Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 17, No. 3, pp. 453–466, 2021

LOW RANK NON-NEGATIVE TRIPLE DECOMPOSITION AND NON-NEGATIVE TENSOR COMPLETION

Chunyan Wang, Min Wang, Haibin Chen* and Yiju Wang

Abstract: Non-negative tensor factorization (NTF) is an important tool in data analysis and signal processing, and non-negative triple decomposition is a new kind of NTF. In this paper, we study the non-negative triple decomposition of third order non-negative tensors. For this purpose, an alternating proximal gradient method is introduced and the global convergence of the algorithm is also established. As an application of the proposed results, we consider the non-negative tensor completion problem. Numerical experiments show that the proposed algorithm offers competitive performance even though the given tensor is highly sparse.

Key words: *high-dimensional data; Non-negative tensor factorization; low rank non-negative tensors*

Mathematics Subject Classification: *65H17, 15A18, 90C30*

1 Introduction

Tensor decomposition is an important tool for data analysis in applications such as chemometrics, biogeochemistry, neuroscience, signal processing, cyber traffic analysis, and many others. Two well-known representations of tensors are the CANDECOMP/PARAFAC (CP) decomposition and Tucker decomposition, which are generally expressed as a sum of outer products of vectors [2, 7, 18]. Recently, a new tensor decomposition named tensor triple decomposition is introduceed in [13]. The new tensor decomposition is applied to third order tensors, which decomposes a third order tensor to three third order factor tensors and each factor tensor has two low dimensions. Third order tensors are very common and useful high order tensors in applications [1, 5, 17, 21, 23, 25, 26, 27, 28].

Generally speaking, tensor rank is an essential definition in tensor decomposition and tensor completion problems. In [13], Qi *et al.* gave the definition of triple rank for third order tensors. It is shown that the triple rank of a third order tensor is not greater than the CP rank and the middle value of the Tucker rank, is strictly less than the CP rank with a substantial probability, and is strictly less than the middle value of the Tucker rank for an essential class of examples. This indicates that practical data can be approximated by low rank triple decomposition as long as it can be approximated by low rank CP or Tucker decomposition.

Through various applications, non-negative tensor decomposition is an effective technique that has proven to be useful for a wide variety of applications [3, 19, 10, 4]. Compared with

*Corresponding author.

non-negative matrix factorization (NMF) [15, 12], tensor factorization can more accurately consider the spatial and temporal correlation [14]. From the calculation point of view, noted that there are negative values in the decomposition results, but negative value elements are often meaningless in some practical problems. For example, it is impossible to have negative pixels in image data, document statistics, negative values cannot be explained.

In this paper, we focus on non-negative tensor factorization for third order non-negative tensors by tensor triple decomposition, which decomposes a third order non-negative tensor to corresponding low rank non-negative tensors in a balanced way. Then, an alternating proximal gradient (APG) method is introduced to apply to solve a kind of non-negative tensor completion problems. Furthermore, we establish the global convergence and the asymptotic convergence rate of the method based on the Kurdyka- Lojasiewicz inequality. The efficiency of the proposed algorithm is tested on tensor decomposition, as well as tensor completion from incomplete observations. Moreover, we also compare the results between the case of nonnegative constraint and unconstraint case. Example shows that the error and relative error for the unconstraint case are less than the nonnegative constraint case.

The remainder of this paper is organized as follows. In Section 2, we recall some basic concepts and preliminary results in the literature. In Section 3, we introduce an APG method and the convergence of the algorithm is also established. Then, we study the third order non-negative tensor completion problem in Section 4. Several numerical examples are given to show the efficiency of the corresponding conclusions in Section 5.

To end this section, we present several useful symbols in the paper. Let $\mathbb{R}^n$ be the n dimensional real Euclidean space. The set of all nonnegative vectors in $\mathbb{R}^n$ is denoted by $\mathbb{R}^n_+$. The set of all positive integers is denoted by $\mathbb{N}$. Denote $[n] = \{1, 2, \cdots, n\}$ for any $n \in \mathbb{N}$. Vectors are denoted by bold lowercase letters i.e. $\mathbf{x}, \ \mathbf{y}, \cdots$, matrices are denoted by capital letters i.e. $A, \ B, \ \cdots$, and tensors are written as calligraphic capitals such as $\mathcal{A}, \mathcal{T}, \cdots$. $\mathcal{A} \geq 0$ means that all elements of $\mathcal{A}$ are nonnegative.

2 Preliminaries

In this section, we recall some useful notations, basic concepts and preliminary results.

We know that a third-order tensor $\mathcal{A}$ has column, row, and tube fibers, which are defined by fixing every index but one and denoted by $a_{:jk}, a_{i:k}$ and $a_{ij:}$, respectively. Correspondingly, we obtain three matricizations of a tensor $\mathcal{A}$:

$$
\begin{aligned}
A_{(1)} &= [a_{:11}, a_{:21}, \ldots, a_{:J1}, a_{:12}, \ldots, a_{:J2}, \ldots, a_{:1K}, \ldots, a_{:JK}] \in R^{I \times JK},\\
A_{(2)} &= [a_{1:1}, a_{2:1}, \ldots, a_{I:1}, a_{1:2}, \ldots, a_{I:2}, \ldots, a_{1:K}, \ldots, a_{I:K}] \in R^{J \times IK},\\
A_{(3)} &= [a_{11:}, a_{21:}, \ldots, a_{I1:}, a_{12:}, \ldots, a_{I2:}, \ldots, a_{1J:}, \ldots, a_{IJ:}] \in R^{K \times IJ}.
\end{aligned}
$$

Let $\mathcal{X} = (x_{ijt}) \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. As in [7], let $\mathcal{X}(i, :, :)$ denote the i-th horizontal slice, $\mathcal{X}(:, j, :)$ to denote the j-th lateral slice; $\mathcal{X}(:, :, t)$ to denote the t-th frontal slice. We say that $\mathcal{X}$ is a third order horizontally square tensor if all of its horizontal slices are square, i.e., $n_2 = n_3$. Similarly, $\mathcal{X}$ is a third order laterally square tensor (resp. frontally square tensor) if all of its lateral slices (resp. frontal slices) are square, i.e., $n_1 = n_3$ (resp. $n_1 = n_2$).

For $m, n \in \mathbb{N}$, assume $\mathcal{D} \subseteq \mathbb{R}^{m \times n}$ is a set of matrices. Let $\delta_{\mathcal{D}}$ denotes the indicator function of $\mathcal{D}$ such that

$$
\delta_{\mathcal{D}}(A) = \begin{cases} 0, & A \in \mathcal{D}, \\ \infty, & \text{otherwise}. \end{cases} \tag{2.1}
$$

Suppose $\mathbf{x}$ is a vector that can be decomposed into three blocks $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$. Then, the function $f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ is called block multiconvex if, for each $i \in [3]$, f is a convex function of $\mathbf{x}_i$ while other two blocks are fixed.

