

## Research Article

# Traffic Volume Data Outlier Recovery via Tensor Model

Huachun Tan,<sup>1</sup> Jianshuai Feng,<sup>2</sup> Guangdong Feng,<sup>3</sup> Wuhong Wang,<sup>1</sup> and Yu-Jin Zhang<sup>4</sup>

<sup>1</sup> Department of Transportation Engineering, Beijing Institute of Technology, Beijing 100081, China

<sup>2</sup> Integrated Information System Research Center, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

<sup>3</sup> Civil Aviation Engineering Consulting Company of China, Beijing 100621, China

<sup>4</sup> Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

Correspondence should be addressed to Huachun Tan; [htan8@wisc.edu](mailto:htan8@wisc.edu)

Received 22 September 2012; Accepted 24 February 2013

Academic Editor: Huimin Niu

Copyright © 2013 Huachun Tan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Traffic volume data is already collected and used for a variety of purposes in intelligent transportation system (ITS). However, the collected data might be abnormal due to the problem of outlier data caused by malfunctions in data collection and record systems. To fully analyze and operate the collected data, it is necessary to develop a validate method for addressing the outlier data. Many existing algorithms have studied the problem of outlier recovery based on the time series methods. In this paper, a multiway tensor model is proposed for constructing the traffic volume data based on the intrinsic multilinear correlations, such as day to day and hour to hour. Then, a novel tensor recovery method, called ADMM-TR, is proposed for recovering outlier data of traffic volume data. The proposed method is evaluated on synthetic data and real world traffic volume data. Experimental results demonstrate the practicability, effectiveness, and advantage of the proposed method, especially for the real world traffic volume data.

## 1. Introduction

In order to alleviate the traffic congestion problem and facilitate the mobility in metropolises, large amounts of traffic information are collected as a part of intelligent transportation system (ITS) such as CVIS (Cooperative Vehicle Infrastructure System) in China. These collected traffic data have wide range of applications. The real time traffic information is provided to travelers to support their decision for making process on the optimal route choice [1]. As shown by the work of Kim et al. [2], the real time information can contribute to reduce the operation cost and maximize resource utilization. In addition to these applications, the collected data could be applied to maximize the utilization of the infrastructure for smooth flow of the traffic. One such application of real time traffic data is traffic information control [3]. On the other hand, several data mining techniques have been applied to mine time related association rules from traffic databases and their results have been used for traffic prediction such as the works of Williams et al. [4] and Xu et al. [5]. From the above discussion, it is concluded that the collected traffic data are essential for many potential applications in ITS.

In real world, the collected data are always corrupted due to noise values, especially outlier value, which may be caused by detector failures, communication problems, or any other hardware/software related problems. The presence of outlier data in the database would degrade significantly the quality as well as reliability of the data and might impede the effectiveness of ITS applications. Therefore, it is essential to fill the gaps caused by outlier data in order to fully explore the applicability of the data and realize the ITS applications.

While many different kinds of traffic data such as traffic volume, speed, and occupancy are collected, the focus of this research is on the traffic volume outlier data recovery. It is supposed that the detectors collecting traffic information are set up at road sections and the collected values represent the traffic volume for those road sections. The aim of this research is to recover the traffic volume outlier data for road sections.

Literature survey in the related field shows that several filtering recovery techniques have been applied to recover the outlier traffic data [6–8]. Filtering methods include techniques such as singular value decomposition, wavelet analysis, immune algorithm, and spectrum subtraction. Filtering methods formulate the traffic volume as time series

model and smooth the traffic waveform. These approaches recover the outlier data of day by day through spectrum analysis and feature information extracting. However, the traffic data through the same location is significantly similar from day to day and these approaches cannot utilize such characteristic. Pei and Ma [6] show that similarity is an important factor impacting on recovery performance. While the above methods consider only one mode similarity, the recovery performance is mainly dependent on the smooth threshold. Unfortunately, the smooth threshold is empirically determined.

In order to improve the recovery performance and consider the multidimension characteristic of traffic data, mining the multimode similarities will make a great contribution for recovering outlier value. Our approach is based on utilizing multimode correlations of traffic data; that is, traffic data have different correlation on different modes, such as week mode, day mode, and hour mode. More concretely, the feature of the proposed method is to recover the outlier value using the traffic volume information of the different modes. But, the problem is not so simple, because traffic volumes from many days might be corrupted by outlier data simultaneously. In order to consider the multiple outlier traffic volumes, we use tensor modeling the traffic volume.

In order to solve the traffic volume outlier data problem, we formulate the traffic volume recovery problem as a data recovery problem based on the assumption that the essential traffic volume is low- $n$ -rank/low rank and the outliers are sparse. That is, the corrupted traffic volumes can be formulated as

$$\mathcal{A} = \mathcal{L} + \mathcal{S}, \quad (1)$$

where  $\mathcal{A}$  is the observed traffic volume which is corrupted,  $\mathcal{L}$  is the recovered traffic volume, and  $\mathcal{S}$  represents the outliers. In the problem, the entrances of corrupted traffic data are unknown. One straight solution is optimizing the following problem under the assumption that the  $n$ -rank of  $\mathcal{L}$  is small and the corrupted outliers are sparse or bounded:

$$\begin{aligned} \min_{\mathcal{L}} \sum_i \mu_i \text{rank}_i(\mathcal{L}) \\ \text{s.t. } \|\mathcal{A} - \mathcal{L} - \mathcal{S}\|_F \leq \delta. \end{aligned} \quad (2)$$

The tensor recovery problem of (2) has been studied in recent years, which will be detailed in Section 3. In this paper, a new data recovery method based on tensor model called Alternating Direction Method of Multipliers for Tensor Recovery (ADMM-TR) is proposed to handle the outlier traffic volumes.

