A Divide-and-conquer Method for Scalable Low-rank Latent Matrix Pursuit: Supplementary Material

Yan Pan, Hanjiang Lai, Cong Liu
Sun Yat-sen University, Guangzhou, China
{panyan5,laihanj,liucong3}@mail.sysu.edu.cn

Shuicheng Yan
National University of Singapore, Singapore
eleyans@nus.edu.sg

This supplementary material accompanies the paper "A Divide-and-Conquer Method for Scalable Low-rank Latent Matrix Pursuit". It provides the proofs of Theorem 1 and Theorem 2 in the main paper.

1. Proof of Theorem 1

To prove Theorem 1, we introduce a key lemma from [2].

Lemma 1. *([2], Theorem 2) Let $L^\star$, $S^\star \in \mathbb{R}^{m_1 \times m_2}$. Suppose that $L^\star$ is (μ, r)-coherent and that the support set of $S^\star$ is uniformly distributed among all sets with cardinality s. Then if $m_1 \le m_2$ and $||M - L^\star - S^\star||_F \le \Delta$, there is a constant c_p such that with probability at least $1 - c_p m_2^{-\beta}$, the minimizer (L', S') of*

$$\min_{L',S'} ||L'||_* + \lambda ||S'||_1 \; s.t. \; ||M - L' - S'||_F \le \Delta \; with \; \lambda = \frac{1}{m_2}$$

satisfies

$$||L' - L^\star||_F^2 + ||S' - S^\star||_F^2 \le c_\epsilon^2 m_1 m_2 \Delta^2,$$

provided that $r \le p_r m_1 \mu^{-1} log^{-2}(m_2)$ and $s \le (1 - p_s \beta) m_1 m_2$ for $\beta \ge 2$ and some positive constants p_r, p_s, c_ϵ.

Next we start to prove Theorem 1.

Proof. We denote $(T, E^{(i)})$ as the minimizer of the general formulation of the LLMP problem ((11) in the main paper)

$$\min_{T, E^{(i)}} ||T||_* + \lambda \sum_{i=1}^{n} ||E^{(i)}||_1$$
$$s.t. \; ||T^{(i)} - T - E^{(i)}||_F \le \delta, i = 1, 2, ..., n,$$

and $(T_0, E_0^{(i)})$ as the ground truth of low-rank and sparse decomposition of $T^{(i)}$. By the assumptions, T_0 is (μ_0, r_0)-coherent and T is (μ_T, r_T)-coherent.

For $i = 1, 2, ..., n$, we denote (T'_i, E'_i) as the minimizer of

$$\min_{T'_i, E'_i} ||T'_i||_* + \lambda ||E'_i||_1$$
$$s.t. \; ||T^{(i)} - T'_i - E'_i||_F \le \delta.$$

For $i = 1, 2, ..., n$, by the assumptions in Theorem 1, we have $||T^{(i)} - T_0 - E_0^{(i)}||_F \le \Delta$. In terms of Lemma 1 (by replacing $(L^\star, S^\star, L', S', \mu, r, s)$ with $(T_0, E_0^{(i)}, T'_i, E'_i, \mu_0, r_0, s_{0,i})$), we have $||T'_i - T_0||_F^2 + ||E'_i - E_0^{(i)}||_F^2 \le c_\epsilon^2 m_1 m_2 \Delta^2$ with probability at least $1 - c_p m_2^{-\beta}$.

Similarly, from the general form of the LLMP problem ((11) in the main paper), we have $||T^{(i)} - T - E^{(i)}||_F \le \delta$ for $i = 1, 2, ..., n$. In terms of Lemma 1 (by replacing $(L^\star, S^\star, L', S', \mu, r, s)$ with $(T, E^{(i)}, T'_i, E'_i, \mu_T, r_T, s_{T,i})$), we have $||T'_i - T||_F^2 + ||E'_i - E^{(i)}||_F^2 \le c_\epsilon^2 m_1 m_2 \Delta^2$ with probability at least $1 - c_p m_2^{-\beta}$.

Denote A_i as the event that $||T'_i - T_0||_F^2 + ||E'_i - E_0^{(i)}||_F^2 \le c_\epsilon^2 m_1 m_2 \Delta^2$ and B_i as the event that $||T'_i - T||_F^2 + ||E'_i - E^{(i)}||_F^2 \le c_\epsilon^2 m_1 m_2 \Delta^2$ $(i = 1, 2, ..., n)$. Then the joint probability of $\bigcap_{A_i, B_i}^{i=1,2,...,n}$ is at least $(1 - c_p m_2^{-\beta})^{2n} \ge 1 - 2nc_p m_2^{-\beta}$.