Next, we present the definition of tensor triple decomposition [13].

Definition 2.1. Let $\mathcal{X} = (x_{ijt}) \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ be a third order tensor. We say that $\mathcal{X}$ is the triple product of a third order horizontally square tensor $\mathcal{A} = (a_{iqs}) \in \mathbb{R}^{n_1 \times r \times r}$, a third order laterally square tensor $\mathcal{B} = (b_{pjs}) \in \mathbb{R}^{r \times n_2 \times r}$ and a third order frontally square tensor $\mathcal{C} = (c_{pqk}) \in \mathbb{R}^{r \times r \times n_3}$, and denotes

$$\mathcal{X} = \mathcal{A}\mathcal{B}\mathcal{C}, \tag{2.2}$$

if for $i = 1, \cdots n_1, j = 1, \cdots n_2$ *and* $t = 1, \cdots n_3$, we have

$$x_{ijt} = \sum_{p,q,s=1}^{r} a_{iqs} b_{pjs} c_{pqt}. \tag{2.3}$$

If

$$r \leq \mathrm{mid}\{n_1, n_2, n_3\}, \tag{2.4}$$

then we call (2.2) a low rank triple decomposition of $\mathcal{X}$.

By Definition 2.1, we mainly focus on that $\mathcal{X}$ is nonnegative in this paper, and $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ are also nonnegative.

$\boxed{3}$ Alternating Proximal Gradient (APG) Algorithm

In this section, we first give an alternating proximal gradient (APG) algorithm, and then the global convergence and convergence rate are also studied.

To continue, we consider a given nonnegative third order tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ with $n_1, n_2, n_3 \geq 1$ and a fixed positive integer $r \leq \mathrm{mid}\ \{n_1, n_2, n_3\}$. Then, we have the following optimization problem:

$$\min_{\mathcal{A}\geq 0, \mathcal{B}\geq 0, \mathcal{C}\geq 0} f(\mathcal{A}, \mathcal{B}, \mathcal{C}), \tag{3.1}$$

where

$$f(\mathcal{A}, \mathcal{B}, \mathcal{C}) = \| \mathcal{X} - \mathcal{A}\mathcal{B}\mathcal{C} \|_F^2 = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{t=1}^{n_3} \left(x_{ijt} - \sum_{p=1}^{r} \sum_{q=1}^{r} \sum_{s=1}^{r} a_{iqs} b_{pjs} c_{pqt} \right)^2,$$

and $\mathcal{A} = (a_{iqs}) \in \mathbb{R}^{n_1 \times r \times r}$, $\mathcal{B} = (b_{pjs}) \in \mathbb{R}^{r \times n_2 \times r}$, $\mathcal{C} = (c_{pqk}) \in \mathbb{R}^{r \times r \times n_3}$. By (3.1), we want to obtain a triple decomposition $\mathcal{A}\mathcal{B}\mathcal{C}$ with triple rank not greater than r, to approximate $\mathcal{X}$.

Generally speaking, $f(\mathcal{A}, \mathcal{B}, \mathcal{C})$ in (3.1) maybe not convex in joint variables $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. However, f is convex when any two variables in $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ are fixed. So f is block multiconvex, which can be solved by the famous block coordinate descent (BCD) method [22, 6, 16, 20].

For the sake of computation in Algorithm 3.1, we present matrices $M_1^{k-1} \in \mathbb{R}^{r^2 \times n_2 n_3}$, $M_2^{k-1} \in \mathbb{R}^{r^2 \times n_1 n_3}$, $M_3^{k-1} \in \mathbb{R}^{r^2 \times n_1 n_2}$ in the k-th iteration with elements:

$$(M_1^{k-1})_{lm} = \sum_{p=1}^{r} b_{pjs}^{k-1} c_{pqt}^{k-1}, \qquad \text{where } l = q + (s-1)r, \quad m = j + (t-1)n_2, \tag{3.2}$$

$$(M_2^{k-1})_{lm} = \sum_{q=1}^{r} a_{iqs}^k c_{pqt}^{k-1}, \qquad \text{where } l = p + (s-1)r, \quad m = i + (t-1)n_1, \tag{3.3}$$

$$(M_3^{k-1})_{lm} = \sum_{s=1}^{r} a_{iqs}^k b_{pjs}^k, \qquad \text{where } l = p + (q-1)r, \quad m = i + (j-1)n_1. \tag{3.4}$$

Let $A_{(1)} \in \mathbb{R}^{n_1 \times r^2}$ be the mode-1 unfolding of the tensor $\mathcal{A}$ and $B_{(2)} \in \mathbb{R}^{n_2 \times r^2}$ be the mode-2 unfolding of the tensor $\mathcal{B}$ and $C_{(3)} \in \mathbb{R}^{n_3 \times r^2}$ be the mode-3 unfolding of the tensor $\mathcal{C}$. Let $X_{(i)}$ be the mode-i unfolding of the tensor $\mathcal{X}$. $\mathfrak{B}^* : \mathbb{R}^{n_i \times r^2} \to \mathcal{T}$ denotes refolding mode-i unfolding of a tensor to the corresponding tensor.

Now, we first show how to update $A_{(1)}$ at k-th iteration. Since $f = \frac{1}{2}\left\|A_{(1)}M_1^{k-1} - X_{(1)}\right\|_F^2$, we know that

$$\nabla_{A_{(1)}} f = (A_{(1)}M_1^{k-1} - X_{(1)})(M_1^{k-1})^T. \tag{3.5}$$

Let

$$L_1^{k-1} = \left\|\left(M_1^{k-1}\right)^\top M_1^{k-1}\right\|, \quad \omega_1^{k-1} = \min\left\{\hat{\omega}_{k-1}, \delta_\omega \sqrt{\frac{L_1^{k-2}}{L_1^{k-1}}}\right\}, \tag{3.6}$$

where $\|A\|$ is the spectral norm of A, $\delta_\omega < 1$ and $\hat{\omega}_{k-1} = \frac{t_{k-1}-1}{t_k}$ with

$$t_0 = 1, \quad t_k = \frac{1}{2}\left(1 + \sqrt{1 + 4t_{k-1}^2}\right).$$

Furthermore, let $\hat{A}_{(1)}^{k-1} = A_{(1)}^{k-1} + \omega_1^{k-1}\left(A_{(1)}^{k-1} - A_{(1)}^{k-2}\right)$. By (3.5), it follows that

$$\hat{G}_1^{k-1} = \left(\hat{A}_1^{k-1} M_1^{k-1} - X_{(1)}\right)(M_1^{k-1})^\top.$$

Then we update $A_{(1)}^k$ as below:

$$A_{(1)}^k = \underset{A_{(1)} \geq 0}{\operatorname{argmin}} \left\langle \hat{G}_1^{k-1}, A_{(1)} - \hat{A}_{(1)}^{k-1} \right\rangle + \frac{L_1^{k-1}}{2}\left\|A_{(1)} - \hat{A}_{(1)}^{k-1}\right\|_F^2,$$

which can be written in the closed form

$$A_{(1)}^k = \max\left\{0, \hat{A}_{(1)}^{k-1} - \hat{G}_1^{k-1}/L_1^{k-1}\right\}. \tag{3.7}$$

Then, $B_{(2)}^k$ and $C_{(3)}^k$ can be updated similarly.