This paper makes three main contributions. (1) We use tensor to model the traffic volume and take advantage of the multiway characteristics of tensor, which could explore the multicorrelations of different modes in traffic data; and (2) we formulate the problem of the traffic volume outlier data recovery as a tensor recovery problem; (3) we proposed ADMM-TR algorithm by extending ADMM from matrix to tensor case to solve the formulated tensor recovery problem for traffic volume, and the convergence of ADMM-TR is

proved. It also should be noted that the proposed ADMM-TR method is different from [9], which reported that extended ADM for tensor recovery is proposed. In fact, they presented ADM for tensor completion, in which the data are missed and the entrances of missing data are known. While in this research, the objective is to recover the data which are corrupted including missing or noised and the entrances of corrupted values are unknown.

The paper is organized as follows. We present the review of tensor model in Section 2. Section 3 briefly reviews the related data recovery methods. In Section 4, tensor model for traffic volume is constructed, traffic data recovery problem is formulated, and an efficient algorithm is proposed to solve the formulation. Also a simple convergence guarantee for the proposed algorithm is given. In Section 5, we evaluated the proposed method on synthetic data and real world traffic volume data. Finally, we provide some concluding remarks in Section 6.

## 2. Notation and Review of Tensor Models

In this section, we adopt the nomenclature of Kolda and Bader's review on tensor decomposition [10] and partially adopt the notation in [11].

A tensor is the generalization of a matrix to higher dimensions. We denote scalars by lowercase letters ( $a, b, c, \dots$ ), vectors as bold-case lowercase letters ( $\mathbf{a}, \mathbf{b}, \mathbf{c}, \dots$ ) and matrices as uppercase letters ( $A, B, C, \dots$ ). Tensors are written as calligraphic letters ( $\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$ ).

An  $n$ -mode tensor is denoted as  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ . Its elements are denoted as  $a_{i_1 \dots i_k \dots i_n}$ , where  $1 \leq i_k \leq I_k$ ,  $1 \leq k \leq N$ . The mode- $n$  unfolding (also called matricization or flattening) of a tensor  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$  is defined as  $\text{unfold}(\mathcal{A}, n) := A_{(n)}$ . The tensor element  $(i_1, i_2, \dots, i_N)$  is mapped to the matrix element  $(i_n, j)$ , where

$$j = 1 + \sum_{\substack{k=1 \\ k \neq n}}^N (i_k - 1) J_k, \quad \text{with } J_k = \prod_{\substack{m=1 \\ m \neq n}}^{k-1} I_m. \quad (3)$$

Therefore,  $A_{(n)} \in \mathbb{R}^{I_n \times J}$ , where  $J = \prod_{k=1, k \neq n}^N I_k$ . Accordingly, its inverse operator fold can be defined as  $\text{fold}(A_{(n)}, n) := \mathcal{A}$ .

The  $n$ -rank of an  $N$ -dimensional tensor  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ , denoted by  $r_n$ , is the rank of the mode- $n$  unfolding matrix  $A_{(n)}$ :

$$r_n = \text{rank}_n(\mathcal{A}) = \text{rank}(A_{(n)}). \quad (4)$$

The inner product of two same-size tensors  $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$  is defined as the sum of the products of their entries, that is,

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1} \sum_{i_2} \dots \sum_{i_N} a_{i_1 \dots i_k \dots i_n} b_{i_1 \dots i_k \dots i_n}. \quad (5)$$

The corresponding Frobenius norm is  $\|\mathcal{A}\|_F = \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}$ . Besides, the  $l_0$  norm of a tensor  $\mathcal{A}$ , denoted by  $\|\mathcal{A}\|_0$ , is the number of nonzero elements in  $\mathcal{A}$  and the  $l_1$  norm is defined

as  $\|\mathcal{A}\|_1 := \sum_{i_1, \dots, i_k, \dots, i_n} |a_{i_1, \dots, i_k, \dots, i_n}|$ . It is clear that  $\|\mathcal{A}\|_F = \|A_{(n)}\|_F$ ,  $\|\mathcal{A}\|_0 = \|A_{(n)}\|_0$ , and  $\|\mathcal{A}\|_1 = \|A_{(n)}\|_1$  for any  $1 \leq n \leq N$ .

The  $n$ -mode (matrix) product of a tensor  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$  with a matrix  $M \in \mathbb{R}^{J \times I_n}$  is denoted by  $\mathcal{A} \times_n M$  and is size  $I_1 \times \dots \times I_{n-1} \times J \times I_{n+1} \times \dots \times I_N$ . In terms of flattened matrix, the  $n$ -mode product can be expressed as

$$\mathcal{Y} = \mathcal{A} \times_n M \iff Y_{(n)} = MA_{(n)}. \quad (6)$$

### 3. Review of Data Recovery Methods

Recently, the problem of recovering the sparse and low-rank components with no prior knowledge about the sparsity pattern of the sparse matrix, or the rank of the low-rank matrix, has been well studied. Authors of [12] proposed the concept of ‘‘rank-sparse incoherence’’ and solved the problem by an interior point solver after being reformulated as a semidefinite problem. However, although interior point methods normally take very few iterations to converge, they have difficulty in handling large matrices. So this limitation prevented the usage of the technique in computer vision and the traffic volume recovery in this research.

To solve the problem for large scale matrices, Wright et al. [13] have adopted the iterative thresholding technique to solve the problem and obtained scalability properties. Lin et al. have proposed an accelerated proximal gradient (APG) algorithm [14] and applied techniques of augmented Lagrange multipliers (ALM) [15] to solve the problem. Yuan and Yang [16] have utilized the alternating direction method (ADM) which can be regarded as a practical version of the classical ALM method to solve the matrix recovery problem. The ADM method has been proved to have a pleasing convergence speed and results in [16] demonstrated its excellent performance.

