The generalized inverses of tensors and an application to linear models

Hongwei Jin* Minru Bai† Julio Benítez‡ Xiaoji Liu§

Abstract

In this paper, we recall and extend some tensor operations. Then, the generalized inverse of tensors is established by using tensor equations. Moreover, we investigate the least-squares solutions of tensor equations. An algorithm to compute the Moore-Penrose inverse of an arbitrary tensor is constructed. Finally, we apply the obtained results to higher order Gauss-Markov theorem.

Keywords: Tensors; Generalized inverses; Moore-Penrose inverse of tensors; linear models

AMS classification: 15A18, 15A69, 62J12, 65F99.

1 Introduction

It is a well know definition that the Moore-Penrose inverse (see e.g. [1]) of a matrix $A \in \mathbb{C}^{m\times n}$ is a matrix $X \in \mathbb{C}^{n\times m}$ which satisfies

$$(1)\ AXA = A \qquad (2)\ XAX = X \qquad (3)\ (AX)^* = AX \qquad (4)\ (XA)^* = XA.$$

The Moore-Penrose inverse of A is unique and it is denoted by $A^\dagger$.

The Moore-Penrose inverse plays an important role in theoretic research and numerical computations in many areas, including singular matrix problems, ill-posed problems, optimization problems, and statistics problems [1, 2, 3, 4, 5, 6, 7, 8].

Operations with tensors, or multiway arrays, have become increasingly prevalent in recent years. A tensor can be regarded as a multidimensional array of data [9], which takes the form

$$\mathcal{A} = (a_{i_1\ldots i_m}) \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_p}. \tag{1.1}$$

*College of Mathematics and Econometrics, Hunan University, 410082, Changsha, P.R. China and Departamento de Matemática Aplicada, Instituto de Matemática Multidisciplinar, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia, Spain. *E-mail address*: hw-jin@hotmail.com.

†College of Mathematics and Econometrics, Hunan University, 410082, Changsha, P.R. China. *E-mail address*: minru-bai@163.com.

‡Corresponding author. Departamento de Matemática Aplicada, Instituto de Matemática Multidisciplinar, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia, Spain. *E-mail address*: jbenitez@mat.upv.es.

§Faculty of Science, Guangxi University for Nationalities, 530006, Nanning, P.R. China. *E-mail address*: xiaojiliu72@126.com.

The ***order*** of a tensor is the number of dimensions. For the tensor $\mathcal{A}$ given in (1.1), its order is p. The dimensions of a tensor also are known as ***ways*** or ***modes***.

It was discovered that some important theoretical and practical problems of higher order tensors are NP-hard [10]. So, it is natural to transform tensors to other simpler objects. Traditionally, the well-known representations of tensors are the CANDECOMP/PARAFAC (CP) [11, 12] and Tucker models [13]. CANDECOMP/PARAFAC (CP) decomposes a tensor as a sum of rank-one tensors, and the Tucker decomposition is a higher-order form of principal component analysis. Each model can be considered an extension of the singular value decomposition (SVD) for matrices. Kolda et al. [14] introduced the *fibers* and *slices* of tensors, which permits a better understanding of third-order tensors. Kilmer et al. [15] explore an alternate representation based on matrix slices and the functions unfold$(\cdot)$ and fold$(\cdot)$ on the third-order tensor, which permits to define several concepts (tensor transpose, inverse, and identity, especially the multiplication of tensors). The multiplication of tensors is a framework for tensor operations, which also leads to the notion of orthogonal tensors, norm of a tensor and factorizations of tensors. Later, Kilmer et al. [16] extended these results in [15] to p order tensors and concluded with two applications. The first application is image deblurring, and the second one is video facial recognition.

Now, a question is natural. Can we extend the Moore-Penrose inverse of matrices to tensors? By using the definitions given in [15, 16, 17] we will see that the answer to the aforementioned question is "yes".

In fact, this work is inspired by the papers [16] and [18]. In [18], the authors proposed an image restoration method, which generalizes image restoration algorithms that are based on the Moore-Penrose solution of certain matrix equations. The approach presented in [18] is based on the usage of least-squares solutions of these matrix equations, wherein an arbitrary matrix of appropriate dimensions is included besides the Moore-Penrose inverse. It is nature to define the Moore-Penrose inverse of higher order tensors by using the t-product constructed in the work [16] and establish the least-squares solutions of tensors in order to tackle the difficult 3-D image deblurring problem.

This work is organized as follows. In Section 2, we provide some preliminaries. We introduce the t-product of two tensors firstly. Then, we show the definitions of the identity tensor, the orthogonal tensor, the symmetric tensor, the f-diagonal tensor and the inverse, the transpose, the Frobenius norm of a tensor. Examples are also given to illustrate these definitions.

In Section 3, we define the Moore-Penrose inverse of the tensors. Then, we prove that the Moore-Penrose inverse of an arbitrary tensor $\mathcal{A}$ exists and is unique by using the technique of fast Fourier transform. Then, we present some properties of the Moore-Penrose inverse of tensors and establish some representations of $\{1\}$-inverses, $\{1,3\}$-inverses and $\{1,4\}$-inverses of tensors.

In Section 4, we study the tensor equations. We give the least-squares solutions of an inconsistent tensor equation, the minimum-norm solution of a consistent tensor equation and the minimum-norm least-squares solution of an arbitrary tensor equation. Furthermore, the relations of the least-squares solutions with $\{1,3\}$-inverses of $\mathcal{A}$, the minimum-norm solutions with $\{1,4\}$-inverses of $\mathcal{A}$ and the minimum-norm least-squares solution of the Moore-Penrose inverse of $\mathcal{A}$ are established.

In Section 5, we construct an algorithm to compute the Moore-Penrose inverse of an arbitrary tensor. Supplementary example is given to test the algorithm.

In Section 6, we derive an application to linear models. We define the random tensor, the expectation and covariance tensor of a random tensor, and then establish the linear model for tensors. In addition, we show how the Moore-Penrose inverse of tensors works for the higher order Gauss-Markov theorem.

2 Preliminaries

Throughout this paper tensors are denoted by Euler script letters (e.g., $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$,...), while capital letters represent matrices, boldface lowercase letters represent vectors, and lowercase letters refer to scalars.

Let $\mathbf{c} \in \mathbb{R}^n$. Recall that if $\mathbf{c} = \begin{bmatrix} c_1 & c_2 & ... & c_n \end{bmatrix}^T$, then

$$\operatorname{circ}(\mathbf{c}) = \begin{bmatrix} c_1 & c_n & \cdots & c_2 \\ c_2 & c_1 & \cdots & c_3 \\ \vdots & \vdots & & \vdots \\ c_n & c_{n-1} & \cdots & c_1 \end{bmatrix}$$

is a circulant matrix. Similarly, if $C_1, ..., C_n$ are $n_1 \times n_2$ real matrices, then

$$\operatorname{circ}(C_1, ..., C_n) = \begin{bmatrix} C_1 & C_n & \cdots & C_2 \\ C_2 & C_1 & \cdots & C_3 \\ \vdots & \vdots & & \vdots \\ C_n & C_{n-1} & \cdots & C_1 \end{bmatrix}.$$

Let $\mathcal{A} = (a_{i_1 \dots i_m}) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$, $p > 1$. For $i = 1, ..., n_p$, denote by $\mathcal{A}_i \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_{p-1}}$, the tensor whose order is $(p-1)$ and is created by holding the pth index of $\mathcal{A}$ fixed at i. For example, let $\mathcal{A}$ be a $2 \times 2 \times 2 \times 3$ tensor. Fixing the 4th index of $\mathcal{A}$. One can get three $2 \times 2 \times 2$ tensors, which are $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$ and with elements

$$\begin{array}{lllllllll} \mathcal{A}_1: & a_{1111} & a_{1211} & a_{2111} & a_{2211} & a_{1121} & a_{1221} & a_{2121} & a_{2221}, \\ \mathcal{A}_2: & a_{1112} & a_{1212} & a_{2112} & a_{2212} & a_{1122} & a_{1222} & a_{2122} & a_{2222}, \\ \mathcal{A}_3: & a_{1113} & a_{1213} & a_{2113} & a_{2213} & a_{1123} & a_{1223} & a_{2123} & a_{2223}, \end{array}$$

respectively.

Define unfold$(\cdot)$ to take an $n_1 \times n_2 \times \cdots \times n_p$ tensor and return an $n_1 n_p \times n_2 \times \cdots \times n_{p-1}$ block tensor in the following way:

$$\operatorname{unfold}(\mathcal{A}) = \begin{bmatrix} \mathcal{A}_1 \\ \mathcal{A}_2 \\ \vdots \\ \mathcal{A}_{n_p} \end{bmatrix}$$

and fold$(\cdot)$ is the inverse operation, which takes an $n_1 n_p \times n_2 \times \cdots \times n_{p-1}$ block tensor and returns an $n_1 \times n_2 \times \cdots \times n_p$ tensor. Then,

$$\operatorname{fold}(\operatorname{unfold}(\mathcal{A})) = \mathcal{A}.$$

Now, one can create a tensor in a block circulant pattern, where each block is a tensor whose order is $(p-1)$:

$$\text{circ}(\text{unfold}(\mathcal{A})) = \begin{bmatrix} \mathcal{A}_1 & \mathcal{A}_{n_p} & \mathcal{A}_{n_p-1} & \cdots & \mathcal{A}_2 \\ \mathcal{A}_2 & \mathcal{A}_1 & \mathcal{A}_{n_p} & \cdots & \mathcal{A}_3 \\ \vdots & \vdots & \vdots & & \vdots \\ \mathcal{A}_{n_p} & \mathcal{A}_{n_p-1} & \mathcal{A}_{n_p-2} & \cdots & \mathcal{A}_1 \end{bmatrix}, \tag{2.1}$$

which is an $n_1 n_p \times n_2 n_p \times \cdots \times n_{p-2} n_p \times n_{p-1}$ tensor.