At the end of iteration k, we check whether $f(\mathcal{A}^k, \mathcal{B}^k, \mathcal{C}^k) \geq f(\mathcal{A}^{k-1}, \mathcal{B}^{k-1}, \mathcal{C}^{k-1})$. If so, we reupdate $A_{(1)}^k, B_{(2)}^k, C_{(3)}^k$ by (3.7) with $\hat{A}_{(1)}^{k-1} = A_{(1)}^{k-1}$, $\hat{B}_{(2)}^{k-1} = B_{(2)}^{k-1}$, $\hat{C}_{(3)}^{k-1} = C_{(3)}^{k-1}$.

Algorithm 3.1. Alternating proximal gradient (APG) method.

Input: Nonnegative tensor $\mathcal{X}$. Choose an integer $1 \le r \le \text{mid}\{n_1, n_2, n_3\}$.
Output: Nonnegative factors $A_{(1)}, B_{(2)}, C_{(3)}$.
Initialization: Choose a positive number $\delta_\omega < 1$ and randomize $A^{-1}_{(1)} = A^0_{(1)}$, $B^{-1}_{(2)} = B^0_{(2)}$, $C^{-1}_{(3)} = C^0_{(3)}$, as nonnegative matrices of appropriate sizes.
for $k = 1, 2, \ldots$ do
  for $n = 1, 2, 3$ do
    Compute L^{k-1}_n and set ω^{k-1}_n according to (3.6).
    Let $\hat{A}^{k-1}_{(1)} = A^{k-1}_{(1)} + \omega^{k-1}_1 \left(A^{k-1}_{(1)} - A^{k-2}_{(1)}\right)$,
      $\hat{B}^{k-1}_{(2)} = B^{k-1}_{(2)} + \omega^{k-1}_2 \left(B^{k-1}_{(2)} - B^{k-2}_{(2)}\right)$,
      $\hat{C}^{k-1}_{(3)} = C^{k-1}_{(3)} + \omega^{k-1}_3 \left(C^{k-1}_{(3)} - C^{k-2}_{(3)}\right)$.
    Update $A^k_{(1)}$, $B^k_{(2)}$, $C^k_{(3)}$ according to (3.7).
  end for
  if $f(\mathcal{A}^k, \mathcal{B}^k, \mathcal{C}^k) \ge f(\mathcal{A}^{k-1}, \mathcal{B}^{k-1}, \mathcal{C}^{k-1})$ then
    Reupdate $A^k_{(1)}$, $B^k_{(2)}$, $C^k_{(3)}$ according to (3.7) with $\hat{A}^{k-1}_{(1)} = A^{k-1}_{(1)}$, $\hat{B}^{k-1}_{(2)} = B^{k-1}_{(2)}$, $\hat{C}^{k-1}_{(3)} = C^{k-1}_{(3)}$.
  end if
  if stopping criterion is satisfied then
    Return $A^k_{(1)}$, $B^k_{(2)}$, $C^k_{(3)}$.
  end if
end for

$\mathcal{A} = \mathfrak{B}^*(A_{(1)})$, $\mathcal{B} = \mathfrak{B}^*(B_{(2)})$ and $\mathcal{C} = \mathfrak{B}^*(C_{(3)})$ are the final required results.

To give the convergence of the algorithm, we first review the KL inequality, which is essential in the following analysis.

Definition 3.1. A function $\psi(\mathbf{x})$ satisfies the Kurdyka-Lojasiewicz (KL) property at point $\overline{\mathbf{x}} \in \text{dom}(\partial\psi)$ if there exists $\theta \in [0, 1)$ such that

$$\frac{|\psi(\mathbf{x}) - \psi(\overline{\mathbf{x}})|^\theta}{\text{dist}(\mathbf{0}, \partial\psi(\mathbf{x}))}, \tag{3.8}$$

is bounded around $\overline{\mathbf{x}}$ under the following notational conventions: $0^0 = 1$, $\infty/\infty = 0/0 = 0$.

In other words, in a certain neighborhood $\mathcal{U}$ of $\overline{\mathbf{x}}$, there exists $\phi(s) = cs^{1-\theta}$ for some $c > 0$ and $\theta \in [0, 1)$ such that the KL inequality holds:

$$\phi'(|\psi(\mathbf{x}) - \psi(\overline{\mathbf{x}})|)\,\text{dist}(\mathbf{0}, \partial\psi(\mathbf{x})) \ge 1 \tag{3.9}$$

for any $\mathbf{x} \in \mathcal{U} \cap \text{dom}(\partial\psi)$ and $\psi(\mathbf{x}) \ne \psi(\overline{\mathbf{x}})$, where $\text{dom}(\partial\psi) \triangleq \{\mathbf{x} : \partial\psi(\mathbf{x}) \ne \emptyset\}$ and $\text{dist}(\mathbf{0}, \partial\psi(\mathbf{x})) \triangleq \min\{\|\mathbf{y}\| : \mathbf{y} \in \partial\psi(\mathbf{x})\}$.

Let $\mathcal{D}_n = \mathbb{R}^{I_n \times r^2}_+$ and $\delta_{\mathcal{D}_n}(\cdot)$ be the indicator function on $\mathcal{D}_n$ for $n = 1, 2, 3$. Then (3.1) is equivalent to

$$\min_{\mathcal{A},\mathcal{B},\mathcal{C}} F(\mathcal{A}, \mathcal{B}, \mathcal{C}) \equiv f(\mathcal{A}, \mathcal{B}, \mathcal{C}) + \delta_{\mathcal{D}_1}\left(A_{(1)}\right) + \delta_{\mathcal{D}_2}\left(B_{(2)}\right) + \delta_{\mathcal{D}_3}\left(D_{(3)}\right). \tag{3.10}$$

Obviously, F is continuous in $\text{dom}(F)$ and $\inf F > -\infty$. ∇f^k_i is Lipschitz continuous, and there exist constants $0 < \ell_i \le L_i < \infty, i = 1, 2, 3$ for parameters L^{k-1}_i obey $\ell_i \le L^{k-1}_i \le L_i$

and

$$f_1^k\left(A_{(1)}^k\right) \le f_1^k\left(\hat{A}_{(1)}^{k-1}\right) + \left\langle \hat{G}_1^k, A_{(1)}^k - \hat{A}_{(1)}^{k-1}\right\rangle + \frac{L_1^{k-1}}{2}\left\|A_{(1)}^k - \hat{A}_{(1)}^{k-1}\right\|^2,$$

$$f_2^k\left(B_{(2)}^k\right) \le f_2^k\left(\hat{B}_{(2)}^{k-1}\right) + \left\langle \hat{G}_2^k, B_{(2)}^k - \hat{B}_{(2)}^{k-1}\right\rangle + \frac{L_2^{k-1}}{2}\left\|B_{(2)}^k - \hat{B}_{(2)}^{k-1}\right\|^2,$$

$$f_3^k\left(C_{(3)}^k\right) \le f_3^k\left(\hat{C}_{(3)}^{k-1}\right) + \left\langle \hat{G}_3^k, C_{(3)}^k - \hat{C}_{(3)}^{k-1}\right\rangle + \frac{L_3^{k-1}}{2}\left\|C_{(3)}^k - \hat{C}_{(3)}^{k-1}\right\|^2.$$

So, F satisfies the Assumption 1 and 2 of [22], the same with [22] we also get the following conclusions.