Inspired by the idea of [16], this paper extends the sparse and low-rank recovery problem to tensor case, which is due to the fact that the multidimensional traffic data can be formulated into the form of tensor.

### 4. ADMM-TR for Traffic Volume Outlier Recovery

In this section, we show the solution of problem (2). The tensor model is firstly constructed for traffic volume in Section 4.1. Then we present the tensor recovery problem in Section 4.2. In Section 4.3, the classical ADMM approach is introduced. In Section 4.4, we convert the original problem into a constrained convex optimization problem which can be solved by the extended ADMM approach and present the details of the proposed algorithm. Also the convergence guarantees of the proposed algorithm are given in this section.

*4.1. Tensor Model for Traffic Volume.* The correlations of traffic volume data are critical for recovering the corrupted traffic volume data. Traditional methods mostly exploit part of correlations, such as historical or temporal neighboring correlations. The classic methods usually utilize the temporal

TABLE 1: The similarity coefficient of four modes.

Mode	Size	Similarity coefficient
Hour	$6 \times 12$	0.9670
Day	$7 \times 288$	0.8654
Week	$7 \times 288$	0.9153
Link	$4 \times 288$	0.8497

correlations of traffic data from day to day. For the single detector data, multiple correlations contain the relations of traffic data from day to day, hour to hour, and so forth. In addition, the spatial correlations exist in multiple detectors data.

In this paper, quantitative analysis of traffic data correlation is analyzed based on the traffic volume data downloaded from <http://pems.dot.ca.gov/>. The correlation coefficient applied to measuring the data correlation is given by [17]

$$s = \frac{\sum_{n \geq i > j \geq 1} R(i, j)}{n(n-1)/2}, \quad (7)$$

where  $n$  refers to the whole data points;  $R(i, j)$  refers to the correlation coefficient matrix. Table 1 gives the results of correlation coefficient of four modes, which is hour, day, week, and month.

Conventional methods usually use day-to-day matrix pattern to model the traffic data. Although each mode of traffic data has a very high similarity, these methods do not utilize the multimode correlations, which are ‘‘Day  $\times$  Hour,’’ ‘‘Week  $\times$  Hour,’’ and ‘‘Link  $\times$  Hour,’’ simultaneously and thus may result in poor recovery performance.

To make full use of the multimode correlation and traffic spatial-temporal information, traffic data need to be constructed into multiway data set. Fortunately, tensor pattern based traffic data can be well used to model the multiway traffic data. This helps keep the original structure and employ enough traffic spatial-temporal information.

*4.2. The Tensor Recovery Problem.* The problem of (2) is NP hard, since it is not convex. Then, we use an approximation formulation, as shown in (8),

$$\begin{aligned} \min_{\mathcal{L}, \mathcal{S}} : & \|\mathcal{L}\|_* + \eta \|\mathcal{S}\|_1, \\ \text{s.t.} \quad & \|\mathcal{A} - \mathcal{L} - \mathcal{S}\|_F \leq \delta, \end{aligned} \quad (8)$$

where  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$  is the given matrix to be recovered;  $\mathcal{L} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$  is the low-rank component of  $\mathcal{A}$ ;  $\mathcal{S} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$  is the sparse component of  $\mathcal{A}$ . Compared with (2), (8) relaxes the constrain for recovering a low- $n$ -rank tensor from a high-dimensional data tensor despite both small entry-wise noise and gross sparse errors.

Recently, Liu et al. [18] have proposed the definition of the nuclear norm of an  $n$ -mode tensor:

$$\|\mathcal{X}\|_* := \frac{1}{n} \sum_{i=1}^n \|X_{(i)}\|_*. \quad (9)$$

Based on this definition, the optimization in (8) can be written as

$$\min_{\mathcal{L}, \mathcal{S}}: \|\mathcal{L}\|_* + \eta \|\mathcal{S}\|_1 \equiv \frac{1}{n} \sum_{i=1}^n \lambda_i \|L_{(i)}\|_* + \frac{1}{n} \sum_{i=1}^n \eta_i \|S_{(i)}\|_1 \quad (10)$$

$$\text{s.t. } \|\mathcal{A} - \mathcal{L} - \mathcal{S}\|_F \leq \delta.$$

In order to recover  $(\widehat{\mathcal{L}}, \widehat{\mathcal{S}})$ , instead of directly solving (10), we solve the following dual problem:

$$\min_{\mathcal{L}, \mathcal{S}}: \frac{1}{2\gamma} \|\mathcal{A} - \mathcal{L} - \mathcal{S}\|_F^2 + \sum_{i=1}^n \lambda_i \|L_{(i)}\|_* + \sum_{i=1}^n \eta_i \|S_{(i)}\|_1. \quad (11)$$

The problem in (11) is still difficult to solve due to the interdependent nuclear norm and  $l_1$  norm constraints. To simplify the problem, the formulation can be reformulated as follows:

$$\min_{\mathcal{L}, \mathcal{S}, M_i, N_i}: \frac{1}{2\gamma} \|\mathcal{A} - \mathcal{L} - \mathcal{S}\|_F^2 + \sum_{i=1}^n \lambda_i \|M_i\|_* + \sum_{i=1}^n \eta_i \|N_i\|_1 \quad (12)$$

$$\text{s.t. } P_i \mathcal{L} = M_i \quad P_i \mathcal{S} = N_i \quad \forall i,$$

where  $P_i$  is the matrix representation of mode- $i$  unfolding (note that  $P_i$  is a permutation matrix; thus  $P_i^T P_i = I$ );  $M_{(i)}$  and  $N_{(i)}$  are additional auxiliary matrices of the same size as the mode- $i$  unfolding of  $\mathcal{L}$  (or  $\mathcal{S}$ ).