The formula (2.1) allows us to define the t-product of two tensors.

Definition 2.1 *[16] Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$ and $\mathcal{B} \in \mathbb{R}^{n_2 \times l \times n_3 \times \cdots \times n_p}$. Then the t-**product** $\mathcal{A} * \mathcal{B}$ is the $n_1 \times l \times n_3 \times \cdots \times n_p$ order-p tensor $(p \geq 3)$ defined recursively as*

$$\mathcal{A} * \mathcal{B} = \text{fold}(\text{circ}(\text{unfold}(\mathcal{A})) * \text{unfold}(\mathcal{B})). \tag{2.2}$$

Notice that the right-hand side in (2.2) involves a t-product of order-$(p-1)$ tensors. Each successive t-product operation therefore involves tensors of one order less. The recursive multiplication structure eventually reduces to standard matrix multiplication of blocks of block circulant matrices.

Example 2.1. Let $\mathcal{A} \in \mathbb{R}^{3\times 3\times 2\times 2}$ and $\mathcal{B} \in \mathbb{R}^{3\times 3\times 2\times 2}$. Then,

$$\begin{aligned}
\mathcal{A} * \mathcal{B} &= \text{fold}\left(\begin{bmatrix} \mathcal{A}_1 & \mathcal{A}_2 \\ \mathcal{A}_2 & \mathcal{A}_1 \end{bmatrix} * \begin{bmatrix} \mathcal{B}_1 \\ \mathcal{B}_2 \end{bmatrix} \right) \\
&= \text{fold}\left(\begin{bmatrix} \mathcal{A}_1 * \mathcal{B}_1 + \mathcal{A}_2 * \mathcal{B}_2 \\ \mathcal{A}_2 * \mathcal{B}_1 + \mathcal{A}_1 * \mathcal{B}_2 \end{bmatrix} \right) \\
&= \text{fold}\left(\begin{bmatrix} \text{fold}\left(\begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{12} & \mathcal{A}_{11} \end{bmatrix} * \begin{bmatrix} \mathcal{B}_{11} \\ \mathcal{B}_{12} \end{bmatrix} \right) + \text{fold}\left(\begin{bmatrix} \mathcal{A}_{21} & \mathcal{A}_{22} \\ \mathcal{A}_{22} & \mathcal{A}_{21} \end{bmatrix} * \begin{bmatrix} \mathcal{B}_{21} \\ \mathcal{B}_{22} \end{bmatrix} \right) \\ \\ \text{fold}\left(\begin{bmatrix} \mathcal{A}_{21} & \mathcal{A}_{22} \\ \mathcal{A}_{22} & \mathcal{A}_{21} \end{bmatrix} * \begin{bmatrix} \mathcal{B}_{11} \\ \mathcal{B}_{12} \end{bmatrix} \right) + \text{fold}\left(\begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{12} & \mathcal{A}_{11} \end{bmatrix} * \begin{bmatrix} \mathcal{B}_{21} \\ \mathcal{B}_{22} \end{bmatrix} \right) \end{bmatrix} \right) \\
&= \text{fold}\left(\begin{bmatrix} \text{fold}\left(\begin{bmatrix} \mathcal{A}_{11} * \mathcal{B}_{11} + \mathcal{A}_{12} * \mathcal{B}_{12} \\ \mathcal{A}_{12} * \mathcal{B}_{11} + \mathcal{A}_{11} * \mathcal{B}_{12} \end{bmatrix} \right) + \text{fold}\left(\begin{bmatrix} \mathcal{A}_{21} * \mathcal{B}_{21} + \mathcal{A}_{22} * \mathcal{B}_{22} \\ \mathcal{A}_{22} * \mathcal{B}_{21} + \mathcal{A}_{21} * \mathcal{B}_{22} \end{bmatrix} \right) \\ \\ \text{fold}\left(\begin{bmatrix} \mathcal{A}_{21} * \mathcal{B}_{11} + \mathcal{A}_{22} * \mathcal{B}_{12} \\ \mathcal{A}_{22} * \mathcal{B}_{11} + \mathcal{A}_{21} * \mathcal{B}_{12} \end{bmatrix} \right) + \text{fold}\left(\begin{bmatrix} \mathcal{A}_{11} * \mathcal{B}_{21} + \mathcal{A}_{12} * \mathcal{B}_{22} \\ \mathcal{A}_{12} * \mathcal{B}_{21} + \mathcal{A}_{11} * \mathcal{B}_{22} \end{bmatrix} \right) \end{bmatrix} \right).
\end{aligned}$$

Obviously, the t-product of $\mathcal{A}$ and $\mathcal{B}$ eventually reduces to some 3×3 matrix multiplications as one can see in the last equality. □

How to compute this new product? Martin et. al., [16] also gave the answer based on the well known fact that block circulant matrices can be block diagonalized by using the Fourier transform. See [16, Algorithm T-MULT] for details.

It is easy to check the following basic properties of the t-product.

Lemma 2.1 *If $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are tensors of adequate size, then the following statements are true:*

(a) (The left distributivity): $\mathcal{A} * (\mathcal{B} + \mathcal{C}) = \mathcal{A} * \mathcal{B} + \mathcal{A} * \mathcal{C}$;

(b) (The right distributivity): $(\mathcal{A} + \mathcal{B}) * \mathcal{C} = \mathcal{A} * \mathcal{C} + \mathcal{B} * \mathcal{C}$;

(c) *(The associativity):* $(\mathcal{A} * \mathcal{B}) * \mathcal{C} = \mathcal{A} * (\mathcal{B} * \mathcal{C})$.

Definition 2.2 *[16] The $n \times n \times n_3 \times \cdots \times n_p$ order-p $(p \geq 3)$* ***identity tensor*** *$\mathcal{I}$ is the tensor such that $\mathcal{I}_1$ is the $n \times n \times n_3 \times \cdots \times n_{p-1}$ order-$(p-1)$ identity tensor and $\mathcal{I}_j$, $j = 2, 3, ..., n_p$ is the $n \times n \times n_3 \times \cdots \times n_{p-1}$ order-$(p-1)$ zero tensor.*

Example 2.2. The $4 \times 4 \times 3 \times 2$ identity tensor $\mathcal{I}$ has the following form:

$$\mathcal{I} = \text{fold}\left(\begin{bmatrix} \mathcal{I}_1 \\ O_2 \end{bmatrix}\right) = \text{fold}\left(\begin{bmatrix} \text{fold}\left(\begin{bmatrix} \mathcal{I}_{11} \\ O_{12} \\ O_{13} \end{bmatrix}\right) \\ O_{22} \\ O_{23} \end{bmatrix}\right),$$

where $\mathcal{I}_{11}$ is the 4×4 identity matrix and O_{ij} is a tensor all of whose components are zero. □

The following result is easy to check.

Lemma 2.2 *[16] Let $\mathcal{I}$ be an $n \times n \times n_3 \times \cdots \times n_p$ order-p $(p \geq 3)$ identity tensor. Then,*

$$\mathcal{I} * \mathcal{A} = \mathcal{A} * \mathcal{I} = \mathcal{A}.$$

Definition 2.3 *[16] Let $\mathcal{A} \in \mathbb{R}^{n \times n \times n_3 \times \cdots \times n_p}$. If there exists an order-p $(p \geq 3)$ tensor $\mathcal{B} \in \mathbb{R}^{n \times n \times n_3 \times \cdots \times n_p}$ such that*

$$\mathcal{A} * \mathcal{B} = \mathcal{I} \quad \text{and} \quad \mathcal{B} * \mathcal{A} = \mathcal{I},$$

then $\mathcal{A}$ is said to be ***invertible***. *Moreover, $\mathcal{B}$ is the* ***inverse*** *of $\mathcal{A}$, which is denoted by $\mathcal{A}^{-1}$.*

In fact, the inverse of an invertible tensor is unique.

Lemma 2.3 *If $\mathcal{A} \in \mathbb{R}^{n \times n \times n_3 \times \cdots \times n_p}$ $(p \geq 3)$ is invertible, then its inverse tensor is unique.*

Similar as the transpose of real matrices, the transpose of tensors can be defined.

Definition 2.4 *[16] If $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$, then the* ***transpose*** *of $\mathcal{A}$, which is denoted by $\mathcal{A}^T$, is the $n_2 \times n_1 \times n_3 \times \cdots \times n_p$ tensor obtained by tensor transposing each $\mathcal{A}_i$, for $i = 1, 2, \ldots, n_p$ and then reversing the order of the $\mathcal{A}_i$ from 2 through n_p, i.e.,*

$$\mathcal{A}^T = \text{fold}\left(\begin{bmatrix} \mathcal{A}_1^T \\ \mathcal{A}_{n_p}^T \\ \mathcal{A}_{n_p-1}^T \\ \vdots \\ \mathcal{A}_2^T \end{bmatrix}\right). \tag{2.3}$$

Example 2.3. Let $\mathcal{A}$ be a $4 \times 4 \times 3 \times 3$ tensor. Then

$$\mathcal{A}^T = \text{fold}\left(\begin{bmatrix} \mathcal{A}_1^T \\ \mathcal{A}_3^T \\ \mathcal{A}_2^T \end{bmatrix}\right) = \text{fold}\left(\begin{bmatrix} \text{fold}\left(\begin{bmatrix} \mathcal{A}_{11}^T \\ \mathcal{A}_{13}^T \\ \mathcal{A}_{12}^T \end{bmatrix}\right) \\ \text{fold}\left(\begin{bmatrix} \mathcal{A}_{31}^T \\ \mathcal{A}_{33}^T \\ \mathcal{A}_{32}^T \end{bmatrix}\right) \\ \text{fold}\left(\begin{bmatrix} \mathcal{A}_{21}^T \\ \mathcal{A}_{23}^T \\ \mathcal{A}_{22}^T \end{bmatrix}\right) \end{bmatrix}\right). \tag{2.4}$$

So, the transposition of the order-4 tensor $\mathcal{A}$ eventually reduces to some 4×4 matrix transpositions as above. □

Lemma 2.4 *[16] Suppose that $\mathcal{A}$, $\mathcal{B}$ are two tensors such that $\mathcal{A} * \mathcal{B}$ and $\mathcal{B}^T * \mathcal{A}^T$ are defined. Then*

$$(\mathcal{A} * \mathcal{B})^T = \mathcal{B}^T * \mathcal{A}^T. \tag{2.5}$$

The following definitions are useful in establishing the main results.

