Research Article

A New Algorithm for Positive Semidefinite Matrix Completion

Fangfang Xu and Peng Pan

College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266590, China

Correspondence should be addressed to Fangfang Xu; xuff@sdust.edu.cn

Positive semidefinite matrix completion (PSDMC) aims to recover positive semidefinite and low-rank matrices from a subset of entries of a matrix. It is widely applicable in many fields, such as statistic analysis and system control. This task can be conducted by solving the nuclear norm regularized linear least squares model with positive semidefinite constraints. We apply the widely used alternating direction method of multipliers to solve the model and get a novel algorithm. The applicability and efficiency of the new algorithm are demonstrated in numerical experiments. Recovery results show that our algorithm is helpful.

1. Introduction

Matrix completion (MC) is the process of recovering the unknown or missing elements of a matrix. Under certain assumptions on the matrix, for example, low-rank or approximately low-rank, the incomplete matrix can be reconstructed very well [1, 2]. Matrix completion is widely applicable in many fields, such as machine learning, statistic analysis, system control, and image and video processing [3], where matrices with low-rank or approximately low-rank are widely used in the model construction.

Recently, there have been extensive research on the problems of low-rank matrix completion (LRMC). The affine rank minimization problem consists of finding a matrix of minimum rank that satisfies a given system of linear constraints. However, it is NP-hard (nondeterministic polynomial-time hard) due to the combinatorial nature of the rank function. References [1, 2] showed that the solution of LRMC could be found by solving a nuclear norm minimization problem under some reasonable conditions. The singular value thresholding (SVT) method [4] and fixed point continuation method using approximate singular value decomposition (FPCA) [5] are two well-known algorithms because of their good recoverability, fast speed, and robustness. SVT applied the linearized Bregman iterations to solve the unconstrained nuclear norm regularized linear least squares problem. FPCA used iterations based on an iterative shrinkage-thresholding algorithm and used the continuation technique together with

an approximate singular value decomposition procedure to accelerate the algorithm. Reference [6] proposed an accelerated proximal gradient singular value thresholding algorithm. A completely different model was developed in LMaFit [7], which was a nonlinear successive overrelaxation algorithm that only requires solving a linear least squares problem per iteration. More details on LRMC can be found in [1, 8–10] and references therein.

In practice, the completed matrix is often required to be positive semidefinite. For example, covariance matrix and its inverse: precision matrix of statistic analysis, are both positive semidefinite. Recently, there have been extensive research on high-dimensional covariance matrix estimation. They all motivate the development of positive semidefinite matrix completion (PSDMC). Reference [11] accomplished the matrix completion task in some special conditions and used the alternating direction method of multipliers (ADMM) [12–15] to solve the model. References [16, 17] proposed new models for nonnegative matrix completion and also used ADMM to solve them.

Our main contribution in this work is the development of an efficient algorithm for PSDMC. First of all, we present the nuclear norm regularized linear least squares model with positive semidefinite constraints. Because of its robustness, we choose it as the model of PSDMC in this paper. The structure of the model suggests an alternating minimization scheme, which is very suitable for solving large-scale problems. We give an exact ADMM-based algorithm, whose subproblems

are solved exactly. We test the new ADMM-based algorithm on two kinds of problems: random matrix completion problems and random low-rank approximation problems. Numerical experiments show that all our proposed algorithm outputs have satisfactory results. The paper is organized as follows. Section 2 presents models and algorithms of PSDMC. Some numerical results are given in Section 3.