**4.3. The Classical ADMM Approach.** The classical alternating direction method of multipliers (ADMM) is for solving structured convex programs of the form:

$$\begin{aligned} \min_{x \in C_x, y \in C_y}: & f(x) + g(y) \\ \text{s.t. } & Ax + By = c, \end{aligned} \quad (13)$$

where  $f$  and  $g$  are convex functions defined on closed subsets  $C_x$  and  $C_y$ ;  $A$ ,  $B$ , and  $c$  are matrices and vector of appropriate sizes. The segmented Lagrangian function of (13) is

$$\begin{aligned} L_A(x, y, w) = & f(x) + g(y) + \langle w, Ax + By - c \rangle \\ & + \frac{\beta}{2} \|Ax + By - c\|_2^2, \end{aligned} \quad (14)$$

where  $w$  is a Lagrangian multiplier vector and  $\beta > 0$  is a penalty parameter.

The approach performs one sweep of alternating minimization with respect to  $x$  and  $y$  individually then updates the multiplier  $w$ ; at the iteration  $k$  the steps are given by [18, Equations (4.79)–(4.81)]:

$$\begin{aligned} x^{(k+1)} & \leftarrow \arg \min_{x \in C_x} L_A(x, y^{(k)}, w^{(k)}), \\ y^{(k+1)} & \leftarrow \arg \min_{y \in C_y} L_A(x^{(k+1)}, y, w^{(k)}), \\ w^{(k+1)} & \leftarrow w^{(k)} + \rho \beta (Ax^{(k+1)} + By^{(k+1)} - c), \end{aligned} \quad (15)$$

where  $\rho$  is the step length. A convergence proof for the above ADMM algorithm was shown as follows.

**Theorem 1** (See [19, Proposition 5.2]). *Assume that the optimal solution set  $X^*$  of (13) is nonempty. Furthermore, assume that  $C_x$  is bounded or else the matrix  $A^* A$  is invertible. Then a sequence  $\{x^{(k)}, y^{(k)}, w^{(k)}\}$  generated by (15) is bounded, and every limit point of  $\{x^{(k)}\}$  is an optimal solution of the original of problem (13).*

**4.4. ADMM Extension to Tensor Recovery.** We observe (12) is well structured in the sense that the separable structure emerges in both the objective function and constraints. Thus, we propose an algorithm based on an extension of the classical ADMM approach for solving the tensor recovery problem by taking advantage of this favorable structure.

The augmented Lagrangian of (12) is

$$\begin{aligned} L_A(\mathcal{L}, \mathcal{S}, M_i, N_i) & = \frac{1}{2\gamma} \|\mathcal{A} - \mathcal{L} - \mathcal{S}\|_F^2 \\ & + \sum_{i=1}^n \left( \lambda_i \|M_i\|_* + \langle Y_i, P_i \mathcal{L} - M_i \rangle + \frac{\alpha_i}{2} \|P_i \mathcal{L} - M_i\|_F^2 \right) \\ & + \sum_{i=1}^n \left( \eta_i \|N_i\|_1 + \langle Z_i, P_i \mathcal{S} - N_i \rangle + \frac{\beta_i}{2} \|P_i \mathcal{S} - N_i\|_F^2 \right), \end{aligned} \quad (16)$$

where  $Y_i, Z_i$  are Lagrangian multipliers and  $\alpha_i, \beta_i > 0$  are penalty parameters.

Then we can now directly apply ADMM with this augmented Lagrangian function.

**Computing  $M_i$ .** The optimal  $M_i$  can be solved with all other variables to be constant by the following subproblem:

$$\min_{M_i}: \lambda_i \|M_i\|_* + \langle Y_i, P_i \mathcal{L} - M_i \rangle + \frac{\alpha_i}{2} \|P_i \mathcal{L} - M_i\|_F^2. \quad (17)$$

As shown in [20], the optimal solution of (17) is given by

$$\widehat{M}_i = U_i D_{\lambda_i/\alpha_i}(\Lambda) V_i^T, \quad (18)$$

where  $U_i \Lambda V_i^T$  is the singular value decomposition given by

$$U_i \Lambda V_i^T = P_i \mathcal{L} + \frac{Y_i}{\alpha_i}, \quad (19)$$

and the ‘‘shrinkage’’ operator  $D_\tau(x)$  with  $\tau > 0$  is defined as

$$D_\tau(x) = \begin{cases} x - \tau & \text{if } x > \tau \\ x + \tau & \text{if } x < -\tau \\ 0 & \text{otherwise.} \end{cases} \quad (20)$$

**Computing  $N_i$ .** The optimal  $N_i$  can be solved with all other variables to be the constants by the following subproblem:

$$\min_{N_i}: \eta_i \|N_i\|_1 + \langle Z_i, P_i \mathcal{S} - N_i \rangle + \frac{\beta_i}{2} \|P_i \mathcal{S} - N_i\|_F^2. \quad (21)$$

