Robust Matrix Completion via Alternating Projection

Xue Jiang

xuejiang@sjtu.edu.cn

Department of Electronic Engineering

Shanghai Jiao Tong University, China

This work was supported in part by the National Natural Science Foundation of China under Grant 61601284, in part by the Huawei Innovation Research Program 2016, and in part by the CityU under Project 7004431.

X. Jiang, Z. Zhong, X. Liu and H. C. So, "Robust matrix completion via alternating projection," *IEEE Signal Processing Letters*, vol. 24, no. 5, pp. 579-583, May 2017.

Outline

- Introduction
- Matrix Completion as a Feasibility Problem
- Alternating Projection Algorithm
- Numerical Examples
- Concluding Remarks
- List of References

Introduction

What is Matrix Completion?

The aim is to recover a low-rank matrix given only a subset of its possibly noisy entries, e.g.,

$$\begin{pmatrix} 1 & ? & ? & 4 & ? \\ ? & 2 & 5 & ? & ? \\ ? & ? & 4 & 5 & ? \\ 5 & ? & ? & ? & 4 \end{pmatrix}$$

Why Matrix Completion is Important?

It is a core problem in many applications including:

- recommendation systems
- Image inpainting and restoration
- Sensor network
- Path loss map reconstruction

Many real-world signals can be approximated by a matrix whose rank is $r \ll \max\{n_1, n_2\}$.

Netflix problem, whose goal was to accurately **predict user preferences** with a database of over 100 million movie ratings made by 480,189 users in 17,770 films, which corresponds to the task of completing a matrix with around 99% missing entries.

How to Recover an Incomplete Matrix?

Matrix completion is to find a matrix $\boldsymbol{X} \in \mathbb{R}^{n_1 \times n_2}$, which is an estimate of $\boldsymbol{M}$, given $\boldsymbol{M}_\Omega$ with the use of low-rank information of $\boldsymbol{M}$, which can be mathematically formulated as a noise-free version:

$$\min_{\boldsymbol{X}} \ \text{rank}(\boldsymbol{X}), \quad \text{s.t. } \boldsymbol{X}_\Omega = \boldsymbol{M}_\Omega.$$

or noisy version:

$$\min_{\boldsymbol{X}} \ \text{rank}(\boldsymbol{X}), \quad \text{s.t. } \|\boldsymbol{X}_\Omega - \boldsymbol{M}_\Omega\|_F^2 \leq \epsilon_F$$

where Ω is a subset of the complete set of entries $[n_1] \times [n_2]$, with $[n]$ being the list $\{1, \cdots, n\}$. But the rank minimization problem is NP-hard.

A popular and practical solution is to replace the nonconvex rank by convex nuclear norm:

$$\min_{\boldsymbol{X}} \ \|\boldsymbol{X}\|_*, \quad \text{s.t. } \boldsymbol{X}_\Omega = \boldsymbol{M}_\Omega$$

or

$$\min_{\boldsymbol{X}} \ \|\boldsymbol{X}\|_*, \quad \text{s.t. } \|\boldsymbol{X}_\Omega - \boldsymbol{M}_\Omega\|_F^2 \le \epsilon_2$$

where $\|\boldsymbol{X}\|_*$ equals the sum of singular values of $\boldsymbol{X}$. However, complexity of nuclear norm minimization is still high and this approach is not robust when $\boldsymbol{M}_\Omega$ contains outliers.

Matrix Completion as a Feasibility Problem

We formulate matrix completion with noise-free entries as:

$$\text{find } \boldsymbol{X}, \quad \text{s.t. } \ \text{rank}(\boldsymbol{X}) \le r, \quad \boldsymbol{X}_\Omega = \boldsymbol{M}_\Omega$$

where an estimate or true value of rank r is needed.

- Low-rank constraint: $\mathrm{rank}(\boldsymbol{X}) \leq r$
- Fidelity constraint: $\boldsymbol{X}_\Omega = \boldsymbol{M}_\Omega$

With Gaussian noise, the fidelity constraint is modified as:

$$\|\boldsymbol{X}_\Omega - \boldsymbol{M}_\Omega\|_F^2 \leq \epsilon_2.$$

To achieve robustness against outliers, the problem is formulated as:

$$\text{find } \boldsymbol{X}, \quad \text{s.t. } \mathrm{rank}(\boldsymbol{X}) \leq r, \quad \|\boldsymbol{X}_\Omega - \boldsymbol{M}_\Omega\|_p^p \leq \epsilon_p.$$