Definition 2.5 *Let $\mathcal{A} \in \mathbb{R}^{n \times n \times n_3 \times \cdots \times n_p}$. We say that $\mathcal{A}$ is* ***symmetric*** *if $\mathcal{A}^T = \mathcal{A}$.*

Definition 2.6 *Let $\mathcal{A} \in \mathbb{R}^{n \times n \times n_3 \times \cdots \times n_p}$ be a symmetric tensor. If there exists a tensor $\mathcal{X} \in \mathbb{R}^{n \times 1 \times n_3 \times \cdots \times n_p}$ such that all the elements of the tensor $\mathcal{Z}^T \mathcal{A} \mathcal{Z}$ are nonnegative, then $\mathcal{A}$ is called* ***positive semi-definite****.*

Definition 2.7 *[16] An $n \times n \times n_3 \times \cdots \times n_p$ order-p tensor $\mathcal{Q}$ is* ***orthogonal*** *if*

$$\mathcal{Q}^T * \mathcal{Q} = \mathcal{Q} * \mathcal{Q}^T = \mathcal{I}.$$

Definition 2.8 *[16] Let $\mathcal{A} = (a_{i_1 \ldots i_p}) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$. Then, the* ***Frobenius norm*** *of $\mathcal{A}$ is*

$$\|\mathcal{A}\|_F^2 = \mathcal{A}^T * \mathcal{A} = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_p=1}^{n_p} a_{i_1 \ldots i_p}^2. \tag{2.6}$$

Definition 2.9 *Let $\mathcal{A} = (a_{i_1 \ldots i_p}) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$. Then, $\mathcal{A}$ is called an* ***f-diagonal*** *tensor if $a_{i_1 i_2 \ldots i_p} = 0$ when $i_1 \neq i_2$. Furthermore, $\{a_{i_1 i_2 \ldots i_p} | i_1 = i_2\}$ are call the t-diagonal entries of $\mathcal{A}$.*

Example 2.4. Let $\mathcal{A} \in \mathbb{R}^{3\times4\times2\times2}$ with the following form:

$$\mathcal{A} = \text{fold}\left(\begin{bmatrix} \mathcal{A}_1 \\ \mathcal{A}_2 \end{bmatrix}\right) = \text{fold}\left(\begin{bmatrix} \text{fold}\left(\begin{bmatrix} \mathcal{A}_{11} \\ \mathcal{A}_{12} \end{bmatrix}\right) \\ \\ \text{fold}\left(\begin{bmatrix} \mathcal{A}_{21} \\ \mathcal{A}_{22} \end{bmatrix}\right) \end{bmatrix}\right)$$

$$= \text{fold}\left(\begin{bmatrix} \text{fold}\left(\begin{bmatrix} a_{1111} & 0 & 0 & 0 \\ 0 & a_{2211} & 0 & 0 \\ 0 & 0 & a_{3311} & 0 \\ a_{1112} & 0 & 0 & 0 \\ 0 & a_{2212} & 0 & 0 \\ 0 & 0 & a_{3312} & 0 \end{bmatrix}\right) \\ \\ \text{fold}\left(\begin{bmatrix} a_{1121} & 0 & 0 & 0 \\ 0 & a_{2221} & 0 & 0 \\ 0 & 0 & a_{3321} & 0 \\ a_{1122} & 0 & 0 & 0 \\ 0 & a_{2222} & 0 & 0 \\ 0 & 0 & a_{3322} & 0 \end{bmatrix}\right) \end{bmatrix}\right).$$

Then, $\mathcal{A}$ is a f-diagonal tensor. □

By the tensor operations constructed and the definition of the linear space, it is easy to get the following result.

Lemma 2.5 *The tensor space $\mathbb{R}^{n_1\times n_2\times\cdots\times n_p}$ is a linear space under the addition of tensors "$+$" and the t-product of tensors "$*$".*

3 The Generalized Inverse of Tensors

The t-product of two tensors presented in Definition 2.1 allows us to obtain the Moore-Penrose inverse of an arbitrary tensor $\mathcal{A}$.

Definition 3.1 *Let $\mathcal{A} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_p}$. If there exists a tensor $\mathcal{X} \in \mathbb{R}^{n_2\times n_1\times n_3\times\cdots\times n_p}$ such that*

$$(1)\ \mathcal{A}*\mathcal{X}*\mathcal{A} = \mathcal{A} \qquad (2)\ \mathcal{X}*\mathcal{A}*\mathcal{X} = \mathcal{X} \qquad (3)\ (\mathcal{A}*\mathcal{X})^T = \mathcal{A}*\mathcal{X} \qquad (4)\ (\mathcal{X}*\mathcal{A})^T = \mathcal{X}*\mathcal{A}, \quad (3.1)$$

*then $\mathcal{X}$ is called the **Moore-Penrose inverse** of the tensor $\mathcal{A}$ and is denoted by $\mathcal{A}^\dagger$.*

For any $\mathcal{A} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_p}$, denote $\mathcal{A}\{i, j, \ldots, k\}$ the set of all $\mathcal{X} \in \mathbb{R}^{n_2\times n_1\times n_3\times\cdots\times n_p}$ which satisfy equations (i), (j), ... , (k) of (3.1). In this case, $\mathcal{X}$ is a $\{i, j, \ldots, k\}$-inverse.

If $\mathcal{A}$ is invertible, it is clear that $\mathcal{X} = \mathcal{A}^{-1}$ trivially satisfies the four equations.

It is worth noting that the order $k\{1\}$ inverse of tensors defined by Sun et al. [19, Definition 2.1] and the Moore-Penrose inverse of tensors defined by Sun et al. [20, Definition 2.2] differ from the topic we focused. This is due to the different products of tensors adopted. Sun et al. [19, Definition 2.1] and Sun et al. [20, Definition 2.2] follow the products of tensors defined by Shao [21] and Einstein [22], respectively while we employ the t-product of tensors defined by Martin et al. [16]. Different definitions on the generalized inverses of tensors may have different applications.

In the following, we will show the existence and uniqueness of the Moore-Penrose inverse of a tensor $\mathcal{A}$.

In the next proof we will use the Kronecker product, symbolized as $\otimes$. Its use in the t-product can be viewed in [15, 16].

Theorem 3.1 *The Moore-Penrose inverse of an arbitrary tensor $\mathcal{A} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_p}$ exists and is unique.*

Proof: We prove the existence of the Moore-Penrose inverse of an arbitrary tensor by construction. For $\mathcal{A} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_p}$, let $\widetilde{A}$ be the $(n_1n_3n_4\cdots n_p \times n_2n_3\cdots n_p)$ matrix at the base level of recursion in [16, Figure 3.2]. Let F_{n_i} be the $n_i \times n_i$ discrete Fourier transform (DFT) matrix and define $F = F_{n_p} \otimes F_{n_{p-1}} \otimes \cdots \otimes F_{n_3}$ and $\rho = n_3 \cdots n_p$. Then there exist matrices $D_1, \ldots, D_\rho$ whose size is $n_1 \times n_2$, possibly with complex entries, such that

$$(F \otimes I_{n_1})\widetilde{A}(F^* \otimes I_{n_2}) = \text{blockdiag}(D_1, \ldots, D_\rho) = \begin{bmatrix} D_1 & & \\ & \ddots & \\ & & D_\rho \end{bmatrix}.$$

Let $D_i = U_i\Sigma_iV_i^T$ be the SVD of each D_i, $i = 1, ..., \rho$ and for each $\Sigma_i = (\sigma^i_{jk})$, we define the matrices $R_i = (r^i_{jk})$, for $i = 1, ..., \rho$, as follows

$$r^i_{jk} = \begin{cases} \frac{1}{\sigma^i_{jk}}, & \text{if } \sigma^i_{jk} \neq 0, \\ 0, & \text{if } \sigma^i_{jk} = 0. \end{cases}$$

Observe that $R_i = \Sigma_i^\dagger$ for $i = 1, \ldots, \rho$. Let $X_i = V_iR_iU_i^T$ for $i = 1, \ldots, \rho$. Now, we have

$$\begin{bmatrix} X_1 & & \\ & \ddots & \\ & & X_\rho \end{bmatrix} = \begin{bmatrix} V_1 & & \\ & \ddots & \\ & & V_\rho \end{bmatrix} \begin{bmatrix} R_1 & & \\ & \ddots & \\ & & R_\rho \end{bmatrix} \begin{bmatrix} U_1^T & & \\ & \ddots & \\ & & U_\rho^T \end{bmatrix}. \quad (3.2)$$

Apply $(F^* \otimes I_{n_1})$ to the left and $(F \otimes I_{n_2})$ to the right of each of the block diagonal matrices in (3.2). One has $\widetilde{X} = \widetilde{V}\widetilde{R}\widetilde{U}^T$, where $\widetilde{X}, \widetilde{U}, \widetilde{R}$ and $\widetilde{V}$ are matrices with same pattern as $\widetilde{A}$. Employ the defined function fold$(\cdot)$ to each matrix in the equality $\widetilde{X} = \widetilde{V}\widetilde{R}\widetilde{U}^T$ in order to have $\mathcal{X} = \mathcal{V} * \mathcal{R} * \mathcal{U}^T$, where $\mathcal{U}$, $\mathcal{V}$ are orthogonal $n_1 \times n_1 \times n_3 \times \cdots \times n_p$, $n_2 \times n_2 \times n_3 \times \cdots \times n_p$ tensors, respectively, and $\mathcal{R}$ is an $n_2 \times n_1 \times n_3 \times \cdots \times n_p$ f-diagonal tensor. One can check that $\mathcal{X}$ satisfies (3.1).