**Input:**  $n$ -mode tensor  $\mathcal{A}$   
**Parameters:**  $\alpha, \beta, \gamma, \lambda, \eta, \rho$   
(1) Initialization:  $\mathcal{L}^{(0)} = \mathcal{A}$ ,  $\mathcal{S}^{(0)} = 0$ ,  $M_i = \mathcal{L}_{(i)}$ ,  $N_i = 0$ ,  $k = 1$   
(2) Repeat until convergence  
(3) for  $i = 1$  to  $n$   
(4)  $M_i^{(k+1)} = U_i D_{\lambda_i/\alpha_i}(\Lambda) V_i^T$   
where  $U_i \Lambda V_i^T = P_i \mathcal{L}^{(k)} + Y_i^{(k)}/\alpha_i$   
(5)  $N_i^{(k+1)} = D_{\eta_i/\beta_i}(P_i \mathcal{S}^{(k)} + Z_i^{(k)}/\beta_i)$   
(6)  $Y_i^{(k+1)} = Y_i^{(k)} + \alpha_i(P_i \mathcal{L}^{(k)} - M_i^{(k+1)})$   
(7)  $Z_i^{(k+1)} = Z_i^{(k)} + \beta_i(P_i \mathcal{S}^{(k)} - N_i^{(k+1)})$   
(8) end for  
(9)  $\mathcal{L}^{(k+1)} = \frac{\mathcal{A} - \mathcal{S}^{(k)} - \gamma \sum_{i=1}^n P_i^T (Y_i^{(k+1)} - \alpha_i M_i^{(k+1)})}{1 + \gamma \sum_{i=1}^n \alpha_i}$   
(10)  $\mathcal{S}^{(k+1)} = \frac{\mathcal{A} - \mathcal{L}^{(k+1)} - \gamma \sum_{i=1}^n P_i^T (Z_i^{(k+1)} - \alpha_i N_i^{(k+1)})}{1 + \gamma \sum_{i=1}^n \beta_i}$   
(11)  $\alpha = \rho\alpha$ ,  $\beta = \rho\beta$   
(12) End  
(13)  $k = k + 1$   
**Output:**  $n$ -mode tensor  $\mathcal{L}, \mathcal{S}$

ALGORITHM 1: ADMM-TR: ADMM for tensor recovery.

By the well-known  $l_1$  minimization [21], the optimal solution of (21) is

$$\widehat{N}_i = D_{\eta_i/\beta_i} \left( P_i \mathcal{S} + \frac{Z_i}{\beta_i} \right), \quad (22)$$

where  $D_\tau$  is the ‘‘shrinkage’’ operation.

*Computing  $\mathcal{L}$ .* Now we fix all variables except  $\mathcal{L}$  and minimize  $L_A$  over  $\mathcal{L}$ . The resulting minimization problem is the minimization of a quadratic function:

$$\begin{aligned} \min_{\mathcal{L}}: L_A(\mathcal{L}) &= \frac{1}{2\gamma} \|\mathcal{A} - \mathcal{L} - \mathcal{S}\|_F^2 \\ &+ \sum_{i=1}^n \left( \langle Y_i, P_i \mathcal{L} - M_i \rangle + \frac{\alpha_i}{2} \|P_i \mathcal{L} - M_i\|_F^2 \right). \end{aligned} \quad (23)$$

The objective function is differentiable, so the minimizer  $\mathcal{L}_{\min}$  is characterized by  $(\partial L_A(\mathcal{L}))/\partial \mathcal{L} = 0$ . Thus, we obtain

$$\mathcal{L}_{\min} = \frac{\{\mathcal{A} - \mathcal{S} - \gamma \sum_{i=1}^n P_i^T (Y_i - \alpha_i M_i)\}}{(1 + \gamma \sum_{i=1}^n \alpha_i)}. \quad (24)$$

*Computing  $\mathcal{S}$ .* Now we fix all variables except  $\mathcal{S}$  and minimize  $L_A$  over  $\mathcal{S}$ . The resulting minimization problem is the minimization of a quadratic function:

$$\begin{aligned} \min_{\mathcal{S}}: L_A(\mathcal{S}) &= \frac{1}{2\gamma} \|\mathcal{A} - \mathcal{L} - \mathcal{S}\|_F^2 \\ &+ \sum_{i=1}^n \left( \langle Z_i, P_i \mathcal{S} - N_i \rangle + \frac{\beta_i}{2} \|P_i \mathcal{S} - N_i\|_F^2 \right). \end{aligned} \quad (25)$$

The objective function is also differentiable, so the minimizer  $\mathcal{S}_{\min}$  is characterized by  $(\partial L_A(\mathcal{S}))/\partial \mathcal{S} = 0$ . Thus, we have

$$\mathcal{S}_{\min} = \frac{\{\mathcal{A} - \mathcal{L} - \gamma \sum_{i=1}^n P_i^T (Z_i - \beta_i N_i)\}}{(1 + \gamma \sum_{i=1}^n \beta_i)}. \quad (26)$$

For comparing with RSTD [22], we also choose the difference of  $\mathcal{L}$  and  $\mathcal{S}$  in successive iterations against a certain tolerance as the stopping criterion. The pseudocode of the proposed ADMM-TR algorithm is summarized in Algorithm 1.

**Theorem 2** (Convergence of ADMM-TR). *Assume that the optimal solution set  $X^*$  of (11) is nonempty. A sequence  $\{\mathcal{L}^{(k)}, \mathcal{S}^{(k)}, M_i^{(k)}, N_i^{(k)}, Y_i^{(k)}, Z_i^{(k)}\}$  generated by our proposed ADMM-TR algorithm is bounded, and every limit point of  $\{\mathcal{L}^{(k)}, \mathcal{S}^{(k)}\}$  is an optimal solution of the original problem (11).*

*Proof.* We check the assumptions of Theorem 1.  $C_x$  is not bounded, but  $P_i^* P_i = I$  is a constant multiple of the identity operator. Thus Theorem 1 can also be applied to ADMM-TR and Theorem 2 can be derived.  $\square$

## 5. Numerical Experiments

This section evaluates the empirical performance of the proposed algorithm on synthetic data and compares the results with RSTD (Rank Sparsity Tensor Decomposition) [22]. Also, experiments on traffic volume data outlier recovery illustrate the efficiency of the proposed method in traffic research filed.

We use the Lanczos algorithm for computing the singular values decomposition and adopt the same rule for predicting

TABLE 2: Comparison of ADMM-TR with RSTD for synthetic data where  $\mathcal{A}_0 \in \mathbb{R}^{40 \times 40 \times 40}$ ,  $n$ -rank = [5, 5, 5].