Lemma 3.2. *Let* $\left\{A_{(1)}^k, B_{(2)}^k, C_{(3)}^k\right\}$ *be the sequence generated by Algorithm* 3.1 *with*

$$0 \le \omega_i^{k-1} \le \delta_\omega \sqrt{L_i^{k-2}/L_i^{k-1}} \text{ for } \delta_\omega < 1.$$

Then

$$\lim_{k\to\infty}\left\|A_{(1)}^k - A_{(1)}^{k+1}\right\|^2 = 0,\ \ \lim_{k\to\infty}\left\|B_{(2)}^k - B_{(2)}^{k+1}\right\|^2 = 0,\ \ \lim_{k\to\infty}\left\|C_{(3)}^k - C_{(3)}^{k+1}\right\|^2 = 0.$$

.

Proof. Let $F_1^k = f_1^k + \delta_{\mathcal{D}_1}(A_{(1)})$. By Lemma 2.1 [22], we know that

$$\begin{aligned} F_1^k\left(A_{(1)}^{k-1}\right) - F_1^k\left(A_{(1)}^k\right) &\ge \frac{L_1^{k-1}}{2}\left\|\hat{A}_{(1)}^{k-1} - A_{(1)}^k\right\|^2 + L_1^{k-1}\left\langle \hat{A}_{(1)}^{k-1} - A_{(1)}^{k-1}, A_{(1)}^k - \hat{A}_{(1)}^{k-1}\right\rangle \\ &\ge \frac{L_1^{k-1}}{2}\left\|A_{(1)}^{k-1} - A_{(1)}^k\right\|^2 - \frac{L_1^{k-2}}{2}\delta_\omega^2\left\|A_{(1)}^{k-2} - A_{(1)}^{k-1}\right\|^2. \end{aligned} \tag{3.11}$$

Similarly, we can get that

$$F_2^k\left(B_{(2)}^{k-1}\right) - F_2^k\left(B_{(2)}^k\right) \ge \frac{L_2^{k-1}}{2}\left\|B_{(2)}^{k-1} - B_{(2)}^k\right\|^2 - \frac{L_2^{k-2}}{2}\delta_\omega^2\left\|B_{(2)}^{k-2} - B_{(2)}^{k-1}\right\|^2, \tag{3.12}$$

$$F_3^k\left(C_{(3)}^{k-1}\right) - F_3^k\left(C_{(3)}^k\right) \ge \frac{L_3^{k-1}}{2}\left\|C_{(3)}^{k-1} - C_{(3)}^k\right\|^2 - \frac{L_3^{k-2}}{2}\delta_\omega^2\left\|C_{(3)}^{k-2} - C_{(3)}^{k-1}\right\|^2. \tag{3.13}$$

Therefore,

$$\begin{aligned} &F\left(A_{(1)}^{k-1}, B_{(2)}^{k-1}, C_{(3)}^{k-1}\right) - F\left(A_{(1)}^k, B_{(2)}^k, C_{(3)}^k\right) \\ &= \sum_{i=1}^3\left(F_i^{k-1}\left(A_{(1)}^{k-1}, B_{(2)}^{k-1}, C_{(3)}^{k-1}\right) - F_i^k\left(A_{(1)}^k, B_{(2)}^k, C_{(3)}^k\right)\right) \\ &\ge \frac{L_1^{k-1}}{2}\left\|A_{(1)}^{k-1} - A_{(1)}^k\right\|^2 - \frac{L_1^{k-2}}{2}\delta_\omega^2\left\|A_{(1)}^{k-2} - A_{(1)}^{k-1}\right\|^2 \\ &+ \frac{L_2^{k-1}}{2}\left\|B_{(2)}^{k-1} - B_{(2)}^k\right\|^2 - \frac{L_2^{k-2}}{2}\delta_\omega^2\left\|B_{(2)}^{k-2} - B_{(2)}^{k-1}\right\|^2 \\ &+ \frac{L_3^{k-1}}{2}\left\|C_{(3)}^{k-1} - C_{(3)}^k\right\|^2 - \frac{L_3^{k-2}}{2}\delta_\omega^2\left\|C_{(3)}^{k-2} - C_{(3)}^{k-1}\right\|^2. \end{aligned}$$

Summing the above inequality over k from 1 to K, we have

$$
\begin{aligned}
&F(A^0_{(1)}, B^0_{(2)}, C^0_{(3)}) - F(A^K_{(1)}, B^K_{(2)}, C^K_{(3)}) \\
&\geq \sum_{k=1}^{K}(\frac{L_1^{k-1}}{2}\left\|A^{k-1}_{(1)} - A^k_{(1)}\right\|^2 - \frac{L_1^{k-2}}{2}\delta_\omega^2\left\|A^{k-2}_{(1)} - A^{k-1}_{(1)}\right\|^2 \\
&+ \frac{L_2^{k-1}}{2}\left\|B^{k-1}_{(2)} - B^k_{(2)}\right\|^2 - \frac{L_2^{k-2}}{2}\delta_\omega^2\left\|B^{k-2}_{(2)} - B^{k-1}_{(2)}\right\|^2 \\
&+ \frac{L_3^{k-1}}{2}\left\|C^{k-1}_{(3)} - C^k_{(3)}\right\|^2 - \frac{L_3^{k-2}}{2}\delta_\omega^2\left\|C^{k-2}_{(3)} - C^{k-1}_{(3)}\right\|^2) \\
&\geq \sum_{k=1}^{K}(\frac{\left(1-\delta_\omega^2\right)L_1^{k-1}}{2}\left\|A^{k-1}_{(1)} - A^k_{(1)}\right\|^2 + \frac{\left(1-\delta_\omega^2\right)L_2^{k-1}}{2}\left\|B^{k-1}_{(2)} - B^k_{(2)}\right\|^2 \\
&+ \frac{\left(1-\delta_\omega^2\right)L_3^{k-1}}{2}\left\|C^{k-1}_{(3)} - C^k_{(3)}\right\|^2).
\end{aligned}
$$

Since F is lower bounded, taking $K \to \infty$ completes the proof. □

Theorem 3.3. *Let $\left\{\mathcal{A}^k, \mathcal{B}^k, \mathcal{C}^k\right\}$ be the sequence generated by Algorithm* 3.1*. Assume that $\left\{\mathcal{A}^k, \mathcal{B}^k, \mathcal{C}^k\right\}$ is bounded and there is a positive constant ℓ such that $\ell \leq L_n^k$ for all k and n. Then $\left\{\mathcal{A}^k, \mathcal{B}^k, \mathcal{C}^k\right\}$ converges to a critical point $\{\overline{\mathcal{A}}, \overline{\mathcal{B}}, \overline{\mathcal{C}}\}$.*

Proof. Obviously, ∇f is Lipschitz continuous on any bounded set. According to [22], we can get that $F(\mathcal{A}, \mathcal{B}, \mathcal{C})$ is a semialgebraic function and satisfies the KL inequality at $\{\overline{\mathcal{A}}, \overline{\mathcal{B}}, \overline{\mathcal{C}}\}$.

