PRODUCTS OF TENSORS

Student Name: Angel Mabuza

Student No: 1615055

Supervisor: Professor Yorick Hardy

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science.

Johannesburg, 2024

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

____
(Signature of candidate)

01 day of April 20 25 at Randburg

Abstract

This dissertation surveys the products of tensors, their respective multilinear structures and their properties. Matrix products arise naturally from the composition of linear operations. In like manner, products of tensors derive from composing multilinear functions. However, there are many ways to compose multilinear functions, hence many forms of the product. We will gain further insight of linear operations and introduce the *tensor-matrix* multiplication (the simplest form of the compositional product). We will consider different types of products building up to the study of the general *tensor-tensor* multiplication defined in [Lin. Alg. Appl. 439, 2350–2366 (2013)]. Finally, we conclude with an investigation of some n-ary products of tensors.

Dedication

To my late grandmother, Nontombi Florence Maho.

Contents

List of Symbols

The following notations will be used in this paper:

$\mathcal{A}$	tensor
A	matrix
a	vector
$a^{(i)}$	i-th column of a matrix A
α	scalar
$\mathbb{F}$	field
$\mathbb{F}^n$	The n-dimensional vector space over the field $\mathbb{F}$
$\mathbb{F}^{m\times n}$	The space of $m \times n$ matrices over the field $\mathbb{F}$
$\otimes$	Tensor product or Kronecker product
$\cong$	Isomorphic
$\odot$	Entrywise product
$\circ$	Composition
$*$	T-product
$\times_i$	i-th modal product
$*_E$	Einstein-Product
A^*	Conjugate transpose of the matrix A
A^+	Moore-Penrose inverse of the matrix A
A^T	Transpose of the matrix A
$C_n^k(\alpha)$	k-conjugate of the complex number α
$A^\dagger$	Adjoint of the matrix A
e_i	Standard basis vector
E_{ij}	Standard basis matrix
δ	Kronecker delta

Chapter 1

Introduction

The study of tensors and tensor analysis dates back to the 19-th century. In 1884, Gibbs introduced the concept of tensor products (squares) of $\mathbb{R}^3$ [6] and the general concept of the tensor product has been subsequently studied extensively (see for example [3, 12]). In the last few years, the study of tensors has interested many mathematicians both for application purposes and for theoretical development of the field. In this dissertation, we investigate the products of tensors and their underlying algebraic properties. Our focus will be on the following papers [4, 7, 9, 10, 13, 16, 17].

We will often investigate the third-order tensors for simplicity, most products considered are binary products. In the following chapters we will investigate the products of tensors described in [4, 7, 9, 10, 13, 16, 17]. We will primarily be concerned with algebraic properties, such as distributivity, multilinearity and associativity. Additionally, all of the products investigated induce a decomposition similar to the singular value decomposition of matrices (excluding [16]). We will describe these decompositions for the higher order tensor products.

In Chapter 2, we consider the *modal* product, a natural extension of the matrix product to the multilinear algebraic setting, We will also discuss the singular value decomposition (SVD) for modal products [4, 9].

Chapter 3 focuses on the *Einstein* product. We will consider the Einstein product as described in [17] and investigate the algebraic properties of the Einstein product. The Einstein product is a binary product and uses tensor unfolding to multiply tensors via matrix multiplication.

The Einstein product may be reduced from multilinear to linear and this is how we investigate the algebraic properties of the Einstein product. The main properties we will investigate are the Moore-Penrose inverse and the singular value decomposition.

In Chapter 4, we study the *T-product*. First introduced in [10], the T-product is a binary product that preserves the tensor order. The T-product is defined by matricizing the tensors using the circ and the MatVec operators.

We will only consider third-order tensors, but the T-product extends to higher-order tensors. We will show the algebraic properties of the T-product and investigate the singular value decomposition.

In Chapter 5, we investigate the *general tensor-tensor product*, which is the first product (in this dissertation) that does not have an underlying matrix product structure. The main paper for this product is [16] and we will show that this product is associative. This chapter also investigate the tensor identity and distributivity conditions of the tensor-tensor product.

In Chapter 6, we will study a *ternary product* for third-order tensors and investigate the product's algebraic properties. The focus will be on the papers [7, 13], a We will conclude this chapter by showing that a third-order tensor may be factored analogous to the spectral decomposition of matrices.

The products defined in each chapter have their own notion of transpose, symmetry, etc. In each case we will use the same notation, since different products are not usually mixed. In general, decompositions and pseudo inverses exist when the underlying field is the field of real numbers or complex numbers, we will consider the real numbers in this dissertation (in most cases the analogous results for complex numbers is straightforward, except for the ternary products in Chapter 6).

The following section introduces the prerequisite for the dissertation. We will study matrix algebra, which naturally follows from linear compositions. Furthermore, we will investigate the algebraic properties of matrices and use these properties to investigate the properties of higher-order tensor products.

1.1 Matrix product

We are interested in studying products of tensors. To define what a tensor product is, let us first define what a product is for matrices (tensor of order $m = 2$). Let

$$L : \mathbb{F}^m \to \mathbb{F}^p \quad \text{and} \quad H : \mathbb{F}^p \to \mathbb{F}^n,$$

be linear maps, where $\mathbb{F}^m, \mathbb{F}^p$ and $\mathbb{F}^n$ are finite dimensional vector spaces. Furthermore, let each linear map be associated with the matrix representations A and B, respectively. Each $u \in \mathbb{F}^m$ has the corresponding column vector $\bar{u}$, similarly for $v \in \mathbb{F}^p$, the corresponding column vector is $\bar{v}$. In other words,

$$L(u) = A\bar{u} \quad \text{and} \quad H(v) = B\bar{v}, \tag{1.1}$$

where $A \in \mathbb{F}^{n\times m}$, $B \in \mathbb{F}^{p\times n}$, $A\bar{u} \in \mathbb{F}^n$ and $B\bar{v} \in \mathbb{F}^p$. The composition $H \circ L$ is a map from U to W. Using (1.1) and the definition of composition, it follows that

$$\begin{aligned}(H \circ L)(u) &= H(L(u)) \\ &= H(A\bar{u}) \\ &= B(A\bar{u}) \\ &= (BA)\bar{u}.\end{aligned} \tag{1.2}$$

So the representation matrix of the composition of linear transformations is a matrix product of the respective representation matrices.

Example 1.1.1. Let $A \in \mathbb{F}^{3\times 2}$ and $B \in \mathbb{F}^{2\times 2}$ where

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

The product of B by A is

$$\begin{aligned} AB &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} \end{pmatrix}. \end{aligned}$$

That is, for $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{n\times p}$ the entries of the product of B by A are

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}, \tag{1.3}$$

where $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, p$. In the following section we will introduce the properties of matrices, which will be fundamental for our investigation of the products of tensors.

1.1.1 Properties of the matrix product

Let A, B, C be matrices such that their respective sizes allow for the operations defined below to be true and α, $\beta \in \mathbb{F}$. The following are algebraic properties of the matrix product (see for example [1]).

1. $A(BC) = (AB)C$, (Multiplicative associative law)
2. $A(B + C) = AB + AC$,
3. $(B + C)A = BA + CA$,
4. $A + B = B + A$, (Addition commutative law)
5. $(A + B) + C = A + (B + C)$, (Addition associative law)
6. $\alpha(B + C) = \alpha B + \alpha C$,
7. $(\alpha + \beta)C = \alpha C + \beta C$,
8. $\alpha(\beta C) = (\alpha\beta)C$,
9. $\alpha(BC) = (\alpha B)C = B(\alpha C)$.

Next, we will study the single value decomposition of matrices. We follow the approach in [1].

1.1.2 Singular value decomposition

In the following section, we will define and prove the singular value decomposition for real matrices.

Theorem 1.1.2. *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix, then there exist an orthogonal matrix $U \in \mathbb{R}^{n\times n}$ and diagonal matrix $D \in \mathbb{R}^{n\times n}$, such that*

$$A = UDU^T. \tag{1.4}$$

Lemma 1.1.3. *Let $A \in \mathbb{R}^{m\times n}$. The eigenvalues of A^TA are non-negative.*

Proof. A^TA is symmetric, that is

$$(A^TA)^T = A^T(A^T)^T = A^TA.$$

Using Theorem 1.1.2, there exists an orthogonal matrix $U \in \mathbb{R}^{n\times n}$ and a diagonal matrix $D \in \mathbb{R}^{n\times n}$ such that

$$A^TA = UDU^T.$$

Let $\{u_1, u_2, \ldots, u_i, \ldots, u_n\}$ be orthonormal columns of U and the eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_i, \ldots, \lambda_n\}$ be diagonal elements of D, then $A^TAu_i = \lambda_i u_i$ and so

$$\begin{aligned}
||Au_i||^2 =& \langle Au_i, Au_i\rangle \\
&= \langle u_i, A^TAu_i\rangle \\
&= \langle u_i, \lambda_i u_i\rangle
\end{aligned}$$

$$\begin{aligned} &= \lambda_i \langle u_i, u_i \rangle \\ &= \lambda_i \geq 0. \end{aligned}$$ □

Definition 1.1.4. If A is an $m \times n$ matrix and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of $A^T A$, then the numbers

$$\sigma_1 = \sqrt{\lambda_1},\ \sigma_2 = \sqrt{\lambda_2},\ \ldots,\ \sigma_n = \sqrt{\lambda_n} \tag{1.5}$$

are called the singular values of A.

Lemma 1.1.5. *Let A be an $m \times n$ matrix, then*

$$\text{rank}(A^T A) = \text{rank}(A). \tag{1.6}$$

Proof. We use the rank-nullity theorem which states that for $A \in \mathbb{R}^{m \times n}$ it holds that

$$\text{rank}(A) + \dim(\text{null}(A)) = n. \tag{1.7}$$

We want to show that $\dim(\text{null}(A)) = \dim(\text{null}(A^T A))$. Let $x \in \text{null}(A)$, then

$$Ax = 0 \Rightarrow A^T Ax = 0 \tag{1.8}$$

$$\Rightarrow x \in \text{null}(A^T A). \tag{1.9}$$

Thus, $\text{null}(A) \subseteq \text{null}(A^T A)$. Now let $x \in \text{null}(A^T A)$ then

$$\begin{aligned} A^T Ax = 0 &\Rightarrow x^T A^T Ax = 0 \\ &\Rightarrow (Ax)^T Ax = 0 \\ &\Rightarrow ||Ax||^2 = 0 \\ &\Rightarrow Ax = 0 \\ &\Rightarrow x \in \text{null}(A). \end{aligned}$$

Thus, $\text{null}(A^T A) \subseteq \text{null}(A)$ and so it follows that

$$\dim(\text{null}(A^T A)) = \dim(\text{null}(A)) \Rightarrow \text{rank}(A^T A) = \text{rank}(A).$$ □

We have shown that the eigenvalues of symmetric matrices are non-negative. We will now consider the singular value decomposition for an $m \times n$ matrix A [1].

Theorem 1.1.6. *Every matrix $A \in \mathbb{R}^{m\times n}$ can be expressed in the form*

$$A = U\Sigma V^T, \tag{1.10}$$

where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal, Σ is an $m \times n$ matrix with entry σ_i in the i-th row and i-th column for $i \in \{1, \dots, \operatorname{rank}(A)\}$ and zero for all the other entries. Here, $\sigma_i > 0$ for all $i = 1, \dots, \operatorname{rank}(A)$.

Proof. Without loss of generality, assume $n \geq m$ (otherwise, we decompose $A^T = V\Sigma U^T$ and obtain (1.10) by transposing A^T). The matrix A^TA is symmetric and thus orthogonally diagonalizable, that is

$$A^TA = VDV^T, \tag{1.11}$$

where the columns v_i of V are the orthonormal unit eigenvectors corresponding to the eigenvalues λ_i of A^TA. D is the diagonal matrix $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. From Lemma 1.1.5 it follows that $\operatorname{rank}(A) = \operatorname{rank}(A^TA) = \operatorname{rank}(D) = r$. Here, $D = \operatorname{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0)$ with $n-r$ zeros, where V and D are chosen such that

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0.$$

Let $\{u_1, \dots, u_r\}$ be a set of column vectors such that

$$u_i = Av_i, \qquad i \in \{1, \dots, r\}. \tag{1.12}$$

Then it follows that

$$\begin{aligned}
\langle u_i, u_j\rangle &= u_j^T u_i \\
&= (Av_j)^T Av_i \\
&= v_j^T A^T A v_i \\
&= v_j^T \lambda_i v_i \\
&= \lambda_i v_j^T v_i \\
&= \begin{cases} 0 & \text{if } i \neq j, \\ \lambda_i & \text{if } i = j. \end{cases}
\end{aligned}$$

Hence,

$$||Av_i|| = \sqrt{\lambda_i} \qquad i \in \{1, \dots, r\}. \tag{1.13}$$

By Lemma 1.1.3, the eigenvalues are non-negative, and $\sqrt{\lambda_i}$ is the non-negative square-root, so $\sigma_i := \sqrt{\lambda_i} > 0$ for $i = 1, \dots, r$. Normalizing (1.12), it follows that

$$u_i = \frac{Av_i}{||Av_i||} = \frac{Av_i}{\sqrt{\lambda_i}}$$

so that $u_i\sqrt{\lambda_i} = Av_i$. The remaining columns $u_{r+1}, \ldots, u_m$ may be chosen such that U is orthogonal. Let Σ be the $m \times n$ matrix with entry σ_i in the i-th row and i-th column for $i \in \{1, \ldots, \text{rank}(A)\}$ and zero for all the other entries. Then,

$$A = U\Sigma V^T. \qquad \square$$

Corollary 1.1.7. *Let $A \in \mathbb{R}^{m\times n}$, then $AA^Tu_i = \lambda u_i$, where $u_i \in U$ are the eigenvectors of AA^T as defined in Theorem 1.1.6.*

Proof. Using (1.12), it follows that

$$\begin{aligned} AA^Tu_i &= AA^TAv_i \\ &= A\lambda_i v_i \\ &= \lambda_i Av_i \\ &= \lambda_i u_i. \end{aligned}$$

Therefore A^TA and AA^T share the same eigenvalues. $\square$

Remark 1.1.8. The singular value decomposition works for complex matrices, but instead of the regular transpose, we use the conjugate transpose.

1.1.3 Moore-Penrose pseudoinverse

Definition 1.1.9. Let $A \in \mathbb{F}^{m\times n}$ and $A^+ \in \mathbb{F}^{n\times m}$. If the following conditions

1. $AA^+A = A$,
2. $A^+AA^+ = A^+$,
3. $(AA^+)^* = AA^+$,
4. $(A^+A)^* = A^+A$,

are satisfied, then A^+ is called the Moore-Penrose pseudoinverse (or simply the Moore-Penrose inverse) of A. The Moore-Penrose inverse generalizes the inverse of a matrix.

The Moore-Penrose inverse always exists given that $A \in \mathbb{R}^{m\times n}$ (or $\mathbb{C}^{m\times n}$), this follows from the fact that $A \in \mathbb{R}^{m\times n}$ can be decomposed such that $A = U\Sigma V^T$.

In the following section the terms "More-Penrose inverse" and "More-Penrose pseudoinverse" are used interchangeably.

Lemma 1.1.10. *Let $A \in \mathbb{F}^{m\times n}$, the More-Penrose inverse is unique.*

Proof. Suppose that $A_1^+$and A_2^+ are Moore–Penrose inverses of A then

$$\begin{aligned}
A_1^+ &= A_1^+ A A_1^+ \\
&= A_1^+ (A A_2^+ A) A_1^+ \\
&= A_1^+ ((A A_2^+ A) A_2^+ A) A_1^+ \\
&= (A_1^+ A)^* (A_2^+ A)^* (A_2^+ A) A_1^+ \\
&= (A_2^+ A A_1^+ A)^* (A_2^+ A) A_1^+ \\
&= A_2^+ (A A_2^+)^* (A A_1^+)^* \\
&= A_2^+ (A A_1^+ A A_2^+)^* \\
&= A_2^+ (A A_2^+)^* \\
&= A_2^+ (A A_2^+) \\
&= A_2^+.
\end{aligned}$$

□

Lemma 1.1.11. *For an invertible matrix A, the Moore-Penrose pseudoinverse is the matrix inverse A^{-1} of A. That is, $A^{-1} = A^+$.*

Proof. Suppose that the inverse of A exists, then

$$\begin{aligned}
A^{-1} &= A^{-1} A A^{-1} \\
&= A^{-1} (A A^+ A) A^{-1} \\
&= (A^{-1} A) A^+ (A A^{-1}) \\
&= A^+.
\end{aligned}$$

□

Lemma 1.1.12. *Let $A \in \mathbb{C}^{m\times n}$ and let $A = U\Sigma V^*$ be its singular value decomposition. Then*

$$A^+ = V\Sigma^+ U^*, \tag{1.14}$$

where Σ is the $m \times n$ matrix with entry σ_i in the i-th row and i-th column for $i \in \{1, \dots, \mathrm{rank}(A)\}$ and zero for all the other entries and Σ^+ is the $n \times m$ matrix with entry $\frac{1}{\sigma_i}$ in the i-th column and i-th row and zero for all the other entries.