$spr$	ADMM-TR				RSTD			
	RSE- $\mathcal{L}_0$ (e-3)	RSE- $\mathcal{S}_0$ (e-3)	# iter	Time (s)	RSE- $\mathcal{L}_0$ (e-3)	RSE- $\mathcal{S}_0$ (e-3)	# iter	Time (s)
0.05	4.3	4.3	133	13.7	4.6	4.6	208	17.8
0.15	9.8	4.2	162	20.8	10.5	4.7	235	33.6
0.25	12.3	4.3	514	44.9	15.9	5.2	676	51.0
0.35	56.5	9.5	654	53.3	67.5	11.4	737	57.5

TABLE 3: Comparison of ADMM-TR with RSTD for synthetic data where  $\mathcal{A}_0 \in \mathbb{R}^{40 \times 40 \times 40}$ ,  $n$ -rank = [10, 10, 10].

$spr$	Algorithm: ADMM-TR				Algorithm: RSTD			
	RSE- $\mathcal{L}_0$ (e-3)	RSE- $\mathcal{S}_0$ (e-3)	# iter	Time (s)	RSE- $\mathcal{L}_0$ (e-3)	RSE- $\mathcal{S}_0$ (e-3)	# iter	Time (s)
0.05	4.4	2.0	236	26.4	4.7	2.5	338	29.9
0.15	4.7	2.1	417	42.1	5.2	2.5	603	50.8
0.25	8.9	2.9	664	60.8	12.0	4.2	663	53.6
0.35	17.3	5.2	981	80.6	21.1	6.5	1110	87.7

the dimension of the principal singular space as [22]. And the parameters are set as  $\alpha = \beta = [I_1/I_{\max}, I_2/I_{\max}, \dots, I_n/I_{\max}]^T$  and  $\gamma = 1/\text{sum}([I_1/I_{\max}, I_2/I_{\max}, \dots, I_n/I_{\max}]^T)$  for all experiments, where  $I_{\max} = \max\{I_i\}$ .  $\eta$  is set to  $1/\sqrt{I_{\max}}$  as suggested in [23].

All the experiments are conducted and timed on the same desktop with an Pentium (R) Dual-Core 2.50 GHz CPU that has 4 GB memory, running on Windows 7 and MATLAB.

**5.1. Synthetic Data.** ADMM-TR and RSTD are tested on the synthetic data of size  $40 \times 40 \times 40$ . We generate the dimension  $r$  of a ‘‘core tensor’’  $\mathcal{C} \in \mathbb{R}^{r \times \dots \times r}$  which we fill with Gaussian distributed entries ( $\sim \mathcal{N}(0, 1)$ ). Then, we generate matrices  $U^{(1)}, \dots, U^{(N)}$ , with  $U^{(i)} \in \mathbb{R}^{n_i \times r}$  whose elements are also i.i.d. Gaussian random variables ( $\sim \mathcal{N}(0, 1)$ ) and set

$$\mathcal{L}_0 = \mathcal{C}_{\times 1} U^{(1)}_{\times 2 \dots \times N} U^{(N)}. \quad (27)$$

The entries of sparse tensor  $\mathcal{S}_0$  are independently distributed, each taking on value 0 with probability  $1-spr$  and each taking on impulsive value with probability  $spr$ . We apply the proposed algorithm to the tensor  $\mathcal{A}_0 = \mathcal{L}_0 + \mathcal{S}_0$  to recover  $\mathcal{L}$  and  $\mathcal{S}$  and compare with RSTD. For these experiments, two cases of  $n$ -rank are investigated,  $n$ -rank = [5, 5, 5] and  $n$ -rank = [10, 10, 10]. Table 1 presents the average results (across 30 instances) for different  $spr$ .

The quality of recovery is measured by the relative square error (RSE) to  $\mathcal{L}_0$  and  $\mathcal{S}_0$ , defined to be

$$\begin{aligned} \text{RSE-}\mathcal{L}_0 &= \frac{\|\widehat{\mathcal{L}} - \mathcal{L}_0\|_F}{\|\mathcal{L}_0\|_F}, \\ \text{RSE-}\mathcal{S}_0 &= \frac{\|\widehat{\mathcal{S}} - \mathcal{S}_0\|_F}{\|\mathcal{S}_0\|_F}. \end{aligned} \quad (28)$$

Tables 2 and 3 show that the proposed algorithm (ADMM-TR) is about 10 percent faster than RSTD proposed

in [22] and achieves better accuracy in terms of relative square error. Though both of the algorithms involve computing a SVD per iteration, we observe that the proposed algorithm take much fewer iterations than RSTD to converge to the optimal solution.

The more impulsive entries are added, that is, for higher value of  $spr$ , the less probable it becomes for the tensor recovery problem. In addition, the problem becomes more sophisticated when the  $n$ -rank is higher for the ground truth tensors of the same size. In Table 1, different  $spr$  is set for two tensor cases. And results show the recovered accuracy decreases as the  $spr$  grows for a certain case. In particular, the recovered accuracy for the tensor  $\mathcal{A}_0 \in \mathbb{R}^{40 \times 40 \times 40}$  with  $n$ -rank = [5, 5, 5] decreases sharply when the  $spr$  is up to 40%, while the phenomenon occurs for tensor  $\mathcal{A}_0 \in \mathbb{R}^{40 \times 40 \times 40}$  with  $n$ -rank = [10, 10, 10] when the  $spr$  is about 25%. This is due to that relative low rank and low sparse ratio are precondition of the tensor recovery problem.

**5.2. Traffic Volume Data.** To evaluate the performances of the proposed method in traffic volume data outlier recovery, a complete traffic volume data set is used as the test data set. We use the data of a fixed point in Sacramento County which is downloaded from <http://pems.dot.ca.gov/>. The traffic volume data are recorded every 5 minutes. Therefore, a daily traffic volume series for a loop detector contains 288 records, and the whole period of the data lasts for 16 days, that is, from August 2 to August 17, 2010.

