Information Sciences 326 (2016) 243–257

Contents lists available at ScienceDirect

Information Sciences

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

Tensor completion using total variation and low-rank matrix factorization

Teng-Yu Ji, Ting-Zhu Huang*, Xi-Le Zhao, Tian-Hui Ma, Gang Liu

School of Mathematical Sciences/Resrarch Center for Image and Vision Computing, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, P R China

A R T I C L E I N F O

Article history:
Received 18 January 2015
Revised 22 May 2015
Accepted 29 July 2015
Available online 5 August 2015

Keywords:
Tensor completion
Total variation
Low-rank matrix factorization
Block coordinate descent

A B S T R A C T

In this paper, we study the problem of recovering a tensor with missing data. We propose a new model combining the total variation regularization and low-rank matrix factorization. A block coordinate decent (BCD) algorithm is developed to efficiently solve the proposed optimization model. We theoretically show that under some mild conditions, the algorithm converges to the coordinatewise minimizers. Experimental results are reported to demonstrate the effectiveness of the proposed model and the efficiency of the numerical scheme.

1. Introduction

As a high-dimensional extension of matrix, tensor is an important big data format, which plays a significant role in a wide range of real-world applications [19,31–36]. Among them, one important problem is to estimate the missing data from the observed incomplete data, e.g., image inpainting [1,17], video inpainting [18], hyperspectral data recovery [22,23,41,45], magnetic resonance imaging (MRI) data recovery [38], high-order web link analysis [16], personalized web search [31], and seismic data reconstruction [19]. In this study, we specially focus on the reconstruction of low-rank tensors with randomly missing data.

Matrix completion can be regarded as the 2-mode tensor completion [4]. One powerful tool for matrix completion is to minimize the matrix rank, which can effectively estimate the missing data exploiting both the local and global information [24]. The model for low-rank matrix completion is formulated as:

$$\begin{aligned} \min_{Y} \quad & \operatorname{rank}(Y) \\ \text{s.t.} \quad & \mathcal{P}_{\Omega}(Y) = F, \end{aligned} \tag{1}$$

where $Y \in \mathbb{R}^{m\times n}$ is the underlying matrix, $F \in \mathbb{R}^{m\times n}$ is the observed matrix, and $\mathcal{P}_{\Omega}(\cdot)$ is the projection operator: see details in Section 2. However, the main difficulty of solving (1) arises from the non-convexity of the rank of matrices, which may prevent one from getting a global solution [40]. To solve the challenging problem of rank minimizing, Fazel et al. [8] and Kurucz et al. [20] proposed to use rank constraint to iteratively estimate the missing values. Another popular and effective approach is to use the trace norm, which is theoretical soundness and can be considered as the approximation for the rank of matrices [3,26,28]. And

* Corresponding author. Tel.: +86 28 61831608; fax: +86 28 61831280.
E-mail addresses: tengyu_j66@126.com (T.-Y. Ji), tingzhuhuang@126.com, tzhuang@uestc.edu.cn (T.-Z. Huang), xlzhao122003@163.com (X.-L. Zhao), nkmth0307@126.com (T.-H. Ma), wd5577@163.com (G. Liu).

http://dx.doi.org/10.1016/j.ins.2015.07.049

under certain conditions [4,5], the problem (1) is converted to the following convex optimization problem:

$$\begin{array}{ll} \min\limits_{Y} & \|Y\|_* \\ \text{s.t.} & \mathcal{P}_\Omega(Y) = F. \end{array} \tag{2}$$

Then the model (2) can be efficiently solved by using some optimization algorithms, such as FPCA [26], APGL [37], LMaFit [40], and the alternating direction method (ADM) [30,43].

For tensor completion, the low-rank based methods have also been widely studied [9,24,25,42,44]. The low-rank tensor completion model can be formulated as:

$$\begin{array}{ll} \min\limits_{\mathcal{Y}} & \text{rank}(\mathcal{Y}) \\ \text{s.t.} & \mathcal{P}_\Omega(\mathcal{Y}) = \mathcal{F}, \end{array} \tag{3}$$

where $\mathcal{Y} \in \mathbb{R}^{I_1\times\cdots\times I_N}$ is the underlyingtensor, and $\mathcal{F}$ is the observed data. However, there is no unique definition for the rank of tensors, such as CP-rank and n-rank [15], and both of the corresponding minimization problems are NP-hard [12]. As the tensor is a generalization of the matrix, one can generalize matrix completion problem (2) to the tensor case:

$$\begin{array}{ll} \min\limits_{\mathcal{Y}} & \|\mathcal{Y}\|_* \\ \text{s.t.} & \mathcal{P}_\Omega(\mathcal{Y}) = \mathcal{F}. \end{array} \tag{4}$$