The robust feasibility problem can be rewritten as:

$$\boxed{\text{find } \boldsymbol{X} \in \mathcal{S}_r \cap \mathcal{S}_p.}$$

where the rank constraint set is:

$$\mathcal{S}_r := \{\boldsymbol{X} | \text{rank}(\boldsymbol{X}) \leq r\}$$

and the fidelity constraint set is:

$$\mathcal{S}_p := \left\{\boldsymbol{X} | \; \|\boldsymbol{X}_\Omega - \boldsymbol{M}_\Omega\|_p^p \leq \epsilon_p\right\}, \quad 0 < p \leq 2$$

where

$$\|\boldsymbol{X}_\Omega\|_p = \left(\sum_{(i,j)\in\Omega} |[\boldsymbol{X}]_{i,j}|^p\right)^{1/p}$$

is element-wise ℓ_p-norm which is robust to outliers if $p < 2$.

Remarks:

- $\mathcal{S}_p := \left\{ \boldsymbol{X} \middle| \, \|\boldsymbol{X}_\Omega - \boldsymbol{M}_\Omega\|_p^p \leq \epsilon_p \right\}$ is a generalization as $\epsilon_p = 0$ reduces to noise-free version while $p = 2$ reduces to conventional scenario of handling Gaussian noise.

Alternating Projection Algorithm

The proposed alternating projection algorithm (APA) is outlined in Algorithm 1:

Algorithm 1 Alternating Projection for Matrix Completion

Input: $\boldsymbol{M}_\Omega$, Ω, and $\epsilon_p > 0$

Initialize: $\boldsymbol{X}^0 = \boldsymbol{M}_\Omega$

for $k = 0, 1, 2 \cdots$ **do**

$\boldsymbol{Y}^k = \Pi_{\mathcal{S}_r}\left(\boldsymbol{X}^k\right)$

$\boldsymbol{X}^{k+1} = \Pi_{\mathcal{S}_p}\left(\boldsymbol{Y}^k\right)$

Stop if a termination condition is satisfied.

end for

Output: $\boldsymbol{X}^{k+1}$

According to Eckart-Young theorem, the projection of $\boldsymbol{Z} \notin \mathcal{S}_r$ onto $\mathcal{S}_r$ can be computed via truncated singular value decomposition (SVD) of $\boldsymbol{Z}$:

$$\Pi_{\mathcal{S}_r}(\boldsymbol{Z}) = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^T$$

where $\{\sigma_i\}_{i=1}^r$, $\{\boldsymbol{u}_i\}_{i=1}^r \in \mathbb{R}^{n_1}$, and $\{\boldsymbol{v}_i\}_{i=1}^r \in \mathbb{R}^{n_2}$ are the r largest singular values and the corresponding left and right singular vectors of $\boldsymbol{Z}$, respectively.

Note that $\mathcal{S}_p$ is an ℓ_p-ball and the projection onto ℓ_p-ball has closed-form solution in the following three cases:

- $\epsilon_p = 0$: the fidelity constraint reduces to equality constraint.

- $p = 2$ and $\epsilon_2 > 0$: the projection is derived as the closed-form expression of the projection onto the ℓ_2-ball.

- $p = 1$ and $\epsilon_1 > 0$: the projection onto ℓ_1-ball can be solved by soft-thresholding operater.

  Note that $1 < p < 2$ also involves the projection onto a convex ℓ_p-ball, which is not difficult to solve but requires an iterative procedure.

  As $p = 1$ is more robust than $1 < p < 2$ in the presence of outliers, the latter case will not be considered.

- We prove that if initial point is close enough to $\mathcal{S}_r \cap \mathcal{S}_p$, then APA locally converges to $\boldsymbol{X} \in \mathcal{S}_r \cap \mathcal{S}_p$ at a linear rate.

Numerical Examples

Noise-free $\boldsymbol{M} \in \mathbb{R}^{n_1 \times n_2}$ of rank r is generated by the product of $\boldsymbol{M}_1 \in \mathbb{R}^{n_1 \times r}$ and $\boldsymbol{M}_2 \in \mathbb{R}^{r \times n_2}$ whose entries satisfy standard Gaussian distribution, where $n_1 = 150$, $n_2 = 300$, and $r = 10$.

45% of the entries of $\boldsymbol{M}$ are randomly selected as the known observations.