Based on multiple correlations of the traffic volume data, we model the data set as a tensor model of size  $16 \times 24 \times 12$  which stands for 16 days, 24 hours in a day, and 12 sample intervals (i.e., recorded by 5 minutes) per hour. The ratios of outlier data are set from 5% to 15% and the outlier data are produced randomly. All the results are averaged by 10 instances.

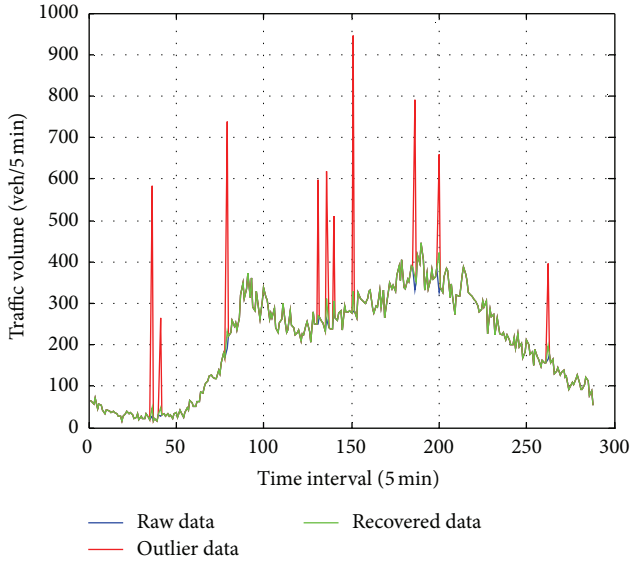


FIGURE 1: Comparisons with raw traffic volume data, data corrupted by outliers with 5% ratio, and data recovered by ADMM-TR.

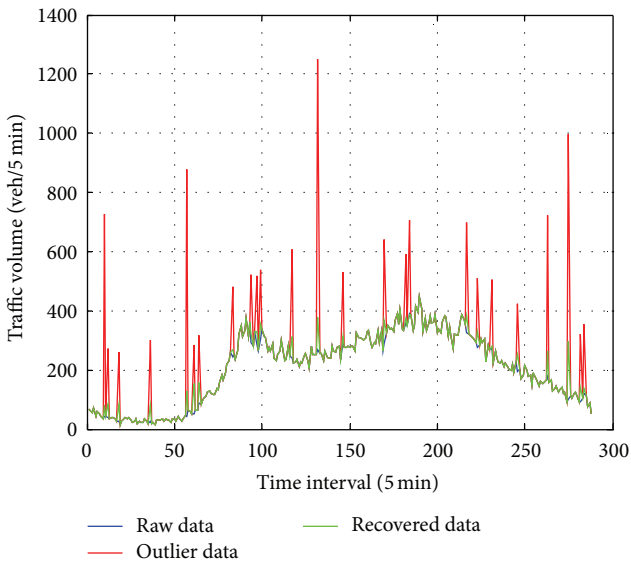


FIGURE 2: Comparisons with raw traffic volume data, data corrupted by outliers with 10% ratio, and data recovered by ADMM-TR.

For the real world data, we mainly pay attention to the correctness of the recovered traffic volume data. Thus the quality of recovery is measured by the relative square error (RSE) to  $\mathcal{L}_0$  and Mean Absolute Percentage Error (MAPE) to  $\mathcal{L}_0$ , defined to be

$$\begin{aligned} \text{RSE}_{\mathcal{L}_0} &= \frac{\|\widehat{\mathcal{L}} - \mathcal{L}_0\|_F}{\|\mathcal{L}_0\|_F}, \\ \text{MAPE}_{\mathcal{L}_0} &= \frac{1}{M} \sum_{m=1}^M \left| \frac{t_r^{(m)} - t_e^{(m)}}{t_r^{(m)}} \right|, \end{aligned} \quad (29)$$

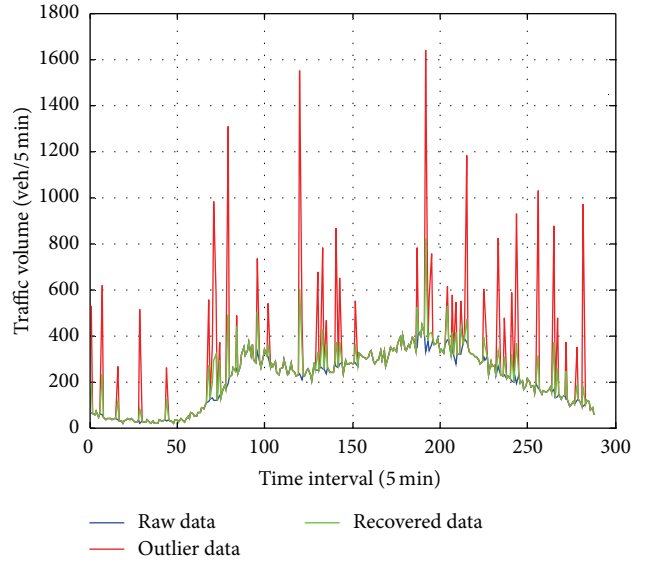


FIGURE 3: Comparisons with raw traffic volume data, data corrupted by outliers with 15% ratio, and data recovered by ADMM-TR.

TABLE 4: Traffic volume outlier data before and after being recovered.

$spr$	Before recovery		After recovery	
	RSE $_{\mathcal{L}_0}$ ( $e-3$ )	MAPE $_{\mathcal{L}_0}$ ( $e-3$ )	RSE $_{\mathcal{L}_0}$ ( $e-3$ )	MAPE $_{\mathcal{L}_0}$ ( $e-3$ )
0.05	0.5005	0.2136	0.0961	0.0314
0.10	0.7066	0.5008	0.1417	0.1027
0.15	0.8850	0.8777	0.2221	0.2351

where  $t_r^{(m)}$  and  $t_e^{(m)}$  are the  $m$ th elements which stand for the known real value and recovered value, respectively.  $M$  denotes the number of recovered traffic volumes.