On the other hand, let $\mathcal{X}_1$ and $\mathcal{X}_2$ be solutions of (3.1). One has

$$\begin{aligned} \mathcal{X}_1 &= \mathcal{X}_1 * \mathcal{A} * \mathcal{X}_1 = \mathcal{X}_1 * (\mathcal{A} * \mathcal{X}_2 * \mathcal{A}) * \mathcal{X}_1 = \mathcal{X}_1 * (\mathcal{A} * \mathcal{X}_2)^T * (\mathcal{A} * \mathcal{X}_1)^T \\ &= \mathcal{X}_1 * (\mathcal{A} * \mathcal{X}_1 * \mathcal{A} * \mathcal{X}_2)^T = \mathcal{X}_1 * (\mathcal{A} * \mathcal{X}_2)^T \\ &= \mathcal{X}_1 * \mathcal{A} * \mathcal{X}_2 \\ &= \mathcal{X}_1 * (\mathcal{A} * \mathcal{X}_2 * \mathcal{A}) * \mathcal{X}_2 = (\mathcal{X}_1 * \mathcal{A})^T * (\mathcal{X}_2 * \mathcal{A})^T * \mathcal{X}_2 \\ &= (\mathcal{X}_2 * \mathcal{A} * \mathcal{X}_1 * \mathcal{A})^T * \mathcal{X}_2 = (\mathcal{X}_2 * \mathcal{A})^T * \mathcal{X}_2 \\ &= \mathcal{X}_2 * \mathcal{A} * \mathcal{X}_2 = \mathcal{X}_2. \end{aligned}$$

Therefore, the Moore-Penrose inverse of $\mathcal{A}$ is unique. □

The following lemma is proved in [16, Theorem 4.1] and called T-SVD of a tensor.

Lemma 3.1 *[16, Theorem 4.1] Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$. Then $\mathcal{A}$ can be decomposed as*

$$\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T \tag{3.3}$$

where $\mathcal{U}$, $\mathcal{V}$ are orthogonal $n_1 \times n_1 \times n_3 \times \cdots \times n_p$, $n_2 \times n_2 \times n_3 \times \cdots \times n_p$ tensors, respectively, and $\mathcal{S}$ is an $n_1 \times n_2 \times \cdots \times n_p$ f-diagonal tensor.

In fact, the tensor $\mathcal{R}$ obtained in the proof of Theorem 3.1 is the Moore-Penrose inverse of the tensor $\mathcal{S}$. So, the following is straightforward.

Corollary 3.1 *Let $\mathcal{A}$ be a tensor and factorized as $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$, where $\mathcal{U}$, $\mathcal{V}$ are orthogonal tensors and $\mathcal{S} = (s_{i_1 \ldots i_p})$ is f-diagonal tensor. Then,*

$$\mathcal{A}^\dagger = \mathcal{V} * \mathcal{S}^\dagger * \mathcal{U}^T.$$

In the following, we will state some properties of the Moore-Penrose inverse of tensors and some representations of $\{1\}$-inverses, $\{1,3\}$-inverses and $\{1,4\}$-inverses of tensors. Since the proofs are similar as matrices, we omit them here. The reader can refer to [1].

Theorem 3.2 *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$. Then, the following statements are true:*

(a) $(\mathcal{A}^\dagger)^\dagger = \mathcal{A}$.

(b) $(\mathcal{A}^T)^\dagger = (\mathcal{A}^\dagger)^T$.

(c) $(\mathcal{A} * \mathcal{A}^T)^\dagger = (\mathcal{A}^T)^\dagger * \mathcal{A}^\dagger, \quad (\mathcal{A}^T * \mathcal{A} * \mathcal{A}^T)^\dagger = (\mathcal{A}^T)^\dagger * \mathcal{A}^\dagger * (\mathcal{A}^T)^\dagger$.

(d) $\mathcal{A}^\dagger = \mathcal{A}^T * (\mathcal{A} * \mathcal{A}^T)^\dagger = (\mathcal{A}^T * \mathcal{A})^\dagger * \mathcal{A}^T$.

(e) $\mathcal{X} \in \mathcal{A}^T\{1\}$ *if and only if* $\mathcal{X}^T \in \mathcal{A}\{1\}$.

Theorem 3.3 *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times l \times n_3 \times \cdots \times n_p}$, $\mathcal{B} \in \mathbb{R}^{m \times k \times n_3 \times \cdots \times n_p}$ and $\mathcal{D} \in \mathbb{R}^{n_1 \times k \times n_3 \times \cdots \times n_p}$. Then the tensor equation*

$$\mathcal{A} * \mathcal{X} * \mathcal{B} = \mathcal{D}$$

is consistent if and only if exist $\mathcal{A}^{(1)} \in \mathcal{A}\{1\}$, $\mathcal{B}^{(1)} \in \mathcal{B}\{1\}$ such that

$$\mathcal{A} * \mathcal{A}^{(1)} * \mathcal{D} * \mathcal{B}^{(1)} * \mathcal{B} = \mathcal{D}$$

in which case the general solution is

$$\mathcal{X} = \mathcal{A}^{(1)} * \mathcal{D} * \mathcal{B}^{(1)} + \mathcal{Y} - \mathcal{A}^{(1)} * \mathcal{A} * \mathcal{Y} * \mathcal{B} * \mathcal{B}^{(1)} \tag{3.4}$$

for arbitrary $\mathcal{Y} \in \mathbb{R}^{l \times m \times n_3 \times \cdots \times n_p}$.

Theorem 3.4 *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$. The set $\mathcal{A}\{1,3\}$ consists of all solutions for $\mathcal{X}$ of*

$$\mathcal{A} * \mathcal{X} = \mathcal{A} * \mathcal{A}^{(1,3)}, \tag{3.5}$$

where $\mathcal{A}^{(1,3)}$ is an arbitrary element of $\mathcal{A}\{1,3\}$.

Theorem 3.5 *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$. The set $\mathcal{A}\{1,4\}$ consists of all solutions for $\mathcal{X}$ of*

$$\mathcal{X} * \mathcal{A} = \mathcal{A}^{(1,4)} * \mathcal{A}, \tag{3.6}$$

where $\mathcal{A}^{(1,4)}$ is an arbitrary element of $\mathcal{A}\{1,4\}$.

Corollary 3.2 *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$, $\mathcal{A}^{(1)} \in \mathcal{A}\{1\}$, $\mathcal{A}^{(1,3)} \in \mathcal{A}\{1,3\}$ and $\mathcal{A}^{(1,4)} \in \mathcal{A}\{1,4\}$. Then, the following statements are true:*

(*a*) $\mathcal{A}\{1\} = \{\mathcal{A}^{(1)} + \mathcal{Z} - \mathcal{A}^{(1)} * \mathcal{A} * \mathcal{Z} * \mathcal{A} * \mathcal{A}^{(1)} : \quad \mathcal{Z} \in \mathbb{R}^{n_2 \times n_1 \times n_3 \times \ldots \times n_p}\}$.

(*b*) $\mathcal{A}\{1,3\} = \{\mathcal{A}^{(1,3)} + (\mathcal{I} - \mathcal{A}^{(1,3)} * \mathcal{A}) * \mathcal{Z} : \quad \mathcal{Z} \in \mathbb{R}^{n_2 \times n_1 \times n_3 \times \cdots \times n_p}\}$.

(*c*) $\mathcal{A} * \mathcal{A}^{(1,3)} = \mathcal{A} * \mathcal{A}^{\dagger}$.

(*d*) $\mathcal{A}\{1,4\} = \{\mathcal{A}^{(1,4)} + \mathcal{Z} * (\mathcal{I} - \mathcal{A} * \mathcal{A}^{(1,4)}) : \quad \mathcal{Z} \in \mathbb{R}^{n_2 \times n_1 \times n_3 \times \cdots \times n_p}\}$.

(*e*) $\mathcal{A}^{(1,4)} * \mathcal{A} = \mathcal{A}^{\dagger} * \mathcal{A}$.

(*f*) $\mathcal{A}^{\dagger} = \mathcal{A}^{(1,4)} * \mathcal{A} * \mathcal{A}^{(1,3)}$.

4 The Least-squares Solutions of Tensor Equations

By Theorem 3.3, the tensor equation $\mathcal{A} * \mathcal{X} - \mathcal{B} = 0$ has a solution if and only if exists $\mathcal{A}^{(1)} \in \mathcal{A}\{1\}$ such that $\mathcal{A} * \mathcal{A}^{(1)} * \mathcal{B} = \mathcal{B}$. However, if

$$\mathcal{R} = \mathcal{A} * \mathcal{X} - \mathcal{B} \neq 0, \tag{4.1}$$

it may be desired to find a tensor $\mathcal{X}$ that minimizes the norm of $\mathcal{R}$. Such tensor $\mathcal{X}$ is said to be a least-squares solutions of $\mathcal{A} * \mathcal{X} = \mathcal{B}$.