Proof. We want to show that the axioms in Definition 1.1.9 hold. Straight forward calculation yields that $\Sigma^+\Sigma = \mathrm{diag}(I_r, 0_{n-r})$ and $\Sigma\Sigma^+ = \mathrm{diag}(I_r, 0_{m-r})$, where 0_{n-r} denotes the $(n-r)\times(n-r)$ matrix and I_r denotes the $r\times r$ identity matrix. It follows by direct calculation that $\Sigma\Sigma^+\Sigma=\Sigma$ and $\Sigma^+\Sigma\Sigma^+=\Sigma^+$.

1. First we want to show that $AA^+A=A$.

$$\begin{aligned} AA^+A &= (U\Sigma V^*V\Sigma^+U^*U\Sigma V^*) \\ &= U\Sigma V^* \\ &= A. \end{aligned}$$

2. We will prove that $A^+AA^+=A^+$.

$$\begin{aligned} A^+AA^+ &= V\Sigma^+U^*U\Sigma V^*V\Sigma^+U^* \\ &= V\Sigma^+U^* \\ &= A^+. \end{aligned}$$

3. We will show that AA^+ is self adjoint, that is, $(AA^+)^*=AA^+$.

$$\begin{aligned} (AA^+)^* &= (U\Sigma V^*V\Sigma^+U^*)^* \\ &= U(\Sigma\Sigma^+)^*U^* \\ &= U\Sigma\Sigma^+U^* \\ &= U\Sigma V^*V\Sigma^+U^* \\ &= AA^+. \end{aligned}$$

4. We will show that $(A^+A)^*=A^+A$.

$$\begin{aligned} (A^+A)^* &= (V\Sigma^+U^*U\Sigma V^*)^* \\ &= V(\Sigma^+\Sigma)^*V^* \\ &= V\Sigma^+\Sigma V^* \\ &= V\Sigma^+U^*U\Sigma V^* \\ &= A^+A. \end{aligned}$$ $\square$

Next we will define products of higher-order tensors and investigate the properties of tensors from such a structure. First, let us introduce the tensor notation.

1.2 Tensor notation

Throughout this dissertation we will assume that all the vector spaces are finite dimensional. In this section and subsequent sections we will follow the notation, definitions and theory as described in [12, Chapter 5]. Tensors may be represented as a multidimensional arrays, or N-way arrays. The concept of tensors is a generalization of vectors and matrices, which are order-one and order-two tensors respectively. Let $\mathbb{F}^{n_1}$ denote an n_1-dimensional vector space over the field $\mathbb{F}$. An m-th order tensor $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2 \times \cdots \times n_m}$ [3], is given by its entries $a_{i_1 i_2 \cdots i_m}$ and may be written in the form

$$\mathcal{A} = \sum a_{i_1 i_2 \cdots i_m} e_{i_1} \otimes \cdots \otimes e_{i_m}, \quad i_j \in \{1, \ldots, n_m\}, j \in \{1, \ldots, m\}, \tag{1.15}$$

where $\{e_1, \ldots, e_n\}$ is the standard basis in $\mathbb{F}^n$ and $e_{i_1} \otimes \cdots \otimes e_{i_m}$ denotes a simple tensor in the tensor product $\mathbb{F}^{n_1 \times \cdots \times n_m} \cong \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_m}$ (defined in (1.16)). Additionally, if $n_1 = n_2 = \cdots = n_m = n$, then we say that $\mathcal{A}$ is an n-dimensional tensor.

1.3 Tensor product

Tensor products play a significant role in multilinear algebra. To define the Kronecker product, first we will first describe tensor products of vector spaces.

Definition 1.3.1. Let $V_1, V_2 \ldots V_m$ and N be vector spaces, then the function $g : V_1 \times V_2 \times \cdots \times V_m \to N$ is a multilinear function if g is linear separately in each component of the vector space $V_1 \times V_2 \times \cdots \times V_m$. In other words, g is a multilinear function if

$$g(v_1, \ldots, a_i + b_i, \ldots, v_m) = g(v_1, \ldots, a_i, \ldots, v_m) + g(v_1, \ldots, b_i, \ldots, v_m),$$

and

$$g(v_1, v_2, \ldots, \alpha a_i, \ldots, v_m) = \alpha g(v_1, v_2, \ldots, a_i, \ldots, v_m),$$

where $V_1 \times V_2 \times \cdots \times V_m = \{(v_1, v_2, \ldots, v_m) : v_i \in V_i, 1 \leq i \leq m\}$ is the Cartesian product of vector spaces.

The Universal Factorization Property (UFP) will aid us in our definition of the tensor product. Moreover, we will use the UFP to show that the tensor product is not uniquely defined.

Definition 1.3.2. (**Universal factorization property** [12])
Let $\mathbf{V} = (V_1, V_2, \ldots, V_m)$ be a sequence of vector spaces, $\mathbb{T}$ a vector space and $\phi : V_1 \times V_2 \times \cdots \times V_m \to \mathbb{T}$ a multilinear function. The pair $(\mathbb{T}, \phi)$ satisfies the universal factorization property if for every vector space K and multilinear function $f : V_1 \times V_2 \times \cdots \times V_m \to K$, there exists a unique linear function $h : \mathbb{T} \to K$, such that $f = h \circ \phi$. The following commutative diagram demonstrates the universal factorization property.

$$\begin{tikzcd}
V_1 \times V_2 \times \cdots \times V_m \arrow[r, "f"] \arrow[d, "\phi"] & K \\
\mathbb{T} \arrow[ur, "h"] &
\end{tikzcd}$$

Figure 1.1

Definition 1.3.3. Suppose that $(\mathbb{S}, x)$ and $(\mathbb{T}, y)$ satisfy the universal factorization property for $V_1, V_2, \ldots, V_m$. The pair $(\mathbb{S}, x)$ and $(\mathbb{T}, y)$ are *isomorphic* if there exists an invertible linear transformation $T : \mathbb{S} \to \mathbb{T}$ such that $y = Tx$.

Theorem 1.3.4. *Suppose that $(\mathbb{S}, x)$ and $(\mathbb{T}, y)$ satisfy the universal factorization property for $V_1, V_2, \ldots, V_m$. If the reach of x and y is all of $\mathbb{S}$ and $\mathbb{T}$, respectively, then $(\mathbb{S}, x)$ and $(\mathbb{T}, y)$ are isomorphic.*

Proof. Using the following commutative diagram

$$\begin{tikzcd}
V_1 \times V_2 \times \cdots \times V_m \arrow[r, "x"] \arrow[d, "y"] & \mathbb{S} \arrow[dl, shift left, "k"] \\
\mathbb{T} \arrow[ur, shift left, "h"] &
\end{tikzcd}$$

Figure 1.2

where h and k are unique maps, it follows that $x = h \circ y$ and $y = k \circ x$. Thus,

$$\begin{aligned}
x &= h \circ y \\
&= h \circ (k \circ x) \\
&= (h \circ k) \circ x.
\end{aligned}$$

Since the reach of x is all of $\mathbb{S}$, $h \circ k$ is the identity of $\mathbb{S}$. Similarly, $y = k \circ (h \circ y)$ and so $k \circ h$ is the identity of $\mathbb{T}$. Thus

$$(\mathbb{S}, x) \cong (\mathbb{T}, y).$$

□

Theorem 1.3.5. *Let $y : V_1 \times V_2 \times \cdots \times V_m \to \mathbb{T}$, if $(\mathbb{T}, y)$ satisfies the universal factorization property, then* $\operatorname{span}(y(V_1 \times V_2 \times \cdots \times V_m)) = \mathbb{T}$.

Proof. Suppose that $\operatorname{span}(y(V_1 \times V_2 \times \cdots \times V_m)) \neq \mathbb{T}$, then there exists a $t \in \mathbb{T}$, such that $t \notin \operatorname{span}(y(V_1 \times V_2 \times \cdots \times V_m))$ for all $(v_1, \ldots, v_m) \in V_1 \times V_2 \times \cdots \times V_m$.

Clearly $t \neq 0$. We must have

$$\text{span}(y(V_1 \times V_2 \times \cdots \times V_m)) \cap \text{span}\{t\} = \{0\}.$$

Suppose that $\alpha : \mathbb{T} \to \mathbb{T}$ is the zero map, $\beta : \mathbb{T} \to \mathbb{T}$ is the projection onto $\text{span}\{t\}$ and $f : V_1 \times V_2 \times \cdots \times V_m \to \mathbb{T}$ is the multilinear zero map.

By the UFP there exists $h : \mathbb{T} \to \mathbb{T}$ such that $h \circ y = f$. From the maps α and β it follows that

$$\alpha \circ y = f \quad \text{and} \quad \beta \circ y = f \ (\text{since span}(y(V_1 \times V_2 \times \cdots \times V_m)) \subseteq \text{span}\{t\}^{\perp}).$$

Therefore

$$\alpha \circ y = \beta \circ y = f,$$

but $\alpha \neq \beta$ (since $\alpha = \beta \iff t = 0$) which contradicts uniqueness, thus $\text{span}(y(V_1 \times V_2 \times \cdots \times V_m)) = \mathbb{T}$. □

Definition 1.3.6. (**Tensor products** [12])
If $(\mathbb{T}, \phi)$ satisfies the *universal factorization property* for $\mathbf{V}$, then $\mathbb{T}$ is the tensor product of $\mathbf{V}$ and we write,

$$\mathbb{T} = V_1 \otimes V_2 \otimes \cdots \otimes V_m. \tag{1.16}$$

1.4 Kronecker product

Let $\mathbb{F}^{m\times n}$ and $\mathbb{F}^{p\times q}$ denote the vector space of $m \times n$ and $p \times q$ (respectively) matrices over the field $\mathbb{F}$.

Definition 1.4.1. Let A be an $m \times n$ matrix over a field $\mathbb{F}$ and let B be a $p \times q$ matrix over the same field $\mathbb{F}$. The Kronecker product of the matrix A and B is an $mp \times nq$ matrix [15], defined by

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}. \tag{1.17}$$

The use of the symbol $\otimes$ is sensible, since the map $(A, B) \to A \otimes B$ defines a tensor product $\mathbb{F}^{m\times n} \times \mathbb{F}^{p\times q} \to \mathbb{F}^{m\times n} \otimes \mathbb{F}^{p\times q}$. To show that this is true let us consider the following lemma

Lemma 1.4.2. *Let $f : \mathbb{F}^m \times \mathbb{F}^n \to K$ be a bilinear map and $\phi : \mathbb{F}^m \times \mathbb{F}^n \to \mathbb{F}^{mn}$ be the Kronecker product. Then there exists a unique linear map $h : \mathbb{F}^{mn} \to K$ such that $h \circ \phi = f$.*

Proof. Let $e_i \in \mathbb{F}^m$ and $e_j \in \mathbb{F}^n$ denote the standard bases in $\mathbb{F}^m$ and $\mathbb{F}^n$ respectively, where $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.
First, we need to show that the set

$$S = \{e_i \otimes e_j : 1 \leq i \leq m,\ 1 \leq j \leq n\}$$

is a basis for $\mathbb{F}^{mn}$. Suppose that the linear combination of $e_i \otimes e_j$ is zero,

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}(e_i \otimes e_j) = 0, \tag{1.18}$$

we can write (1.18) as

$$\sum_{i=1}^{m} \sum_{j=1}^{n} (a_{ij} e_i) \otimes e_j = 0.$$

Since, e_i and e_j are linearly independent in $\mathbb{F}^m$ and $\mathbb{F}^n$ respectively, we must have $a_{ij} = 0$. Thus, the set S is linearly independent. Given an arbitrary vector $b \in \mathbb{F}^{mn}$, we can write b as a linear combination of $e_i \otimes e_j$. That is,

$$b = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}(e_i \otimes e_j),$$

this shows that the set S spans $\mathbb{F}^{mn}$. Thus, the set S is a basis for $\mathbb{F}^{mn}$. By the definition of ϕ we have $\phi(e_i, e_j) = e_i \otimes e_j = e_{(i-1)n+j}$ and ϕ is a bijection between $\{(e_i, e_j) : 1 \leq i \leq m,\ 1 \leq j \leq n\}$ and $\{e_k : 1 \leq k \leq mn\}$. Thus $h(e_{(i-1)n+j}) = f(e_i, e_j)$ defines a unique linear map. Therefore $e_i \otimes e_j$ is a tensor product. $\square$

Let $A \in \mathbb{F}^{s\times m}$ and $B \in \mathbb{F}^{t\times n}$ we will show that the map $(A, B) \to A \otimes B$ defines a tensor product $\mathbb{F}^{m\times n} \times \mathbb{F}^{p\times q} \to \mathbb{F}^{m\times n} \otimes \mathbb{F}^{p\times q}$.

$$\begin{tikzcd}
\mathbb{F}^m \times \mathbb{F}^n \arrow[r, "(A,B)"] \arrow[d, "\otimes"'] \arrow[dr, "\psi\circ(A,B)"] & \mathbb{F}^s \times \mathbb{F}^t \arrow[d, "\psi"] \\
\mathbb{F}^m \otimes \mathbb{F}^n \arrow[r, "A\otimes B"'] & \mathbb{F}^s \otimes \mathbb{F}^t
\end{tikzcd}$$

Let $\psi : \mathbb{F}^s \times \mathbb{F}^t \to \mathbb{F}^s \otimes \mathbb{F}^t$ be the Kronecker product and $(A, B) : \mathbb{F}^m \times \mathbb{F}^n \to \mathbb{F}^s \times \mathbb{F}^t$ a linear map defined by $(A, B)(a, b) = (Aa, Bb)$. Then $\psi \circ (A, B)$ is a bilinear map since for $a_1 \in \mathbb{F}^n$, $a_2 \in \mathbb{F}^m$ and scalars α, $\beta \in \mathbb{F}$

$$\begin{aligned}
\psi \circ (A, B)(\alpha a_1 + \beta a_2, b) &= \psi(A(\alpha a_1 + \beta a_2), Bb) \\
&= \psi(\alpha A a_1 + \beta A a_2, Bb) \\
&= \alpha\psi(Aa_1, Bb) + \beta\psi(Aa_2, Bb) \\
&= \alpha\psi \circ (A, B)(a_1, b) + \beta\psi \circ (A, B)(a_2, b).
\end{aligned}$$

Similarly, $\psi \circ (A, B)(a, \alpha b_1 + \beta b_2) = \alpha\psi \circ (A, B)(a, b_1) + \beta\psi \circ (A, B)(a, b_2)$. Since $\psi \circ (A, B)$ is bilinear, $A \otimes B$ is the unique linear map such that

$$\psi \circ (A, B) = (A \otimes B) \circ \otimes,$$

then by the universal factorization property the map $A \otimes B$ is uniquely defined by (A, B). Now consider the diagram

$$\begin{tikzcd}
\mathbb{F}^{s\times m} \times \mathbb{F}^{t\times n} \arrow[r, "(a,b)"] \arrow[d, "k"] \arrow[dr, "\psi\circ(a,b)"] \arrow[dd, bend right, "f"'] & \mathbb{F}^s \times \mathbb{F}^t \arrow[d, "\psi"] \\
\mathbb{F}^{st\times mn} \arrow[r, "h"'] \arrow[d] & \mathbb{F}^s \otimes \mathbb{F}^t \\
V &
\end{tikzcd}$$

where ψ is the Kronecker product, $a \in \mathbb{F}^m$, $b \in \mathbb{F}^n$ and $(a, b) : \mathbb{F}^{s\times m} \times \mathbb{F}^{t\times n} \to \mathbb{F}^s \times \mathbb{F}^t$ is a linear map given by $(a, b)(A, B) = (Aa, Bb)$. Then $\psi \circ (a, b)$ is bilinear since for $A \in \mathbb{F}^{s\times m}$, $B \in \mathbb{F}^{s\times m}$,$C \in \mathbb{F}^{t\times n}$ and $\alpha, \beta \in \mathbb{F}$

$$\begin{aligned}
\psi \circ (a, b)(\alpha A + \beta B, C) &= \psi \circ ((\alpha A + \beta B)a, Cb) \\
&= \psi(\alpha Aa + \beta Ba, Cb) \\
&= \psi(\alpha Aa, Cb) + \psi(\beta Ba, Cb) \\
&= \alpha\psi(Aa, Cb) + \beta\psi(Ba, Cb) \\
&= \alpha\psi \circ (a, b)(A, C) + \beta\psi \circ (a, b)(B, C).
\end{aligned}$$