The following notations will be used throughout this paper. Uppercase (lowercase) letters are used for matrices (column vectors). All vectors are column vectors; the subscript $(\cdot)^T$ denotes matrix and vector transposition. Diag(x) denotes a diagonal matrix with x on its main diagonal. $\mathbf{0}$ is a matrix of all zeros of proper dimension; I_n stands for the $n \times n$ identity matrix. The trace of $X \in \mathbb{R}^{m \times n}$, that is, the sum of the diagonal elements of X, is denoted by $\mathrm{tr}(X)$. The Frobenius norm of $X \in \mathbb{R}^{m \times n}$ is defined as $\|X\|_F = \sqrt{\sum_{i,j} |X_{i,j}|^2}$. The Euclidean inner product between two matrices $X \in \mathbb{R}^{m \times n}$ and $Z \in \mathbb{R}^{m \times n}$ is defined as $\langle X, Z \rangle = \sum_{i,j}(X_{i,j} Z_{i,j}) = \mathrm{tr}(X^\top Z)$. The inequality $X \succeq \mathbf{0}$ means that X is semidefinite positive. The equality $X = Z$ means that $X_{ij} = Z_{ij}$ for all entries (i, j).

2. ADMM-Based Methods for PSDMC

2.1. The Model of PSDMC.

The matrix completion problem of recovering a positive semidefinite low-rank matrix from a subset of its entries is

$$
\begin{aligned}
\min \quad & \mathrm{rank}(X) \\
\text{s.t.} \quad & X_{ij} = M_{ij}, \quad \forall (i, j) \in \Omega \\
& X \succeq \mathbf{0},
\end{aligned}
\tag{1}
$$

where $X \in \mathbb{R}^{m \times n}$ is the decision variable and Ω is the index set of p known elements of X.

Let $\mathscr{P}$ be the projection onto the subspace of sparse matrices with nonzeros restricted to the index set Ω; that is,

$$
\mathscr{P}_\Omega(X)_{ij} = \begin{cases} X_{ij}, & \text{if } (i, j) \in \Omega, \\ 0, & \text{otherwise.} \end{cases}
\tag{2}
$$

From the definition of $\mathscr{P}_\Omega$, we can reformulate the equality constraint in model (1) in terms of $\mathscr{P}_\Omega(X) = \mathscr{P}_\Omega(M)$. Due to the combinational property of the objective function rank($\cdot$), model (1) is NP-hard in general. Inspired by the success of matrix completion under nuclear norm in [1, 2, 10], we use the nuclear norm as an approximation to rank(X) to estimate the optimal solution X^* of model (1) from the following model:

$$
\begin{aligned}
\min \quad & \|X\|_* \\
\text{s.t.} \quad & \mathscr{P}_\Omega(X) = \mathscr{P}_\Omega(M) \\
& X \succeq 0,
\end{aligned}
\tag{3}
$$

where the nuclear norm $\|X\|_*$ of X is defined as the summation of the singular values of X; that is,

$$
\|X\|_* = \sum_{i=1}^{\min(m,n)} \sigma_i(X),
\tag{4}
$$

where $\sigma_i(X)$ is the ith largest singular value. Moreover, for a positive semidefinite matrix, $\|X\|_* = \langle I, X \rangle = \mathrm{tr}(X)$.

If the known elements of the matrix X are noise-free, that is to say, $\mathscr{P}_\Omega(M)$ is reliable, we will directly solve model (3) to conduct SDPMC. On the contrary, if the vector of known elements $\mathscr{P}_\Omega(M)$ is contaminated by noise, the constraints $\mathscr{P}_\Omega(X) = \mathscr{P}_\Omega(M)$ must be relaxed, resulting in the following problem:

$$
\begin{aligned}
\min_X \quad & \mathrm{tr}(X), \\
\text{s.t.} \quad & \|\mathscr{P}_\Omega(X) - \mathscr{P}_\Omega(M)\|_F \leqslant \delta, \\
& X \succeq \mathbf{0},
\end{aligned}
\tag{5}
$$

or the nuclear norm regularized linear least squares model with positive semidefinite constraints:

$$
\begin{aligned}
\min_X \quad & \mu\,\mathrm{tr}(X) + \frac{1}{2}\|\mathscr{P}_\Omega(X) - \mathscr{P}_\Omega(M)\|_F^2, \\
\text{s.t.} \quad & X \succeq \mathbf{0}.
\end{aligned}
\tag{6}
$$

Here, δ and μ are given parameters, whose values should be set according to the noise level. When the values of μ and δ are set properly, (5) and (6) are equivalent. Model (6) is usually preferred over (5) for the case of noisy observations. Our algorithms can be extended to treat (5) with minor modifications.

