EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

Matrix Completion via Successive Low-rank Matrix Approximation

Jin Wang[1,*], Zeyao Mo[2]

[1]Institute of Applied Physics and Computational Mathematics, Beijing 100094, PR China
[2]China Academy of Engineering Physics, Mianyang 621900, PR China

Abstract

In this paper, a successive low-rank matrix approximation algorithm is presented for the matrix completion (MC) based on hard thresholding method, which approximate the optimal low-rank matrix from rank-one matrix step by step. The algorithm enables the distance between the matrix with the observed elements and the projection on low-rank manifold to be minimum. The optimal low-rank matrix with observed elements is obtained when the distance is zero. In theory, convergence and convergent error of the new algorithm are analyzed in detail. Furthermore, some numerical experiments show that the algorithm is more effective in CPU time and precision than the orthogonal rank-one matrix pursuit(OR1MP) algorithm and the augmented Lagrange multiplier (ALM) method when the sampling rate is low.

Received on 24 November 2022; accepted on 15 December 2022; published on 04 January 2023

Keywords: matrix completion, low-rank matrix approximation, hard thresholding

doi:10.4108/eetsis.v10i3.2878

1. Introduction

The study of recovering an unknown low-rank or approximately low-rank matrix from incomplete samples of its entries has aroused great interest in machine learning([1], [2], [3]), imaging inpainting([4],[5], [6]), computer vision([7]), control([8],[9]), and so on. This problem is well known as the MC problem, mathematically,

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

where $rank(X)$ denotes the rank of matrix X. where $\mathbf{X}, \mathbf{D} \in R^{m\times n}$, $\mathcal{P}_\Omega(\mathbf{D})$ is the observed matrix, Ω is the set of all index pairs (i,j) of observed entries, $\mathcal{P}_\Omega$ is the orthogonal projector onto Ω.

Since the problem (1) is discontinuous and generally NP-hard, continuous optimization models instead of (1) were presented such as the convex optimization models (see [10–15]) and non-convex optimization model (see [16–19]).

For the problem (1), *wang* presented the orthogonal rank-one matrix pursuit algorithm, which only the top singular vector pair is calculated in each iteration and can obtain an ϵ-feasible solution in only $O(\log(1/\epsilon))$ iterations with less computation (see [20]). But the algorithm is not satisfied in its precision. This motivates that the new algorithm is designed to improve the precision. In this paper, we present a simple and efficient algorithms to solve the low-rank matrix completion problem (1). Repeat iteration to decrease the error between the matrix with the known elements and its corresponding elements of the projection matrix on {r}-dimensional manifold generated by hard thresholding method. If the error is not reduced and not less than the given precision, more better approximation on {r+1}-dimensional manifold is completed. Thus, repeat the operation till convergence.

The rest of the paper is organized as follows. The successive low-rank matrix approximation algorithms is presented in Section 2. Convergence and convergent error are discussed in Section 3. Section 4 shows the performance of the algorithm compared to the OR1MP and ALM algorithms and a conclusion remark.

Notation. Let $\mathbf{X} = (\mathbf{x}_1, ..., \mathbf{x}_n) \in R^{m\times n}$ be an $m \times n$ real matrix, where $\mathbf{x}_i \in R^{m\times 1}$. $\mathbf{X}_{ij}$ denotes the (i,j)th entry of $\mathbf{X}$. The nuclear norm of a matrix $\mathbf{X}$ is denoted by $\|\mathbf{X}\|_*$, the Frobenius norm by $\|\mathbf{X}\|_F$. $r(\mathbf{X})$ denotes

*Corresponding author. Email: wangjin910103@163.com

the rank of a matrix $\mathbf{X}$. $\Omega \subset \{1,...,m\} \times \{1,...,n\}$ is the indices of the observed entries of a matrix $\mathbf{X}$, $\bar{\Omega}$ is the complementary set of Ω. $\mathcal{P}_\Omega$ is the projection operator onto Ω. $[\mathbf{U}, \Sigma, \mathbf{V}] = lansvd(\mathbf{Y}, r, L)$ denotes computing the top-r singular pairs of the matrix $\mathbf{Y}$ by using the Lanczos method.