Similarly, $\psi \circ (a, b)(C, \alpha A + \beta B) = \alpha\psi \circ (a, b)(C, A) + \beta\psi \circ (a, b)(C, B)$. The composition, $\psi \circ (a,b)(A, B)$, gives us

$$\begin{aligned}
\psi \circ (a,b)(A, B) &= \psi(Aa, Bb) \\
&= (Aa) \otimes (Bb).
\end{aligned}$$

To show that the map k is the Kronecker product of matrices let $a = e_i$, $b = e_j$

and noting that $\psi((A,B)(a,b)) = \psi((a,b)(A,B))$ so that

$$(A \otimes B)(a \otimes b) = (Aa) \otimes (Bb),$$

thus

$$(A \otimes B)(e_i \otimes e_j) = (Ae_i) \otimes (Be_j).$$

Hence, using Lemma 1.4.2 and the fact that

$$\begin{pmatrix} a_1 \\ \vdots \\ a_s \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ \vdots \\ b_t \end{pmatrix} = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 \\ \vdots \\ \vdots \\ a_s b_t \end{pmatrix},$$

it follows that

$$\begin{aligned} A \otimes B &= (Ae_1 \otimes Be_1 \ Ae_1 \otimes Be_2 \cdots Ae_m \otimes Be_n) \\ &= \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & \cdots & \cdots & a_{1m}b_{1n} \\ a_{11}b_{21} & a_{11}b_{22} & \cdots & \cdots & a_{1m}b_{2n} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{s1}b_{t1} & a_{s1}b_{t2} & \cdots & \cdots & a_{sm}b_{tn} \end{pmatrix}. \end{aligned}$$

Therefore k is the Kronecker product. Now we want to establish a bijection similar to Lemma 1.4.2. Let $E_{ij} = e_i e_j^T$, it follows that

$$h(E_{ij} \otimes E_{kl}) = (E_{ij}a) \otimes (E_{kl}b).$$

Thus, h is uniquely defined (by its action on the standard basis).

Example 1.4.3. Suppose that $A = \begin{pmatrix} 2 & 5 \\ 4 & 3 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ and $B = \begin{pmatrix} 3 & 7 \\ 8 & 9 \end{pmatrix} \in \mathbb{R}^{2\times 2}$, then

$$A \otimes B = \begin{pmatrix} 2\begin{pmatrix} 3 & 7 \\ 8 & 9 \end{pmatrix} & 5\begin{pmatrix} 3 & 7 \\ 8 & 9 \end{pmatrix} \\ 4\begin{pmatrix} 3 & 7 \\ 8 & 9 \end{pmatrix} & 3\begin{pmatrix} 3 & 7 \\ 8 & 9 \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} 6 & 14 & 15 & 35 \\ 16 & 18 & 40 & 45 \\ 12 & 28 & 9 & 21 \\ 32 & 36 & 24 & 27 \end{pmatrix}.$$

We now present some properties of the Kronecker product (see for example [15]). For $A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{p\times q}$, $C \in \mathbb{F}^{r\times s}$ and $\alpha \in \mathbb{F}$,

(a) multiplication of the Kronecker product by a scalar,

$$(\alpha A) \otimes B = A \otimes (\alpha B) = \alpha(A \otimes B),$$

(b) transpose of the Kronecker product,

$$(A \otimes B)^T = A^T \otimes B^T,$$

(c) the Kronecker product is associative,

$$(A \otimes B) \otimes C = A \otimes (B \otimes C),$$

(d) complex conjugate (when $\mathbb{F} = \mathbb{C}$) of the Kronecker product,

$$(A \otimes B)^* = A^* \otimes B^*,$$

(e) left and right distributive law of the Kronecker product,

$$A \otimes (B + C) = A \otimes B + A \otimes C,$$

$$(A + B) \otimes C = A \otimes C + B \otimes C,$$

(f) interchange law of the Kronecker product,

$$(A \otimes B)(C \otimes D) = AC \otimes BD.$$

The following chapter describes a matrix-tensor product (the modal product) and its calculation in terms of Kronecker products. The modal product follows from the composition of a multilinear map and a linear map. The resulting product will be between a higher-order tensor and a second-order tensor.

Chapter 2

Modal product

The simplest multiplication of a tensor by a matrix is called a modal product. To understand this product let us first consider the *matrix-matrix product*, which has two modes (rows and columns). To see this, let A be a matrix defined by

$$A = \sum_{ij} a_{ij} e_i e_j^T. \tag{2.1}$$

Now let B be a matrix, the product of B and the first mode of A is

$$\begin{aligned} B \times_1 A &= \sum_{ij} a_{ij} \left(B e_i\right) e_j^T \\ &= B \sum_{ij} a_{ij} e_i e_j^T \\ &= BA. \end{aligned}$$

The product of B and the second mode of A is

$$\begin{aligned} B \times_2 A &= \sum_{ij} a_{ij} e_i \left(B e_j\right)^T \\ &= \sum_{ij} a_{ij} e_i e_j^T B^T \\ &= AB^T. \end{aligned}$$

Thus, for matrices, modal products are right and left multiplication. We can also see this by composing three linear functions. Let A, B, C and D be finite dimensional vector spaces, such that

$$A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D,$$

Figure 2.1

and let $[f], [g]$ and $[h]$ be matrix representations of f, g and h, respectively. The matrix representation of the composition of f, g and h is

$$[h \circ g \circ f] = [h][g][f]. \tag{2.2}$$

From (2.2) we see that there are two modes of multiplication with $[g]$, i.e, $[h][g]$ and $[g][f]$. Modal products are simply matrix products, with a matrix unfolding of a tensor. Higher-order tensor-tensor multiplications reduce to normal matrix multiplication. Matrix unfolding is a matrix representation of a tensor. Let us first discuss the column space of the matrix unfolding. We will do this by considering the following example.

Example 2.0.1. Consider $\mathbb{R}^2 \otimes \mathbb{R}^3$. Let $e_i \in \mathbb{R}^2$ denote the elements of the standard basis in $\mathbb{R}^2$ and similarly for $e_j \in \mathbb{R}^3$. When we write $e_1 \otimes e_2 \in \mathbb{R}^2 \otimes \mathbb{R}^3$, we mean $e_1 \in \mathbb{R}^2$ and $e_2 \in \mathbb{R}^3$ (although, in general, this notation is ambiguous). Then

$$e_1 \otimes e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad e_1 \otimes e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad e_1 \otimes e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$e_2 \otimes e_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 \otimes e_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_2 \otimes e_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

where $e_i \in \mathbb{R}^2$, $e_j \in \mathbb{R}^3$ and i indicates whether we are in the first or second half of the tensor and j indicates the position of the nonzero element. Thus $e_i \otimes e_j = e_{3(i-1)+j}$. We can generalize this for $e_{i_1} \in \mathbb{F}^{I_1}, e_{i_2} \in \mathbb{F}^{I_2}, \ldots, e_{i_n} \in \mathbb{F}^{I_n}$

$$e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_n} = e_{(i_1-1)I_2 \cdots I_n + \cdots + (i_2-1)I_3 \cdots I_n + \cdots + i_n}.$$

2.1 Matrix unfolding

Tensor unfolding is the process of flattening a tensor into a matrix. The goal of flattening a tensor is to use the algebraic properties of a matrix for a higher order tensor. We define matrix unfolding for higher-order tensors [4], as follows.

Definition 2.1.1. Assume an m-th order tensor $\mathcal{A} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_m}$. The n-th unfolding of $\mathcal{A}$, is the matrix $A_{(n)} \in \mathbb{F}^{I_n \times (I_{n+1} I_{n+2} \ldots I_m I_1 I_2 \ldots I_{n-1})}$, where

$$A_{(n)} = \sum_{(i_1,\ldots,i_m) \in I_1 \times \cdots \times I_m} a_{i_1 \ldots i_m} e_{i_n} (e_{i_{n+1}} \otimes e_{i_{n+2}} \otimes \cdots \otimes e_{i_m} \otimes e_{i_1} \otimes \cdots \otimes e_{i_{n-1}})^T, \tag{2.3}$$

and contains the element $a_{i_1 i_2 \ldots i_m}$ at the row position i_n and column position

$$(i_{n+1} - 1) I_{n+2} I_{n+3} \ldots I_m I_1 I_2 \ldots I_{n-1} + (i_{n+2} - 1) I_{n+3} I_{n+4} \ldots I_m I_1 I_2 \ldots I_{n-1} + \cdots + (i_m - 1) I_1 I_2 \ldots I_{n-1} + (i_1 - 1) I_2 I_3 \ldots I_{n-1} + (i_2 - 1) I_3 I_4 \ldots I_{n-1} \cdots + i_{n-1}.$$

Example 2.1.2. Define a tensor $\mathcal{A} \in \mathbb{R}^{2 \times 2 \times 3}$ by $a_{111} = a_{223} = a_{212} = a_{221} = 1$, $a_{121} = a_{122} = a_{213} = 0$, $-a_{112} = a_{211} = 3$, $-a_{113} = a_{222} = a_{123} = 2$. Thus the matrix unfoldings $A_{(1)} \in \mathbb{R}^{2 \times 6}$, $A_{(2)} \in \mathbb{R}^{2 \times 6}$ and $A_{(3)} \in \mathbb{R}^{3 \times 4}$ are given by

$$A_{(1)} = \left(\begin{array}{ccc|ccc} 1 & -3 & -2 & 0 & 0 & 2 \\ 3 & 1 & 0 & 1 & 2 & 1 \end{array}\right), \quad A_{(2)} = \left(\begin{array}{cc|cc|cc} 1 & 0 & -3 & 0 & -2 & 2 \\ 3 & 1 & 1 & 2 & 0 & 1 \end{array}\right),$$

$$A_{(3)} = \left(\begin{array}{cc|cc} 1 & 0 & 3 & 1 \\ -3 & 0 & 1 & 2 \\ -2 & 2 & 0 & 1 \end{array}\right).$$

We have established that matrix products have two modes of multiplication. Similarly, m-th order tensors have m-modes of multiplication. The general modal product is defined in terms of the mode of multiplication [4]. We will now define a operation that transforms the unfolded tensor (second-order tensor) back into a higher order tensor.

Definition 2.1.3. Let $A_{(n)} \in \mathbb{F}^{I_n \times (I_{n+1} I_{n+2} \ldots I_m I_1 I_2 \ldots I_{n-1})}$ be the n-th unfolding of $\mathcal{A} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_m}$, then matrix folding is the operation

$$\text{fold}(A_{(n)}) = \mathcal{A}. \tag{2.4}$$

where $\mathcal{A} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_m}$.

2.2 The n-mode product

The tensor-matrix product (modal product) is the product between a tensor of order $m \geq 2$ and a matrix (a tensor of order $k = 2$). We have shown that a matrix has two modes of multiplication and that the matrix-matrix product is compositional. For higher-order tensors, let us first consider an example of a trilinear map (fourth-order tensor). Suppose that A, B, C, P, Q, W, U, and V are vector spaces such that, for a trilinear map $g : A \times B \times C \to P$, linear maps $f : U \to B,\ k : W \to A,\ h : V \to C$ and $y : P \to Q$

$$\begin{tikzcd}
 & & & U \arrow[d, "f"] & & \\
 & & & B & & \\
W \times U \times V \arrow[r] \arrow[uurrr] \arrow[ddrrr] \arrow[rrr, dashed, bend left] & W \arrow[r, "k"] & A & A \times B \times C \arrow[l, "\pi_1"] \arrow[u, "\pi_2"'] \arrow[d, "\pi_3"] \arrow[r, "g"] & P \arrow[r, "y"] & Q \\
 & & & C & & \\
 & & & V \arrow[u, "h"] & &
\end{tikzcd}$$

Figure 2.2

and $\pi_1(a,b,c) = a,\ \pi_2(a,b,c) = b,\ \pi_3(a,b,c) = c$. From the above diagram, the representation tensor, $[g]$, is multiplied by the representation matrices $[f], [k], [h]$ and $[y]$, we have four modes, namely, up, down, left and right. The following is a compact form of the diagram above,

$$\begin{tikzcd}
 & U & \\
W & W \times U \times V \arrow[l] \arrow[u] \arrow[d] \arrow[r, "y \circ g \circ (f,k,h)"] & Q \\
 & V &
\end{tikzcd}$$

Figure 2.3

Therefore, fourth-order tensors have 4-modes of multiplication and the above diagram demonstrates that the tensor-matrix multiplication (for higher-order tensors) is compositional.

For arbitrary-order tensors let us consider the following definition of the modal product.

Definition 2.2.1. The n-mode product of a tensor $\mathcal{A} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_m}$ by a matrix $U \in \mathbb{F}^{J_n \times I_n}$ is an $(I_1 \times I_2 \times \cdots \times I_{n-1} \times J_n \times I_{n+1} \times \cdots \times I_m)$-dimensional tensor, given by

$$\mathcal{A} \times_n U = \sum_{i_1,\ldots,i_m} a_{i_1 \ldots i_m} e_{i_1} \otimes \cdots \otimes e_{i_{n-1}} \otimes U e_{i_n} \otimes e_{i_{n+1}} \otimes \cdots \otimes e_{i_m}. \tag{2.5}$$

Proposition 2.2.2. *The modal product satisfies the following properties.*

(1) For a tensor $\mathcal{A} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_m}$ and matrices $A \in \mathbb{F}^{J_n \times I_n}$ and $B \in \mathbb{F}^{J_p \times I_p}$, where $p \neq n$ and $1 \leq p, n \leq m$

$$(\mathcal{A} \times_n A) \times_p B = (\mathcal{A} \times_p B) \times_n A = \mathcal{A} \times_n A \times_p B. \tag{2.6}$$

(2) For a tensor $\mathcal{A} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_m}$ and matrices $A \in \mathbb{F}^{J_n \times I_n}$ and $B \in \mathbb{F}^{K_n \times J_n}$

$$(\mathcal{A} \times_n A) \times_n B = \mathcal{A} \times_n (BA). \tag{2.7}$$

Proof. ***(1)***

$$\begin{aligned}
&(\mathcal{A} \times_n A) \times_p B \\
&\quad = \left(\sum_{i_1, \ldots, i_m} a_{i_1 \ldots i_m} e_{i_1} \otimes \cdots \otimes A e_{i_n} \otimes \cdots \otimes e_{i_m} \right) \times_p B \\
&\quad = \sum_{i_1, \ldots, i_m} a_{i_1 \ldots i_m} e_{i_1} \otimes \cdots \otimes A e_{i_n} \otimes \cdots \otimes B e_{i_p} \otimes \cdots \otimes e_{i_m} \\
&\quad = \left(\sum_{i_1, \ldots, i_m} a_{i_1 \ldots i_m} e_{i_1} \otimes \cdots \otimes B e_{i_p} \otimes \cdots \otimes e_{i_m} \right) \times_n A \\
&\quad = (\mathcal{A} \times_p B) \times_n A.
\end{aligned}$$

(2)

$$\begin{aligned}
&(\mathcal{A} \times_n A) \times_n B \\
&\quad = \left(\sum_{i_1, \ldots, i_m} a_{i_1 \ldots i_m} e_{i_1} \otimes \cdots \otimes A e_{i_n} \otimes \cdots \otimes e_{i_m} \right) \times_n B \\
&\quad = \left(\sum_{i_1, \ldots, i_m} a_{i_1 \ldots i_m} \sum_{j_n} (A)_{j_n i_n} e_{i_1} \otimes \cdots \otimes e_{j_n} \otimes \cdots \otimes e_{i_m} \right) \times_n B \\
&\quad = \left(\sum_{i_1, \ldots, i_m} a_{i_1 \ldots i_m} \sum_{j_n} (A)_{j_n i_n} e_{i_1} \otimes \cdots \otimes B e_{j_n} \otimes \cdots \otimes e_{i_m} \right) \\
&\quad = \sum_{i_1, \ldots, i_m} a_{i_1 \ldots i_m} e_{i_1} \otimes \cdots \otimes B \sum_{j_n} (A_{j_n i_n}) e_{j_n} \otimes \cdots \otimes e_{i_m} \\
&\quad = \sum_{i_1, \ldots, i_m} a_{i_1 \ldots i_m} e_{i_1} \otimes \cdots \otimes B A e_{i_n} \otimes \cdots \otimes e_{i_m} \\
&\quad = \mathcal{A} \times_n (BA).
\end{aligned}$$

$\square$

2.3 Rank of a higher-order tensor

A vector space may be a tensor product in more than one way and each tensor product has an associated rank.