Actually, model (6) is especially useful in practice. The reason is that the known information is usually gotten from large surveys and contaminated by sampling error inevitably. In this paper, we choose model (6) as the model to conduct NMC.

2.2. An ADMM-Based Method for Model (6).

In this subsection, we present an algorithm developed for model (6). To facilitate an efficient use of ADMM, we introduce one new matrix (splitting) variable Y and consider an equivalent form of model (6):

$$
\begin{aligned}
\min_{X,Y} \quad & \mu\,\mathrm{tr}(X) + \frac{1}{2}\|\mathscr{P}_\Omega(Y) - \mathscr{P}_\Omega(M)\|_F^2, \\
\text{s.t.} \quad & X = Y, \\
& X \succeq \mathbf{0},
\end{aligned}
\tag{7}
$$

where $X, Y \in \mathbb{R}^{n \times n}$. The augmented Lagrangian function of model (7) is

$$
\begin{aligned}
\mathscr{L}(X, Y, \Pi) = {}& \mu\,\mathrm{tr}(X) + \frac{1}{2}\|\mathscr{P}_\Omega(Y) - \mathscr{P}_\Omega(M)\|_F^2 \\
& + \langle \Pi, X - Y \rangle + \frac{\rho}{2}\|X - Y\|_F^2,
\end{aligned}
\tag{8}
$$

where $\Pi \in \mathbb{R}^{n \times n}$ is a Lagrangian multiplier. $\rho > 0$ is a penalty parameter.

---

(1) Input $\mathscr{P}_\Omega(M)$, $I_m \geq 0$, and tol ≥ 0.
(2) Set μ, γ, and $\rho \geq 0$. Set X_0 and Y_0 as random matrices, and Π_0 as zero matrices of appropriate sizes.
(3) **while** *not converge* **do**
(4)   update X_k, Y_k, Π_k by the formulas (16a), (16b) and (16c).

ALGORITHM 1: An exact ADMM-based algorithm of PSDMC.

The alternating direction method of multipliers for model (7) is derived by successively minimizing $\mathscr{L}$ with respect to X and Y in an alternating fashion; namely,

$$X_{k+1} := \arg \min_{X \geq 0} \mathscr{L}(X, Y_k, \Pi_k), \tag{9a}$$

$$Y_{k+1} := \arg \min_{Y} \mathscr{L}(X_{k+1}, Y, \Pi_k), \tag{9b}$$

$$\Pi_{k+1} := \Pi_k + \gamma\rho(X_{k+1} - Y_{k+1}), \tag{9c}$$

where $\gamma \in (0, 1.618)$.

By rearranging the terms of (9a), it is equivalent to

$$\min_{X \geq 0} \mu \text{tr}(X) + \frac{\rho}{2} \|X - G_k\|_F^2, \tag{10}$$

where $G_k = Y_k - (1/\rho)\Pi_k$. Let its Eigenvalue Value Decomposition (EVD) be $G_k = U_k \text{Diag}(w_k)U_k^T$, where $U_k \in \mathbb{R}^{n \times n}$ and $w_k \in R^n$. Define the shrinkage operator $S_\nu(\cdot)$ as

$$S_\nu(w) = \overline{w}, \quad \text{with } \overline{w}_i = \begin{cases} w_i - \nu, & \text{if } w_i - \nu > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

Then

$$X_{k+1} := U_k \text{Diag}\left(S_{\mu/\rho}(w_k)\right)U_k^T \tag{12}$$

is an optimal solution of problem (9a).

