Signal Processing 121 (2016) 153–159

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Fast communication

Alternating strategies with internal ADMM for low-rank matrix reconstruction

Kezhi Li [a,b,*], Martin Sundin [a], Cristian R. Rojas [a], Saikat Chatterjee [a], Magnus Jansson [a]

[a] ACCESS Linnaeus Center, Electrical Engineering, KTH Royal Institute of Technology, S-100 44 Stockholm, Sweden
[b] Medical Research Council, Imperial College London, White City, London, W12 0NN, United Kingdom

ARTICLE INFO

Article history:
Received 19 March 2015
Received in revised form
24 August 2015
Accepted 10 November 2015
Available online 27 November 2015

Keywords:
Low-rank matrix reconstruction
Alternating strategies
Least squares
ADMM

ABSTRACT

This paper focuses on the problem of reconstructing low-rank matrices from underdetermined measurements using alternating optimization strategies. We endeavour to combine an alternating least-squares based estimation strategy with ideas from the alternating direction method of multipliers (ADMM) to recover low-rank matrices with linear parameterized structures, such as Hankel matrices. The use of ADMM helps to improve the estimate in each iteration due to its capability of incorporating information about the direction of estimates achieved in previous iterations. We show that merging these two alternating strategies leads to a better performance and less consumed time than the existing alternating least squares (ALS) strategy. The improved performance is verified via numerical simulations with varying sampling rates and real applications.

1. Introduction

The low-rank matrix reconstruction problem arises naturally in many fields, such as system identification [1–4], computer vision [5,6] and quantum state tomography [7]. Suppose an r-rank matrix $\mathbf{X}$ has size $n_1 \times n_2$, $r \ll \min(n_1, n_2)$; the objective is to recover $\mathbf{X}$ from the noisy measurements that satisfy the equation

$$\mathbf{y} = \mathcal{A}(\mathbf{X}) + \mathbf{e}, \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^{m \times 1}$ is the measurement vector, $\mathcal{A}$ denotes a known sensing function $\mathbb{R}^{n_1 \times n_2} \to \mathbb{R}^{m \times 1}$, and $\mathbf{e}$ is assumed to be zero-mean noise with known covariance $E(\mathbf{e}\mathbf{e}^T) = \mathbf{C} \in \mathbb{R}^{m \times m}$. Usually $m < n_1 \times n_2$, that is, the number of coefficients of $\mathbf{X}$ is larger than the number of measurements, and hence (1) is underdetermined. Specifically we consider the case where $\mathcal{A}$ is a linear operator so that (1) can be rewritten equivalently as the product of a matrix $\mathbf{A}$ and the vectorization of a low-rank $\mathbf{X}$

$$\mathcal{A}(\mathbf{X}) = \begin{bmatrix} \langle \mathbf{X}, \mathbf{A}_1 \rangle \\ \vdots \\ \langle \mathbf{X}, \mathbf{A}_m \rangle \end{bmatrix} = \mathbf{A}\mathrm{vec}(\mathbf{X}), \tag{2}$$

where $\mathbf{A} \in \mathbb{R}^{m \times n_1 n_2}$ consists of vectorized $\mathbf{A}_i$ as its ith row, $i = 1, \dots, m$. This operator arises in many applications, for instance, in quantum state tomography [7,8]; $\mathbf{X}$ then corresponds to the nearly pure density matrix of the unknown quantum state, and $\mathcal{A}$ represents a series of the measurement bases (e.g., tensor product of Pauli matrices).

Different techniques have been developed to solve such underdetermined problems, such as nuclear norm minimization [9,10] or alternating approaches [11–15]. Compared to the nuclear norm minimization [9], an *alternating technique* solution provides faster computation, higher accuracy and is hence useful for solving large-scale underdetermined problems based on different criteria such as maximum likelihood (ML) [13] or least squares (LS)

* Corresponding author.
E-mail addresses: kezhi@kth.se (K. Li), masundi@kth.se (M. Sundin), crro@kth.se (C.R. Rojas), sach@kth.se (S. Chatterjee), janssonm@kth.se (M. Jansson).

http://dx.doi.org/10.1016/j.sigpro.2015.11.002
0165-1684/

[12,16]. Typically, alternating approaches provide locally optimal solutions. Each iteration leads to the best solution of a set of variables given another set of fixed variables found in the previous iteration. Here our hypothesis is that, if the *updating directions* of previous iterations are also considered in each iteration, the reconstruction will be improved in both aspects of accuracy and efficiency because the potential feasible set of solutions is narrowed in each iteration. In this regard, few relevant attempts have been made in matrix completion [11], or to update solutions using a gradient descent method [17,14].