2. Successive Low-rank Matrix Approximation Algorithm

In first, the distance is introduced between the matrix with the observed elements and its projection onto the r-dimensional manifold as following,

$$d(\mathbf{Y}, r) = \min_{dim(\mathbf{X})=r} \|\mathbf{Y} - \mathbf{X}\|_F^2, \quad \text{s.t. } \mathcal{P}_\Omega(\mathbf{Y}) = \mathcal{P}_\Omega(\mathbf{D}) \tag{2}$$

$$d(\mathbf{Y}^*, r) = \min_{Y} d(\mathbf{Y}, r), \quad \text{s.t. } \mathcal{P}_\Omega(\mathbf{Y}) = \mathcal{P}_\Omega(\mathbf{D}) \tag{3}$$

Obviously, if $d(\mathbf{Y}^*, r) = 0$, Y^* is the optimal solution of (1).

Y^* is researched by repeated iterations in the matrix set with the observed elements until $d(\mathbf{Y}, r)$ to be minimum. Thus, $d(\mathbf{Y}^*, r) = 0$ can be obtained as rank r raises. The key steps of the algorithm are given in Algorithm 1 in the following.

Algorithm 1 (Successive low-rank matrix approximation (SLRMA) algorithm)

Initialize: Set an initial matrix $\mathcal{P}_\Omega(\mathbf{D})$, a tolerance parameter ϵ, $0 < c < 1$, $\mathbf{Y}_0 = \mathcal{P}_\Omega(\mathbf{D})$, $\mathbf{X}_{-1} = 0$, $r := 1$, $k := 0$;

repeat [Step 1:]

1. $[\mathbf{U}_k, \Sigma_k, \mathbf{V}_k] = lansvd(\mathbf{Y}_k, r, L)$;
2. Set $\mathbf{X}_k = \mathbf{U}_k \Sigma_k \mathbf{V}_k^T$;
3. If $\|\mathcal{P}_\Omega(\mathbf{D}) - \mathcal{P}_\Omega(\mathbf{X}_k)\|_F \le c \|\mathcal{P}_\Omega(\mathbf{D}) - \mathcal{P}_\Omega(\mathbf{X}_{k-1})\|_F$, $r := r$; otherwise $r := r + 1$;
4. Set $\mathbf{Y}_{k+1} = \mathcal{P}_\Omega(\mathbf{D}) + \mathcal{P}_{\bar{\Omega}}(\mathbf{X}_k)$, $k := k + 1$;

until $\|\mathcal{P}_\Omega(\mathbf{D}) - \mathcal{P}_\Omega(\mathbf{X}_k)\|_F / \|\mathcal{P}_\Omega(\mathbf{D})\|_F \le \epsilon$

Output: $\mathbf{X}_k$

3. Convergence Analysis

In this section, the iteration error and convergence of the new algorithm are discussed.

Lemma 1 Let $\mathbf{X}, \mathbf{Y} \in R^{n \times n}$, $r(\mathbf{X}) = r$, $r(\mathbf{Y}) \ge r$. *Then*

$$\min_{r(\mathbf{X})=r} \|\mathbf{Y} - \mathbf{X}\| = \left\|\mathbf{Y} - \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T - \cdots - \sigma_r \mathbf{u}_r \mathbf{v}_r^T\right\| \tag{4}$$

where $\|\cdot\|$ is $\|\cdot\|_F$ or $\|\cdot\|_2$, $\mathbf{Y} = \mathbf{U}\Sigma\mathbf{V}^\mathbf{T} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T + \cdots + \sigma_n \mathbf{u}_n \mathbf{v}_n^T$. Let

$$\mathcal{R}(\mathbf{Y}, r) = \mathbf{Y} - \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \tag{5}$$

obviously,

$$d(\mathbf{Y}, r) = \|\mathcal{R}(\mathbf{Y}, r)\|_F^2 = (\sigma_{r+1}^2 + \cdots + \sigma_n^2).$$

Lemma 2 Let $\mathbf{Y}^*$ be the optimal solution of (3). Then $\mathcal{P}_{\bar{\Omega}}(\mathcal{R}(\mathbf{Y}^*, r)) = 0$.

Proof. Because of

$$\frac{\partial d(\mathbf{Y}, r)}{\partial \mathbf{Y}} = 2\left(\sum_{i=r+1}^{n} \sigma_i \mathbf{u}_i \mathbf{v}_i^T\right) = 2(\mathcal{R}(\mathbf{Y}, r)).$$

According to KKT condition, if Y^* is the optimal solution, $\frac{\partial d(\mathbf{Y}^*, r)}{\partial \mathbf{Y}} = \mathcal{P}_\Omega(U^*)$, where U^* is the Lagrange multiplier. This implies $\mathcal{P}_{\bar{\Omega}}(\mathcal{R}(\mathbf{Y}^*, r)) = 0$.