By rearranging the terms of (9b), it is equivalent to

$$\min_{Y \in \mathbb{R}^{n \times n}} \frac{1}{2}\left\|\mathscr{P}_\Omega(Y) - \mathscr{P}_\Omega(M)\right\|_F^2 + \frac{\rho}{2}\left\|Y - \left(X_k + \frac{1}{\rho}\Pi_k\right)\right\|_F^2. \tag{13}$$

Model (13) can be split into two subproblems:

$$\min \quad \frac{1}{2}\left\|\mathscr{P}_\Omega(Y) - \mathscr{P}_\Omega(M)\right\|_F^2 + \frac{\rho}{2}\left\|\mathscr{P}_\Omega(Y) - \mathscr{P}_\Omega\left(X_k + \frac{1}{\rho}\Pi_k\right)\right\|_F^2, \tag{14}$$

$$\min \quad \left\|\mathscr{P}_{\widehat{\Omega}}(Y) - \mathscr{P}_{\widehat{\Omega}}\left(X_k + \frac{1}{\rho}\Pi_k\right)\right\|_F^2,$$

where $\widehat{\Omega}$ is the complement of Ω. Finally, we can get the solution of (13) by solving the above two subproblems.

$$(Y_{k+1})_\Omega = \frac{1}{\rho + 1}\mathscr{P}_\Omega(M + \rho E_k),$$
$$(Y_{k+1})_{\widehat{\Omega}} = \mathscr{P}_{\widehat{\Omega}}(E_k), \tag{15}$$

where $E_k = X_k + (1/\rho)\Pi_k$.

In short, ADMM applied to model (7) yields the iteration:

$$X_{k+1} := U_k \text{Diag}\left(S_{\mu/\rho}(w_k)\right)U_k^T, \tag{16a}$$

$$(Y_{k+1})_\Omega := \frac{1}{\rho + 1}\mathscr{P}_\Omega(M + \rho E_k),$$
$$(Y_{k+1})_{\widehat{\Omega}} := \mathscr{P}_{\widehat{\Omega}}(E_k), \tag{16b}$$

$$\Pi_{k+1} := \Pi_k + \gamma\rho(X_{k+1} - Y_{k+1}). \tag{16c}$$

From the above considerations, we arrive at Algorithm 1.

The convergence of ADMM and its variants for convex problems has been studied extensively. The interested reader is referred to [12–15, 18–20], and the references therein. Since model (6) is convex, the convergence of Algorithm 1 to a global optimal solution is guaranteed.

3. Numerical Results

In this section, we report on the application of our proposed ADMM-based algorithm to a series of matrix problems to demonstrate its ability. To illustrate the performance of our algorithmic approaches combined with different procedures, we test the following two solvers.

(1) ADMM-VI; Algorithm 2 in [11].

(2) ADMM-SDP; Algorithm 1, the exact ADMM-based method for model (6).

We implement our algorithms in MATLAB. All the experiments are performed on a 2.20 GHz Intel Pentium PC with 6.0 GHz of memory and MATLAB 2012b.

3.1. Implementation and Parameters of the Two Solvers. We test the above two solvers on random positive semidefinite matrix problems. We do numerical experiments by the following procedure. Firstly, we create a low-rank positive semidefinite matrix $M \in \mathcal{S}_+^n$. Secondly, we select a subset of p elements uniformly at random from the $n \times n$ elements of M and denote their index set as Ω. The index set of unknown elements is denoted as $\widehat{\Omega}$. The ratio p/n^2 between the number of measurements and the number of entries in the matrix is denoted as "SR" (sampling ratio). We use $\mathscr{P}_\Omega(M)$ to complete X; the result is denoted as $\widehat{X}$. Finally, we will check the differences between $\widehat{X}$ and its actual value M.

$$\text{err} = \left\|\widehat{X} - M\right\|_F. \tag{17}$$

The most important algorithmic parameters in Algorithm 1 are μ, γ, tol, ρ, and the maximal number of iterations I_m.

TABLE 1: Numerical results of different values of μ ($n = 2000$, $r = 20$, $\sigma = 0.01$).