Tensor products are not unique, therefore, there are multiple ways of representing tensor products due to the following isomorphism

$$\begin{aligned} \mathbb{F}^m \otimes \mathbb{F}^n \cong \mathbb{F}^{mn} \cong \mathbb{F}^n \otimes \mathbb{F}^m, \\ (\mathbb{F}^m \otimes \mathbb{F}^n) \otimes \mathbb{F}^p \cong \mathbb{F}^m \otimes (\mathbb{F}^n \otimes \mathbb{F}^p), \\ \mathbb{F}^m \otimes \mathbb{F}^p \cong \mathbb{F}^{m \times p}. \end{aligned} \tag{2.8}$$

Example 2.3.1. Using (2.8), let $m = n = p = 2$, so $\mathcal{A} \in \mathbb{F}^{2\times2\times2}$. If $a_{111} = a_{122} = a_{212} = a_{221} = a_{222} = 0,\ a_{112} = a_{121} = a_{211} = 1$, we have

$$\mathcal{A} = \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix}$$

suppose that $\mathcal{A}$ is a rank one tensors, then

$$\begin{pmatrix}a\\b\end{pmatrix} \otimes \begin{pmatrix}c\\d\end{pmatrix} \otimes \begin{pmatrix}e\\f\end{pmatrix} = \begin{pmatrix}ace\\acf\\ade\\adf\\bce\\bcf\\bde\\bdf\end{pmatrix} = \begin{pmatrix}0\\1\\1\\0\\1\\0\\0\\0\end{pmatrix}$$

$$\therefore ace = 0 \Rightarrow e = 0 \quad (\text{since } acf = 1 \text{ and } ac \neq 0)$$
$$bce = 1 \text{ is a contradiction since } e = 0.$$

thus the rank of $\mathcal{A}$ is either 2 or 3 and this corresponds with the left-hand side of (2.8). The right-hand side of (2.8) is the rank of a 2×4 matrix and the rank is 2, we can formally define this as the modal rank.

Definition 2.3.2. The n-rank of a tensor $\mathcal{A}$, denoted $R_{(n)}(\mathcal{A})$, is the dimension of a vector space spanned by the columns of the n-th unfolding of $\mathcal{A}$

$$R_{(n)}(\mathcal{A}) = \text{rank}(A_{(n)}). \tag{2.9}$$

The n-rank of a tensor is not the same as the rank of the tensor. As suggested by example 2.3.1 we shall see that matrix rank can be defined by decomposing the matrix into outer products. Similarly, we can define the tensor rank by writing it as an outer product [10]. In this dissertation we take the Kronecker product as the outer product.

Definition 2.3.3. A tensor $\mathcal{A} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_m}$ is rank one if it can be written as the outer product of m vectors, that is

$$\mathcal{A} = v^{(1)} \otimes v^{(2)} \otimes \cdots \otimes v^{(m)} \tag{2.10}$$

with the corresponding entries,

$$a_{i_1 i_2 \cdots i_m} = v^{(1)}_{i_1} v^{(2)}_{i_2} \cdots v^{(m)}_{i_m} \quad \text{for all } 1 \leq i_m \leq I_m.$$

CANDECOMP-PARAFAC (CP) decomposition gives an insight into tensor rank. The decomposition factorizes a tensor into a sum of rank-one tensors [10, 9]. Rank-one tensors are those tensors which can be expressed as a single outer product (Kronecker product) $v^{(1)} \otimes \cdots \otimes v^{(m)}$. Generally the CP decomposition is not unique. However, the conditions for the unique definition of the CP decomposition are discussed in [11].

Definition 2.3.4. CANDECOMP-PARAFAC decomposition of a tensor $\mathcal{A} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_m}$ is a sum of rank-one tensors up to the rank R, that is

$$\mathcal{A} = \sum_{r=1}^{R} \sigma_r (v^{(1)}_r \otimes v^{(2)}_r \otimes \cdots \otimes v^{(m)}_r), \tag{2.11}$$

where R is minimal.

Definition 2.3.5. The rank of a tensor $\mathcal{A}$, denoted by $R = \text{rank}(\mathcal{A})$, is defined as the minimum number of rank-one tensors needed to express $\mathcal{A}$ (as a linear combination of rank-one tensors).

Example 2.3.6. Let $\mathcal{A} \in \mathbb{F}^{2\times 2\times 2}$, with entries $a_{111} = a_{221} = a_{112} = 1, a_{211} = a_{121} = a_{212} = a_{122} = a_{222} = 0$. The CP decomposition of $\mathcal{A}$ is

$$\mathcal{A} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Thus, rank($\mathcal{A}$)=2. For the modal rank, we take a look at n unfoldings of $\mathcal{A}$

$$A_{(1)} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$A_{(2)} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$
$$A_{(3)} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Thus $R_{(1)}(\mathcal{A}) = R_{(2)}(\mathcal{A}) = R_{(3)}(\mathcal{A}) = 2$. In this example, the rank of this tensor equals the n-rank, this is not necessarily true for all tensors. In fact $R_{(n)}(\mathcal{A}) \leq \operatorname{rank}(\mathcal{A})$ (see [4]).

2.4 Multilinear single value decomposition

Our discussion of the higher-order SVD begins with the matrix SVD. We will use the matrix unfolding of the tensor and apply the matrix SVD on every mode of the tensor (see for example [4]).

Lemma 2.4.1. *Let $\mathcal{A} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_m}$ be a tensor and $X_1 \in \mathbb{F}^{J_1 \times I_1}, \ldots, X_m \in \mathbb{F}^{J_m \times I_m}$ be matrices. If $\mathcal{A} = \mathcal{Y} \times_1 X_1 \times_2 \cdots \times_m X_m$, then we can express $\mathcal{A}$ as*

$$\mathcal{A} = (X_1 \otimes \cdots \otimes X_m)\mathcal{Y}.$$

Proof.

$$\begin{aligned}
\mathcal{A} &= \mathcal{Y} \times_1 X_1 \times_2 \cdots \times_m X_m \\
&= \sum_{i_1,\ldots,i_m} y_{i_1 \ldots i_m} X_1 e_{i_1} \otimes \cdots \otimes X_m e_{i_m} && \text{(by definition 2.2.1)} \\
&= (X_1 \otimes \cdots \otimes X_m) \sum_{i_1,\ldots,i_m} y_{i_1 \ldots i_m} e_{i_1} \otimes \cdots \otimes e_{i_m} && \text{(by the interchange law)} \\
&= (X_1 \otimes \cdots \otimes X_m)\mathcal{Y}.
\end{aligned}$$

Thus, $\mathcal{A} = (X_1 \otimes \cdots \otimes X_m)\mathcal{Y}$. □

We will use the short-hand notation

$$U_{k+1} \otimes \cdots \otimes U_{k-1} = U_{k+1} \otimes \cdots \otimes U_m \otimes U_1 \otimes \cdots \otimes U_{k-1}$$

Lemma 2.4.2. *Let $\mathcal{A} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_m}$ be a tensor and $U_1 \in \mathbb{F}^{J_1 \times I_1}, \ldots, U_m \in \mathbb{F}^{J_m \times I_m}$ be matrices, then*

$$[(U_1 \otimes \cdots \otimes U_m)\, \mathcal{A}]_{(k)} = U_k \mathcal{A}_{(k)} (U_{k+1} \otimes \cdots \otimes U_{k-1})^T.$$

Proof. From the left side

$$\begin{aligned}
(U_1 \otimes \cdots \otimes U_m)\,\mathcal{A} &= (U_1 \otimes \cdots \otimes U_m) \sum_{i_1,\ldots,i_m} a_{i_1\cdots i_m} e_{i_1} \otimes \cdots \otimes e_{i_m} \\
&= \sum_{i_1,\ldots,i_m} a_{i_1\cdots i_m} \left(U_1 e_{i_1} \otimes \cdots \otimes U_m e_{i_m}\right) \quad \text{(by the interchange law)} \\
&= \sum_{\substack{i_1,\ldots,i_m \\ j_1,\ldots,j_m}} a_{i_1\cdots i_m} \left((U_1)_{j_1 i_1} e_{j_1} \otimes \cdots \otimes (U_m)_{j_m i_m} e_{j_m}\right).
\end{aligned}$$

For the k-th mode unfolding we have

$$\begin{aligned}
&[(U_1 \otimes \cdots \otimes U_m)\,\mathcal{A}]_{(k)} \\
&= \sum_{\substack{i_1,\ldots,i_m \\ j_1,\ldots,j_m}} a_{i_1\cdots i_m} \left((U_k)_{j_k i_k} \cdots (U_{k-1})_{j_{k-1} i_{k-1}}\right) e_{i_k} \left(e_{i_{k+1}} \otimes \cdots \otimes e_{i_{k-1}}\right)^T \\
&= \sum_{\substack{i_1,\ldots,i_m \\ j_1,\ldots,j_m}} a_{i_1\cdots i_m} (U_k)_{j_k i_k} e_{i_k} \left[(U_{k+1})_{j_{k+1} i_{k+1}} e_{i_{k+1}} \otimes \cdots \otimes (U_{k-1})_{j_{k-1} i_{k-1}} e_{i_{k-1}}\right]^T \\
&= \sum_{i_1,\ldots,i_m} a_{i_1\cdots i_m} U_k e_{i_k} \left[U_{k+1} e_{i_{k+1}} \otimes \cdots \otimes U_{k-1} e_{i_{k-1}}\right]^T \\
&= U_k \sum_{i_1,\ldots,i_m} a_{i_1\cdots i_m} e_{i_k} \otimes \cdots \otimes e_{i_{k-1}} \left[U_{k+1} \otimes \cdots \otimes U_{k-1}\right]^T \text{ (by the interchange law)} \\
&= U_k \mathcal{A}_{(k)} \left[U_{k+1} \otimes \cdots \otimes U_{k-1}\right]^T.
\end{aligned}$$

Therefore, $[(U_1 \otimes \cdots \otimes U_m)\,\mathcal{A}]_{(k)} = U_k \mathcal{A}_{(k)} (U_{k+1} \otimes \cdots \otimes U_{k-1})^T$. □

Theorem 2.4.3. *Every tensor $\mathcal{A} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_m}$ can be written as the product*

$$\mathcal{A} = (U_1 \otimes U_2 \otimes \cdots \otimes U_m)\mathcal{S} \tag{2.12}$$

where $U_1, U_2, \ldots, U_m$ are orthogonal matrices and $\mathcal{S} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_m}$.

Proof. We want to show that such a tensor, $\mathcal{S}$, exists and that it is unique. Let $\mathcal{A}$ be an m-th order tensor, with unfoldings $\mathcal{A}_{(1)}, \ldots, \mathcal{A}_{(m)}$. The j-th mode singular value decomposition of $\mathcal{A}$ is

$$\mathcal{A}_{(j)} = U_j \Sigma_j V_j^T. \tag{2.13}$$

The modal products by U_j are invertible, and since unfolding is bijective. There exists an m-th order tensor $\mathcal{S}$ such that

$$\mathcal{A}_{(j)} = U_j \mathcal{S}_{(j)} (U_{j+1} \otimes \cdots \otimes U_{j-1})^T, \tag{2.14}$$

where

$$\mathcal{S}_{(j)} = \Sigma_j V_j^T (U_{j+1} \otimes \cdots \otimes U_{j-1}).$$

We note that by Lemma 2.4.2, $\mathcal{A} = (U_1 \otimes \cdots \otimes U_m)\mathcal{S}$. So that

$$\mathcal{A}_{(k)} = U_k \mathcal{S}_{(k)} (U_{k+1} \otimes \cdots \otimes U_{k-1})^T = U_k \Sigma_k V_k^T,$$

by Lemma 2.4.2 and (2.13). Hence, since $U_1, \ldots, U_m$ are orthogonal,

$$\mathcal{S}_{(k)} = \Sigma_k V_k^T (U_{k+1} \otimes \cdots \otimes U_{k-1}),$$

so that (2.14) holds for all modes j. It follows that $\mathcal{S}$ is given by

$$\mathcal{S} = \text{fold}(\mathcal{S}_{(j)}) \quad \text{for any } j = 1, \ldots, m. \qquad \square$$

Theorem 2.4.4. *The subtensor, $\mathcal{S}_{i_j}$, of $\mathcal{S} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_m}$ is obtained by fixing the j-th index to a constant, that is, $\mathcal{S}_{i_j = \alpha}$ for $1 \leq i_j \leq I_j$. The higher-order SVD has the following properties:*

(1) The Frobenius-norms of the subtensors are defined to be the j-th mode singular values of the matrix $A_{(j)}$, that is $||\mathcal{S}_{i_j=k}|| = \sigma_k(A_{(k)})$.

(2) The Frobenius-norms of the subtensors are decreasing, that is

$$||\mathcal{S}_{i_j=1}|| \geq ||\mathcal{S}_{i_j=2}|| \geq \cdots \geq ||\mathcal{S}_{i_j=I_j}||.$$

Proof. **(1)** Following the notation and singular value decompositions in the proof of theorem 2.4.3, we have

$$\begin{aligned} ||\mathcal{S}_{i_j=k}|| &= ||(\mathcal{S}_{(j)})_{k-th \text{ row}}|| \\ &= ||e_k^T \Sigma_j V_j^T (U_{j+1} \otimes \cdots \otimes U_{j-1})^T|| \\ &= \sigma_k ||e_k^T V_j^T (U_{j+1} \otimes \cdots \otimes U_{j-1})^T|| \\ &= \sigma_k, \end{aligned}$$

where σ_k is the k-th singular value of $\mathcal{A}_{(j)}$. The last equality follows by orthogonality of $V_j^T (U_{j+1} \otimes \cdots \otimes U_{j-1})^T$.

(2) The proof follows from Theorem 1.1.6. $\square$

In the following chapter, we look at the first binary product between higher-order tensors in this dissertation. We use a family of vector space isomorphisms, fold, to redefine a tensor as a matrix and unfold the matrix back into a tensor (similar to the fold and unfold operation studied above).

Chapter 3

Einstein product

3.1 Einstein product

The Einstein product uses the isomorphism between the vector space and the matrix space through tensor unfolding. For $\mathcal{A} \in \mathbb{F}^{I_1 \times \cdots \times I_p \times K_1 \times \cdots \times K_m}$ we have

$$\mathbb{F}^{I_1} \otimes \cdots \otimes \mathbb{F}^{I_p} \otimes \mathbb{F}^{K_1} \otimes \cdots \otimes \mathbb{F}^{K_m} \cong \mathbb{F}^{(I_1 \times I_2 \times \cdots \times I_p)(K_1 \times K_2 \times \cdots \times K_m)},$$

This isomorphism is called an unfolding. Using the tensor, $\mathcal{A}$, the Einstein product is defined using [17]. In the following, we will assume that tensor unfolding unfolds around the different letters used for indices, in order to simplify the notation.