Impulsive noise is modelled by two-term Gaussian mixture model (GMM) whose PDF is

$$p_v(v) = \sum_{i=1}^{2} \frac{c_i}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{v^2}{2\sigma_i^2}\right), \quad c_1 = 0.9,\ c_2 = 0.1,\ \sigma_2^2 = 100\sigma_1^2$$

Normalized root mean square error (RMSE) is defined as:

$$\mathrm{RMSE}(\boldsymbol{X}) = \sqrt{\mathrm{E}\left\{\frac{\|\boldsymbol{X}-\boldsymbol{M}\|_F^2}{\|\boldsymbol{M}\|_F^2}\right\}}$$

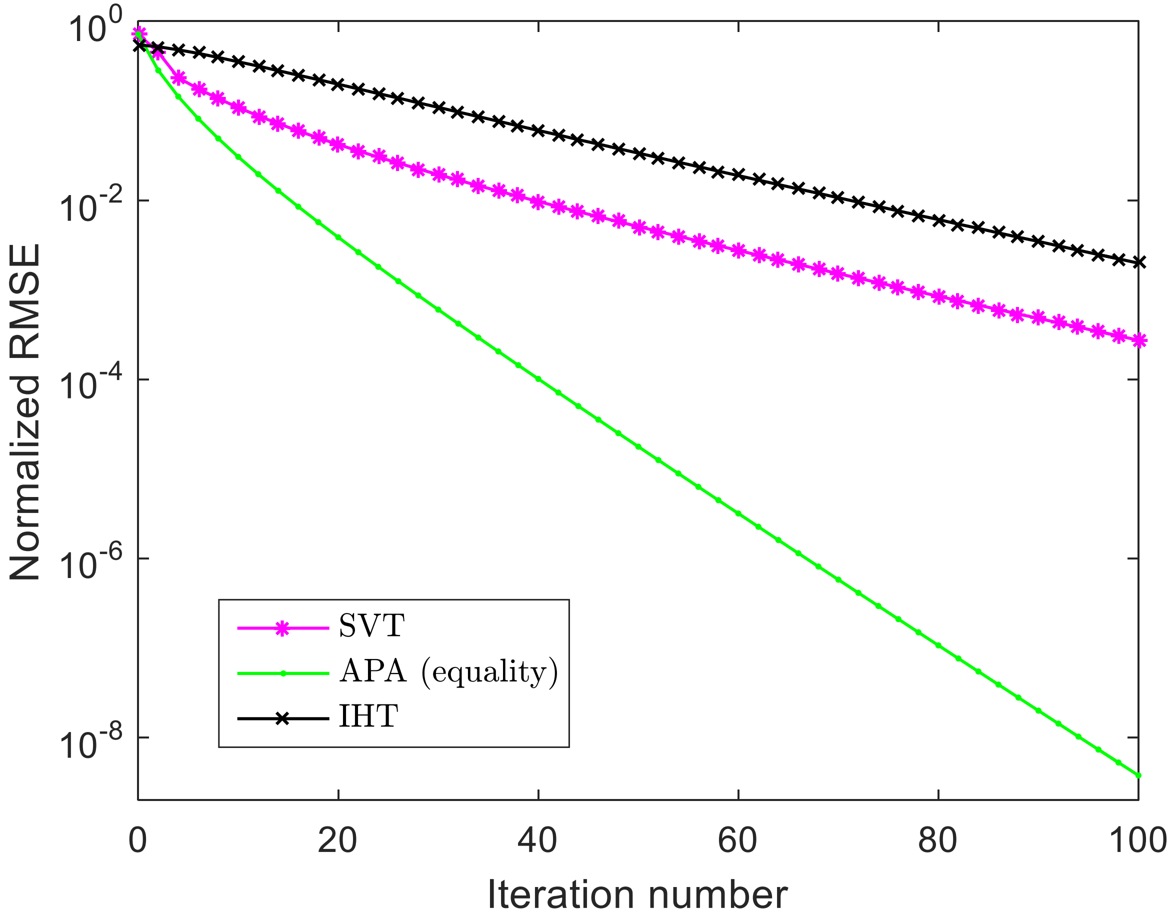

Figure 1: Normalized RMSE versus iteration number.