Definition 4.1 *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$ and $\mathcal{B} \in \mathbb{R}^{n_1 \times 1 \times n_3 \times \cdots \times n_p}$. We say that $\mathcal{X}_0 \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}$ is a* ***least-squares solution*** *of the tensor equation $\mathcal{A} * \mathcal{X} = \mathcal{B}$ if*

$$\|\mathcal{A} * \mathcal{X}_0 - \mathcal{B}\|_F = \min\{\|\mathcal{A} * \mathcal{X} - \mathcal{B}\|_F : \mathcal{X} \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}\}.$$

The following theorem shows that $\|\mathcal{A} * \mathcal{X} - \mathcal{B}\|_F$ is minimized by choosing $\mathcal{X} = \mathcal{A}^{(1,3)} * \mathcal{B}$, where $\mathcal{A}^{(1,3)} \in \mathcal{A}\{1,3\}$. Thus a relation between the $\{1,3\}$-inverses of tensors and the least-squares solutions of $\mathcal{A} * \mathcal{X} = \mathcal{B}$ is established.

Theorem 4.1 *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$, $\mathcal{X}_0 \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}$, $\mathcal{B} \in \mathbb{R}^{n_1 \times 1 \times n_3 \times \cdots \times n_p}$. Let $\mathcal{A}^{(1,3)}$ be an arbitrary element of $\mathcal{A}\{1,3\}$. Then $\mathcal{X}_0$ is a least-squares solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$ if and only if*

$$\mathcal{A} * \mathcal{X}_0 = \mathcal{A} * \mathcal{A}^{(1,3)} * \mathcal{B}.$$

Proof: Let $\mathcal{B}_1 = \mathcal{A} * \mathcal{A}^{(1,3)} * \mathcal{B}$ and $\mathcal{B}_2 = \mathcal{B} - \mathcal{B}_1$. Let $\mathcal{X}$ be an arbitrary tensor of $\mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}$. It is easy to check

$$\mathcal{A}^T * \mathcal{B}_1 = \mathcal{A}^T * \mathcal{A} * \mathcal{A}^{(1,3)} * \mathcal{B} = \mathcal{A}^T * (\mathcal{A} * \mathcal{A}^{(1,3)})^T * \mathcal{B} = (\mathcal{A} * \mathcal{A}^{(1,3)} * \mathcal{A})^T * \mathcal{B} = \mathcal{A}^T * \mathcal{B}.$$

Therefore, $\mathcal{A}^T * \mathcal{B}_2 = 0$, which yields $(\mathcal{B}_1 - \mathcal{A} * \mathcal{X})^T * \mathcal{B}_2 = 0 = \mathcal{B}_2^T * (\mathcal{B}_1 - \mathcal{A} * \mathcal{X})$. Now,

$$\begin{aligned}
&(\mathcal{B} - \mathcal{A} * \mathcal{X})^T * (\mathcal{B} - \mathcal{A} * \mathcal{X}) = (\mathcal{B}_2 + \mathcal{B}_1 - \mathcal{A} * \mathcal{X})^T * (\mathcal{B}_2 + \mathcal{B}_1 - \mathcal{A} * \mathcal{X}) \\
&= \mathcal{B}_2^T * \mathcal{B}_2 + (\mathcal{B}_1 - \mathcal{A} * \mathcal{X})^T * (\mathcal{B}_1 - \mathcal{A} * \mathcal{X}) + (\mathcal{B}_1 - \mathcal{A} * \mathcal{X})^T * \mathcal{B}_2 + \mathcal{B}_2^T * (\mathcal{B}_1 - \mathcal{A} * \mathcal{X}) \\
&= \mathcal{B}_2^T * \mathcal{B}_2 + (\mathcal{B}_1 - \mathcal{A} * \mathcal{X})^T * (\mathcal{B}_1 - \mathcal{A} * \mathcal{X}),
\end{aligned}$$

that is,

$$\|\mathcal{B} - \mathcal{A} * \mathcal{X}\|_F^2 = \|\mathcal{B}_2\|_F^2 + \|\mathcal{B}_1 - \mathcal{A} * \mathcal{X}\|_F^2, \qquad \forall\, \mathcal{X} \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}. \tag{4.2}$$

Assume that $\mathcal{A} * \mathcal{X}_0 = \mathcal{A} * \mathcal{A}^{(1,3)} * \mathcal{B}$, or equivalently, $\mathcal{A} * \mathcal{X}_0 = \mathcal{B}_1$. Using (4.2) we get $\|\mathcal{A} * \mathcal{X}_0 - \mathcal{B}\|^2 = \|\mathcal{B}_2\|^2 \le \|\mathcal{A} * \mathcal{X} - \mathcal{B}\|^2$ for arbitrary $\mathcal{X} \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}$, which means that $\mathcal{X}_0$ is a least-squares solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$.

Assume that $\mathcal{X}_0$ is a least-squares solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$. Theorem 3.3 implies that the tensor equation $\mathcal{A} * \mathcal{X} = \mathcal{A} * \mathcal{A}^{(1,3)} * \mathcal{B}$ is consistent, and so, exists $\mathcal{Y} \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}$ such that $\mathcal{A} * \mathcal{Y} = \mathcal{B}_1$. Since $\mathcal{X}_0$ is a least-squares solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$, we get

$$\|\mathcal{A} * \mathcal{X}_0 - \mathcal{B}\| \le \|\mathcal{A} * \mathcal{Y} - \mathcal{B}\|.$$

Applying (4.2) we get $\|\mathcal{B}_2\|^2 + \|\mathcal{B}_1 - \mathcal{A} * \mathcal{X}_0\|^2 \le \|\mathcal{B}_2\|^2$, and therefore, $\mathcal{B}_1 = \mathcal{A} * \mathcal{X}_0$. □

Remark 4.1 *(a) Notice that the system*

$$\mathcal{A} * \mathcal{X} = \mathcal{A} * \mathcal{A}^{(1,3)} * \mathcal{B} \tag{4.3}$$

is always consistent.

In fact, using Theorem 3.3, one has $\mathcal{A} * \mathcal{X} = \mathcal{B}$ is consistent if and only if $\mathcal{A} * \mathcal{A}^{(1)} * \mathcal{B} = \mathcal{B}$, where $\mathcal{A}^{(1)} \in \mathcal{A}\{1\}$. Applying this to (4.3), it is trivial to see (4.3) is consistent because $\mathcal{A} * \mathcal{A}^{(1)} * \mathcal{A} * \mathcal{A}^{(1,3)} * \mathcal{B} = \mathcal{A} * \mathcal{A}^{(1,3)} * \mathcal{B}$.

(b) Again using Theorem 3.3, we can get the general least-squares solutions of $\mathcal{A} * \mathcal{X} = \mathcal{B}$ is

$$\mathcal{X} = \mathcal{A}^{(1,3)} * \mathcal{B} + (\mathcal{I} - \mathcal{A}^{(1,3)} * \mathcal{A}) * \mathcal{Y} \tag{4.4}$$

where $\mathcal{A}^{(1,3)} \in \mathcal{A}\{1,3\}$, $\mathcal{Y} \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}$ is arbitrary. □

Next, we will show some equivalent conditions for a tensor $\mathcal{X}_0$ being a least-squares solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$. We need the following elementary fact: if $\mathcal{X}$ is a tensor such that $\mathcal{X} * \mathcal{Y} = O$ for any tensor $\mathcal{Y}$ such that $\mathcal{X} * \mathcal{Y}$ is defined, then $\mathcal{X} = O$, where O means a tensor all of whose elements are zero.

Theorem 4.2 *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$, $\mathcal{G} \in \mathbb{R}^{n_2 \times n_1 \times n_3 \times \cdots \times n_p}$. Then, for all $\mathcal{B} \in \mathbb{R}^{n_1 \times 1 \times n_3 \times \cdots \times n_p}$, $\mathcal{X}_0 = \mathcal{G} * \mathcal{B}$ is a least-squares solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$ if and only if $\mathcal{G} \in \mathcal{A}\{1,3\}$.*

Proof: (⇐) The proof follows by choosing $\mathcal{Y} = 0$ in the general solution given in (4.4).