Definition 3.1.1. Let $\mathcal{A} \in \mathbb{F}^{I_1 \times \cdots \times I_p \times K_1 \times \cdots \times K_m}$ be a tensor the unfold of $\mathcal{A}$ is defined as follow

$$\text{unfold}_{I_1 \times \cdots \times I_p}(\mathcal{A}) = \sum_{\substack{i_1, \ldots, i_p \\ k_1, \ldots, k_m}} a_{i_1 \cdots i_p k_1 \cdots k_m} e_{i_1} \otimes \cdots \otimes e_{i_p} (e_{k_1} \otimes \cdots \otimes e_{k_m})^T \tag{3.1}$$

For a tensor $\mathcal{A} \in \mathbb{F}^{I_1 \times \cdots \times I_p \times K_1 \times \cdots \times K_m}$ the following notation will be used to represent the fold and unfold operations

$$\begin{aligned} \text{unfold}_{I_1 \times \cdots \times I_p}(\mathcal{A}) &:= \text{unfold}_I(\mathcal{A}) \\ \text{fold}_{K_1 \times \cdots \times K_m}(\mathcal{A}) &:= \text{fold}_K(\mathcal{A}) \end{aligned}$$

Definition 3.1.2. Let $\mathcal{A} \in \mathbb{F}^{I_1 \times \cdots \times I_p \times K_1 \times \cdots \times K_m}$ and $\mathcal{B} \in \mathbb{F}^{K_1 \times K_2 \times \cdots \times K_m \times J_1 \times J_2 \times \cdots \times J_q}$, the Einstein product is

$$\mathcal{A} *_E \mathcal{B} = \text{fold}_J((\text{unfold}_I(\mathcal{A}))(\text{unfold}_K(\mathcal{B}))) \tag{3.2}$$

then, $\mathcal{A} *_E \mathcal{B} \in \mathbb{F}^{I_1 \times I_2 \times \cdots \times I_p \times J_1 \times J_2 \times \cdots \times J_q}$. The Einstein product is simply the matrix product under the fold and unfold operation. The following figure demonstrates this.

$$\begin{tikzcd}
(\mathcal{A}, \mathcal{B}) \arrow[rr, "\text{Einstein product}"] \arrow[d, shift right, "\text{(unfold, unfold)}"'] & & \mathcal{A} *_E \mathcal{B} \arrow[d, shift right, "\text{unfold}"'] \\
(A, B) \arrow[rr, "\text{matrix product}"] \arrow[u, shift right, "\text{(fold, fold)}"'] & & AB \arrow[u, shift right, "\text{fold}"']
\end{tikzcd}$$

Figure 3.1

In the following corollary, we will show that Definition 3.1.2 is equivalent to the Einstein product in [17].

Corollary 3.1.3. *Let* $\mathcal{A} \in \mathbb{F}^{I_1 \times \cdots \times I_p \times K_1 \times \cdots \times K_m}$ *and* $\mathcal{B} \in \mathbb{F}^{K_1 \times \cdots \times K_p \times J_1 \times \cdots \times J_q}$. *The Einstein product is defined by the operator* $*_E$ *and*

$$(\mathcal{A} *_E \mathcal{B})_{i_1 \ldots i_m j_1 \ldots j_q} = \sum_{k_1, \ldots, k_m} a_{i_1 \ldots i_p k_1 \ldots k_m} b_{k_1 \ldots k_m j_1 \ldots j_q},$$

is equivalent to the product defined in Definition 3.1.2.

In the proof below we will use the short-hand notation for the entries of tensor $\mathcal{A}$

$$a_{i_1 \ldots i_p k_1 \ldots k_m} := a_{IK}$$

Proof. From the right-hand side of (3.2) and by Definition 3.1.1 we have

$$\begin{aligned}
&\text{fold}_J(\text{unfold}_I(\mathcal{A}))(\text{unfold}_K(\mathcal{B}))) \\
&= \text{fold}_J\Big(\sum_{\substack{i_1, \ldots, i_p \\ k_1, \ldots, k_m}} \sum_{\substack{k_1, \ldots, k_m \\ j_1, \ldots, j_q}} a_{IK} b_{KJ} (e_{i_1} \otimes \cdots \otimes e_{i_p})(e_{k_1} \otimes \cdots \otimes e_{k_p})^T \\
&\qquad (e_{k_1} \otimes \cdots \otimes e_{k_p})(e_{j_1} \otimes \cdots \otimes e_{j_q})^T \Big) \\
&= \text{fold}_J\Big(\sum_{\substack{i_1, \ldots, i_p \\ k_1, \ldots, k_m}} \sum_{\substack{k_1, \ldots, k_m \\ j_1, \ldots, j_q}} a_{IK} b_{KJ} (e_{i_1} \otimes \cdots \otimes e_{i_p})(e_{k_1}^T e_{k_p} \otimes \cdots \otimes e_{k_p}^T e_{k_p})(e_{j_1} \otimes \cdots \otimes e_{j_q})^T \Big) \\
&= \text{fold}_J\Big(\sum_{\substack{i_1, \ldots, i_p \\ k_1, \ldots, k_m}} \sum_{\substack{k_1, \ldots, k_m \\ j_1, \ldots, j_q}} a_{IK} b_{KJ} (e_{i_1} \otimes \cdots \otimes e_{i_p})(e_{j_1} \otimes \cdots \otimes e_{j_q})^T \Big) \\
&= \sum_{k_1, \ldots, k_m} a_{IK} b_{KJ} \\
&= \sum_{k_1, \ldots, k_m} a_{i_1 \ldots i_p k_1 \ldots k_m} b_{k_1 \ldots k_m j_1 \ldots j_q}.
\end{aligned}$$

Thus, $\text{fold}_J(\text{unfold}_I(\mathcal{A}))(\text{unfold}_K(\mathcal{B}))) = \sum_{k_1,\dots,k_m} a_{i_1\dots i_p k_1\dots k_m} b_{k_1\dots k_m j_1\dots j_q}$. $\square$

Lemma 3.1.4. *Let $\mathcal{A} \in \mathbb{F}^{I_1\times I_2\times\cdots\times I_p\times K_1\times\cdots\times K_m}$, $\mathcal{B} \in \mathbb{F}^{K_1\times K_2\times\cdots\times K_m\times J_1\times\cdots\times J_q}$ and $\mathcal{C} \in \mathbb{F}^{J_1\times J_2\times\cdots\times J_m\times L_1\times L_2\times\cdots\times L_n}$, then*

$$(\mathcal{A} *_E \mathcal{B}) *_E \mathcal{C} = \mathcal{A} *_E (\mathcal{B} *_E \mathcal{C}). \tag{3.3}$$

Proof.

$$\begin{aligned}
(\mathcal{A} *_E \mathcal{B}) *_E \mathcal{C} &= \text{fold}_J((\text{unfold}_I(\mathcal{A}))(\text{unfold}_K(\mathcal{B}))) *_E \mathcal{C} \\
&= \text{fold}_L[\text{unfold}_I(\text{fold}_J[(\text{unfold}_I(\mathcal{A}))(\text{unfold}_K(\mathcal{B}))]\text{unfold}_J(\mathcal{C})] \\
&= \text{fold}_L[((\text{unfold}_I(\mathcal{A})(\text{unfold}_K(\mathcal{B})))\text{unfold}_J(\mathcal{C})] \\
&= \text{fold}_L[(\text{unfold}_I(\mathcal{A})((\text{unfold}_K(\mathcal{B}))\text{unfold}_J(\mathcal{C}))] \\
&= \text{fold}_L[(\text{unfold}_I(\mathcal{A})\text{unfold}_K(\text{fold}_L[\text{unfold}_K(\mathcal{B})\text{unfold}_J(\mathcal{C})])] \\
&= \mathcal{A} *_E (\mathcal{B} *_E \mathcal{C}).
\end{aligned}$$

$\square$

Definition 3.1.5. Let $\mathcal{A} \in \mathbb{F}^{I_1\times I_2\times\cdots\times I_n\times J_1\times J_2\times\cdots\times J_m}$, the transpose of $\mathcal{A}$ is a tensor $\mathcal{A}^T \in \mathbb{F}^{J_1\times J_2\times\cdots\times J_m\times I_1\times I_2\times\cdots\times I_n}$ given by

$$\mathcal{A}^T = \text{fold}_I(\text{unfold}_J(\mathcal{A})^T). \tag{3.4}$$

Proposition 3.1.6. *Let $\mathcal{A} \in \mathbb{F}^{I_1\times\cdots\times I_n\times J_1\times\cdots\times J_m}$ and $\mathcal{B} \in \mathbb{F}^{J_1\times\cdots\times J_m\times I_1\times\cdots\times I_p}$, then*

1. $$(\mathcal{A} *_E \mathcal{B})^T = \mathcal{B}^T *_E \mathcal{A}^T. \tag{3.5}$$

2. *Let $\mathcal{I}_m \in \mathbb{F}^{J_1\times\cdots\times J_m\times J_1\times\cdots\times J_m}$ and $\mathcal{I}_n \in \mathbb{F}^{I_1\times\cdots\times I_n\times I_1\times\cdots\times I_n}$ be identity tensors then*
$$\mathcal{I}_n *_E \mathcal{A} = \mathcal{A} *_E \mathcal{I}_m = \mathcal{A} \tag{3.6}$$
where $\mathcal{I}_n = \text{fold}_I(\text{unfold}_I(\mathcal{I}_n))$, is a tensor that unfolds to an identity matrix.

Proof. **1.** We will show that $(\text{unfold}_J(\mathcal{A} *_E \mathcal{B}))^T = \text{unfold}_I(\mathcal{B}^T *_E \mathcal{A}^T)$.

$$\begin{aligned}
(\text{unfold}_J(\mathcal{A} *_E \mathcal{B}))^T &= [\text{unfold}_I(\mathcal{A})\text{unfold}_J(\mathcal{B})]^T \\
&= \text{unfold}_I(\mathcal{B})^T\text{unfold}_J(\mathcal{A})^T \\
&= \text{unfold}_I(\mathcal{B}^T *_E \mathcal{A}^T).
\end{aligned}$$

After applying the fold operation $(\mathcal{A} *_E \mathcal{B})^T = \mathcal{B}^T *_E \mathcal{A}^T$.

2.

$$\begin{aligned}\mathcal{I}_n *_E \mathcal{A} &= \text{fold}_J[\text{unfold}_I(\mathcal{I}_n)\text{unfold}_I(\mathcal{A})] \\ &= \text{fold}_J[\text{unfold}_I(\mathcal{A})] \\ &= \mathcal{A}. \qquad \square\end{aligned}$$

Similarly, $\mathcal{A} *_E \mathcal{I}_m = \mathcal{A}$.

3.2 Einstein product Moore-Penrose pseudo inverse

Since the Einstein product is defined (through unfolding) by the matrix product, the matrix algebraic properties reflect to the setting of tensors. Thus, the notion of a Moore-Penrose pseudo inverse of a matrix yields the same definition for tensors (via the Einstein product).

Definition 3.2.1. Let $\mathcal{A} \in \mathbb{F}^{I_1 \times \cdots \times I_m \times K_1 \times \cdots \times K_p}$. The Moore-Penrose pseudo inverse of $\mathcal{A}$ is the tensor, $\mathcal{A}^+ \in \mathbb{F}^{K_1 \times \cdots \times K_p \times I_1 \times \cdots \times I_m}$, that satisfies the following conditions:

1. $\mathcal{A} *_E \mathcal{A}^+ *_E \mathcal{A} = \mathcal{A}$,
2. $\mathcal{A}^+ *_E \mathcal{A} *_E \mathcal{A}^+ = \mathcal{A}^+$,
3. $(\mathcal{A} *_E \mathcal{A}^+)^* = \mathcal{A} *_E \mathcal{A}^+$,
4. $(\mathcal{A}^+ *_E \mathcal{A})^* = \mathcal{A}^+ *_E \mathcal{A}$.

The singular value decomposition of matrices yields a higher-order singular value decomposition via the Einstein product.

Definition 3.2.2. Let $\mathcal{A} \in \mathbb{F}^{I_1 \times \cdots \times I_m \times J_1 \times \cdots \times J_m}$, then $\mathcal{A}$ is orthogonal if

$$\mathcal{A}^* *_E \mathcal{A} = \mathcal{A} *_E \mathcal{A}^* = \mathcal{I}_m. \tag{3.7}$$

Lemma 3.2.3. *Let $\mathcal{A} \in \mathbb{F}^{I_1 \times \cdots \times I_m \times J_1 \times \cdots \times J_m}$, the singular value decomposition of $\mathcal{A}$ has the form*

$$\mathcal{A} = \mathcal{U} *_E \mathcal{S} *_E \mathcal{V}^* \tag{3.8}$$

where $\mathcal{U} \in \mathbb{F}^{I_1 \times \cdots \times I_m \times I_1 \times \cdots \times I_m}$, $\mathcal{V} \in \mathbb{F}^{J_1 \times \cdots \times J_m \times J_1 \times \cdots \times J_m}$ are orthogonal tensors and $\mathcal{S} \in \mathbb{F}^{I_1 \times \cdots \times I_m \times J_1 \times \cdots \times J_m} = 0$ when $I_1 \times \cdots \times I_m \neq J_1 \times \cdots \times J_m$.

The singular value decomposition allows us to define the Moore-Penrose inverse for any real tensor.

Theorem 3.2.4. *Let $\mathcal{A} \in \mathbb{F}^{I_1\times\cdots\times I_m\times J_1\times\cdots\times J_m}$, be a tensor. The Moore-Penrose inverse in the Einstein product exists and it is uniquely given by*

$$\mathcal{A}^+ = \mathcal{V} *_E \mathcal{S}^+ *_E \mathcal{U}^* \tag{3.9}$$

where $\mathcal{U} \in \mathbb{F}^{I_1\times\cdots\times I_m\times I_1\times\cdots\times I_m}$, $\mathcal{V} \in \mathbb{F}^{J_1\times\cdots\times J_m\times J_1\times\cdots\times J_m}$, are orthogonal tensors and $\mathcal{S}^+ \in \mathbb{F}^{J_1\times\cdots\times J_m\times I_1\times\cdots\times I_m}$

Proof. First we need to show that (3.9) satisfies Definition 3.2.1.

(1) $\mathcal{A} *_E \mathcal{A}^+ *_E \mathcal{A} = \mathcal{A}$.

$$\begin{aligned}\mathcal{A} *_E \mathcal{A}^+ *_E \mathcal{A} &= (\mathcal{U} *_E \mathcal{S} *_E \mathcal{V}^*) *_E (\mathcal{V} *_E \mathcal{S}^+ *_E \mathcal{U}^*) *_E (\mathcal{U} *_E \mathcal{S} *_E \mathcal{V}^*)\\ &= \mathcal{U} *_E \mathcal{S} *_E \mathcal{S}^+ *_E \mathcal{S} *_E \mathcal{V}^*\\ &= \mathcal{U} *_E \mathcal{S} *_E \mathcal{V}^*\\ &= \mathcal{A}.\end{aligned}$$

Therefore $\mathcal{A} *_E \mathcal{A}^+ *_E \mathcal{A} = \mathcal{A}$.

(2) $\mathcal{A}^+ *_E \mathcal{A} *_E \mathcal{A}^+ = \mathcal{A}^+$.

$$\begin{aligned}\mathcal{A}^+ *_E \mathcal{A} *_E \mathcal{A}^+ &= (\mathcal{V} *_E \mathcal{S}^+ *_E \mathcal{U}^*) *_E (\mathcal{U} *_E \mathcal{S} *_E \mathcal{V}^*) *_E (\mathcal{V} *_E \mathcal{S}^+ *_E \mathcal{U}^*)\\ &= \mathcal{V} *_E \mathcal{S}^+ *_E \mathcal{S} *_E \mathcal{S}^+ *_E \mathcal{U}^*\\ &= \mathcal{V} *_E \mathcal{S}^+ *_E \mathcal{U}^*\\ &= \mathcal{A}^+.\end{aligned}$$

Therefore $\mathcal{A}^+ *_E \mathcal{A} *_E \mathcal{A}^+ = \mathcal{A}^+$.

(3) $(\mathcal{A} *_E \mathcal{A}^+)^* = \mathcal{A} *_E \mathcal{A}^+$

$$\begin{aligned}(\mathcal{A} *_E \mathcal{A}^+)^* &= [(\mathcal{U} *_E \mathcal{S} *_E \mathcal{V}^*) *_E (\mathcal{V} *_E \mathcal{S}^+ *_E \mathcal{U}^*)]^*\\ &= [\mathcal{U} *_E \mathcal{S} *_E \mathcal{S}^+ *_E \mathcal{U}^*]^*\\ &= \mathcal{U} *_E \mathcal{S}^+ *_E \mathcal{V}^* *_E \mathcal{V} *_E \mathcal{S} *_E \mathcal{U}^*\\ &= \mathcal{A} *_E \mathcal{A}^+.\end{aligned}$$

Therefore $(\mathcal{A} *_E \mathcal{A}^+)^* = \mathcal{A} *_E \mathcal{A}^+$.