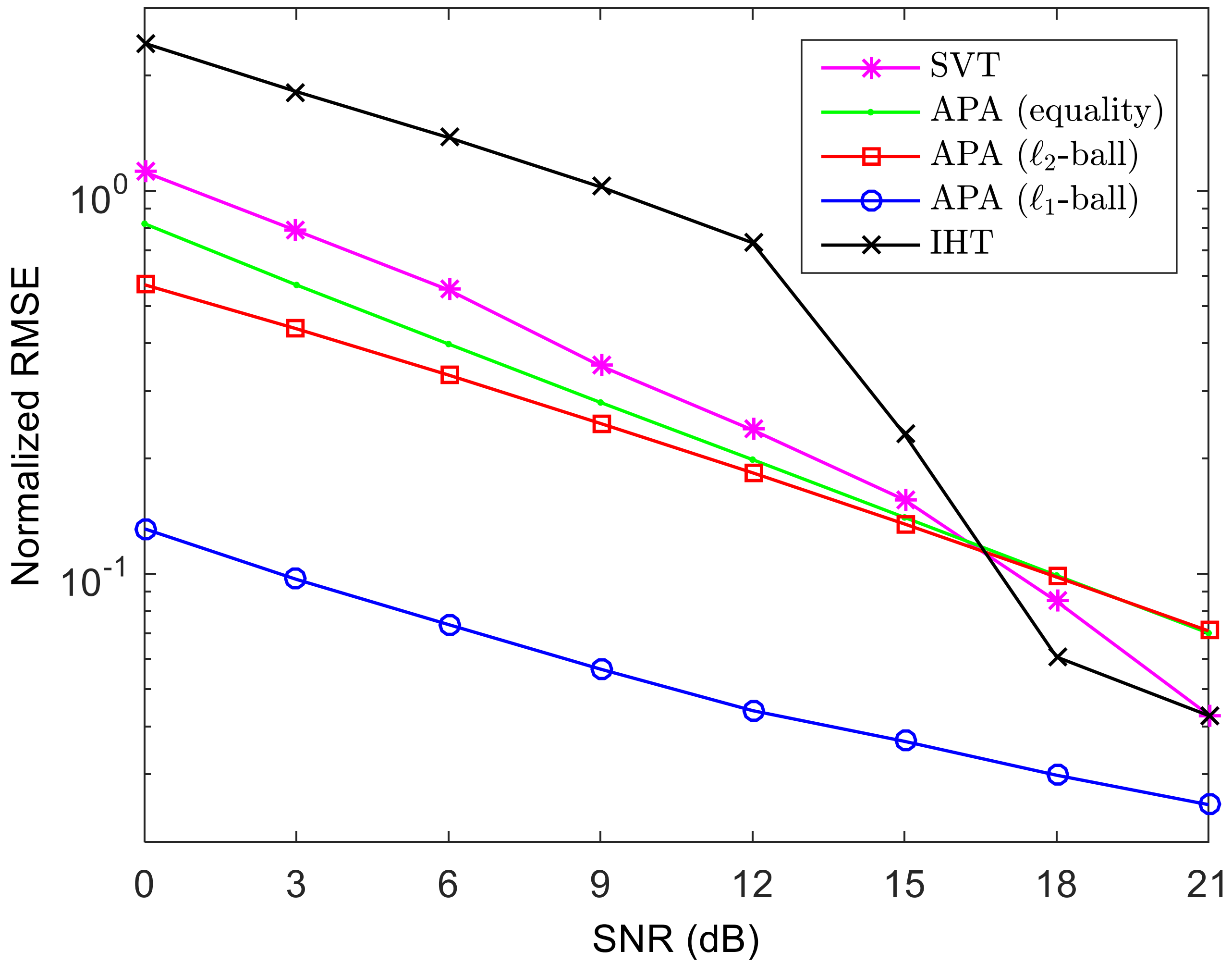

Figure 2: Normalized RMSE versus SNR.