In this paper we develop two algorithms based on the alternating technique. First we incorporate suitable modifications to the alternating least squares (ALS) algorithm of [12] to derive a new algorithm called alternating linear estimator (ALE) for low rank matrices with a linear structure. Then based on the ALS and ALE, we develop a novel algorithm called alternating direction least squares (ADLS) that endeavours to validate our hypothesis on updating directions by fusing two alternating strategies. It utilizes the alternating strategy with the help of an updating direction for structured matrix reconstruction. Inspired by the ALS, the proposed framework is based on running the LS estimation to update the low rank component matrices $\mathbf{L}$, $\mathbf{R}$ and $\mathbf{X}$ iteratively, where $\mathbf{LR}=\mathbf{X}$. In our new approach, the new $\mathbf{L}$, $\mathbf{R}$ are calculated by solving optimization problems involving the augmented Lagrangian to incorporate direction update knowledge. This method is able to push variables converging to solutions more efficiently, as in the standard alternating direction method of multipliers (ADMM) [18,11]. The new algorithm also inherits the capability of ALS of handling structured matrices, e.g., with Hankel structure. The simulation results are compared with the performance of ALS and Cramér–Rao bounds (CRBs). Besides the signal-to-reconstruction error ratio (SRER), we also compare their processing time to show the effectiveness and the efficiency of the proposed approach. Real applications in system identification and inpainting are also demonstrated.

This paper is organized as follows. We review the ALS method in Section 2. In Section 3 we propose the ALE algorithm for reconstructing low rank matrices with linear operators. Then by combining ALE and ADMM, the ADLS algorithm is proposed and analyzed in Section 4. Numerical simulations are shown in Section 5 and finally Section 6 concludes the paper.

Notations: Bold letters are used to denote a vector or a matrix. For vectors, $||\cdot||_1$, $||\cdot||_2$, $||\cdot||_\infty$ represent the l_1, l_2 and l_∞ norms, respectively. For matrices, $\mathbf{A}^T$ and $\mathbf{A}^\dagger$ denote the transpose and Moore–Penrose pseudoinverse of $\mathbf{A}$. Moreover, $||\cdot||_F$ represents the Frobenius norm and $||\mathbf{x}||_\mathbf{W} \triangleq \sqrt{\mathbf{x}^T\mathbf{W}\mathbf{x}}$. $\chi_r \triangleq \{\mathbf{A}\in\mathbb{R}^{n_1\times n_2}: \mathrm{rank}(\mathbf{A})=r\}$ denotes the set of rank r matrices. vec($\mathbf{A}$) represents the column vector of concatenated columns of $\mathbf{A}$, and (mat_{n_1,n_2}) is (vec)'s inverse operation to convert a vector to a matrix of size $n_1\times n_2$. $\otimes$ is the Kronecker product, and $\nabla_\mathbf{X}\{f\}$ denotes the partial derivative of the function f with respect to $\mathbf{X}$. Finally we use p.s.d. as the short form for positive semidefinite.

2. ALS for low-rank matrix reconstruction

The alternating least-squares approach was developed in [12,13]. For an r-rank matrix $\mathbf{X}$ satisfying (1) with noise covariance $\mathbf{C}$, the weighted least-squares estimator is

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}\in\chi_r} ||\mathbf{y}-\mathcal{A}(\mathbf{X})||^2_{\mathbf{C}^{-1}}. \tag{3}$$

To rewrite (3) in terms of the standard 2-norm, the measurements and sensing operator can be prewhitened by forming $\overline{\mathbf{y}}=\mathbf{C}^{-1/2}\mathbf{y}$ and $\overline{\mathbf{A}}=\mathbf{C}^{-1/2}\mathbf{A}$. Expressing $\mathbf{X}=\mathbf{LR}$ where $\mathbf{L}\in\mathbb{R}^{n_1\times r}$ and $\mathbf{R}\in\mathbb{R}^{r\times n_2}$, the square of residuals becomes

$$\mathbf{J}(\mathbf{L},\mathbf{R}) \triangleq ||\overline{\mathbf{y}}-\overline{\mathcal{A}}(\mathbf{LR})||_2^2 = ||\overline{\mathbf{y}}-\overline{\mathbf{A}}(\mathbf{I}_{n_1}\otimes\mathbf{L})\mathrm{vec}(\mathbf{R})||_2^2 = ||\overline{\mathbf{y}}-\overline{\mathbf{A}}(\mathbf{R}^T\otimes\mathbf{I}_{n_2})\mathrm{vec}(\mathbf{L})||_2^2. \tag{4}$$