Table 4 presents the relative errors of traffic volume outlier data before and after being recovered by ADMM-TR. The results show that the RSE $_{\mathcal{L}_0}$  and MAPE $_{\mathcal{L}_0}$  for traffic volume data corrupted by outlier data are about 5 times than the data after being recovered by ADMM-TR. Figures 1, 2, and 3 present the profiles of traffic volume data for a day. The results show that ADMM-TR could recover the traffic volume outlier data with perfect performance.

## 6. Conclusions

In this paper, we concentrate on the mathematical problem in traffic volume outlier data recovery and proposed a novel tensor recovery method based on alternating direction method of multipliers (ADMM). The proposed algorithm can automatically separate the low- $n$ -rank tensor data and sparse part. The experiments show that the proposed method is more stable and accurate in most cases and has excellent convergence rate. Experiments on real world traffic volume data demonstrate the practicability and effectiveness of the proposed method in traffic research domain.

In the future, we would like to investigate how to automatically choose the parameters in our algorithm and explore

additional applications of our method in traffic research domain.

## Acknowledgments

This research was supported by NSFC (Grant No. 61271376, 61171118, and 91120015), Beijing Natural Science Foundation (4122067). The authors would like to thank Professor Bin Ran from the University of Wisconsin-Madison and Yong Li from the University of Notre Dame for the suggestive discussions.

## References

- [1] K. Tamura and M. Hirayama, "Toward realization of VICS—vehicle information and communications system," in *Proceedings of the IEEE-IEE Vehicle Navigation and Informations Systems Conference*, pp. 72–77, October 1993.
- [2] S. Kim, M. E. Lewis, and C. C. White III, "Optimal vehicle routing with real-time traffic information," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 2, pp. 178–188, 2005.
- [3] P. Mirchandani and L. Head, "A real-time traffic signal control system: architecture, algorithms, and analysis," *Transportation Research Part C*, vol. 9, no. 6, pp. 415–432, 2001.
- [4] B. M. Williams, P. K. Durvasula, and D. E. Brown, "Urban freeway traffic flow prediction: application of seasonal autoregressive integrated moving average and exponential smoothing models," *Transportation Research Record*, no. 1644, pp. 132–141, 1998.
- [5] J. Xu, X. Li, and H. Shi, "Short-term traffic flow forecasting model under missing data," *Journal of Computer Applications*, vol. 30, no. 4, pp. 1117–1120, 2010.
- [6] Y. Pei and J. Ma, "Real-time traffic data screening and reconstruction," *China Civil Engineering Journal*, vol. 36, no. 7, pp. 78–83, 2003.
- [7] S. Chen, W. Wang, and W. Li, "Noise recognition and noise reduction of real-time traffic data," *Dongnan Daxue Xuebao*, vol. 36, no. 2, pp. 322–325, 2006.
- [8] B.-G. Daniel, J. D.-P. Francisco, G.-O. David et al., "Wavelet-based denoising for traffic volume time series forecasting with self-organizing neural networks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 25, no. 7, pp. 530–545, 2010.
- [9] S. Gandy, B. Recht, and I. Yamada, "Tensor completion and low-rank tensor recovery via convex optimization," *Inverse Problems*, vol. 27, no. 2, Article ID 025010, 2011.
- [10] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.
- [11] A. S. Lewis and G. Knowles, "Image compression using the 2-D wavelet transform," *IEEE Transactions of Image Processing*, vol. 1, no. 2, pp. 244–250, 1992.
- [12] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky, "Rank-sparsity incoherence for matrix decomposition," *SIAM Journal on Optimization*, vol. 21, no. 2, pp. 572–596, 2011.
- [13] J. Wright, Y. Peng, Y. Ma, and A. Ganesh, "Robust principal component analysis: exact recovery of corrupted low-rank matrices by convex optimization," in *Proceedings of the 23rd Annual Conference on Neural Information Processing Systems (NIPS '09)*, pp. 2080–2088, Vancouver, Canada, 2009.
- [14] A. Ganesh, Z. Lin, J. Wright, L. Wu, M. Chen, and Y. Ma, "Fast algorithms for recovering a corrupted low-rank matrix," in *Proceedings of the 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP '09)*, pp. 213–216, December 2009.
- [15] Z. Lin, M. Chen, L. Wu, and Y. Ma, "The augmented Lagrange multiplier method for exact recovery of a corrupted low-rank matrices," *Mathematical Programming*, 2009.
- [16] X. Yuan and J. Yang, "Sparse and low-rank matrix decomposition via alternating direction methods," Tech. Rep., Department of Mathematics, Hong Kong Baptist University, 2009.
- [17] L. Qu, Y. Zhang, J. Hu, L. Jia, and L. Li, "A BPCA based missing value imputing method for traffic flow volume data," in *Proceedings of the IEEE Intelligent Vehicles Symposium (IV '08)*, pp. 985–990, June 2008.
- [18] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," in *Proceedings of the International Conference on Computer Vision (ICCV '09)*, 2009.
- [19] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation*, Prentice-Hall, 1989.
- [20] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.
- [21] E. T. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation for  $\ell_1$ -minimization: methodology and convergence," *SIAM Journal on Optimization*, vol. 19, no. 3, pp. 1107–1130, 2008.
- [22] Y. Li, J. Yan, Y. Zhou, and J. Yang, "Optimum subspace learning and error correction for tensors," in *Proceedings of the 11th European Conference on Computer Vision (ECCV '10)*, Crete, Greece, 2010.
- [23] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM*, vol. 58, no. 3, article 11, 2011.





# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