SR	$\mu = 10^{-2}$			$\mu = 10^{-1}$			$\mu = 1$			$\mu = 10$		
	Iter	Rank	Err	Iter	Rank	Err	Iter	Rank	Err	Iter	Rank	Err
0.25	67	297	$6.54e-03$	57	20	$2.93e-03$	44	20	$7.79e-03$	72	20	$8.80e-02$
0.35	50	274	$5.61e-03$	38	20	$2.37e-03$	32	20	$7.17e-03$	51	20	$5.91e-02$
0.45	39	194	$3.90e-03$	28	20	$1.97e-03$	26	20	$5.05e-03$	39	20	$4.62e-02$
0.55	33	137	$2.91e-03$	22	20	$1.77e-03$	21	20	$3.53e-03$	31	20	$3.69e-02$
0.65	29	97	$2.26e-03$	17	20	$1.58e-03$	17	20	$3.15e-03$	26	20	$2.97e-02$
0.75	26	55	$1.71e-03$	15	20	$1.43e-03$	14	20	$2.87e-03$	22	20	$2.62e-02$
0.85	23	20	$1.46e-03$	12	20	$1.37e-03$	11	20	$2.75e-03$	19	20	$2.25e-02$

TABLE 2: Numerical results on medium randomly created matrix completion problems.

(n, r, SR)	ADMM-VI			ADMM-SDP		
	Iter	Rank	Err	Iter	Rank	Err
(500, 10, 0.45)	24	10	$2.43e-03$	29	10	$2.62e-03$
(500, 10, 0.65)	16	10	$2.12e-03$	19	10	$2.09e-03$
(500, 10, 0.85)	10	10	$1.90e-03$	13	10	$1.88e-03$
(2000, 10, 0.45)	27	10	$1.86e-03$	26	10	$1.53e-03$
(2000, 10, 0.65)	18	10	$1.19e-03$	17	10	$1.17e-03$
(2000, 10, 0.85)	11	10	$1.24e-03$	13	10	$1.04e-03$
(2000, 20, 0.45)	27	20	$2.12e-03$	28	20	$2.20e-03$
(2000, 20, 0.65)	16	20	$1.84e-03$	18	20	$1.69e-03$
(2000, 20, 0.85)	11	20	$1.46e-03$	13	20	$1.38e-03$

In our implementation, we set $\gamma = 1.618$, tol $= 2 \times 10^{-4}$, $\rho = 2.5/n$, and $I_m = 300$. However, the value of μ is very difficult to set since it can be neither too large nor too small. μ is usually chosen to be a moderate value. Three sets of results are computed when μ is equal to three different values. They are shown in Table 1.

We can use continuation technique employed in [5, 6] to accelerate the convergence of ADMM-SDP. If model (6) is to be solved with the target parameter value $\mu = \overline{\mu}$, we propose solving a sequence of model (6) by a decreasing sequence $\{\mu^0, \mu^1, \ldots\}$. When a new problem, associated with μ^{j+1}, is to be solved, the approximate solution for the current problem with $\mu = \mu^j$ is used as the starting point. In our numerical experiments in Section 3.2, we set the initial $\mu^0 = 1$ and update $\mu^k = \mu^{k-1}/1.01$ at iteration k. A stopping criterion of Algorithm 1 is met as long as all the following three conditions are satisfied:

$$\left\| X_k - X_{k-1} \right\|_F \leq \text{tol}, \tag{18a}$$

$$\left\| X_k - Y_k \right\|_F \leq \text{tol}, \tag{18b}$$

$$\left\| \mathscr{P}_\Omega \left(X_k \right) - \mathscr{P}_\Omega \left(M \right) \right\|_F \leq \text{tol}. \tag{18c}$$

3.2. Experiments on Random Matrix Completion Problems. The matrix $M \in \mathcal{S}_+^{m \times n}$ with rank r in this subsection is created randomly by the following procedure (see also [5, 7]): two random matrices $M_L \in \mathbb{R}^{n \times r}$ and $M_R \in R^{n \times r}$ with i.i.d. standard Gaussian entries are first generated, $\Sigma \in \mathbb{R}^{n \times n}$ is the

noise matrix, and $\sigma > 0$. Then $M = M_L M_R^T + \sigma \Sigma$ is assembled. From the creation process of M, its rank is no larger than r.