The cost function $\mathbf{J}(\mathbf{L},\mathbf{R})$ is minimized alternatingly by

$$\hat{\mathbf{R}} = \arg\min_{\mathbf{R}} ||\overline{\mathbf{y}}-\overline{\mathbf{A}}(\mathbf{I}_{n_1}\otimes\hat{\mathbf{L}})\mathrm{vec}(\mathbf{R})||_2^2,$$
$$\hat{\mathbf{L}} = \arg\min_{\mathbf{L}} ||\overline{\mathbf{y}}-\overline{\mathbf{A}}(\hat{\mathbf{R}}^T\otimes\mathbf{I}_{n_2})\mathrm{vec}(\mathbf{L})||_2^2. \tag{5}$$

The iterations of estimating $\hat{\mathbf{R}}$ and $\hat{\mathbf{L}}$ continue until the residual $||\overline{\mathbf{y}}-\overline{\mathcal{A}}(\mathbf{LR})||_2^2$ no longer decreases. Specifically, we calculate the analytical solutions $\mathrm{vec}(\hat{\mathbf{R}})=[\overline{\mathbf{A}}(\mathbf{I}_{n_1}\otimes\mathbf{L})]^\dagger\overline{\mathbf{y}}$ given $\mathbf{L}$ and $\mathrm{vec}(\hat{\mathbf{L}})=[\overline{\mathbf{A}}(\mathbf{R}^T\otimes\mathbf{I}_{n_2})]^\dagger\overline{\mathbf{y}}$ given $\mathbf{R}$. ALS is also capable of recovering structured low-rank matrices such as Hankel, Toeplitz, as well as p.s.d. matrices. In this case a projection step $\hat{\mathbf{X}}\triangleq\mathcal{P}(\mathbf{L}\hat{\mathbf{R}})$ is added after updating $\hat{\mathbf{R}}$ using a "lift and project" approach, whose core steps are the truncated singular value decomposition with full explanation in [19]. Then a new $\overline{\mathbf{R}}$ is calculated by the least-squares estimator

$$\overline{\mathbf{R}} = \min_{\mathbf{R}} ||\mathbf{LR}-\hat{\mathbf{X}}||_F^2. \tag{6}$$

$\overline{\mathbf{L}}$ can be updated likewise. ALS has been shown to be effective for recovering low-rank matrices of large sizes.

3. Alternating linear estimator

In this section we develop a simple but efficient algorithm for low rank matrices with linear structure, called the alternating linear estimator (ALE). We assume that the low rank matrix has linear structure, which means that $\mathbf{X}\in\mathbb{R}^{n_1\times n_2}$ can be decomposed as

$$\mathbf{X} = \mathcal{S}_\chi(\mathbf{h}) \tag{7}$$

where $\mathbf{h}\in\mathbb{R}^p$ is a parametrization of $\mathbf{X}$ and $\mathcal{S}_\chi: \mathbb{R}^p\to\mathbb{R}^{n_1\times n_2}$ is a linear map parameterizing the linear structure of $\mathbf{X}$. Taking $\mathbf{X}$ as a Hankel matrix for instance, $\mathbf{h}$ can contain the first column and last row of $\mathbf{X}$. We denote the pseudoinverse of $\mathcal{S}_\chi$ by $\mathcal{T}_\chi$, i.e.

$$\mathcal{T}_\chi(\mathbf{X}) = \arg\min_{\mathbf{h}} ||\mathbf{X}-\mathcal{S}_\chi(\mathbf{h})||_F^2. \tag{8}$$

The idea of ALE is to iteratively update $\mathbf{L}$, $\mathbf{R}$ and apply the lift-and-project (or composite mapping) method to project the matrix to its linear structure. As an initial least squares estimate we set

$$\mathbf{h}_0 = \arg\min_{\mathbf{h}} ||\overline{\mathbf{y}}-\overline{\mathcal{A}}(\mathcal{S}_\chi(\mathbf{h}))||_2^2. \tag{9}$$