Lemma 3 The $d(\mathbf{Y}, r)$ obeys $d(\mathbf{Y}, r) \le (1 - \frac{1}{n-r+1}) d(\mathbf{Y}, r-1)$.

Proof. $d(\mathbf{Y}, r-1) = \sigma_r^2 \cdots + \sigma_{r(\mathbf{Y})}^2 = d(\mathbf{Y}, r) + \sigma_r^2$. Since $\sigma_r \ge \sigma_{r+1} \ge \cdots \ge \sigma_{r(\mathbf{Y})}$,

$$\begin{aligned} d(\mathbf{Y}, r) &= d(\mathbf{Y}, r-1) - \sigma_r^2 \\ &\le d(\mathbf{Y}, r-1) - \frac{1}{r(\mathbf{Y}) - r + 1} d(\mathbf{Y}, r-1) \\ &= (1 - \frac{1}{r(\mathbf{Y}) - r + 1}) d(\mathbf{Y}, r-1) \\ &\le (1 - \frac{1}{n - r + 1}) d(\mathbf{Y}, r-1). \end{aligned}$$

The lemma has been proved.

Theorem 1 Let $\{\mathbf{Y}_k\}$ be the matrix sequence generated by Algorithm 1. Then

$$\|\mathbf{Y}_{k+1} - \mathbf{X}_{k+1}\|_F^2 \le \|\mathbf{Y}_k - \mathbf{X}_k\|_F^2 - \|\mathbf{Y}_k - \mathbf{Y}_{k+1}\|_F^2. \tag{6}$$

Proof. By direct computing, it is

$$\begin{aligned} \|\mathbf{Y}_{k+1} - \mathbf{X}_{k+1}\|_F^2 &\le \|\mathbf{Y}_{k+1} - \mathbf{X}_k\|_F^2 \\ &= \|\mathcal{P}_\Omega(\mathbf{Y}_{k+1} - \mathbf{X}_k)\|_F^2 \\ &= \|\mathcal{P}_\Omega(\mathbf{Y}_k - \mathbf{X}_k)\|_F^2 \\ &= \|\mathbf{Y}_k - \mathbf{X}_k\|_F^2 - \left\|\mathcal{P}_{\bar{\Omega}}(\mathbf{Y}_k - \mathbf{X}_k)\right\|_F^2 \\ &= \|\mathbf{Y}_k - \mathbf{X}_k\|_F^2 - \|\mathbf{Y}_k - \mathbf{Y}_{k+1}\|_F^2. \end{aligned}$$

Thus, (6) holds true.

Theorem 2 If $\{\mathbf{Y}_k\}$ satisfies $\mathcal{P}_{\bar{\Omega}}(\mathcal{R}(\mathbf{Y}_k, r_k)) \ne 0$, then there is a positive constant $0 < c_k < 1$ such that

$$\|\mathcal{P}_\Omega(\mathbf{D}) - \mathcal{P}_\Omega(\mathbf{X}_k)\|_F \le c_k \|\mathcal{P}_\Omega(\mathbf{D}) - \mathcal{P}_\Omega(\mathbf{X}_{k-1})\|_F. \tag{7}$$

Proof. Suppose that $\mathbf{Y}_k$ is generated by Algorithm **??**,

$$\begin{aligned} \|\mathcal{P}_\Omega(\mathbf{D} - \mathbf{X}_k)\|_F^2 &= \|\mathcal{P}_\Omega(\mathbf{Y}_{k+1} - \mathbf{X}_k)\|_F^2 \\ &= \|\mathbf{Y}_{k+1} - \mathbf{X}_k\|_F^2 - \left\|\mathcal{P}_{\bar{\Omega}}(\mathbf{Y}_k - \mathbf{X}_k)\right\|_F^2 \\ &\le \|\mathbf{Y}_k - \mathbf{X}_{k-1}\|_F^2 - \left\|\mathcal{P}_{\bar{\Omega}}(\mathbf{Y}_k - \mathbf{X}_k)\right\|_F^2 \\ &= \|\mathcal{P}_\Omega(\mathbf{Y}_k - \mathbf{X}_{k-1})\|_F^2 - \left\|\mathcal{P}_{\bar{\Omega}}(\mathcal{R}(\mathbf{Y}_k, r_k))\right\|_F^2 \\ &= \|\mathcal{P}_\Omega(\mathbf{D} - \mathbf{X}_{k-1})\|_F^2 - \left\|\mathcal{P}_{\bar{\Omega}}(\mathcal{R}(\mathbf{Y}_k, r_k))\right\|_F^2, \end{aligned}$$