The computational results of positive semidefinite matrix completion are presented in Table 2. We can observe that our proposed solver performs as well as ADMM-VI. Both of them can output satisfactory results.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

References

[1] E. J. Candés and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[2] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Review*, vol. 52, no. 3, pp. 471–501, 2010.

[3] H. Ji, C. Liu, Z. Shen, and Y. Xu, "Robust video denoising using Low rank matrix completion," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 1791–1798, IEEE, San Francisco, Calif, USA, June 2010.

[4] J.-F. Cai, E. J. Candés, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[5] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and Bregman iterative methods for matrix rank minimization," *Mathematical Programming*, vol. 128, no. 1-2, pp. 321–353, 2011.

[6] K.-C. Toh and S. Yun, "An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems," *Pacific Journal of Optimization*, vol. 6, no. 3, pp. 615–640, 2010.

[7] Z. Wen, W. Yin, and Y. Zhang, "Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm," *Mathematical Programming Computation*, vol. 4, no. 4, pp. 333–361, 2012.

[8] M. Fazel, *Matrix rank minimization with application [Ph.D. thesis]*, Stanford University, Stanford, Calif, USA, 2002.

[9] D. Goldfarb, S. Ma, and Z. Wen, "Solving low-rank matrix completion problems efficiently," in *Proceedings of the 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton '09)*, pp. 1013–1020, Monticello, Ill, USA, September 2009.

[10] E. J. Candés and T. Tao, "The power of convex relaxation: near-optimal matrix completion," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2053–2080, 2010.

[11] C. Chen, B. He, and X. Yuan, "Matrix completion via an alternating direction method," *IMA Journal of Numerical Analysis*, vol. 32, no. 1, pp. 227–245, 2012.

[12] M. Fortin and R. Glowinski, "Augmented Lagrangian methods," in *Applications to the Numerical Solution of Boundary Value Problems*, Studies in Mathematics and its Applications, North-Holland, Amsterdam, The Netherlands, 1983, Translated from the French by B. Hunt and D. C. Spicer.

[13] B. S. He, H. Yang, and S. L. Wang, "Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities," *Journal of Optimization Theory and Applications*, vol. 106, no. 2, pp. 337–356, 2000.

[14] Z. Wen, D. Goldfarb, and W. Yin, "Alternating direction augmented Lagrangian methods for semidefinite programming," *Mathematical Programming Computation*, vol. 2, no. 3-4, pp. 203–230, 2010.

[15] W. Deng and W. Yin, "On the global and linear convergence of the generalized alternating direction method of multipliers," Tech. Rep., Rice University CAAM, 2012.

[16] Y. Xu, W. Yin, Z. Wen, and Y. Zhang, "An alternating direction algorithm for matrix completion with nonnegative factors," *Frontiers of Mathematics in China*, vol. 7, no. 2, pp. 365–384, 2012.

[17] F. Xu and G. He, "New algorithms for nonnegative matrix completion," *Pacific Journal of Optimization*, vol. 11, no. 3, pp. 459–469, 2015.

[18] L. M. Briceño-Arias and P. L. Combettes, "A monotone+skew splitting model for composite monotone inclusions in duality," *SIAM Journal on Optimization*, vol. 21, no. 4, pp. 1230–1250, 2011.

[19] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, vol. 9 of *SIAM Studies in Applied Mathematics*, Society for Industrial and Applied Mathematics, Philadelphia, Pa, USA, 1989.

[20] B. He and X. Yuan, "Linearized alternating direction method of multipliers with Gaussian back substitution for separable convex programming," *Numerical Algebra, Control and Optimization*, vol. 3, no. 2, pp. 247–260, 2013.