(⇒) If $\mathcal{G} * \mathcal{B}$ is a least-squares solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$, then, by Theorem 3.4, $\mathcal{A} * \mathcal{X}_0 = \mathcal{A} * \mathcal{A}^{(1,3)} * \mathcal{B}$, which implies that $\mathcal{A} * \mathcal{G} * \mathcal{B} = \mathcal{A} * \mathcal{A}^{(1,3)} * \mathcal{B}$, for all $\mathcal{B}$. Hence, $\mathcal{A} * \mathcal{G} = \mathcal{A} * \mathcal{A}^{(1,3)}$. By Theorem 3.4, $\mathcal{G} \in \mathcal{A}\{1,3\}$. □

Theorem 4.3 *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$, $\mathcal{X}_0 \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}$, $\mathcal{B} \in \mathbb{R}^{n_1 \times 1 \times n_3 \times \cdots \times n_p}$. Then $\mathcal{X}_0$ is a least-squares solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$ if and only if*

$$\mathcal{A}^T * \mathcal{A} * \mathcal{X}_0 = \mathcal{A}^T * \mathcal{B}. \tag{4.5}$$

PROOF: By Theorem 4.1 and Corollary 3.2 (c), we only need to prove that

$$\mathcal{A} * \mathcal{X}_0 = \mathcal{A} * \mathcal{A}^\dagger * \mathcal{B} \Leftrightarrow \mathcal{A}^T * \mathcal{A} * \mathcal{X}_0 = \mathcal{A}^T * \mathcal{B}.$$

If $\mathcal{A} * \mathcal{X}_0 = \mathcal{A} * \mathcal{A}^\dagger * \mathcal{B}$, premultiplication by $\mathcal{A}^T$ on both sides gives

$$\mathcal{A}^T * \mathcal{A} * \mathcal{X}_0 = \mathcal{A}^T * \mathcal{A} * \mathcal{A}^\dagger * \mathcal{B} = \mathcal{A}^T * (\mathcal{A} * \mathcal{A}^\dagger)^T * \mathcal{B} = \mathcal{A}^T * (\mathcal{A}^T)^\dagger * \mathcal{A}^T * \mathcal{B} = \mathcal{A}^T * \mathcal{B}.$$

If $\mathcal{A}^T * \mathcal{A} * \mathcal{X}_0 = \mathcal{A}^T * \mathcal{B}$, premultiplication by $(\mathcal{A}^\dagger)^T$ on both sides leads to

$$(\mathcal{A}^\dagger)^T * \mathcal{A}^T * \mathcal{A} * \mathcal{X}_0 = (\mathcal{A}^\dagger)^T * \mathcal{A}^T * \mathcal{B},$$

which is $\mathcal{A} * \mathcal{X}_0 = \mathcal{A} * \mathcal{A}^\dagger * \mathcal{B}$. □

Suppose that the tensor equation $\mathcal{A} * \mathcal{X} = \mathcal{B}$ is consistent. Then, by Theorem 3.3, the general solution are

$$\mathcal{X} = \mathcal{A}^{(1)} * \mathcal{B} + (\mathcal{I} - \mathcal{A}^{(1)} * \mathcal{A}) * \mathcal{Y}, \tag{4.6}$$

where $\mathcal{A}^{(1)} \in \mathcal{A}\{1\}$, $\mathcal{Y} \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}$ is arbitrary.

Among these solutions, it is interesting to find one whose norm is minimum. So, it is natural to give the following definition.

Definition 4.2 *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$ and $\mathcal{B} \in \mathbb{R}^{n_1 \times 1 \times n_3 \times \cdots \times n_p}$. We say that $\mathcal{X}_0 \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}$ is a* ***minimum-norm solution*** *of the consistent tensor equation $\mathcal{A} * \mathcal{X} = \mathcal{B}$ if $\mathcal{X}_0$ is a solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$ and*

$$\|\mathcal{X}_0\|_F \le \|\mathcal{W}\|_F,$$

where $\mathcal{W}$ is an arbitrary solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$.

Notice that the minimum-norm solution of a consistent tensor equation is always unique. In the following, we will relate the minimum-norm solution with the $\{1, 4\}$-inverses of a tensor $\mathcal{A}$.

Theorem 4.4 *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$, $\mathcal{G} \in \mathcal{A}\{1\}$, $\mathbb{H} = \{\mathcal{A} * \mathcal{Z} | \mathcal{Z} \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}\}$. Then, for all $\mathcal{B} \in \mathbb{H}$, $\mathcal{X}_0 = \mathcal{G} * \mathcal{B}$ is the minimum-norm solution of the consistent system $\mathcal{A} * \mathcal{X} = \mathcal{B}$ if and only if $\mathcal{G} \in \mathcal{A}\{1, 4\}$.*

PROOF: $(\Leftarrow)$: According to (4.6), $\mathcal{X}_0 = \mathcal{G} * \mathcal{B}$ is a solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$ and hence, the general solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$ can be written as $\mathcal{X} = \mathcal{X}_0 + (\mathcal{I} - \mathcal{G} * \mathcal{A}) * \mathcal{Y}$, where $\mathcal{Y} \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}$ is arbitrary.

Since $\mathcal{B} \in \mathbb{H}$, then $\mathcal{B} = \mathcal{A} * \mathcal{Z}$ for some $\mathcal{Z}$, now $(\mathcal{G} * \mathcal{B})^T = (\mathcal{G} * \mathcal{A} * \mathcal{Z})^T = \mathcal{Z}^T * (\mathcal{G} * \mathcal{A})^T = \mathcal{Z}^T * \mathcal{G} * \mathcal{A}$, which implies that $(\mathcal{G} * \mathcal{B})^T * (\mathcal{I} - \mathcal{G} * \mathcal{A}) = 0$. Therefore, if $\mathcal{X}$ is any solution of the tensor equation $\mathcal{A} * \mathcal{X} = \mathcal{B}$, then

$$\|\mathcal{X}\|_F^2 = \|\mathcal{X}_0 + (\mathcal{I} - \mathcal{G} * \mathcal{A}) * \mathcal{Y}\|_F^2 = \|\mathcal{X}_0\|_F^2 + \|(\mathcal{I} - \mathcal{G} * \mathcal{A}) * \mathcal{Y}\|_F^2 \ge \|\mathcal{X}_0\|_F^2, \tag{4.7}$$

which means $\mathcal{X}_0$ is the minimum-norm solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$.

$(\Rightarrow)$: Suppose that for all $\mathcal{B} \in \mathbb{H}$, $\mathcal{X}_0 = \mathcal{G} * \mathcal{B}$ is the minimum-norm solution of the consistent system $\mathcal{A} * \mathcal{X} = \mathcal{B}$. Let $\overline{\mathcal{A}}_i \in \mathbb{R}^{n_1 \times 1 \times n_3 \times \cdots \times n_p}$, $i = 1, 2, \ldots, n_2$ the order-p tensor with $\overline{a}_{i_1 1 i_3 \ldots i_p} = a_{i_1 i i_3 \ldots i_p}$, where $i_1 = 1, 2, \ldots, n_1, i_3 = 1, 2, \ldots, n_3, \cdots, i_p = 1, 2, \ldots, n_p$.

Choose $\mathcal{B} = \overline{\mathcal{A}}_i$, for some $i = 1, 2, \ldots, n_2$. Then, $\mathcal{G} * \overline{\mathcal{A}}_i$ is the minimum-norm solution of $\mathcal{A} * \mathcal{X} = \overline{\mathcal{A}}_i$. Notice that $\mathcal{A}^{(1,4)} * \overline{\mathcal{A}}_i$ is also the minimum-norm solution of $\mathcal{A} * \mathcal{X} = \overline{\mathcal{A}}_i$.

This means $\mathcal{G} * \overline{\mathcal{A}}_i = \mathcal{A}^{(1,4)} * \overline{\mathcal{A}}_i$ for the uniqueness of the minimum-norm solution. Hence, $\mathcal{G} * \overline{\mathcal{A}}_i = \mathcal{A}^{(1,4)} * \overline{\mathcal{A}}_i$ is true for all $i = 1, 2, \dots, n_2$, which implies $\mathcal{G} * \mathcal{A} = \mathcal{A}^{(1,4)} * \mathcal{A}$. Then, we have $\mathcal{G} \in \mathcal{A}\{1,4\}$ by Theorem 3.5. □

In general, the solution of the least square equations is not unique. It is necessary to for us to find a minimum-norm solution among the least-squares solutions when settling some practical problems.

Theorem 4.5 *Let* $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$, $\mathcal{G} \in \mathbb{R}^{n_2 \times n_1 \times n_3 \times \cdots \times n_p}$. *Then, for all* $\mathcal{B} \in \mathbb{R}^{n_1 \times 1 \times n_3 \times \cdots \times n_p}$, $\mathcal{X}_0 = \mathcal{G} * \mathcal{B}$ *is the minimum-norm least-squares solution of* $\mathcal{A} * \mathcal{X} = \mathcal{B}$ *if and only if* $\mathcal{G} = \mathcal{A}^{\dagger}$.

Proof: $(\Leftarrow)$: By Theorem 4.1, the least-squares solutions of $\mathcal{A} * \mathcal{X} = \mathcal{B}$ coincide with the solutions of

$$\mathcal{A} * \mathcal{X} = \mathcal{A} * \mathcal{A}^{(1,3)} * \mathcal{B}. \tag{4.8}$$

Hence, the minimum-norm least-squares solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$ is the minimum-norm solution of (4.8). By Theorem 4.4,

$$\mathcal{X}_0 = \mathcal{A}^{(1,4)} * \mathcal{A} * \mathcal{A}^{(1,3)} \mathcal{B} = \mathcal{A}^{\dagger} * \mathcal{B}, \tag{4.9}$$

which means $\mathcal{G} = \mathcal{A}^{\dagger}$.

$(\Rightarrow)$: If $\mathcal{G} = \mathcal{A}^{\dagger}$, then $\mathcal{X}_0 = \mathcal{A}^{\dagger} * \mathcal{B}$. Hence, it follows $\mathcal{X}_0$ is the minimum-norm least-square solution of $\mathcal{A} * \mathcal{X} = \mathcal{B}$ by Theorem 4.2 and Theorem 4.4. □

5 An Algorithm for Computing the Moore-Penrose Inverse of a Tensor

According to the proof of Theorem 3.1, we propose the following algorithm to compute the Moore-Penrose inverse of an arbitrary tensor. Before that, we declare that fft(·) and ifft(·) are Matlab and Octave functions, which implement the fast Fourier transform and the inverse fast Fourier transform of a matrix, respectively. Also note that pinv(·) is a Matlab (and Octave) built-in function which computes the Moore-Penrose inverse of an arbitrary complex matrix.