If $F\left(\mathcal{A}^{k_0}, \mathcal{B}^{k_0}, \mathcal{C}^{k_0}\right) = F(\overline{\mathcal{A}}, \overline{\mathcal{B}}, \overline{\mathcal{C}})$ at some k_0, then $\{\mathcal{A}^k, \mathcal{B}^k, \mathcal{C}^k\} = \{\mathcal{A}^{k_0}, \mathcal{B}^{k_0}, \mathcal{C}^{k_0}\} = \{\overline{\mathcal{A}}, \overline{\mathcal{B}}, \overline{\mathcal{C}}\}$ for all $k \geq k_0$. It remains to consider $F\left(\mathcal{A}^k, \mathcal{B}^k, \mathcal{C}^k\right) > F(\overline{\mathcal{A}}, \overline{\mathcal{B}}, \overline{\mathcal{C}})$ for all $k \geq 0$. Since $\{\overline{\mathcal{A}}, \overline{\mathcal{B}}, \overline{\mathcal{C}}\}$ is a limit point and $F\left(\mathcal{A}^k, \mathcal{B}^k, \mathcal{C}^k\right) \to F(\overline{\mathcal{A}}, \overline{\mathcal{B}}, \overline{\mathcal{C}})$, there must exist an integer k_0 such that $\{\mathcal{A}^{k_0}, \mathcal{B}^{k_0}, \mathcal{C}^{k_0}\}$ is sufficiently close to $\{\overline{\mathcal{A}}, \overline{\mathcal{B}}, \overline{\mathcal{C}}\}$ as required in Lemma 2.6 [22]. Hence, the entire sequence $\left\{\mathcal{A}^k, \mathcal{B}^k, \mathcal{C}^k\right\}$ converges according to Lemma 2.6 [22]. Since $\{\overline{\mathcal{A}}, \overline{\mathcal{B}}, \overline{\mathcal{C}}\}$ is a limit point of $\{\mathcal{A}^k, \mathcal{B}^k, \mathcal{C}^k\}$, we have $\{\mathcal{A}^k, \mathcal{B}^k, \mathcal{C}^k\} \to \{\overline{\mathcal{A}}, \overline{\mathcal{B}}, \overline{\mathcal{C}}\}$. □

Similar with the proof of [22], we obtain the convergence rate of Algorithm 3.1. For the sake of completeness, we give the proof accordingly.

Theorem 3.4. *(Convergence rate). Let $\left\{\mathcal{A}^k, \mathcal{B}^k, \mathcal{C}^k\right\}$ be the sequence generated by Algorithm* 3.1 *and converges to a critical point $\{\overline{\mathcal{A}}, \overline{\mathcal{B}}, \overline{\mathcal{C}}\}$. For the convenience, let $\mathcal{X}^k = \left\{\mathcal{A}^k, \mathcal{B}^k, \mathcal{C}^k\right\}$. Then the following hold:*
1. If $\theta = 0, \mathcal{X}$ converges to $\overline{\mathcal{X}}$ in finitely many iterations.
2. If $\theta \in \left(0, \frac{1}{2}\right], \left\|\mathcal{X}^k - \overline{\mathcal{X}}\right\| \leq C\tau^k$ for all $k \geq k_0$, for certain $k_0 > 0,\ C > 0, \tau \in [0, 1)$.
3. If $\theta \in \left(\frac{1}{2}, 1\right), \left\|\mathcal{X}^k - \overline{\mathcal{X}}\right\| \leq Ck^{-(1-\theta)/(2\theta-1)}$ for all $k \geq k_0$, for certain $k_0 > 0, C > 0$.

Proof. If $\theta = 0$, we must have $F\left(\mathcal{X}^{k_0}\right) = F(\overline{\mathcal{X}})$ for some k_0. Otherwise, $F\left(\mathcal{X}^k\right) > F(\overline{\mathcal{X}})$ for all sufficiently large k. The KL inequality gives $c \cdot \operatorname{dist}\left(\mathbf{0}, \partial F\left(\mathcal{X}^k\right)\right) \geq 1$ for all $k \geq 0$, which is impossible since $\mathcal{X}^k \to \overline{\mathcal{X}}$ and $\mathbf{0} \in \partial F(\overline{\mathcal{X}})$. The finite convergence now follows from the fact that $F\left(\mathcal{X}^{k_0}\right) = F(\overline{\mathcal{X}})$ implies $\mathcal{X}^k = \mathcal{X}^{k_0} = \overline{\mathcal{X}}$ for all $k \geq k_0$.

For $\theta \in (0, 1)$, we assume $F\left(\mathcal{X}^k\right) > F(\overline{\mathcal{X}}) = 0$ and define $S_k = \sum_{i=k}^{\infty}\left\|\mathcal{X}^i - \mathcal{X}^{i+1}\right\|$. Then according to [22] we can get

$$
S_k \leq C_1\phi\left(F_k\right) + \left(\frac{3\delta_\omega}{1-\delta_\omega}\sqrt{\frac{L}{\ell}} + 2\right)\left(S_{k-2} - S_k\right) \quad \text{for } k \geq 2, \tag{3.14}
$$

where $\ell = \min_i \ell_i$, $L = \max_i L_i$ and $C_1 = \frac{9(L+sL_G)}{2\ell(1-\delta_\omega)^2}$. Since $S_{k-2} - S_{k-1} \geq 0$. Using $\phi(s) = cs^{1-\theta}$, we have from $(A.8)$ [22] for sufficiently large k that

$$c(1-\theta)\left(F_k\right)^{-\theta} \geq \left(L+sL_G\right)^{-1}\left(\left\|\mathcal{X}^k - \mathcal{X}^{k-1}\right\| + \left\|\mathcal{X}^{k-1} - \mathcal{X}^{k-2}\right\|\right)^{-1}, \tag{3.15}$$

or, equivalently, $\left(F_k\right)^{\theta} \leq c(1-\theta)\left(L+sL_G\right)\left(S_{k-2}-S_k\right)$. Then,

$$\phi\left(F_k\right) = c\left(F_k\right)^{1-\theta} \leq c\left(c(1-\theta)\left(L+sL_G\right)\left(S_{k-2}-S_k\right)\right)^{\frac{1-\theta}{\theta}}. \tag{3.16}$$