If $\mathcal{P}_{\bar{\Omega}}(\mathcal{R}(\mathbf{Y}_k, r_k)) \neq 0$, let $\epsilon_k > 0$ and $\left\|\mathcal{P}_{\bar{\Omega}}(\mathcal{R}(\mathbf{Y}_k, r_k))\right\|_F^2 \geq \epsilon_k \left\|\mathcal{P}_{\Omega}(\mathbf{D} - \mathbf{X}_{k-1})\right\|_F^2$. Then $\left\|\mathcal{P}_{\Omega}(\mathbf{D} - \mathbf{X}_k)\right\|_F^2 \leq (1 - \epsilon_k)\left\|\mathcal{P}_{\Omega}(\mathbf{D} - \mathbf{X}_{k-1})\right\|_F^2$. Let $c_k = \sqrt{1 - \epsilon_k}$. Then (7) is true.

Theorem 3 Let $\mathbf{Y}_k$ be the matrix sequence generated by Algorithm 1 and $0 < c = \max_k c_k < 1$. Then the following inequality holds

$$\left\|\mathcal{P}_{\Omega}(\mathcal{R}(\mathbf{Y}_k, r))\right\|_F^2 \leq \prod_{i=1}^{r}(1 - \frac{1}{n-i+1})(c^2)^{m_i}\left\|\mathcal{P}_{\Omega}(\mathbf{Y}_0)\right\|_F^2, \tag{8}$$

where $\sum_{i=1}^{r} m_i = k - r$.

Proof. Suppose that (8) is true for $k - m_k$ and $r - 1$. From Algorithm 1 it is

$$\left\|\mathcal{P}_{\Omega}(\mathcal{R}(\mathbf{Y}_k, r))\right\|_F^2 = (c^2)^{m_k}\left\|\mathcal{P}_{\Omega}(\mathcal{R}(\mathbf{Y}_{k-m_k}, r))\right\|_F^2.$$

On the other hand, from Lemma 3 it has

$$\left\|\mathcal{P}_{\Omega}(\mathcal{R}(\mathbf{Y}_{k-m_k}, r))\right\|_F^2 \leq (1 - \frac{1}{n-r+1})\left\|\mathcal{P}_{\Omega}(\mathcal{R}(\mathbf{Y}_{k-m_k-1}, r-1))\right\|_F^2.$$

Thus, (8) holds true for k and r.

Remark: The proof of theorem 3 is divided to two parts, one is the distance shrinks $(c^2)^{m_i}$ by repeatedly iteration on the $\{i-1\}$- dimensional manifold, another is the distance which shrinks $(1 - \frac{1}{n-i+1})$ on the $\{i\}$-dimensional manifold.

4. Experiments

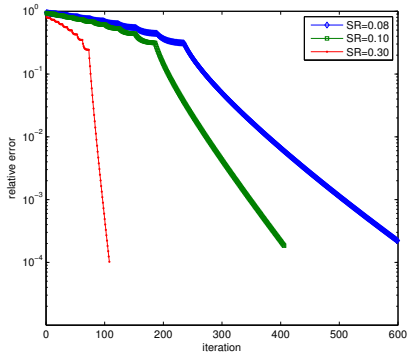
In this section, the new algorithm with the OR1MP algorithm (see [20]) and the ALM algorithm (see [14]) is compared. This is because the rank is not known for the model (1) and convex model in [[10–15]], while the rank is usually known in the non-convex model [[16–19]]. All the experiments are conducted on the same workstation.