(4) $(\mathcal{A}^+ *_E \mathcal{A})^* = \mathcal{A}^+ *_E \mathcal{A}$

$$\begin{aligned}(\mathcal{A}^+ *_E \mathcal{A})^* &= [(\mathcal{V} *_E \mathcal{S}^+ *_E \mathcal{U}^*) *_E (\mathcal{U} *_E \mathcal{S} *_E \mathcal{V}^*)]^*\\ &= [\mathcal{V} *_E \mathcal{S}^+ *_E \mathcal{S} *_E \mathcal{V}^*]^*\\ &= \mathcal{V} *_E \mathcal{S}^+ *_E \mathcal{U}^* *_E \mathcal{U} *_E \mathcal{S} *_E \mathcal{V}^*\\ &= \mathcal{A}^+ *_E \mathcal{A}.\end{aligned}$$

Therefore $(\mathcal{A}^+ *_E \mathcal{A})^* = \mathcal{A}^+ *_E \mathcal{A}$.

For uniqueness, suppose that $\mathcal{A}^+, \mathcal{B}^+ \in \mathbb{F}^{J_1\times\cdots\times J_m\times I_1\times\cdots\times I_m}$ are Moore-Penrose pseudo inverses of $\mathcal{A}$, then

$$\begin{aligned}
\mathcal{A}^+ &= \mathcal{A}^+ *_E \mathcal{A} *_E \mathcal{A}^+ \\
&= \mathcal{A}^+ *_E (\mathcal{A} *_E \mathcal{A}^+)^* && \text{(Definition 3.2.1 (3))} \\
&= \mathcal{A}^+ *_E (\mathcal{A}^+)^* *_E \mathcal{A}^* \\
&= \mathcal{A}^+ *_E (\mathcal{A}^+)^* *_E \mathcal{A}^* *_E (\mathcal{B}^+)^* *_E \mathcal{A}^* \\
&= \mathcal{A}^+ *_E (\mathcal{A} *_E \mathcal{A}^+)^* *_E (\mathcal{A} *_E \mathcal{B}^+)^* \\
&= \mathcal{A}^+ *_E \mathcal{A} *_E \mathcal{A}^+ *_E \mathcal{A} *_E \mathcal{B}^+ \\
&= \mathcal{A}^+ *_E \mathcal{A} *_E \mathcal{B}^+ \\
&= \mathcal{A}^+ *_E \mathcal{A} *_E \mathcal{B}^+ *_E \mathcal{A} *_E \mathcal{B}^+ \\
&= \mathcal{A}^* *_E (\mathcal{A}^+)^* *_E \mathcal{A}^* *_E (\mathcal{B}^+)^* *_E \mathcal{B}^+ \\
&= \mathcal{A}^* *_E (\mathcal{B}^+)^* *_E \mathcal{B}^+ \\
&= \mathcal{B}^+ *_E \mathcal{A} *_E \mathcal{B}^+ \\
&= \mathcal{B}^+.
\end{aligned}$$

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

The Moore-Penrose pseudo inverse has the following properties [17].

Proposition 3.2.5. *Let $\mathcal{A} \in \mathbb{F}^{I_1\times\cdots\times I_m\times K_1\times\cdots\times K_m}$ and $a \in \mathbb{F}$, then*

1. $(\mathcal{A}^+)^+ = \mathcal{A}$,
2. $(\mathcal{A}^+)^* = (\mathcal{A}^*)^+$,
3. $(a\mathcal{A})^+ = a^+\mathcal{A}^+$,
4. $(\mathcal{A}^* *_E \mathcal{A})^+ = \mathcal{A}^+ *_E (\mathcal{A}^*)^+$,
5. $(\mathcal{X} *_E \mathcal{A} *_E \mathcal{Y})^+ = \mathcal{Y}^* *_E \mathcal{A}^+ *_E \mathcal{X}^*$, *where $\mathcal{X}$ and $\mathcal{Y}$ are orthogonal tensors*
6. $\mathcal{A}^+ *_E (\mathcal{I} - \mathcal{A} *_E \mathcal{A}^+) = 0$ *and* $(\mathcal{I} - \mathcal{A} *_E \mathcal{A}^+) *_E \mathcal{A}^+ = 0$,

where

$$a^+ = \begin{cases} \frac{1}{a} & a \neq 0, \\ 0 & a = 0. \end{cases}$$

Proof. **1.** Using Lemma 3.2.3 and Theorem 3.2.4, $\mathcal{A}^+$ is uniquely given by $\mathcal{V} *_E \mathcal{S}^+ *_E \mathcal{U}^*$, where $\mathcal{U}$ and $\mathcal{V}$ are orthogonal, hence

$$\begin{aligned}(\mathcal{A}^+)^+ &= (\mathcal{V} *_E \mathcal{S}^+ *_E \mathcal{U}^*)^+ \\ &= \mathcal{U} *_E \mathcal{S} *_E \mathcal{V}^* \\ &= \mathcal{A}.\end{aligned}$$

2. Since $\mathcal{A}^+$ is the Moore-Penrose pseudo inverse of $\mathcal{A}$, then $(\mathcal{A}^+)^*$ is the Moore-Penrose pseudo inverse of $\mathcal{A}^*$.

(I)

$$\begin{aligned}\mathcal{A} *_E \mathcal{A}^+ *_E \mathcal{A} = \mathcal{A} &\iff (\mathcal{A} *_E \mathcal{A}^+ *_E \mathcal{A})^* = (\mathcal{A})^* \\ &\iff \mathcal{A}^* *_E (\mathcal{A}^+)^* *_E \mathcal{A}^* = \mathcal{A}^*.\end{aligned}$$

(II)

$$\begin{aligned}\mathcal{A}^+ *_E \mathcal{A} *_E \mathcal{A}^+ = \mathcal{A}^+ &\iff (\mathcal{A}^+ *_E \mathcal{A} *_E \mathcal{A}^+)^* = (\mathcal{A}^+)^* \\ &\iff (\mathcal{A}^+)^* *_E \mathcal{A}^* *_E (\mathcal{A}^+)^* = (\mathcal{A}^+)^*.\end{aligned}$$

(III)

$$(\mathcal{A} *_E \mathcal{A}^+)^* = \mathcal{A} *_E \mathcal{A}^+ \iff ((\mathcal{A} *_E \mathcal{A}^+)^*)^* = (\mathcal{A} *_E \mathcal{A}^+)^*.$$

(IV)

$$(\mathcal{A}^+ *_E \mathcal{A})^* = \mathcal{A}^+ *_E \mathcal{A} \iff ((\mathcal{A}^+ *_E \mathcal{A})^*)^* = (\mathcal{A}^+ *_E \mathcal{A})^*.$$

Thus by uniqueness of the Moore-Penrose pseudo inverse of $\mathcal{A}^*$, $(\mathcal{A}^*)^+ = (\mathcal{A}^+)^*$.

3. Using the definition of a^+, we get

$$\begin{aligned}aA &= (e^{i\arg(a)}U)(|a|)\Sigma V^T \\ (aA)^+ &= V\left(\frac{1}{|a|}\Sigma\right)(e^{i\arg(a)}U)^T \\ &= \frac{1}{a}V(\Sigma)(U)^T \text{ when } a \neq 0 \\ &= a^+A^+\end{aligned}$$

4.

$$\begin{aligned}(\mathcal{A}^* *_E \mathcal{A})^+ &= (\text{unfold}(\mathcal{A}^*)\text{unfold}(\mathcal{A}))^+\\ &= \text{unfold}(\mathcal{A})^+\text{unfold}(\mathcal{A}^*)^+\\ &= \mathcal{A}^+ *_E (\mathcal{A}^*)^+\end{aligned}$$

5. $\mathcal{X}$ and $\mathcal{Y}$ are orthogonal tensors, so it follows that $\mathcal{X}^+ = \mathcal{X}^*$ and $\mathcal{Y}^+ = \mathcal{Y}^*$,

$$\begin{aligned}(\mathcal{X} *_E \mathcal{A} *_E \mathcal{Y})^+ &= (\text{unfold}(\mathcal{X})\text{unfold}(\mathcal{A})\text{unfold}(\mathcal{Y}))^+\\ &= (\text{unfold}(\mathcal{Y}))^+(\text{unfold}(\mathcal{A}))^+(\text{unfold}(\mathcal{X}))^+\\ &= \text{unfold}(\mathcal{Y})^*\text{unfold}(\mathcal{A})^+\text{unfold}(\mathcal{X})^*\\ &= \mathcal{Y}^* *_E \mathcal{A}^+ *_E \mathcal{X}^*.\end{aligned}$$

6. Proof follows from using Proposition 3.1.6 and Definition 3.2.1. □

Chapter 4

T-product

The T-product defined in [9], is a binary operation that preserves the order of the tensor, i.e., the T-product of third-order tensors is a third-order tensor. For simplicity, we will only consider third-order tensors. Before the formal definition of the T-product is defined, we will develop some notation.

The notation $t^{(i,j)}$ represents a *tube* of a third-order tensor. Here, i and j are fixed and the third index is varied.

A *slice* of a tensor is a two dimensional section (i.e., a matrices), defined by fixing all but two indices, for a third order tensor, $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2 \times n_3}$. The horizontal, lateral and frontal slices are $A \in \mathbb{F}^{n_2 n_1 \times n_3}$, $A \in \mathbb{F}^{n_1 n_2 \times n_3}$ and $A \in \mathbb{F}^{n_1 n_3 \times n_2}$, respectively.

The frontal slices of a tensor $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2 \times n_3}$ are two-dimensional sections, denoted by $\mathrm{frnt}(\mathcal{A})$ and is presented as

$$\mathrm{frnt}(\mathcal{A}) = \sum_{j=1}^{n_3} e_j^T \otimes A_j, \tag{4.1}$$

where e_j is an element of the standard basis in $\mathbb{F}^{n_3}$ and A_j is the j-th frontal slice of $\mathcal{A}$. We will write $\mathrm{frnt}(\mathcal{A}) = (A_1|A_2|\cdots|A_{n_3})$.

Definition 4.0.1. The circulant of a tensor $\mathcal{A}$ is a square block matrix, defined by shifting the elements of each row to the right relative to the preceding row. We use the notation $\mathrm{circ}(\mathcal{A})$ to represent the circulant of a tensor $\mathcal{A}$ and

$$\mathrm{circ}(\mathcal{A}) = \begin{pmatrix} A_1 & A_{n_3} & A_{n_3-1} & \cdots & A_2 \\ A_2 & A_1 & A_{n_3} & \cdots & A_3 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ A_{n_3} & A_{n_3-1} & \ddots & A_2 & A_1 \end{pmatrix},$$

where $A_1, \cdots, A_{n_3}$ are frontal slices of $\mathcal{A}$.

Let P denote the permutation matrix

$$P = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

it follows that $Pe_j = e_{j\text{mod}n+1}$. In the following example we will show that $P^2 e_j = e_{j\text{mod}4+1}$.

Example 4.0.2. Let P be a 4×4 permutation matrix, such that

$$P = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \Rightarrow P^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

then

$$\begin{aligned} P^2 e_1 &= e_3, \\ P^2 e_2 &= e_4, \\ P^2 e_3 &= e_1, \\ P^2 e_4 &= e_2. \end{aligned}$$

Therefore, for a 4×4 permutation matrix $P^2 e_j = e_{j\text{mod}4+1}$.

From the example above we can see that for $P^k e_i$, we get

$$P^k e_i = e_{[(i-1)+k]\text{mod}n+1} \tag{4.2}$$

Definition 4.0.3. Let $a \in \mathbb{F}^n$ be a column vector, where

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

The circulant of the vector a is a $n \times n$ matrix where each row is the right cyclic shift of the vector a i.e.,

$$\operatorname{circ}(a) = \begin{pmatrix} a_1 & a_n & a_{n-1} & \cdots & a_2 \\ a_2 & a_1 & a_n & \cdots & a_3 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_n & a_{n-1} & \ddots & a_2 & a_1 \end{pmatrix}.$$

The circulant of the tensor $\mathcal{A}$ with frontal slices A_j can be written as follow

$$\operatorname{circ}(\mathcal{A}) = \sum_{j,k=1}^{n_3} (P^{k-1} e_j) e_k^T \otimes A_j. \tag{4.3}$$

Definition 4.0.4. The MatVec operation converts a tensor into a block matrix, that is, an $n_1 \times n_2 \times n_3$ tensor is converted into an $n_1 n_3 \times n_2$ block matrix.

$$\operatorname{MatVec}(\mathcal{A}) = \sum_{j=1}^{n_3} e_j \otimes \ A_j. \tag{4.4}$$

MatVec is a vector space isomorphism. The fold operation takes the MatVec of a tensor and converts it back into a tensor, i.e

$$\begin{aligned} \operatorname{fold}(\operatorname{MatVec}(\mathcal{A})) &= \operatorname{fold}\left(\sum_{j=1}^{n_3} e_j \otimes A_j \right) \\ &= \mathcal{A}. \end{aligned}$$

Here, we note that fold is defined differently to Chapter 3, since the literature uses different definitions of fold. Here we follow the definition in [9]. Now we will define the T-product and state, then prove some lemmas and theorems.

Definition 4.0.5. Let $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2 \times n_3}$ and $\mathcal{B} \in \mathbb{F}^{n_2 \times l \times n_3}$ be third-order tensors. Then the T-product $\mathcal{A} * \mathcal{B} \in \mathbb{F}^{n_1 \times l \times n_3}$ is a third-order tensor

$$\mathcal{A} * \mathcal{B} = \operatorname{fold}(\operatorname{circ}(\mathcal{A}) \cdot \operatorname{MatVec}(\mathcal{B})). \tag{4.5}$$

Using our definition of the circulant and MatVec, the T-product simplifies to

$$\begin{aligned} \mathcal{A} * \mathcal{B} &= \operatorname{fold}(\operatorname{circ}(\mathcal{A}) \cdot \operatorname{MatVec}(\mathcal{B})) \\ &= \operatorname{fold}\left(\left(\sum_{j,k=1}^{n_3} (P^{k-1} e_j) e_k^T \otimes A_j \right) \cdot \left(\sum_{k=1}^{n_3} e_k \otimes B_k \right) \right) \\ &= \operatorname{fold}\left(\sum_{j,k=1}^{n_3} (P^{k-1} e_j) e_k^T e_k \otimes A_j B_k \right) \end{aligned}$$

$$= \text{fold}\left(\sum_{j,k=1}^{n_3}(P^{k-1}e_j)\otimes A_jB_k\right)$$
$$= \text{fold}\left(\sum_{j,k=1}^{n_3}(e_{[(j+k-2) \bmod n_3]+1})\otimes A_jB_k\right).$$

Let us consider the following example to show how the t-product works.

Example 4.0.6. Suppose that $\mathcal{A}$ is a $2\times 2\times 3$ tensor, with frontal slices

$$\text{frnt}(\mathcal{A}) = \left(\begin{array}{cc|cc|cc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{array}\right),$$

and $\mathcal{B}$ a $2\times 3\times 3$ tensor with frontal slices

$$\text{frnt}(\mathcal{B}) = \left(\begin{array}{ccc|ccc|ccc} 1 & 0 & 0 & 1 & 1 & 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right).$$

Then, the circulant of $\mathcal{A}$ is a 6×6 block matrix. That is,

$$\text{circ}(\mathcal{A}) = \begin{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\ \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix}.$$

The MatVec of $\mathcal{B}$ is a 6×3 block matrix, given by

$$\text{MatVec}(\mathcal{B}) = \begin{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{pmatrix}.$$

The product between circ($\mathcal{A}$) and MatVec($\mathcal{B}$) is the usual matrix product, thus

$$\text{circ}(\mathcal{A})\cdot\text{MatVec}(\mathcal{B}) = \begin{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\ \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix} \cdot \begin{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} 4 & 1 & 1 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 0 & 1 \\ 3 & 2 & 3 \\ 2 & 1 & 1 \end{pmatrix}.$$

It follows that the frontal slices of the T-product between the tensors $\mathcal{A}$ and $\mathcal{B}$, yields

$$\begin{aligned} \operatorname{frnt}(\mathcal{A} * \mathcal{B}) &= \operatorname{frnt}(\operatorname{fold}(\operatorname{circ}(\mathcal{A}) \cdot \operatorname{MatVec}(\mathcal{B}))) \\ &= \operatorname{frnt}\left(\operatorname{fold}_{n_3} \begin{pmatrix} 4 & 1 & 1 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 0 & 1 \\ 3 & 2 & 3 \\ 2 & 1 & 1 \end{pmatrix} \right) \\ &= \left(\begin{array}{ccc|ccc|ccc} 4 & 1 & 1 & 2 & 2 & 1 & 3 & 2 & 3 \\ 2 & 2 & 1 & 1 & 0 & 1 & 2 & 1 & 1 \end{array} \right). \end{aligned}$$

4.1 Properties of the T-product

We already know that the matrix product is associative. Tensors are multilinear objects and we will show in the first lemma that the T-product is associative.