Letting $C_2 = C_1 c\left(c(1-\theta)\left(L+sL_G\right)\right)^{\frac{1-\theta}{\theta}}$ and $C_3 = \frac{3\delta_\omega}{1-\delta_\omega}\sqrt{\frac{L}{\ell}} + 2$, we have from (3.14) and (3.16) that

$$S_k \leq C_2\left(S_{k-2}-S_k\right)^{\frac{1-\theta}{\theta}} + C_3\left(S_{k-2}-S_k\right). \tag{3.17}$$

When $\theta \in \left(0, \frac{1}{2}\right]$, i.e., $\frac{1-\theta}{\theta} \geq 1$, (3.17) implies that $S_k \leq \left(C_2 + C_3\right)\left(S_{k-2}-S_k\right)$ for sufficiently large k since $S_{k-2} - S_k \to 0$, and thus $S_k \leq \frac{C_2+C_3}{1+C_2+C_3} S_{k-2}$. Note that $\left\|\mathcal{X}^k - \overline{\mathcal{X}}\right\| \leq S_k$. Therefore, item 2 holds with $\tau = \sqrt{\frac{C_2+C_3}{1+C_2+C_3}} < 1$ and sufficiently large C.

When $\theta \in \left(\frac{1}{2}, 1\right)$, i.e., $\frac{1-\theta}{\theta} < 1$. Since $S_k \to 0$ as $k \to \infty$, we deduce from (3.17) that there exist an integer $N_1 \geq N_0$ and a positive constant C_4 such that

$$S_k^{\frac{\theta}{1-\theta}} \leq C_4\left(S_{k-2}-S_k\right), \tag{3.18}$$

for all $k \geq N_1$. Define $h : (0, +\infty) \to \mathbb{R}$ by $h(s) = s^{-\frac{\theta}{1-\theta}}$ and let $R \in (1, +\infty)$. Take $k \geq N_1$ and assume first that $h\left(S_k\right) \leq Rh\left(S_{k-2}\right)$. By rewriting (3.18) as

$$1 \leq \frac{C_4\left(S_{k-2}-S_k\right)}{S_k^{\frac{\theta}{1-\theta}}},$$

we obtain that

$$\begin{aligned} 1 &\leq C_4\left(S_{k-2}-S_k\right) h\left(S_k\right) \leq RC_4\left(S_{k-2}-S_k\right) h\left(S_{k-2}\right) \\ &\leq RC_4 \int_{S_k}^{S_{k-2}} h(s)\mathrm{d}s \leq RC_4 \frac{1-\theta}{1-2\theta}\left[S_{k-2}^{\frac{1-2\theta}{1-\theta}} - S_k^{\frac{1-2\theta}{1-\theta}}\right]. \end{aligned} \tag{3.19}$$

Thus if we set $\mu = \frac{2\theta-1}{(1-\theta)RC_4} > 0$ and $\nu = \frac{1-2\theta}{1-\theta} < 0$ one obtains that

$$S_k^\nu - S_{k-2}^\nu \geq \mu > 0, \quad S_k^\nu + S_{k-1}^\nu - S_{k-1}^\nu - S_{k-2}^\nu \geq \mu > 0. \tag{3.20}$$

Assume now that $h\left(S_k\right) > Rh\left(S_{k-2}\right)$ and set $q = \left(\frac{1}{R}\right)^{\frac{1-\theta}{\theta}} \in (0,1)$. It follows immediately that $S_k \leq qS_{k-2}$ and furthermore recalling that ν is negative - we have

$$S_k^\nu \geq q^\nu S_{k-2}^\nu, \quad S_k^\nu - S_{k-2}^\nu \geq \left(q^\nu - 1\right) S_{k-2}^\nu.$$

Since $q^\nu - 1 > 0$ and $S_p \to 0^+$ as $p \to +\infty$, there exists $\bar{\mu} > 0$ such that $\left(q^\nu - 1\right) S_{p-1}^\nu > \bar{\mu}$ for all $p \geq N_1$. Therefore we obtain that

$$S_k^\nu + S_{k-1}^\nu - S_{k-1}^\nu - S_{k-2}^\nu \geq \bar{\mu}. \tag{3.21}$$

If we set $\hat{\mu} = \min\{\mu, \bar{\mu}\} > 0$, one can combine (3.20) and (3.21) to obtain that

$$S_k^\nu + S_{k-1}^\nu - S_{k-1}^\nu - S_{k-2}^\nu \geq \hat{\mu} > 0,$$

for all $k \geq N_1$. By summing those inequalities from N_1 to some N greater than N_1 we obtain that $S_N^\nu + S_{N-1}^\nu - S_{N_1+1}^\nu - S_{N_1}^\nu \geq \hat{\mu}\,(N - N_1)$ and

$$S_N \leq \left(\frac{1}{2}\left(S_{N_1+1}^\nu + S_{N_1}^\nu + \hat{\mu}(N-N_1)\right)\right)^{\frac{1}{\nu}} \leq CN^{-\frac{1-\theta}{2\theta-1}},$$

for sufficiently large C and N. This completes the proof. □

4 Non-negative Tensor Completion

In this section, we consider the non-negative tensor completion problem:

$$\min \left\{ \|\mathcal{P}_\Omega(\mathcal{ABC} - \mathcal{M})\|_F^2 : \begin{array}{l} \mathcal{A} \in \mathcal{R}^{n_1\times r\times r}, \mathcal{B} \in \mathcal{R}^{r\times n_2\times r}, \mathcal{C} \in \mathcal{R}^{r\times r\times n_3} \\ \mathcal{A} \geq 0, \mathcal{B} \geq 0, \mathcal{C} \geq 0 \end{array} \right\}. \tag{4.1}$$

To solve (4.1), we consider its equivalent form:

$$\begin{array}{ll} \min & G(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{Z}) = \|\mathcal{ABC} - \mathcal{Z}\|_F^2 \\ \text{s.t.} & \mathcal{A} \geq 0, \mathcal{B} \geq 0, \mathcal{C} \geq 0, \mathcal{P}_\Omega(\mathcal{Z} - \mathcal{M}) = 0, \end{array} \tag{4.2}$$

where Ω indexes the known entries of $\mathcal{M}$ and $\mathcal{P}_\Omega(\mathcal{A})$ returns a copy of $\mathcal{A}$ that zeros out the entries not in Ω. Our algorithm shall cycle through the decision variables $\mathcal{A}, \mathcal{B}, \mathcal{C}$ and $\mathcal{Z}$. It should be noted that Algorithm 3.1. is modified at k-th iteration such that $\mathcal{M} = \mathcal{Z}^{k-1}$ wherever $\mathcal{M}$ is referred to. Then, $\mathcal{Z}$ is updated as (3.10)

$$\mathcal{Z}^k = \mathcal{P}_\Omega(\mathcal{M}) + \mathcal{P}_{\Omega^c}\left(\mathcal{ABC}\right), \tag{4.3}$$

where Ω^c is the complement of Ω. Note that for a fixed $\mathcal{A}, \mathcal{B}, \mathcal{C}, G(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{Z})$ is a strongly convex function of $\mathcal{Z}$ with modulus 1. Hence, according to Theorem 3.3, the convergence result for Algorithm 3.1 still holds for this algorithm with extra update (4.3).