In the experiments, "iteration" denotes the number of iterations, "rank" denotes the rank of the recovered solution, sr denotes the sampling ratio and $sr = p/n^2$, where p is the number of observed entries. Let $sr = 0.1, 0.3, 0.5$, respectively. The relative error is $rel.err = \frac{\|\mathbf{Y}_k - \mathbf{D}\|_F}{\|\mathbf{D}\|_F}$. For each (n, r), repeat the following procedure 5 times, and the reported results are for a mean. The rank-r matrix is generated by $\mathbf{D} = \mathbf{L}\mathbf{R}^T$, where $\mathbf{L} \in R^{n\times r}, \mathbf{R} \in R^{n\times r}$ are independent matrices with independent identically distributed (i.i.d.) Gaussian entries. For Algorithms 1, the empirical parameter set $c = 0.95$. For Table 1, the algorithms stops once the termination criterion $\epsilon = \frac{\|\mathcal{P}_{\Omega}(\mathbf{X}) - \mathcal{P}_{\Omega}(\mathbf{D})\|_F}{\|\mathcal{P}_{\Omega}(\mathbf{D})\|_F}$ is less than 10^{-4}. Since the OR1MP algorithm does not perform well, the maximum number of iterations of OR1MP algorithm is set to 100 (this is because the rank of the matrix generated by OR1MP method is up to 100).

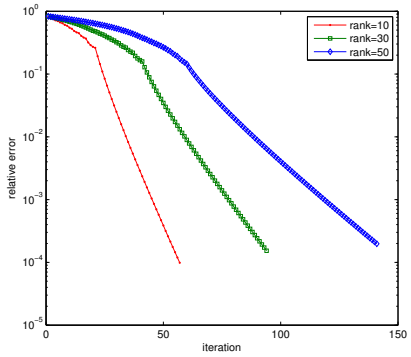
A brief comparison of the algorithms is presented in Table 1. We can see that the SLRMA algorithm achieve better than the OR1MP in the accuracy. The CPU time of the SLRMA algorithm is less than other algorithms. Although ALM method has the least iterations, it is not dominant in the CPU time. As the sampling rate declines, the SLRMA algorithm appears to converge well especially for large-scale matrix completion with less CPU time. Furthermore, except for the OR1MP algorithm, all the other algorithms can recover the rank exactly.

For further comparison, two experiments on a 1000×1000 matrix are tested the efficiency of the SLRMA algorithm (see Figure 1). It is shown from the figures that when the rank is fixed, the smaller the sampling ratio is, the less effectiveness the SLRMA algorithm runs; and when the sampling ratio is fixed, the higher the rank is, the less effectiveness the SLRMA algorithm runs.

For the paining image, the SLRMA algorithm has the same effect as the ALM algorithm (see Figure 2), but the SLRMA algorithm costs less time than the ALM algorithm.



a. m=n=1000, rank=10



b. m=n=1000, sr=0.3

Figure 1. convergence behavior of the relative error in SLRMA runs.

Table 1. Numerical results on random matrix completion problem

problem		SLRMA				ALM				OR1MP			
n	r	iteration	rank	time(s)	rel.err	iteration	rank	time(s)	rel.err	iteration	rank	time(s)	rel.err
sr=0.1													
1000	10	280	10	11.54	1.95e-4	114	10	58.68	1.46e-4	100	100	17.08	2.58e-1
2000	20	265	20	57.32	1.72e-4	118	20	586.87	1.02e-4	100	100	61.22	1.59e-1
3000	30	265	30	176.05	1.71e-4	123	30	2608.08	8.27e-5	100	100	140.13	1.57e-1
4000	40	268	40	357.01	1.70e-4	118	40	7311.72	1.01e-4	100	100	263.28	9.43e-1
sr=0.3													
1000	10	56	10	3.45	1.15e-4	43	10	9.07	8.10e-5	100	100	20.02	1.62e-2
2000	20	63	20	21.25	1.11e-4	43	20	52.87	7.93e-5	100	100	83.95	2.08e-2
3000	30	71	30	62.86	1.10e-4	42	30	122.33	9.75e-5	100	100	190.64	3.32e-2
4000	40	79	40	137.03	1.04e-4	42	40	242.51	9.92e-5	100	100	353.47	8.37e-1
sr=0.5													
1000	10	34	10	2.35	7.94e-5	24	10	6.52	5.85e-5	100	100	24.56	3.61e-2
2000	20	44	20	18.41	6.25e-5	24	20	29.66	6.48e-5	100	100	106.24	7.36e-2
3000	30	52	30	52.81	7.70e-5	24	30	76.22	6.02e-5	100	100	237.27	1.41e-2
4000	40	62	40	126.47	6.23e-5	24	40	147.79	5.91e-5	100	100	438.50	7.07e-1