With $||T'_i - T_0||_F^2 + ||E'_i - E_0^{(i)}||_F^2 \le c_\epsilon^2 m_1 m_2 \Delta^2$ and $||T'_i - T||_F^2 + ||E'_i - E^{(i)}||_F^2 \le c_\epsilon^2 m_1 m_2 \Delta^2$, we have:

$$\begin{aligned} & ||T - T_0||_F^2 + ||E^{(i)} - E_0^{(i)}||_F^2 \\ \le & ||T'_i - T_0||_F^2 + ||E'_i - E_0^{(i)}||_F^2 \\ & + ||T'_i - T||_F^2 + ||E'_i - E^{(i)}||_F^2 \\ \le & 2c_\epsilon^2 m_1 m_2 \Delta^2, \end{aligned} \quad (1)$$

where the first inequality uses the triangle inequality.

Finally, by summing up (1) with $i = 1, 2, ..., n$ and then averaging by n, we have the desired result: $||T - T_0||_F^2 + \frac{1}{n} \sum_{i=1}^{n} ||E^{(i)} - E_0^{(i)}||_F^2 \le 2c_\epsilon^2 m_1 m_2 \Delta^2$ with probability at least $1 - 2nc_p m_2^{-\beta}$. □

2. Proofs of Theorem 2

The proof is identical to that of the second part of Theorem 12 in [1].

Proof. Remind that we define

$$\mathbf{T} = \begin{bmatrix} T_S & T_A \\ T_B & T_C \end{bmatrix}, \mathbf{R} = \begin{bmatrix} T_S \\ T_B \end{bmatrix}, \mathbf{C} = \begin{bmatrix} T_S & T_A \end{bmatrix}.$$

Correspondingly, we also define

$$\mathbf{T_0} = \begin{bmatrix} T_{0,S} & T_{0,A} \\ T_{0,B} & T_{0,C} \end{bmatrix}, \mathbf{R_0} = \begin{bmatrix} T_{0,S} \\ T_{0,B} \end{bmatrix},$$

$$\mathbf{C_0} = \begin{bmatrix} T_{0,S} & T_{0,A} \end{bmatrix}, \tilde{\mathbf{T}} = \begin{bmatrix} T_S & T_A \\ T_B & T_{0,C} \end{bmatrix}.$$

Define K as the event $||\tilde{T} - T||_F \leq (1+\epsilon)^2 ||T_0 - \tilde{T}||_F$, and $A(X)$ as the event that a matrix X is $(\frac{r_0 \mu_0^2}{1-\epsilon/2}, r_0)$-coherent.

When K holds, by the triangle inequality, we have

$$\begin{aligned} &||T_0 - T||_F \\ \leq &||T_0 - \tilde{T}||_F + ||T - \tilde{T}||_F \\ \leq &(2 + 2\epsilon + \epsilon^2)||T_0 - \tilde{T}||_F \\ \leq &(2+3\epsilon)||T_0 - \tilde{T}||_F \\ = &(2+3\epsilon)\sqrt{||T_{0,S} - T_S||_F^2 + ||T_{0,A} - T_A||_F^2 + ||T_{0,B} - T_B||_F^2} \\ = &(2+3\epsilon)\sqrt{||T_{0,S} - T_S||_F^2 + 2||T_{0,A} - T_A||_F^2}, \end{aligned}$$

where the last equality uses the facts $T_{0,B} = -T_{0,A}^T$ and $T_B = -T_A^T$.

Since $k \geq cr_0 \mu_0 log(m_2) log(1/\delta)/\epsilon^2 \geq cr_0 \mu_0 log(m_1) log(1/\delta)/\epsilon^2$, by Lemma 4 in [1], we have that A_C and A_R hold with probability at least $1 - \delta/(2m_2)$ and $1 - \delta/(2m_2)$, respectively. By Corollary 7 in [1], the event K holds with probability at least $(1-\delta/2)(0.8-\delta/2)$.

Denote X^c as the complementary event of X. Then we have:

$$\begin{aligned} &P(K \cap A_C \cap A_R) \geq 1 - P(K^c) - P(A_C^c) - P(A_R^c) \\ \geq &1 - (1 - (1-\delta/2)(0.8-\delta/2)) - \delta/(2m_2) - \delta/(2m_2) \\ \geq &(1-\delta/2)(0.8-\delta/2) - \delta/2 \\ \geq &(1-\delta)(0.8-\delta) \end{aligned}$$

for all $m_2 \geq 2$ and $\delta \leq 0.8$. □

References

[1] L. Mackey, A. Talwalkar, and M. Jordan. Divide-and-Conquer Matrix Factorization. *arXiv:1107.0789*, 2011. 2

[2] Z. Zhou, X. Li, J. Wright, E. J. Candes, and Y. Ma. Stable principal component pursuit. In IEEE International Symposium on Information Theory Proceedings (ISIT), 1518-1522, 2010. 1