5 Numerical Examples

In this section, we investigate some data to show that they can be approximated by nonnegative triple decomposition of low triple rank very well.

Example 5.1. Let $\overline{\mathcal{X}} \in \mathbb{R}^{10\times 15\times 100}$ be a randomly generated nonnegative tensor. Set the corresponding triple rank $r = 15$. A noise term is added such that $\mathcal{X} = \overline{\mathcal{X}} + \lambda\varepsilon$, $\varepsilon = randn(10, 15, 100)$, where $\lambda = 1e^{-3}$ is the parameter to control the noise term. The experimental result is shown in Figure 1.

We compute the triple decomposition approximation $\mathcal{ABC}$ by Algorithm 3.1 and calculate the relative error of nonnegative low triple rank approximation

$$\text{RelativeError } = \frac{\|\mathcal{X} - \mathcal{ABC}\|_F}{\|\mathcal{X}\|_F}.$$

Figure 1 illustrates the relative error of the low triple rank approximation and the relative error is about 1.7%.

Example 5.2. Let $\overline{\mathcal{X}} \in \mathbb{R}^{N_1\times N_2\times N_3}$ be a randomly generated nonnegative tensor, where $N_1, N_2, N_3 \in \mathbb{N}$. The triple rank is considered in three different cases. A noise term is added such that $\mathcal{X} = \overline{\mathcal{X}} + \lambda\varepsilon$, $\varepsilon = randn(N_1, N_2, N_3)$, where $\lambda = 1e^{-3}$ is the parameter to control the noise term. Experimental results are shown in Table 1.

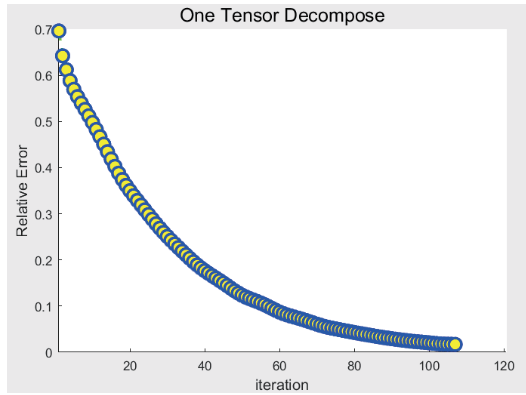

Figure 1: Relative error of low nonnegative triple rank approximation of the $\mathcal{X} \in \mathbb{R}^{10\times 15\times 100}$

Table 1: The numerical results of the problem in Example 5.1

N_1	N_2	N_3	r	relerr	Time	IT	$error$
10	20	100	10	2.2%	3.42	250	2.5307
			15	1%	42.4	1000	0.6675
			20	0.71%	113	968	0.3766
15	30	100	15	1.03%	54.1	1000	1.4881
			25	0.62%	392	1000	0.6311
			30	0.60%	100	733	0.7422
200	200	200	10	0.60%	70	136	61.3114
			15	0.45%	312	282	46.1830
			20	0.36%	977	449	34.0303

In Table 1, "relerr" denotes the relative error, "IT" denotes the number of iterations, "Time" denotes the CPU time in seconds and $error = \|\mathcal{X} - \mathcal{ABC}\|_F$. This result shows clearly the three order nonnegative tensor can be approximated by low rank triple decomposition very well and we can see that the rank increases, the relative error of the tensor by this method decreases.

Example 5.3. In this test, we randomly generate a nonnegative tensor $\overline{\mathcal{X}} \in \mathbb{R}^{10\times 30\times 100}$, and set three different triple rank of the tensor respectively. In this experiment, we compare the test results with and without non-negative constraints. Experimental results are shown in Table 2.

Table 2: The numerical results of the problem in Example 5.3

	r	relerr	Time	IT	$error$
constrained	10	2.2%	8.32	501	3.8206
	15	1.56%	32.1	613	2.4464
	30	1.37%	92.6	985	2.0827
unconstrained	10	1.97%	17.47	1000	3.0881
	15	1.02%	46.46	904	1.0352
	30	0.71%	100	738	0.6121

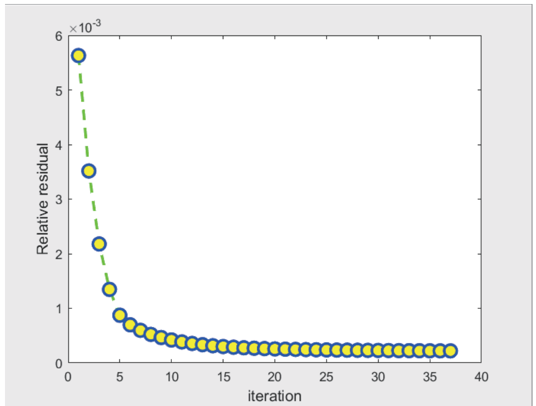

Figure 2: Relative error of the tensor recovery of $\mathcal{X} \in \mathbb{R}^{200\times200\times200}$

In the theory point of view, its apparent that the rank of the nonnegative triple decomposition for a third order nonnegative tensor is greater than or equal to the rank associated to the unconstrained triple decomposition. From Table 2, for given fixed rank, the error and relative error for the unconstraint case are less than the nonnegative constraint case.

Example 5.4. Let $\mathcal{X} \in \mathbb{R}^{200\times200\times200}$ be a randomly generated nonnegative tensor. Set the corresponding triple rank $r = 15$. We sample fifty percent of elements of this tensor. Experimental result is shown in Figure 2.

Figure 2 illustrates the relative error of the tensor recovery via nonnegative triple decomposition and the relative error is about 0.02202%.

Example 5.5. Let $\mathcal{X} \in \mathbb{R}^{N_1\times N_2\times N_3}$ be a randomly generated nonnegative tensor, where $N_1, N_2, N_3 \in \mathbb{N}$. The sampling rate is considered in three different cases. Experimental results are shown in Table 3.

Table 3: The numerical results of the problem in Example 5.5

N_1	N_2	N_3	r	sr	relerr	Time	IT
10	20	100	20	0.3	6.9309×10^{-4}	11.7	94
				0.5	5.6981×10^{-4}	10.5	90
				0.9	4.2789×10^{-4}	10.6	83
15	30	100	30	0.3	3.4973×10^{-4}	93.5	76
				0.5	2.9805×10^{-4}	82.9	68
				0.9	2.1418×10^{-4}	89.3	73
200	200	200	15	0.3	2.8398×10^{-4}	89.7	81
				0.5	2.3535×10^{-4}	65.7	59
				0.9	2.2021×10^{-4}	42.8	37

In Table 2, "sr" denotes sampling rate. As the sampling rate increases, the relative error decreases. Clearly, through the proposed method can achieve very good result.