5. Conclusion

In this paper, we first propose an low-rank matrix approximation algorithm. In the algorithm, by introducing the distance between the feasible matrix and its projection onto the r-dimensional manifold (where r is less than or equal to the rank of the completion matrix), the optimal low-rank matrix can be found by multi-step iterations until the minimum distance equal to zero. Meanwhile, Numerical experiments show that the new algorithms achieve much better performance in precision than the OR1MP algorithm and faster than the ALM algorithm, especially for large-scale matrix completion problem. As the observation ratio decreasing, the SLRMA algorithm consumes the least CPU time than the other algorithms. Besides, the SLRMA algorithm can recover the rank of the matrix being recovered exactly.

a. original image

b. 50% mask image

c. SLRMA(time 45.45s)

d. ALM(time 77.28s)

Figure 2. Image Inpainting

References

[1] Amit, Y., Fink, M., Srebro, N. and Ullman, S.(2007) *Uncovering shared structures in multiclass classification.* in: Processings of the 24th International Conference on Machine Learing, ACM, 17–24

[2] Argyriou, A.,Evgeniou, T., Pontil, M. (2007) *Multi-task feature learing* Adv. Neural Inf. Process Syst., 19, 41–48

[3] Sarki, R., Ahmed, K., Wang,H., Zhang,Y. (2020) *Automated detection of mild and multi-class diabetic eye diseases using deep learning*,Health Information Science and Systems 8(1),1–9.

[4] Bertalmio, M., Sapiro, G., Caselles, V., and Ballester, C. (2000) *Image inpainting*. Image inpainting, Comput. Graph, 34, 417–424

[5] Supriya,S., Siuly S., H Wang and Y Zhang (2020) *automated epilepsy detection techniques from electroencephalogram signalsa review study*, Health Information Science and Systems 8(1),1–15

[6] Pandey, D., H Wang, X Yin, K Wang, Y Zhang and J Shen (2022) *Automatic breast lesion segmentation in phase preserved DCE-MRIs*, Health Information Science and Systems 10(1),1–19

[7] Tomasi, C. and Kanade, T. (1992) *Shape and motion from image streams under orthography: a factorization method*, Int. J. Comput. Vis, 9, 137–154

[8] Mesbahi, M. and Papavassilopoulos, G. P. (1997) *On the rank minimization problem over a positive semidefinite linear matrix inequality*, IEEE Trans. Automat. Control, 42, 239–243

[9] J. Y. Li, Zhan,Z.H., H. Wang and J. Zhang (2020) *Data-driven evolutionary algorithm with perturbation-based ensemble surrogates*, IEEE Transations on Cybernetics, 51(8), 3925–3937

[10] Candès, E. J. and Recht, B. (2009) *Exact matrix completion via convex optimization*. Found. Comput. Math., 9, 717–772

[11] Ma, S., Goldfarb D. and Chen L. (2011) *Fixed point and bregman iterative methods for matrix rank minimization*, Math. Program, 128, 321–353

[12] Cai, J. F., Candès, E. J. and Shen, Z. (2010) *A singular value thresholding algorithm for matrix completion*, SIAM J. Optim., 20, 1956–1982

[13] Toh, K. C. and Yun S. (2010) *An accelerated proximal gradient algorithm for nuclear norm regularized least squares problems*, Pac. J. Optim., 6, 615–640

[14] Lin, Z., Chen, M., Wu, L. and Ma, Y. (2013) *The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices*, arXiv:1009.5055v3

[15] Chen, C., He, B. and Yuan, X. (2012) *Matrix completion via an alternating direction method*, IMA J. Numer. Anal., 32, 227–245

[16] Vandereycken, B. (2013) *Low-rank matrix completion by Riemannian optimization*, SIAM J. Optim., 23(2), 1214–1236

[17] Mishra, B. and Apuroop, K. A. (2012) *Sepulchre, R.: A Riemannian geometry for low-rank matrix completion*, arXiv:1211.1550

[18] Boumal, N. and Absil, P. A. (2011) *RTRMC: A Riemannian trust-region method for low-rank matrix completion*, Neural Information Processing Systems conference, NIPS

[19] Tanner, J. and Wei, K. *Low rank matrix completion by alternating steepest descent methods*, Appl. Comput. Harmon. Anal., 40, 417–429 (2016)

[20] Wang, Z., Lai, M., Lu, Z. and Fan, W., Davulcu, H. and Ye, J.(2014) *Orthogonal rank-one matrix pursuit for low rank matrix completion*, SIAM J. Sci. Comput., 37, 488–514