Lemma 4.1.1. *Let $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2 \times n_3}$, $\mathcal{B} \in \mathbb{F}^{n_2 \times l \times n_3}$ and $\mathcal{C} \in \mathbb{F}^{l \times k \times n_3}$ be third-order tensors, then*

$$(\mathcal{A} * \mathcal{B}) * \mathcal{C} = \mathcal{A} * (\mathcal{B} * \mathcal{C}). \tag{4.6}$$

Proof. We use the fact that the product of powers of P is commutative, that is, $P^{l-1}P^{k-1} = P^{k-1}P^{l-1} = P^{k+l-2}$ and the fact that matrix multiplication is associative. Using the MatVec operation on the tensor product between $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ we get

$$\begin{aligned} (\mathcal{A} * \mathcal{B}) * \mathcal{C} &= \operatorname{fold}(\operatorname{MatVec}(\mathcal{A} * \mathcal{B}) * \mathcal{C}) \\ &= \operatorname{fold}\left(\sum_{j,k,l=1}^{n_3} P^{l+k-2} e_j \otimes (A_j B_k) C_l \right) \\ &= \operatorname{fold}\left(\sum_{k,j,l=1}^{n_3} P^{k+l-2} e_j \otimes A_j (B_k C_l) \right) \\ &= \operatorname{fold}\left(\operatorname{MatVec}\left(\mathcal{A} * (\mathcal{B} * \mathcal{C})\right)\right) \end{aligned}$$

$$= \mathcal{A} * (\mathcal{B} * \mathcal{C}). \qquad \square$$

Therefore, by associativity we can write the T-product of three tensors as follows,

$$(\mathcal{A} * \mathcal{B}) * \mathcal{C} = \mathcal{A} * (\mathcal{B} * \mathcal{C}) = \mathcal{A} * \mathcal{B} * \mathcal{C}.$$

Definition 4.1.2. The $n \times n \times l$ identity tensor, $\mathcal{I}_{nnl}$, is a tensor whose frontal slices are identity and zero matrices, i.e.,

$$\mathrm{frnt}(\mathcal{I}_{nnl}) = (I|0|0|\cdots|0)$$

The T-product of a $n \times n \times l$ identity tensor $\mathcal{I}_{nnl}$ with a $n \times n \times l$ tensor, $\mathcal{A}$ is given by

$$\mathcal{A} * \mathcal{I}_{nnl} = \mathcal{A}, \quad \text{and} \quad \mathcal{I}_{nnl} * \mathcal{A} = \mathcal{A}. \tag{4.7}$$

Definition 4.1.3. Let $\mathcal{A}$ be an $n \times n \times l$ tensor and $\mathcal{B}$ be an $n \times l \times n$ tensor, then $\mathcal{B}$ is an inverse of $\mathcal{A}$ if

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

Lemma 4.1.4. *If $T(\mathcal{X}) = \mathcal{A} * \mathcal{X}$ where $\mathcal{A} \in \mathbb{F}^{n_1 \times m \times n_3}$ and $\mathcal{X} \in \mathbb{F}^{m \times n_2 \times n_3}$, then*

$$T : \mathbb{F}^{m \times n_2 \times n_3} \to \mathbb{F}^{n_1 \times n_2 \times n_3}, \tag{4.8}$$

is linear.

Proof. Let $\mathcal{B} \in \mathbb{F}^{m \times n_2 \times n_3}$ and $\mathcal{C} \in \mathbb{F}^{m \times n_2 \times n_3}$ be third-order tensors and $\alpha, \beta \in \mathbb{F}$ be scalars, then

$$\begin{aligned}
T(\alpha\mathcal{B} + \beta\mathcal{C}) &= \mathcal{A} * (\alpha\mathcal{B} + \beta\mathcal{C}) \\
&= \mathrm{fold}\left(\sum_{j,k=1}^{n_3} P^{k-1} e_j \otimes A_j(\alpha B_k + \beta C_k)\right) \\
&= \mathrm{fold}\left(\sum_{j,k=1}^{n_3} \alpha P^{k-1} e_j \otimes A_j B_k\right) + \mathrm{fold}\left(\sum_{j,k=1}^{n_3} \beta P^{k-1} e_j \otimes A_j C_k\right) \\
&= \mathrm{fold}\left(\alpha \mathrm{MatVec}(\mathcal{A} * \mathcal{B})\right) + \mathrm{fold}\left(\beta \mathrm{MatVec}(\mathcal{A} * \mathcal{C})\right) \\
&= \alpha\mathcal{A} * \mathcal{B} + \beta\mathcal{A} * \mathcal{C} \\
&= \alpha T(\mathcal{B}) + \beta T(\mathcal{C}).
\end{aligned} \qquad \square$$

Lemma 4.1.5. *If $T(\mathcal{A}) = \mathcal{A} * \mathcal{X}$ where $\mathcal{A} \in \mathbb{F}^{n_1 \times m \times n_3}$ and $\mathcal{X} \in \mathbb{F}^{m \times n_2 \times n_3}$,* *then*

$$T : \mathbb{F}^{n_1\times m\times n_3} \to \mathbb{F}^{n_1\times n_2\times n_3}, \tag{4.9}$$

is linear.

Proof. The proof is similar to the proof of Lemma 4.1.4. □

We can represent the T-product in terms of the Kronecker product and thus explore the rank, invertibility and null space of the T-product.

Let $x^{(i)} \in \mathbb{F}^{n_3} \cong x^{(i)} \in \mathbb{F}^{n_3\times 1\times 1}$, then the circulant of the tensor $x^{(i)} \in \mathbb{F}^{n_3\times 1\times 1}$ is give by

$$\begin{aligned}\mathrm{circ}(x^{(i)}) &= P^{k-1}e_j e_k^T x^{(i)}\\ &= (x^{(i)}\, Px^{(i)} \cdots P^{n_3-1}x^{(i)}).\end{aligned}$$

Theorem 4.1.6. *Let $\mathcal{A} \in \mathbb{F}^{n_1\times n_2\times n_3}$ and $\mathcal{B} \in \mathbb{F}^{n_2\times l\times n_3}$ be rank r_A and r_B tensors, respectively, defined by*

$$\mathcal{A} = \sum_{i=1}^{r_A} u^{(i)} \otimes v^{(i)} \otimes (w^{(i)})^T, \quad \mathcal{B} = \sum_{j=1}^{r_B} x^{(j)} \otimes y^{(j)} \otimes (z^{(j)})^T. \tag{4.10}$$

Where $u^{(i)} \in \mathbb{F}^{n_1}$, $v^{(i)} \in \mathbb{F}^{n_2}$, $w^{(i)} \in \mathbb{F}^{n_3}$, $x^{(j)} \in \mathbb{F}^{n_2}$, $y^{(j)} \in \mathbb{F}^{l}$ and $z^{(j)} \in \mathbb{F}^{n_3}$ are the columns of $U \in \mathbb{F}^{n_1\times r_A}$, $V \in \mathbb{F}^{n_2\times r_A}$, $W \in \mathbb{F}^{n_3\times r_A}$, $X \in \mathbb{F}^{n_2\times r_B}$, $Y \in \mathbb{F}^{l\times r_B}$ and $Z \in \mathbb{F}^{n_3\times r_B}$ respectively. Let $t^{(i,j)}$ denote a vector circ$(w^{(i)})z^{(j)}$. Define the scalars $d_{ij} = (v^{(i)})^T x^{(j)}$. Then

$$\mathcal{A} * \mathcal{B} = \sum_{i=1}^{r_A}\sum_{j=1}^{r_B} d_{ij}(u^{(i)} \otimes y^{(j)} \otimes t^{(i,j)}). \tag{4.11}$$

Proof. The frontal slices of $\mathcal{A}$ and $\mathcal{B}$ are $\sum_{i=1}^{r_A} u^{(i)}(v^{(i)})^T w^{(i)}$ and $\sum_{j=1}^{r_A} x^{(j)}(y^{(j)})^T z^{(j)}$ respectively, thus

$$\begin{aligned}\mathrm{circ}(\mathcal{A}) &= \sum_{k,m}(P^{m-1}e_k)e_m^T \otimes \sum_{i=1}^{r_A} u^{(i)}(v^{(i)})^T w_k^{(i)}, \text{ and}\\ \mathrm{MatVec}(\mathcal{B}) &= \sum_k e_k \otimes \sum_{j=1}^{r_A} x^{(j)}(y^{(j)})^T z_k^{(j)}.\end{aligned}$$

Then,

$$\mathcal{A} * \mathcal{B} = \mathrm{fold}(\mathrm{circ}(\mathcal{A}) \cdot \mathrm{MatVec}(\mathcal{B}))$$

$$= \sum_{i=1}^{r_A}\sum_{j=1}^{r_B} u^{(i)}(v^{(i)})^T x^{(j)}(y^{(j)})^T \otimes \sum_{k,m}\sum_{k}(P^{m-1}e_k)(e_m^T w_k^{(i)})z_k^{(j)}$$
$$= \sum_{i=1}^{r_A}\sum_{j=1}^{r_B} d_{ij}(u^{(i)} \otimes y^{(j)} \otimes \mathrm{circ}(w^{(i)})z^{(j)})$$
$$= \sum_{i=1}^{r_A}\sum_{j=1}^{r_B} d_{ij}(u^{(i)} \otimes y^{(j)} \otimes t^{(i,j)}). \qquad \square$$

Corollary 4.1.7. *Let T denote the $r_A \times r_B$ matrix formed by taking the norms of the tubes $t^{(i,j)}$ of $\mathcal{T}$ and let Δ be the matrix with entries d_{ij}. Then the* $\mathrm{rank}(\mathcal{A}*\mathcal{B}) = nnz(\Delta \odot T) \leq r_A r_B$, *where nnz means number of nonzero entries of the tensor and $\odot$ stands for Hadamard (entrywise) product. Furthermore,* $\mathcal{B} \in \mathrm{null}(\mathcal{A})$ *if* $||\Delta \odot T|| = 0$.

Example 4.1.8. Let $\mathcal{A}$ and $\mathcal{B}$ be third-order tensors such that $\mathrm{rank}(\mathcal{A}) = \mathrm{rank}(\mathcal{B}) = 2$, and

$$\mathcal{A} = \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}1\\1\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}1\\1\end{pmatrix},$$
$$\mathcal{B} = \begin{pmatrix}1\\1\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}1\\1\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix}.$$

Then

$$\mathcal{A}*\mathcal{B} = \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix}$$
$$+ \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix}$$

thus $\mathcal{A}*\mathcal{B}$ is a rank 4 tensor and

$$\mathrm{rank}(\mathcal{A}*\mathcal{B}) = 4 = nnz\left(\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix} \odot \begin{pmatrix}\sqrt{2} & \sqrt{2}\\ \sqrt{2} & \sqrt{2}\end{pmatrix}\right).$$

Definition 4.1.9. If $\mathcal{A}$ is an $n_1 \times n_2 \times n_3$ tensor, then $\mathcal{A}^T$ is an $n_2 \times n_1 \times n_3$ transpose tensor obtained by transposing each of the frontal slices of $\mathcal{A}$ and reversing the order of the transposed frontal slices from the 2^{nd} entry through n_3, i.e, if $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2 \times n_3}$ then (indices of A^T are restricted by modulo n_3)

$$\mathrm{MatVec}(\mathcal{A}^T) = \sum_{j=1}^{n_3} e_j \otimes A^T_{[n_3-j+2]\mathrm{mod} n_3} = \begin{bmatrix}A_1^T\\A_{n_3}^T\\ \vdots \\ A_2^T\end{bmatrix}. \qquad (4.12)$$

Lemma 4.1.10. *Let $\mathcal{A}$ and $\mathcal{B}$ be tensors such that $\mathcal{A} * \mathcal{B}$ and $\mathcal{B}^T * \mathcal{A}^T$ are defined, then*

$$(\mathcal{A} * \mathcal{B})^T = \mathcal{B}^T * \mathcal{A}^T. \tag{4.13}$$

Proof.

$$\begin{aligned}
\mathrm{MatVec}(\mathcal{B}^T * \mathcal{A}^T) &= \mathrm{circ}(\mathcal{B}^T)\mathrm{MatVec}(\mathcal{A}^T) \\
&= \left(\sum_{j,k=1}^{n_3} P^{k-1} e_j e_k^T \otimes B_{n_3-j+2}^T\right)\left(\sum_{l=1}^{n_3} e_l \otimes A_{n_3-l+2}^T\right) \\
&= \sum_{j,k,l=1}^{n_3} P^{k-1} e_j e_k^T e_l \otimes B_{n_3-j+2}^T A_{n_3-l+2}^T \\
&= \sum_{j,k,l=1}^{n_3} P^{k-1} e_j e_k^T e_l \otimes (A_{n_3-l+2} B_{n_3-j+2})^T \\
&= \sum_{j,k=1}^{n_3} P^{k-1} e_j \otimes (A_{n_3-k+2} B_{n_3-j+2})^T \\
&= \sum_{j,k=1}^{n_3} e_{j+k-1} \otimes (A_{n_3-k+2} B_{n_3-j+2})^T.
\end{aligned}$$

Let $u = j + k - 1$ and $v = n_3 - j + 2$ then $v = n_3 - u + k + 1$. Thus $n_3 - k + 2 = 2n_3 - u - v + 3$ and $2n_3 - u - v + 3 \equiv n_3 - u - v + 3 (\mathrm{mod} n_3)$ so

$$\begin{aligned}
\sum_{j,k=1}^{n_3} e_{j+k-1} \otimes (A_{n_3-k+2} B_{n_3-j+2})^T &= \sum_{u=1}^{n_3} e_u \otimes \left(\sum_{v=1}^{n_3} A_{n_3-u-v+3} B_v\right)^T \\
&= \mathrm{MatVec}((\mathcal{A} * \mathcal{B})^T).
\end{aligned}$$

□

We will investigate the Frobenius norm of the T-product following the approach taken from [9].

Definition 4.1.11. Suppose that $\mathcal{A} = (a_{i_1 i_2 \dots i_m}) \in \mathbb{F}^{n_1 \times n_2 \times \cdots \times n_m}$ is an m-th order n-dimensional tensor. The Frobenius norm of $\mathcal{A}$ is

$$||\mathcal{A}||_F^2 = \sum_{i_1 i_2 \dots i_m} |a_{i_1 i_2 \dots i_m}|^2. \tag{4.14}$$

The relation between the Frobenius norm of a tensor $\mathcal{A}$ and trace is as follows [9]

$$||\mathcal{A}||_F^2 = \sum_j \mathrm{trace}(A_j^T A_j), \tag{4.15}$$

where A_j are frontal slices.

Example 4.1.12. Let $\mathcal{A}$ be a $3 \times 3 \times 2$ tensor with entries, $-a_{331} = a_{111} =$

$a_{232} = a_{132} = 1$, $-a_{121} = a_{221} = a_{212} = 4$, $a_{311} = a_{231} = a_{312} = a_{222} = 0$, $-a_{322} = a_{211} = a_{321} = a_{112} = 2$, $a_{131} = a_{122} = a_{332} = 3$. Then

$$A_1 = \begin{pmatrix} 1 & -4 & 3 \\ 2 & 4 & 0 \\ 0 & 2 & -1 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 2 & 3 & 1 \\ 4 & 0 & 1 \\ 0 & -2 & 3 \end{pmatrix}$$

$$A_1^T A_1 = \begin{pmatrix} 5 & 4 & 3 \\ 4 & 36 & -14 \\ 3 & -14 & 10 \end{pmatrix}, \quad A_2^T A_2 = \begin{pmatrix} 20 & 6 & 9 \\ 6 & 13 & -3 \\ 6 & -3 & 11 \end{pmatrix}.$$

So that,

$$\begin{aligned} \sum_{j=1}^{2} \text{trace}(A_j^T A_j) &= \text{trace}(A_1^T A_1) + \text{trace}(A_2^T A_2) \\ &= (5+36+10) + (20+13+11) \\ &= 95 \end{aligned}$$

and hence

$$\begin{aligned} ||\mathcal{A}||_F^2 &= ||A_1||_F^2 + ||A_2||_F^2 \\ &= \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij1}^2 + \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij2}^2 \\ &= 52 + 43 \\ &= 95. \end{aligned}$$

Definition 4.1.13. A $n \times n \times l$ tensor, $\mathcal{Q}$, is orthogonal if $\mathcal{Q}^T * \mathcal{Q} = \mathcal{Q} * \mathcal{Q}^T = \mathcal{I}_{nnl}$.