6 Conclusions

In this paper, we introduced a nonnegative triple decomposition for third order nonnegative tensors, which decomposes a third order nonnegative tensor to three third order low rank nonnegative tensors in a balanced way. A nonnegative tensor completion method was proposed based on such low rank nonnegative triple decomposition. Furthermore, an alternating proximal gradient algorithm was provided and its convergence was also established. Numerical experiments confirmed the efficiency of the method.

Acknowledgments

We are very grateful to the editor and the anonymous reviewers for their constructive comments and meaningful suggestions on our manuscript.

References

[1] E. Acar, T.G. Kolda, D.M. Dunlavy and M. Morup, Scalable tensor factorizations for incomplete data, *Chemometrics and Intelligent Laboratory Systems* 106 (2010) 41–56.

[2] J. Carroll and J. Chang, Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-young" decomposition, *Psychometrika* 35 (1970) 283–319.

[3] H. Chen, Y. Chen, G. Li and L. Qi, A semidefinite program approach for computing the maximum eigenvalue of a class of structured tensors and its applications in hypergraphs and copositivity test, *Numerical Linear Algebra with applications* 25 (2018) e2125.

[4] S. Hassanzadeh and A. Karami, Compression and noise reduction of hyperspectral images using non-negative tensor decomposition and compressed sensing, *European Journal of Remote Sensing* 49 (2016) 587–598.

[5] M. Kilmer, K. Braman, N. Hao and R. Hoover, Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging, *SIAM Journal on Matrix Analysis and Applications* 34 (2013) 148–172.

[6] J. Kim, Y. He and H.Park, Algorithms for nonnegative matrix and tensor factorizations: a unified view based on block coordinate descent framework, *Journal of Global Optimization* 58 (2014) 285–319.

[7] T. Kolda and B. Bader, Tensor decompositions and applications, *SIAM Rev.* 51 (2009) 455–500.

[8] D. Lee and H.S. Seung, Learning the parts of objects by non-negative matrix factorization, *Nature* 401 (1999) 788–791.

[9] D.D. Lee and H.S. Seung, Algorithms for non-negative matrix factorization, in: International Conference on Neural Information Processing Systems MIT Press, 2000.

[10] L. Liang, H. Wen, F. Liu, G. Li and M. Li, Feature extraction of impulse faults for vibration signals based on sparse non-negative tensor factorization, *Applied Sciences* 9 (2019) 3642.

[11] H. Liu, Z. Wu, D. Cai and T.S.Huang, Constrained nonnegative matrix factorization for image representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34 (2012) 1299–1311.

[12] P. Paatero and U. Tapper, Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, *environmetrics* 5 (1994) 111–126.

[13] L. Qi, Y. Chen, M. Bakshi and X. Zhang, Triple decomposition and tensor recovery of third order tensors, *SIAM Journal on Matrix Analysis and Applications* 42 (2021) 299–329.

[14] A. Shashua and A. Levin, Linear image coding for regression and classification using the tensor-rank principle, Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on IEEE, 2001.

[15] J. Shen and G.W. Israël, A receptor model using a specific non-negative transformation technique for ambient aerosol, *Atmospheric Environment* 23 (1989) 2289–2298.

[16] Q. Shi, H. Sun, S. Lu and M. Hong, Inexact block coordinate descent methods for symmetric nonnegative matrix factorization, *IEEE Transactions on Signal Processing* 65 (2017) 5995–6008.

[17] H. Tan, Z. Yang, G. Feng, W. Wang and B. Ran, Correlation analysis for tensorbased traffic data imputation method, *Procedia-Social and Behavioral Sciences* 96 (2013) 2611–2620.

[18] L. Tucker, Some mathematical notes on three-mode factor analysis, *Psychometrika* 31 (1966) 279–311.

[19] V.V. Vesselinov, M.K. Mudunuru, S. Karra, D. O'Malley and B.S. Alexandrov, Unsupervised machine learning based on non-negative tensor factorization for analyzing reactive-mixing, *Journal of Computational Physics* 395 (2019) 85–104.

[20] X. Wang, W. Zhang, J. Yan and X. Yuan, On the flexibility of block coordinate descent for large-scale optimization, *Neurocomputing* 272 (2018) 471–480.

[21] K. Xie, L. Wang, X. Wang, G. Xie, J. Wen and G. Zhang, Accurate recovery of internet traffic data: A tensor completion approach, IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications, 2016.

[22] Y. Xu and W. Yin, A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion, *SIAM Journal on Imaging Sciences* 6 (2015) 1758–1789.

[23] L. Yang, Z. Huang, S. Hu and J. Han, An iterative algorithm for third-order tensor multi-rank minimization, *Computational Optimization and Applications* 63 (2016) 169–202.

[24] J. Zhang, A.K. Saibaba, M.E. Kilmer and S. Aeron, A randomized tensor singular value decomposition based on the t-product, *Numerical Linear Algebra with Applications* 25 (2018) e2179.

[25] Z. Zhang and S. Aeron, Exact tensor completion using t-SVD, *IEEE Transactions on Signal Processing* 65 (2017) 1511–1526.

[26] Z. Zhang, G. Ely, S. Aeron, N. Hao and M. Kilmer, Novel methods for multilinear data completion and de-noising based on tensor-SVD, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, ser. CVPR'14*, 2014, pp. 3842–3849.

[27] H. Zhou, D. Zhang, K. Xie and Y. Chen, Spatio-temporal tensor completion for imputing missing internet traffic data, in: *2015 IEEE 34th International Performance Computing and Communications Conference (IPCCC)*, 2015.

[28] P. Zhou, C. Lu, Z. Lin, C. Zhang, Tensor factorization for low-rank tensor completion, *IEEE Transactions on Image Processing* 27 (2018) 1152–1163.

Manuscript received 16 March 2021
revised 19 May 2021
accepted for publication 20 May 2021

Chunyan Wang
School of Management Science, Qufu Normal University
Address: No. 80, Yantai Road, Donggang District, Rizhao Shandong
276800, P.R. China
E-mail address: wchunyanwang562@163.com

Min Wang
School of Management Science, Qufu Normal University
Address: No. 80, Yantai Road, Donggang District
Rizhao Shandong, 276800, P.R. China
E-mail address: 782091135@qq.com

Haibin Chen
School of Management Science, Qufu Normal University
Address: No. 80, Yantai Road, Donggang District
Rizhao Shandong, 276800, P.R. China
E-mail address: chenhaibin508@163.com

Yiju Wang
School of Management Science, Qufu Normal University
Address: No. 80, Yantai Road, Donggang District
Rizhao Shandong, 276800, P.R. China
E-mail address: wyijumail@163.com