Algorithm 5.1: COMPUTE THE MOORE-PENROSE INVERSE OF A TENSOR $\mathcal{A}$

Input: $n_1 \times n_2 \times \cdots \times n_p$ tensor $\mathcal{A}$

Output: $n_2 \times n_1 \times n_3 \times \cdots \times n_p$ tensor $\mathcal{X}$

1. for $i = 3, \ldots, p$

   $\mathcal{D} = \text{fft}(\mathcal{A}, [\ ], i);$

   end

2. $N = n_3 n_4 \cdots n_p$

   for $i = 1, \ldots, N$

   $\mathcal{G}(:, :, i) = \text{pinv}(\mathcal{D}(:, :, i));$ where $\text{pinv}(\mathcal{D}(:, :, i))$ is the Moore-Penrose inverse of $\mathcal{D}(:, :, i)$,

   end

3. for $i = p, \ldots, 3$

   $\mathcal{X} = \text{ifft}(\mathcal{G}, [\ ], i);$

   end

The strategy of this algorithm is using $\text{fft}(\cdot)$ to some objects and then calculate the Moore-Penrose inverse of each result matrix from $\text{fft}(\mathcal{A})$. Finally, employing $\text{ifft}(\cdot)$ to $\mathcal{D}(:, :, i)^\dagger$ as in the Algorithm to get the Moore-Penrose inverse of $\mathcal{A}$. Next, we will test the construct Algorithm by using the following example.

Example 5.5. Let $\mathcal{A}$ be a $5 \times 4 \times 2 \times 2$ tensor with the following form:

$$\mathcal{A} = \left[\underbrace{\begin{matrix} 1 & 2 & 5 & 4 \\ 4 & 3 & 3 & 1 \\ 5 & 8 & 2 & 6 \\ 6 & 2 & 2 & 4 \\ 8 & 2 & 2 & 4 \end{matrix}}_{a_{ij11}} \left| \underbrace{\begin{matrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 2 & 1 \\ 6 & 4 & 2 & 9 \\ 0 & 0 & 2 & 5 \\ 5 & 3 & 3 & 3 \end{matrix}}_{a_{ij21}} \right| \underbrace{\begin{matrix} 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{matrix}}_{a_{ij12}} \left| \underbrace{\begin{matrix} 2 & 2 & 2 & 8 \\ 2 & 2 & 2 & 8 \\ 6 & 3 & 2 & 9 \\ 9 & 1 & 0 & 1 \\ 1 & 1 & 1 & 10 \end{matrix}}_{a_{ij22}} \right.\right].$$

Implement Algorithm 5.1 on $\mathcal{A}$, we have

$$\mathcal{A}^\dagger(:, :, 1, 1) = \begin{bmatrix} -0.0511 & 0.0776 & -0.0422 & 0.0670 & 0.0065 \\ -0.0186 & 0.0963 & 0.0655 & -0.0266 & -0.0782 \\ 0.2238 & 0.0221 & -0.0439 & 0.0680 & -0.1490 \\ -0.0265 & -0.0478 & 0.0250 & -0.0196 & 0.0432 \end{bmatrix},$$

$$\mathcal{A}^\dagger(:, :, 2, 1) = \begin{bmatrix} -0.0582 & 0.0179 & 0.0112 & 0.0504 & 0.0039 \\ 0.0404 & -0.0858 & 0.0013 & 0.0285 & 0.0081 \\ -0.1851 & 0.0830 & -0.0299 & 0.0662 & 0.1304 \\ 0.0346 & -0.0218 & 0.0334 & -0.0859 & 0.0319 \end{bmatrix},$$

$$\mathcal{A}^\dagger(:, :, 1, 2) = \begin{bmatrix} 0.0128 & 0.0490 & -0.0317 & -0.0030 & -0.0270 \\ -0.0984 & 0.1554 & -0.0528 & 0.0458 & -0.0380 \\ 0.0543 & 0.0357 & 0.0243 & -0.0788 & -0.1262 \\ 0.0447 & -0.0837 & 0.0022 & -0.0459 & 0.0113 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,2,2) = \begin{bmatrix} -0.0021 & -0.0186 & 0.0024 & 0.0492 & -0.0045 \\ -0.0188 & -0.0265 & 0.0294 & -0.0399 & 0.0488 \\ -0.0288 & 0.0249 & -0.0764 & 0.0123 & 0.1128 \\ 0.0545 & -0.0688 & 0.0306 & 0.0070 & -0.0163 \end{bmatrix}.$$

6 Applications to Higher Order Gauss-Markov Theorem

In statistics, ***linear regression*** is an approach for modelling the relationship between a scalar dependent variable and one ore more independent variables by fitting a linear equation to observed data. Commonly, the relationships are modelled by using linear predictor functions whose unknown model parameters are estimated from the data. Such models are called ***linear models***.

Recall the ***Gauss-Markov theorem***, named after Carl Friedrich Gauss and Andrey Markov. This theorem states that in a linear model if the errors have expectation zero, are uncorrelated, and have equal variances, then the estimators of the parameters in the model produced by least squares estimation are better than any other unbiased linear estimator. The reader can consult, e.g., [23, Chapter 5].

In this part, we will construct a linear model for tensors and then establish the higher order Gauss-Markov theorem by using the Moore-Penrose inverse of tensors and the least-squares solutions of tensor equations.

Firstly, the following definitions are necessary.

Definition 6.1 *A* ***random tensor*** *is a tensor-valued random variable, that is, a tensor all of whose elements are random variables.*

Definition 6.2 *The* ***mean*** *or* ***expectation*** *of a random tensor* $\mathcal{X} = (X_{i_1 i_2 \dots i_n}) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$ *is defined as* $E[\mathcal{X}] = (E[X_{i_1 i_2 \dots i_n}]) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$.

Definition 6.3 *The* ***covariance tensor*** *of a random tensor* $\mathcal{Y} = (Y_{i_1 i_2 \dots i_n}) \in \mathbb{R}^{n_1 \times 1 \times n_3 \times \cdots \times n_p}$ *is defined as* $Cov(\mathcal{Y}) = E[(\mathcal{Y} - E(\mathcal{Y})) * (\mathcal{Y} - E(\mathcal{Y}))^T] \in \mathbb{R}^{n_1 \times n_1 \times n_3 \times \cdots \times n_p}$.

The random tensor, the expectation of a random tensor, and the covariance tensor are generalizations of the notions of the random vector, the expectation of a random matrix, and the covariance matrix. The covariance matrix plays a significant role in statistics and probability theory. The expectation and the covariance tensor of a random tensor is very useful in some practical problems. For example, a model has two (or more) independent random vector. We can view the two independent random vectors c_1 and c_2 as a $n \times 1 \times 2$ random tensor $\mathcal{C}$, that is

$$\mathcal{C} = \left[\begin{array}{c|c} c_{111} & c_{112} \\ \vdots & \vdots \\ c_{n11} & c_{n12} \end{array} \right].$$
$$\underbrace{\phantom{c_{n11}}}_{c_{i11}} \quad \underbrace{\phantom{c_{n12}}}_{c_{i12}}$$

Then, it is not difficult to apply Definition 6.2 and Definition 6.3 to $\mathcal{C}$ and research some significant problems.

Next, we will establish a linear model for tensors. We call the model of tensors the linear model due to the fact that the tensor space is a linear space under the addition of tensors "$+$" and the t-product of tensors "$*$". See Lemma 2.5.

The ***linear model for tensors*** postulates

$$\mathcal{Y} = \mathcal{X} * \mathcal{P} + \mathcal{E}, \tag{6.1}$$

where $\mathcal{Y} \in \mathbb{R}^{n_1 \times 1 \times n_3 \times \cdots \times n_p}$ is observed or measured in some experimental set-up, $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$ is given, the parameters $\mathcal{P} \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}$ are unknown, and $\mathcal{E} \in \mathbb{R}^{n_1 \times 1 \times n_3 \times \cdots \times n_p}$ a random tensor representing the errors of observing $\mathcal{Y}$ and with

$$E[\mathcal{E}] = O, \quad Cov(\mathcal{E}) = \mathcal{V}^2.$$

The tensor $\mathcal{V}$, assumed known, is positive semi-definite. We denote this model by $(\mathcal{Y}, \mathcal{X}*\mathcal{P}, \mathcal{V}^2)$.

Now, we turn to the problem of estimating a linear function of the parameters $\mathcal{P}$ from the observed $\mathcal{Y}$. A linear function of $\mathcal{P}$ has the form $\mathcal{D} * \mathcal{P}$ for a given tensor $\mathcal{D}$. A ***linear estimator*** of $\mathcal{D} * \mathcal{P}$ is $\mathcal{A} * \mathcal{Y}$, for some $\mathcal{A} \in \mathbb{R}^{n_2 \times n_1 \times n_3 \times \cdots \times n_p}$. The linear estimator $\mathcal{A} * \mathcal{Y}$ is a ***linear unbiased estimator*** if

$$E[\mathcal{A} * \mathcal{Y}] = \mathcal{D} * \mathcal{P}, \quad \text{for all } \mathcal{D},$$

and it is the ***best linear unbiased estimator*** if its variance is minimal among all linear unbiased estimators.

The function $\mathcal{D} * \mathcal{P}$ is called ***estimable*** if it has an linear unbiased estimator, i.e., if there is a tensor $\mathcal{A} \in \mathbb{R}^{n_2 \times n_1 \times n_3 \times \cdots \times n_p}$ such that $E[\mathcal{A} * \mathcal{Y}] = \mathcal{D} * \mathcal{P}$ holds.

Now, we state the higher order Gauss-Markov theorem. Before that a new multilinear rank of a tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$ is needed.

Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$. Create an $\widetilde{A}$ matrix by using the method in [16, Figure 3.2] and apply the discrete Fourier transform to $\widetilde{A}$. One has

$$(F \otimes I_{n_1})\widetilde{A}(F^* \otimes I_{n_2}) = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_\rho \end{bmatrix},$$

where $\rho = n_3 \cdots n_p$. Then the ρ-tuple $\big(rank(A_1), rank(A_2), \ldots, rank(A_\rho)\big)$ is called the ***multilinear rank*** of $\mathcal{A}$. The reader must not be confused with the n-tuple of mode-n ranks defined in [24], which is the number of linearly independent mode-n vectors. For two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{n_1 \times \cdots \times n_p}$, we write $\mathcal{A} \leq \mathcal{B}$ when $a_{i_1,\ldots,i_p} \leq b_{i_1,\ldots,i_p}$ for all $i_1, \ldots, i_p$. The definition of $\mathcal{A} \geq \mathcal{B}$ is similar.

Theorem 6.1 *Let $(\mathcal{Y}, \mathcal{X} * \mathcal{P}, \mathcal{V}^2)$ be a linear model. Suppose that the multilinear rank of $\mathcal{X}$ satisfies*

$$\big(rank(X_1), rank(X_2), \ldots, rank(X_\rho)\big) > \big(\max\{n_1, n_2\}, \max\{n_1, n_2\}, \ldots, \max\{n_1, n_2\}\big).$$

Then:

(a) The linear functional $\mathcal{D} * \mathcal{P}$ has a unique best linear unbiased estimator $\mathcal{D} * \widetilde{\mathcal{P}}$, where

$$\widetilde{\mathcal{P}} = \mathcal{X}^\dagger * (\mathcal{I} - (\mathcal{V} - \mathcal{V} * \mathcal{X}^\dagger * \mathcal{X})^\dagger * \mathcal{V})^T * \mathcal{Y}.$$

(b) $\widetilde{\mathcal{P}} \in \mathbb{K}$*, where* $\mathbb{K} = \{\mathcal{X}^T * \mathcal{Z} | \mathcal{Z} \in \mathbb{R}^{n_2 \times 1 \times n_3 \times \cdots \times n_p}\}$*, and if* $\mathcal{P}^*$ *is any other linear unbiased estimators that belongs to* $\mathbb{K}$*, then*

$$Cov(\mathcal{P}) \leq Cov(\mathcal{P}^*).$$

Proof: Employing the base level of recursion in [16, Figure 3.2] for the tensors $\mathcal{Y}$, $\mathcal{X}$, $\mathcal{P}$, $\mathcal{V}$ and $\mathcal{D}$, respectively, to obtain fives matrices $\widetilde{Y}$, $\widetilde{X}$, $\widetilde{P}$, $\widetilde{V}$, and $\widetilde{D}$. Using the same method as in the proof of Theorem 3.1, we can construct the block diagonal matrices of $\widetilde{Y}$, $\widetilde{X}$, $\widetilde{P}$, $\widetilde{V}$, and $\widetilde{D}$. Specifically, one has

$$(F \otimes I_{n_1})\Psi(F^* \otimes I_{n_2}) = \text{blockdiag}(\psi_1, \ldots, \psi_\rho),$$

where $F = F_{n_p} \otimes F_{n_{p-1}} \otimes \cdots \otimes F_{n_3}$, $\rho = n_3 \cdots n_p$, $\Psi = \widetilde{Y}, \widetilde{X}, \widetilde{P}, \widetilde{V}, \widetilde{D}$ and $\psi = Y, X, P, V, D$.

Imposing [1, Section 8.2, Theorem 2] on each matrix linear model $(Y_i, X_i P_i, V_i^2)$, $i = 1, \ldots, \rho$, one has

$$\begin{bmatrix} \widetilde{P}_1 & & \\ & \ddots & \\ & & \widetilde{P}_\rho \end{bmatrix} = \begin{bmatrix} X_1^\dagger (I_1 - (V_1 - V_1 X_1^\dagger X_1)^\dagger V_1)^T Y_1 & & \\ & \ddots & \\ & & X_\rho^\dagger (I_\rho - (V_\rho - V_\rho X_\rho^\dagger X_\rho)^\dagger V_\rho)^T Y_\rho \end{bmatrix}.$$

Apply $(F^* \otimes I_{n_1})$ to the left and $(F \otimes I_{n_2})$ to the right of the block diagonal matrices in the equality above and then the defined function fold$(\cdot)$ to the obtained equality in the aforementioned step, one has

$$\widetilde{\mathcal{P}} = \mathcal{X}^\dagger * (\mathcal{I} - (\mathcal{V} - \mathcal{V} * \mathcal{X}^\dagger * \mathcal{X})^\dagger * \mathcal{V})^T * \mathcal{Y}.$$

The proof of the item (b) follows similarly. □

Remark 6.1 *If* $\mathcal{V}^2$ *is nonsingular, then* $\widetilde{\mathcal{P}}$ *is reduced to*

$$\widetilde{\mathcal{P}} = (\mathcal{X}^T * \mathcal{V}^{-2} * \mathcal{X})^\dagger * \mathcal{X}^T * \mathcal{V}^{-2} * \mathcal{Y}.$$

For the model $(\mathcal{Y}, \mathcal{X} * \mathcal{P}, \sigma^2 \mathcal{I})$*, where* σ *is a positive real number, the best linear unbiased estimator reduces to*

$$\widetilde{\mathcal{P}} = \mathcal{X}^\dagger * \mathcal{Y},$$

which can be called the ***least-squares estimator****.* □

The generalized inverse of tensors can be very useful in other fields, such as the Bott-Duffin inverse of tensors to higher order electrical networks or hypergraphs theory, the group inverse of tensors to higher order Markov chain and the least-square solutions of the tensor equation in 3-D image deblurring, etc. We will continue these researches in the future.

References

[1] A. Ben-Israel, T.N.E. Greville, Generalized Inverses: Theory and Applications, 2nd Edition, Springer Verlag, New York, 2003.

[2] S. L. Campbell, C. D. Meyer, Generalized Inverse of Linear Transformations, Pitman, London, 1979; Dover, New York, 1991.

[3] G. Wang, Y. Wei, S. Qiao, Generalized Inverses: Theory and Computations, Science Press, Beijing, 2004.

[4] Banachiewicz, T. Zur Berechnung der Determinanten, wie auch der In- versen, und zur darauf basierten Auflosung der Systeme linearer Gleichun- gen. Acta Astronomica, Série C, 3 (1937), pp. 41-67.

[5] C. R. Rao, A Note on a Generalized Inverse of a Matrix with Applications to Problems in Mathematical Statistics, Journal of the Royal Statistical Society, 24(1) (1962), pp. 152-158.

[6] J. J. Hunter, Generalized Inverses and Their Application to Applied Probability Problems, Linear Algebra Appl., 45 (1982), pp. 157-198.

[7] A. L. Puri, C. T. Russell, Convergence of Generalized Inverses with Applications to Asymptotic Hypothesis Testing, The Indian Journal of Statistics, 46(2) (1984), pp. 277-286.

[8] T. D. Tran, Spectral sets and the Drazin inverse with applications to second order differential equations, Applications of Mathematics, 47 (2002), pp. 1-8.

[9] H.A. Kiers, Towards a standardized notation and terminology in multiway analysis, J. Chemom. 14 (2000), pp. 105-122.

[10] C. Hillar, L. H. Lim. Most tensor problems are NP-hard, Journal of the ACM, 60 (2013), no. 6.

[11] J. Carroll, J. Chang, Analysis of individual differences in multidimensional scaling via an n-way generalization of Eckart-Young decomposition, Psychometrika, 35 (1970), pp. 283-319.

[12] R. Harshman, Foundations of the PARAFAC procedure: Model and conditions for an explanatory multi-mode factor analysis, UCLA Working Papers in Phonetics, 16 (1970), pp. 1-84.

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

[14] T. G . Kolda, B. W. Bader, Tensor decompositions and applications, SIAM Rev., 51 (2009), pp. 455-500.

[15] M. E. Kilmer and C. D. Martin, Factorization strategies for third-order tensors, Linear Algebra Appl., 435 (2011), pp. 641-658.

[16] C. D. Martin, R. Shafer and B. LaRue, An Order-p Tensor Factorization with Applications in Imaging. SIAM J. Scientific Computing, 35 (2013), pp. 474-490.

[17] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, Third order tensors as operators on matrices: A theoretical and computational framework with applications in imaging, SIAM J. Matrix Anal. Appl. 34 (2013), pp. 148-172.

[18] PS. Stanimirović, S. Chountasis, D. Pappas, I. Stojanoviá, Removal of blur in images based on least squares solutions. Math Methods Appl Sci. doi:10.1002/mma.2751.

[19] L. Sun, B. Zheng, C. Bu, and Y. Wei, Some results on the generalized inverse of tensors and idempotent tensors, 2014, arXiv:1412.7430.

[20] L. Sun, B. Zheng, C. Bu, and Y. Wei, Moore-Penrose inverse of tensors via Einstein product, Linear and Multilinear Algebra, 2015, http: //dx.doi.org/10.1080-/03081087.2015.1083933.

[21] J. Shao, A general product of tensors with applications, Linear Algebra and its Applications, 439 (2013), pp. 2350-2366.

[22] A. Einstein, The foundation of the general theory of relativity, In: Kox AJ, Klein MJ, Schulmann R, editors. The collected papers of Albert Einstein. Vol. 6. Princeton (NJ): Princeton University Press, 2007.

[23] C. D. Meyer, Matrix Analysis and Applied Linear Algebra, SIAM: Society for Industrial and Applied Mathematics, 2001.

[24] F. L. Hitchcock, Multiple invariants and generalized rank of a p-way matrix or tensor, Journal of Mathematical Physics, 7(1) (1927), pp. 39-79.