1000	10	105	1.2453	9.4364e-07	9.1561e-06	55	0.7470	8.7139e-07	1.9067e-06
2000	5	106	4.8286	9.5538e-07	9.1589e-06	54	2.7135	8.1807e-07	1.6965e-06
2000	10	106	5.5055	9.4669e-07	9.1173e-06	55	3.2375	9.1553e-07	2.1377e-06
3000	5	106	11.5044	9.8756e-07	9.4532e-06	54	6.4046	9.3919e-07	1.5129e-06
3000	10	106	13.7442	9.8234e-07	9.4367e-06	56	8.0697	9.0278e-07	1.3223e-06
4000	5	107	21.7207	9.0776e-07	8.6792e-06	54	12.1955	8.7975e-07	1.7054e-06
4000	10	106	23.7111	9.9751e-07	9.5604e-06	55	13.6618	9.2709e-07	1.8582e-06
5000	5	107	32.7348	9.1750e-07	8.7693e-06	54	18.1193	8.5652e-07	1.6327e-06
5000	10	107	38.3070	9.1271e-07	8.7402e-06	55	21.3278	8.3023e-07	1.9938e-06
10000	5	107	133.2689	9.3464e-07	8.9221e-06	54	77.3902	8.3205e-07	1.2710e-06
10000	10	107	151.2460	9.3303e-07	8.9165e-06	55	89.3874	8.7968e-07	1.1562e-06

Table 3. Comparison of the results of ADMM and IADMM(p=0.5)

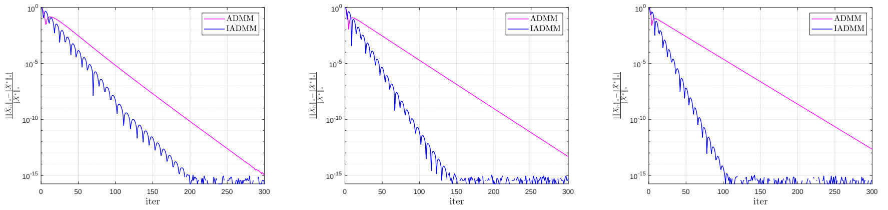
$n \times n$	r	ADMM				IADMM			
		iter	time/s	$\frac{\lVert X_n - X_{n-1} \rVert_F}{\lVert X_n \rVert_F}$	$\frac{\lVert X_n - M \rVert_F}{\lVert M \rVert_F}$	iter	time/s	$\frac{\lVert X_n - X_{n-1} \rVert_F}{\lVert X_n \rVert_F}$	$\frac{\lVert X_n - M \rVert_F}{\lVert M \rVert_F}$
1000	5	107	1.1359	9.7795e-07	1.0020e-05	43	0.5173	7.7578e-07	1.3779e-06
1000	10	107	1.2637	9.7308e-07	9.9414e-06	45	0.5874	9.6398e-07	1.8609e-06
2000	5	108	4.8649	9.6447e-07	9.9125e-06	42	2.0723	9.4963e-07	1.3847e-06
2000	10	108	5.5684	9.6620e-07	9.9220e-06	44	2.5034	9.6617e-07	1.3622e-06
3000	5	108	11.6689	9.9390e-07	1.0216e-05	42	5.2396	7.5947e-07	1.1191e-06
3000	10	108	13.7358	9.9429e-07	1.0215e-05	43	6.0485	8.9154e-07	1.3206e-06
4000	5	109	21.7999	9.2061e-07	9.4656e-06	41	8.8231	9.5879e-07	1.4500e-06
4000	10	109	23.8881	9.1929e-07	9.4475e-06	42	10.0619	9.0891e-07	1.7445e-06
5000	5	109	32.7849	9.2855e-07	9.5478e-06	41	13.6018	9.2695e-07	1.3892e-06
5000	10	109	38.5100	9.2819e-07	9.5429e-06	42	16.3592	8.0248e-07	1.6367e-06
10000	5	109	131.8586	9.4495e-07	9.7175e-06	41	57.8750	9.3215e-07	7.3885e-07
10000	10	109	152.2248	9.4488e-07	9.7162e-06	42	66.6169	8.1913e-07	7.2038e-07



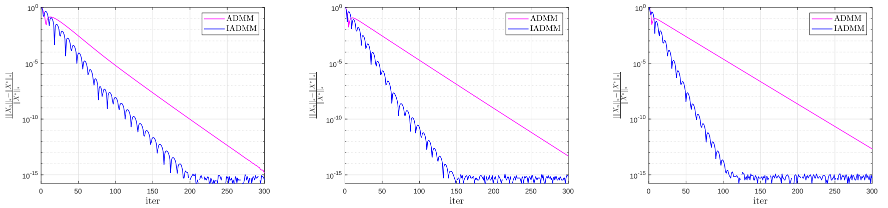
(a) $n = 5000,\ p = 0.3$ (c) $n = 5000,\ p = 0.4$ (e) $n = 5000,\ p = 0.5$

