Journal of Computational Mathematics
Vol.38, No.2, 2020, 337–354.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1809-m2018-0106

ROBUST INEXACT ALTERNATING OPTIMIZATION FOR MATRIX COMPLETION WITH OUTLIERS*

Ji Li
Beijing Computational Science Research Center, Beijing 100193, China
Email: keelee@csrc.ac.cn
Jian-Feng Cai
Department of Mathematics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong
Email: jfcai@ust.hk
Hongkai Zhao
Department of Mathematics, University of California, Irvine, CA, USA
Email: zhao@uci.edu

Abstract

We investigate the problem of robust matrix completion with a fraction of observation corrupted by sparsity outlier noise. We propose an algorithmic framework based on the ADMM algorithm for a non-convex optimization, whose objective function consists of an ℓ_1 norm data fidelity and a rank constraint. To reduce the computational cost per iteration, two inexact schemes are developed to replace the most time-consuming step in the generic ADMM algorithm. The resulting algorithms remarkably outperform the existing solvers for robust matrix completion with outlier noise. When the noise is severe and the underlying matrix is ill-conditioned, the proposed algorithms are faster and give more accurate solutions than state-of-the-art robust matrix completion approaches.

Mathematics subject classification: 65K05, 90C06, 93C41.
Key words: Matrix completion, ADMM, Outlier noise, Inexact projection.

1. Introduction

The problem of matrix completion refers to completing a matrix with many missing entries, and it arises from various applications in statistics, machine learning, and computer vision. This problem is possible to solve only if the underlying matrix is "simple", because otherwise the matrix contains too much information to infer from the limited observed entries. A commonly used notion of simplicity for matrices is *low rank* [11, 13, 17, 32, 33], which provides redundancy of matrix entries. The low rank matrix model is remarkably successful in many applications in machine learning, such as collaborative filtering [38] and the Netflix prize problem [4].

Let $\boldsymbol{M} \in \mathbb{R}^{m\times n}$ be the underlying low-rank matrix we would like to estimate. Let $\Omega \subseteq [m] \times [n]$ be a set of indices with $|\Omega| \ll mn$, where $[m] = \{1, 2, \ldots, m\}$ and the same for $[n]$. In the problem of matrix completion, only the entries $\{M_{ij} : (i, j) \in \Omega\}$ are observed and the other entries are missing. One would like to reconstruct the underlying low-rank matrix $\boldsymbol{M}$ from $\{M_{ij} : (i, j) \in \Omega\}$. The approaches of low-rank matrix completion can be divided into two categories, namely, convex and non-convex optimization based approaches.

* Received May 31, 2018 / Revised version received August 2, 2018 / Accepted September 25, 2018 / Published online February 13, 2019 /

Convex optimization based approaches are formulated from the rank minimization. With sufficiently many observed entries and some mild assumptions, $\boldsymbol{M}$ is the only low-rank matrix in the set $\{\boldsymbol{X} \in \mathbb{R}^{m\times n} \mid X_{ij} = M_{ij}, (i,j) \in \Omega\}$ of matrices that are consistent with observed entries. In this case, the low-rank matrix completion can be reconstructed by solving the following constrained rank minimization:

$$\begin{aligned} \min_{\boldsymbol{X}\in\mathbb{R}^{m\times n}} \quad & \operatorname{rank}(\boldsymbol{X}) \\ \text{s.t.} \quad & X_{ij} = M_{ij}, \quad (i,j) \in \Omega. \end{aligned} \tag{1.1}$$

However, (1.1) is non-convex and, more critically, NP-hard. Therefore, it is computationally intractable. To overcome these, a popular strategy is to replace the rank function in (1.1) by its convex relaxation, the nuclear norm [11–13], to solve the following convex optimization

$$\begin{aligned} \min_{\boldsymbol{X}\in\mathbb{R}^{m\times n}} \quad & \|\boldsymbol{X}\|_* \\ \text{s.t.} \quad & X_{ij} = M_{ij}, \quad (i,j) \in \Omega. \end{aligned} \tag{1.2}$$

Problem (1.2) can be reformulated as a special case of semi-definite programming (SDP) [37], for which polynomial time solvers exist. It was proved in [11, 13, 22, 36] that the unique solution of (1.2) is $\boldsymbol{M}$ under suitable assumptions. Thus, one can complete a low-rank matrix in polynomial time with theoretical guarantee. In real applications, the observed entries are usually corrupted by noise. Under this circumstance, it is natural to consider the nuclear norm regularized optimization

$$\min_{\boldsymbol{X}\in\mathbb{R}^{m\times n}} \quad \frac{1}{2}\sum_{(i,j)\in\Omega} (X_{ij} - \tilde{M}_{ij})^2 + \lambda\|\boldsymbol{X}\|_*, \tag{1.3}$$

where $\tilde{M}_{ij}$, $(i,j) \in \Omega$, are noisy observations. When there is only a small amount of noise in the observed entries, the model (1.3) is provably accurate with the reconstruction error proportional to the noise level [11]. Though off-the-shelf SDP solvers can be applied to solve (1.2) and (1.3), numerically, they are not the most efficient, especially when the matrix size is moderately large. Customized first-order algorithms (e.g., [8, 29, 31, 41]) are developed for solving (1.2) and (1.3). Most of them invoke the singular value thresholding (SVT) operator [8] at each iteration. The most expensive part of these algorithms is the computation of SVT in each iteration. The usual strategy is to compute the singular value decomposition (SVD) followed by the soft-thresholding on the singular values.

To improve the performance of nuclear norm optimization based matrix completion, we may consider non-convex optimization based approaches. Assume that $\operatorname{rank}(\boldsymbol{M}) = r$ is known, then the matrix completion problem can be reformulated as the following constrained least-squares problem

$$\begin{aligned} \min_{\boldsymbol{X}\in\mathbb{R}^{m\times n}} \quad & \sum_{(i,j)\in\Omega} (X_{ij} - \tilde{M}_{ij})^2 \\ \text{s.t.} \quad & \operatorname{rank}(\boldsymbol{X}) = r. \end{aligned} \tag{1.4}$$

Since (1.4) is a non-convex optimization, the challenge here is how to find the global minimum with a provable guarantee. In the past a few years, there is a burst of research works on the design and analysis of provable non-convex matrix completion algorithms by solving (1.4) and its variants. There are two types of such numerical algorithms. One type of algorithms treat the