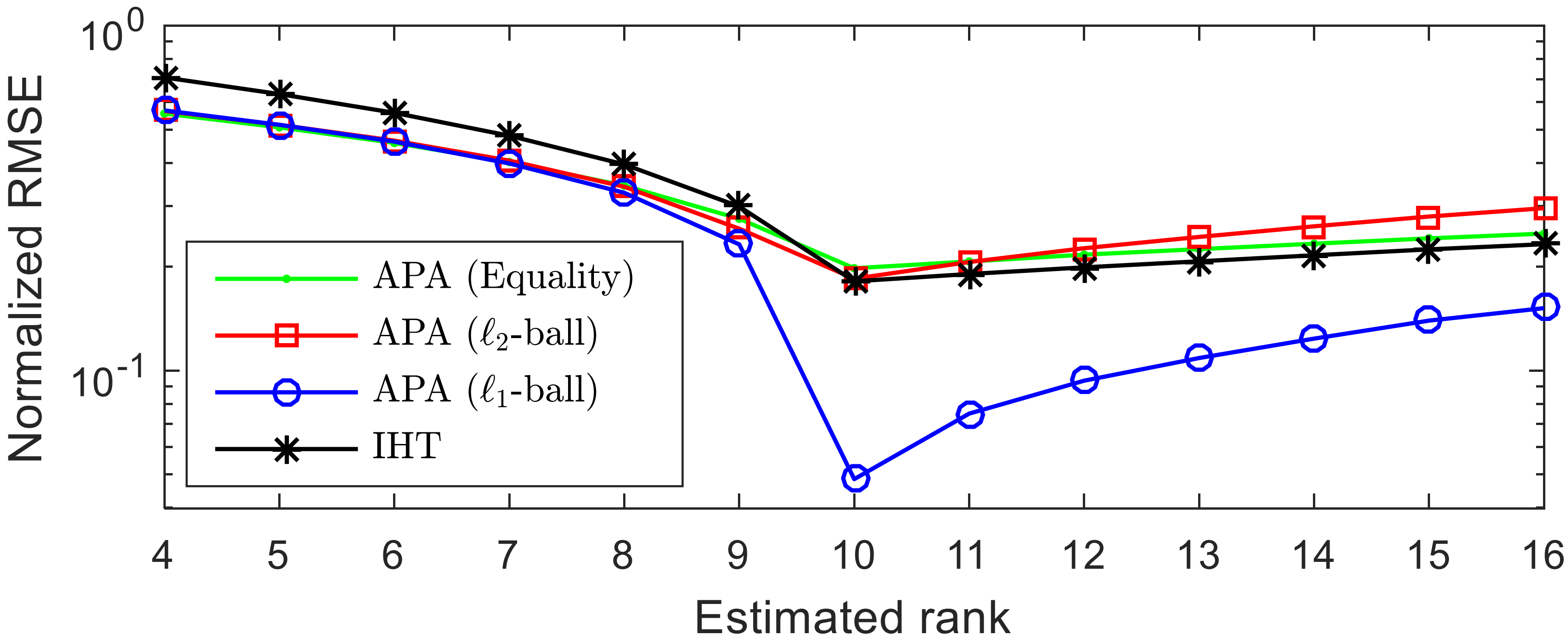

Figure 3: Normalized RMSE versus estimated rank.

Concluding Remarks

- The key idea is to formulate matrix completion as a feasibility problem, where a common point of the low-rank constraint set and fidelity constraint set is found by alternating projection.

- The fidelity constraint set is modelled as an ℓ_p-ball, where $p = 1$ or $p = 2$, which results in closed-form projection.

- The APA achieves robustness against Gaussian noise and outliers, with $p = 2$ and $p = 1$, respectively.

- There is no stepsize within the framework of APA.

- The APA is conceptually simpler and computationally more efficient than the popular nuclear norm minimization approaches.

References

[1] http://www.netflixprize.com/

[2] E. J. Candès and Y. Plan, "Matrix completion with noise," *Proc. IEEE*, vol. 98, no. 6, pp. 925-936, Jun. 2010.

[3] M. A. Davenport and J. Romberg, "An overview of low-rank matrix recovery from incomplete observations," *IEEE J. Sel. Top. Signal Process.*, vol. 10, no. 4, pp. 608-622, Jun. 2016.

[4] E. J. Candès and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2053-2080, May 2010.

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

[6] J.-F. Cai, E. J. Candès and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Opt.*, vol. 20, no. 4, pp. 1956-1982, 2010.

[7] P. Jain, R. Meka and I. S. Dhillon, "Guaranteed rank minimization via singular value projection," in *Adv. Neural Inf. Process. Syst. (NIPS)*, pp. 937-945, 2010.

[8] L. Condat, "Fast projection onto the simplex and the L1-ball," *Math. Program. Ser. A*, vol. 158, no. 1, pp. 575-585, Jul. 2016.

[9] A. S. Lewis, D. R. Luke and J. Malick, "Local linear convergence for alternating and averaged nonconvex projections," *Found. Comp. Math.*, vol. 9, no. 4, pp 485-513, Aug. 2009.