(b) $n = 10000,\ p = 0.3$ (d) $n = 10000,\ p = 0.4$ (f) $n = 10000,\ p = 0.5$

Figure 1. Comparison of the results of ADMM and IADMM



(a) $n = 5000, \ p = 0.3$ (c) $n = 5000, \ p = 0.4$ (e) $n = 5000, \ p = 0.5$



(b) $n = 10000, \ p = 0.3$ (d) $n = 10000, \ p = 0.4$ (f) $n = 10000, \ p = 0.5$

Figure 2. Comparison of the results of ADMM and IADMM

convergence is proved. Finally, through the results of numerical experiments, we can see that compared with ADMM, the new algorithm has fewer iterations, shorter time and smaller errors, which shows the efficiency of the algorithm.

Note that the gold combination factor ψ of IADMM plays a vital role. It may be interesting to expand the value range of ϕ, which we leave as future work.

Acknowledgment

This work was supported by the special fund for Science and Technology Innovation Teams of Shanxi Province(202204051002018), Research and Teaching Research Funding Project for Returned Students in Shanxi Province(2022-170).

References

[1] Argyriou A, Evgeniou T, Pontil M, Multi-task feature learning, *Advances in Neural Information Processing Systems*, **19**, 2007, 41-48.

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

[3] Candès E J, Tao T, The power of convex relaxation: near-optimal matrix completion, *IEEE Transactions on Information Theory*, **56**, 5, 2010, 2053-2080.

[4] Chen C, Chan R H, Ma S, et al, Inertial proximal ADMM for linearly constrained separable convex optimization, *SIAM Journal on Imaging Sciences*, **8**, 4, 2015, 2239-2267.

[5] Chen C, He B, Yuan X, Matrix completion via an alternating direction method, *IMA Journal of Numerical Analysis*, **32**, 1, 2012, 227-245.

[6] Eckstein J, Some saddle-function splitting methods for convex programming, *Optimization Methods and Software*, **4**, 1, 1994, 75-83.

[7] Fazel M, Matrix rank minimization with applications, *PhD Thesis, Stanford University*, 2002.

[8] Fazel M, Hindi H, Boyd S P, Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices, *Proceedings of the 2003 American Control Conference*, 2003.

[9] Fazel M, Pong T K, Sun D, et al, Hankel matrix rank minimization with applications to system identification and realization, *SIAM Journal on Matrix Analysis and Applications*, **34**, 3, 2013, 946-977.

[10] Gabay D, Mercier B, A dual algorithm for the solution of nonlinear variational problems via finite element approximation, *Computers & Mathematics with Applications*, **2**, 1, 1976, 17-40.

[11] Glowinski R, Numerical methods for nonlinear variational problems, *New York: Springer*, 1984.

[12] Glowinski R, Le Tallec P, Augmented Lagrangian and operator-splitting methods in nonlinear mechanics, *Society for Industrial and Applied Mathematics*, 1989.

[13] Hale E T, Yin W, Zhang Y, Fixed-point continuation for l_1-minimization: methodology and convergence, *SIAM Journal on Optimization*, **19**, 3, 2008, 1107-1130.

[14] He B, Ma F, Yuan X, Optimally linearizing the alternating direction method of multipliers for convex programming, *Computational Optimization and Applications*, **75**, 2020, 361-388.

[15] Lin Z, Chen M, Ma Y, The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices, *ArXiv:1009.5055*, 2010.

[16] Ma S, Goldfarb D, Chen L, Fixed point and Bregman iterative methods for matrix rank minimization, *Mathematical Programming*, **128**, 2011, 321-353.

[17] Malitsky Y, Golden ratio algorithms for variational inequalities, *Mathematical Programming*, **184**, 2020, 383-410.

[18] Ouyang Y, Chen Y, Lan G, et al, An accelerated linearized alternating direction method of multipliers, *SIAM Journal on Imaging Sciences*, **8**, 1, 2015, 644-681.

[19] Shefi R, Teboulle M, Rate of convergence analysis of decomposition methods based on the proximal method of multipliers for convex minimization, *SIAM Journal on Optimization*, **24**, 1, 2014, 269–297.

[20] Spyromitros-Xioufis E, Tsoumakas G, Groves W, et al, Multi-target regression via input space expansion: treating targets as inputs, *Machine Learning*, **104**, 2016, 55-98.

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

[22] Tomasi C, Kanade T, Shape and motion from image streams: a factorization method, *International Journal of Computer Vision*, **9**, 2, 1992, 137-154.

[23] Wang X, Yuan X, The linearized alternating direction method of multipliers for Dantzig selector, *SIAM Journal on Scientific Computing*, **34**, 5, 2012, A2792-A2811.

[24] Xue H Y, Zhang S, Cai D, Depth image inpainting: improving low rank matrix completion with low gradient regularization, *IEEE Transactions on Image Processing*, **26**, 9, 2017, 4311-4320.

Received 04.05.2023, Accepted 16.11.2023