A naive method is to unfold the tensor into a matrix, and thus to solve the matrix completion (2). However, the method only utilizes low-rankness to one mode of the tensor, and it cannot recover the tensor well [24,42]. Thus, it is necessary to develop methods considering low-rankness to the all mode of the tensor. Recently, Liu et al. [24] developed a theoretical framework for low-rank tensor completion and established a definition of the trace norm for tensors as a surrogate for the tensor rank:

$$\|\mathcal{Y}\|_* := \sum_{n=1}^{N} \alpha_n \|Y_{(n)}\|_*, \tag{5}$$

where $Y_{(n)}$ is the mode-n unfolding of $\mathcal{Y}$: see details in Section 2. Then low-rank tensor completion problem (4) is rewritten as:

$$\begin{array}{ll} \min\limits_{\mathcal{Y}} & \sum_{n=1}^{N} \alpha_i \|Y_{(n)}\|_* \\ \text{s.t.} & \mathcal{P}_\Omega(\mathcal{Y}) = \mathcal{F}. \end{array} \tag{6}$$

Problem (6) can be solved by some optimization methods, such as FaLRTC [24] and the Douglas–Rachford splitting method [9]. Because the information of the all mode is considered, these methods [9,24] outperform the naive method. However they have to calculate singular value decomposition (SVD) for N matrices, which is expensive in term of time and memory. Considering this difficulty, Xu et al. [42] applied low-rank matrix factorization to the all-mode matricizations of the tensor as an alternative of the tensor trace norm,

$$\begin{array}{ll} \min\limits_{\mathcal{Y},X,A} & \sum\limits_{n=1}^{N} \dfrac{\alpha_n}{2} \|Y_{(n)} - A_n X_n\|_F^2 \\ \text{s.t.} & \mathcal{P}_\Omega(\mathcal{Y}) = \mathcal{F}, \end{array} \tag{7}$$

where $A = (A_1, \ldots, A_N)$, $X = (X_1, \ldots, X_N)$, and $\alpha_n, n = 1, \ldots, N$ are positive weights satisfying $\sum_{n=1}^{N} \alpha_n = 1$. Their method (called TMac) has shown to obtain better results and take less time than FaLRTC [24].

Note that Xu et al. [42] only consider the low-rank prior. However, many real-world data exhibit the piecewise smooth prior. In particular, as one of characterizing piecewise smooth functions, the total variation (TV) norm [29] has been shown to preserve edges well in image restoration [11,21,46]. Recently, other TV based regularization methods have received great success in image processing problems, such as the image segmentation [7,39], the reconstruction for video [6], hyperspectral image [22,45] and MRI [38]. Particularly, the authors in [22,45] considered to apply TV regularization to material identification and unmixing for hyperspectral images, with the aim of exploiting the spatial contextual information presented in the hyperspectral images. Inspired by the former works, we consider to introduce the TV regularization into the tensor completion problem (7).

The contributions of this paper are mainly two folds. First, we propose a new model for low-rank tensor completion with randomly missing data. More precisely, our model is:

$$\begin{array}{ll} \min\limits_{\mathcal{Y},X,A} & \sum\limits_{n=1}^{N} \dfrac{\alpha_n}{2} \|Y_{(n)} - A_n X_n\|_F^2 + \mu \text{TV}(X_3) \\ \text{s.t.} & \mathcal{P}_\Omega(\mathcal{Y}) = \mathcal{F}, \end{array} \tag{8}$$

where μ is the regularization parameter, $A = (A_1, \ldots, A_N)$, $X = (X_1, \ldots, X_N)$, and $\text{TV}(X_3)$ is the total variation of X_3. Similar to [2,45], A_n represents a library (each column contains a signature of the nth mode direction), and X_n is called an encoding. For example, in the unmixing problem for hyperspectral image [2,45], each column of A_3 contains a spectral signature, and each row of X_3 contains the fractional abundances of a given endmember. This interpretation is also valid for the mode-3 unfolding of video