Lemma 4.1.14. *Let $\mathcal{A}$ be a tensor, then*

$$||\mathcal{A}||_F^2 = \text{trace}(\text{MatVec}(\mathcal{A})^T \text{MatVec}(\mathcal{A})).$$

Proof. Since

$$(\text{MatVec}(\mathcal{A}))^T \text{MatVec}(\mathcal{A}) = \left(\sum_j e_j^T \otimes A_j^T\right)\left(\sum_k e_k \otimes A_k\right)$$

$$= \sum_j (e_j^T \otimes A_j^T)(e_k \otimes A_k)$$
$$= \sum_{jk} e_j^T e_k \otimes A_j^T A_k$$
$$= \sum_j A_j^T A_j.$$

We have using (4.15),

$$\begin{aligned}\text{trace}(\text{MatVec}(\mathcal{A}))^T \text{MatVec}(\mathcal{A}) &= \sum_j \text{trace}(A_j^T A_j) \\ &= ||\mathcal{A}||_F^2. \end{aligned}$$ □

Definition 4.1.15. The Euclidean inner product $\langle \mathcal{A}, \mathcal{B} \rangle$ of two real tensors $\mathcal{A}$ and $\mathcal{B}$ is

$$\langle \mathcal{A}, \mathcal{B} \rangle = \text{trace}((\text{MatVec}(\mathcal{B}))^T \text{MatVec}(\mathcal{A})). \tag{4.16}$$

Note that $||\mathcal{A}||_F = \langle \mathcal{A}, \mathcal{A} \rangle$.

Lemma 4.1.16. *Let $\mathcal{Q}$, $\mathcal{B}$ and $\mathcal{C}$ be tensors, then*

$$\langle \mathcal{Q} * \mathcal{B}, \mathcal{C} \rangle = \langle \mathcal{B}, \mathcal{Q}^T * \mathcal{C} \rangle.$$

Proof.

$$\begin{aligned}\langle \mathcal{Q} * \mathcal{B}, \mathcal{C} \rangle &= \text{trace}((\text{MatVec}(\mathcal{C}))^T \text{circ}(\mathcal{Q}) \text{MatVec}(\mathcal{B})) \\ &= \text{trace}((\text{MatVec}(\mathcal{C}))^T (\text{circ}(\mathcal{Q})^T)^T \text{MatVec}(\mathcal{B})) \\ &= \text{trace}((\text{circ}(\mathcal{Q})^T \text{MatVec}(\mathcal{C}))^T \text{MatVec}(\mathcal{B})) \\ &= \langle \mathcal{B}, \mathcal{Q}^T * \mathcal{C} \rangle, \end{aligned}$$

where,

$$\begin{aligned}\text{circ}(\mathcal{Q})^T &= \left(\sum_{j,k=1}^{n_3} (P^{k-1} e_j) e_k^T \otimes \mathcal{Q}_j \right)^T \\ &= \sum_{j,k=1}^{n_3} ((P^{k-1} e_j) e_k^T)^T \otimes \mathcal{Q}_j^T \\ &= \sum_{j,k=1}^{n_3} ((e_k) e_j^T P^{1-k}) \otimes \mathcal{Q}_j^T \\ &= \sum_{j,k=1}^{n_3} (e_k e_{j+k-1}^T) \otimes \mathcal{Q}_j^T \\ &= \sum_{j,k=1}^{n_3} (P^{j+k-2} e_{2-j} e_{j+k-1}^T) \otimes \mathcal{Q}_j^T \end{aligned}$$

$$= \sum_{j,k=1}^{n_3} (P^{j+k-2} e_{n_3+2-j} e_{j+k-1}^T) \otimes \mathcal{Q}_j^T.$$

Let $j+k-1=s$ and $n_3+2-j=t$, where $j,k \in \{1,2,\dots n_3\}$. Then

$$\text{circ}(\mathcal{Q})^T = \sum_{t=1}^{n_3} \sum_{s=1}^{n_3} (P^{s-1} e_t) e_s^T \otimes Q_{n_3+2-t}^T = \text{circ}(\mathcal{Q}^T). \qquad \square$$

The following lemma shows that orthogonal tensors preserve the Frobenius norm.

Lemma 4.1.17. *If $\mathcal{Q}$ is an orthogonal tensor, then*

$$||\mathcal{Q} * \mathcal{A}||_F = ||\mathcal{A}||_F. \tag{4.17}$$

Proof.

$$\begin{aligned} ||\mathcal{Q} * \mathcal{A}||_F^2 &= \langle \mathcal{Q} * \mathcal{A}, \mathcal{Q} * \mathcal{A} \rangle \\ &= \langle \mathcal{A}, \mathcal{Q}^T * (\mathcal{Q} * \mathcal{A}) \rangle \\ &= \langle \mathcal{A}, (\mathcal{Q}^T * \mathcal{Q}) * \mathcal{A} \rangle \\ &= \langle \mathcal{A}, \mathcal{I} * \mathcal{A} \rangle \\ &= \langle \mathcal{A}, \mathcal{A} \rangle \\ &= ||\mathcal{A}||_F^2. \end{aligned} \qquad \square$$

4.2 Singular value decomposition

Definition 4.2.1. A tensor $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2 \times \cdots \times n_m}$ is said to be diagonal if $a_{i_1 i_2 \dots i_m} \neq 0$ only when $i_1 = i_2 = \cdots = i_m$.

Remark 4.2.2. A tensor is said to be f-diagonal if all of its frontal slices are diagonal.

Theorem 4.2.3. *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ be a third-order tensor. Then $\mathcal{A}$ can be factored as*

$$\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T, \tag{4.18}$$

where $\mathcal{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$, $\mathcal{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ are orthogonal and $\mathcal{S} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is an f-*diagonal tensor.*

Proof. First let

$$F_{n_3} = \frac{1}{\sqrt{n_3}} \begin{pmatrix} w^0 & w^0 & \cdots & w^0 \\ w^0 & w^1 & \cdots & w^{n_3-1} \\ \vdots & \vdots & \ddots & \vdots \\ w^0 & w^{n_3-1} & \cdots & w^{(n_3-1)^2} \end{pmatrix} \tag{4.19}$$

where $w = e^{i\frac{2\pi}{n_3}}$. Since circulant matrices are diagonalizable using the discrete Fourier transform domain [8], it follows that for $\text{circ}(\mathcal{A})$

$$(F_{n_3} \otimes I)\text{circ}(\mathcal{A})(F_{n_3}^* \otimes I) = \begin{pmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_{n_3} \end{pmatrix} \tag{4.20}$$

where D_i for $i = 1, 2, \ldots, n_3$ are matrices. Using (4.20) and Theorem 1.1.6 it follows that

$$\begin{aligned} \begin{pmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_{n_3} \end{pmatrix} &= \begin{pmatrix} U_1\Sigma_1V_1^T & 0 & \cdots & 0 \\ 0 & U_2\Sigma_2V_2^T & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & U_{n_3}\Sigma_{n_3}V_{n_3}^T \end{pmatrix} \\ &= \begin{pmatrix} U_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & U_{n_3} \end{pmatrix} \begin{pmatrix} \Sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Sigma_{n_3} \end{pmatrix} \begin{pmatrix} V_1^T & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & V_{n_3}^T \end{pmatrix}. \end{aligned}$$

Now applying $(F_{n_3}^* \otimes I)$ on the left and $(F_{n_3} \otimes I)$ on the right of (4.20) it follows that

$$\begin{aligned} (F_{n_3}^* \otimes I) \begin{pmatrix} U_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & U_{n_3} \end{pmatrix} (F_{n_3} \otimes I) &= \text{circ}(\mathcal{U}), \\ (F_{n_3}^* \otimes I) \begin{pmatrix} \Sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Sigma_{n_3} \end{pmatrix} (F_{n_3} \otimes I) &= \text{circ}(\mathcal{S}), \\ (F_{n_3}^* \otimes I) \begin{pmatrix} V_1^T & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & V_{n_3}^T \end{pmatrix} (F_{n_3} \otimes I) &= \text{circ}(\mathcal{V}^T). \end{aligned}$$

Thus,

$$\text{circ}(\mathcal{A}) = \text{circ}(\mathcal{U})\text{circ}(\mathcal{S})\text{circ}(\mathcal{V}^T). \tag{4.21}$$

There is a bijection between $\mathcal{A}$ and circ$(\mathcal{A})$, thus

$$\begin{aligned}\text{circ}^{-1}(\text{circ}(\mathcal{A})) &= \mathcal{A} \\ &= \text{fold}(\text{circ}(\mathcal{A})(e_1 \otimes I)) \\ &= \text{fold}(\text{MatVec}(\mathcal{A})). \end{aligned} \tag{4.22}$$

Now using (4.21) and (4.22),

$$\begin{aligned}\mathcal{A} &= \text{fold}[\text{circ}(\mathcal{U})\text{circ}(\mathcal{S})(\text{circ}(\mathcal{V}^T)(e_1 \otimes I))] \\ &= \text{fold}[\text{circ}(\mathcal{U})\text{circ}(\mathcal{S})(\text{MatVec}(\mathcal{V}^T)].\end{aligned}$$

Let MatVec$(\mathcal{T})$ = circ$(\mathcal{S})$(MatVec$(\mathcal{V}^T)$. It follows that

$$\begin{aligned}\mathcal{A} &= \text{fold}[\text{circ}(\mathcal{U})\text{MatVec}(\mathcal{T})] \\ &= \mathcal{U} * \mathcal{T}. && \text{(by the definition of the T-product)}\end{aligned}$$

Thus, MatVec$(\mathcal{T})$ = circ$(\mathcal{S})$(MatVec$(\mathcal{V}^T) \Rightarrow \mathcal{T}$ = fold[circ$(\mathcal{S})$(MatVec$(\mathcal{V}^T)$)]. Therefore, by the definition of the T-product

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

4.2.1 Moore-Penrose pseudo inverse

Similar to the Einstein product, the Moore-Penrose pseudo inverse must satisfy the axioms given in the following theorem.

Theorem 4.2.4. *The Moore-Penrose pseudo inverse exists for the T-product and it is unique.*

To prove that the Moore-Penrose inverse exists we need to show that there is a tensor $\mathcal{A}^+$ which satisfies the following axioms,

1. $\mathcal{A} * \mathcal{A}^+ * \mathcal{A} = \mathcal{A}$.
2. $\mathcal{A}^+ * \mathcal{A} * \mathcal{A}^+ = \mathcal{A}^+$.
3. $(\mathcal{A}^+ * \mathcal{A})^T = \mathcal{A}^+ * \mathcal{A}$.
4. $(\mathcal{A} * \mathcal{A}^+)^T = \mathcal{A} * \mathcal{A}^+$.

Proof. First, let $\mathcal{A}^+ \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ be a third order tensor such that

$$(F_{n_3} \otimes I)\text{circ}(\mathcal{A}^+)(F_{n_3}^* \otimes I) = \begin{pmatrix} D_1^+ & 0 & \cdots & 0 \\ 0 & D_2^+ & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_{n_3}^+ \end{pmatrix}, \tag{4.23}$$

where D_i^+ is the Moore-Penrose inverse of D_i.

1. We will show that $\mathcal{A} * \mathcal{A}^+ * \mathcal{A} = \mathcal{A}$.

   In (4.20), we have the diagonal block matrix such that $D_i = U_i \Sigma_i V_i^T$ for all $i = 1, \ldots, n_3$. Here we let $D_i^+ = V_i \Sigma_i^+ U_i^T$, it follows that, applying (4.20) and (4.23) we get

$$\begin{aligned} &[(F_{n_3} \otimes I)\text{circ}(\mathcal{A})\text{circ}(\mathcal{A}^+)\text{circ}(\mathcal{A})(F_{n_3}^* \otimes I)] \\ &= \begin{pmatrix} D_1 D_1^+ D_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & D_{n_3} D_{n_3}^+ D_{n_3} \end{pmatrix}. \end{aligned} \tag{4.24}$$

   Thus, for all i we have

$$(F_{n_3} \otimes I)\text{circ}(\mathcal{A})\text{circ}(\mathcal{A}^+)\text{circ}(\mathcal{A})(F_{n_3}^* \otimes I) = \text{diag}(D_i D_i^+ D_i). \tag{4.25}$$

   Since D_i^+ is the Moore-Penrose inverse of D_i, it holds that

$$D_i D_i^+ D_i = D_i,$$

   thus, (4.25) becomes

$$(F_{n_3} \otimes I)\text{circ}(\mathcal{A})\text{circ}(\mathcal{A}^+)\text{circ}(\mathcal{A})(F_{n_3}^* \otimes I) = \text{diag}(D_i),$$

   and hence,

$$\begin{aligned} \text{circ}(\mathcal{A})\text{circ}(\mathcal{A}^+)\text{circ}(\mathcal{A}) &= (F_{n_3}^* \otimes I)\text{diag}(D_i)(F_{n_3} \otimes I) \\ &= \text{circ}(\mathcal{A}). \end{aligned}$$

   Let $\text{MatVec}(\mathcal{K}) = \text{circ}(\mathcal{A}^+)\text{MatVec}(\mathcal{A})$, i.e., $\mathcal{K} = \mathcal{A}^+ * \mathcal{A}$, by the definition of the T-product. Using (4.22) it follows that

$$\begin{aligned} \mathcal{A} &= \text{fold}[\text{circ}(\mathcal{A})\text{circ}(\mathcal{A}^+)\text{circ}(\mathcal{A})(e_1 \otimes I)] \\ &= \text{fold}[\text{circ}(\mathcal{A})\text{circ}(\mathcal{A}^+)\text{MatVec}(\mathcal{A})] \end{aligned}$$

$$
\begin{aligned}
&= \text{fold}[\text{circ}(\mathcal{A})\text{MatVec}(\mathcal{K})] \\
&= \mathcal{A} * \mathcal{K} \\
&= \mathcal{A} * \mathcal{A}^{+} * \mathcal{A}. \qquad (4.26)
\end{aligned}
$$

2. Similarly, we prove that $\mathcal{A}^{+} * \mathcal{A} * \mathcal{A}^{+} = \mathcal{A}^{+}$. Using (4.23) it follows that

$$(F_{n_3} \otimes I)\text{circ}(\mathcal{A}^{+})\text{circ}(\mathcal{A})\text{circ}(\mathcal{A}^{+})(F_{n_3}^{*} \otimes I) = \text{diag}(D_i^{+} D_i D_i^{+}). \quad (4.27)$$

Since

$$D_i^{+} D_i D_i^{+} = D_i^{+},$$

we find that

$$(F_{n_3} \otimes I)\text{circ}(\mathcal{A}^{+})\text{circ}(\mathcal{A})\text{circ}(\mathcal{A}^{+})(F_{n_3}^{*} \otimes I) = \text{diag}(D_i^{+}).$$

so that,

$$
\begin{aligned}
\text{circ}(\mathcal{A}^{+})\text{circ}(\mathcal{A})\text{circ}(\mathcal{A}^{+}) &= (F_{n_3}^{*} \otimes I)\text{diag}(D_i^{+})(F_{n_3} \otimes I) \\
\text{circ}(\mathcal{A}^{+})\text{circ}(\mathcal{A})\text{circ}(\mathcal{A}^{+}) &= \text{circ}(\mathcal{A}^{+}) \\
\text{fold}[\text{circ}(\mathcal{A}^{+})\text{circ}(\mathcal{A})\text{circ}(\mathcal{A}^{+})(e_1 \otimes I)] &= \mathcal{A}^{+} \\
\text{fold}[\text{circ}(\mathcal{A}^{+})\text{circ}(\mathcal{A})\text{MatVec}(\mathcal{A}^{+})] &= \mathcal{A}^{+} \\
\text{fold}[\text{circ}(\mathcal{A}^{+})\text{MatVec}(\mathcal{U})] &= \mathcal{A}^{+} \\
\mathcal{A}^{+} * \mathcal{U} &= \mathcal{A}^{+} \\
\mathcal{A}^{+} * \mathcal{A} * \mathcal{A}^{+} &= \mathcal{A}^{+}.
\end{aligned}
$$

where $\mathcal{U} = \text{fold}(\text{circ}(\mathcal{A})\text{MatVec}(\mathcal{A}^{+}))$.

3. We also show that $(\mathcal{A}^{+} * \mathcal{A})^{T} = \mathcal{A}^{+} * \mathcal{A}$. As before we use (4.23) and we get

$$
\begin{aligned}
(F_{n_3} \otimes I)\text{circ}(\mathcal{A}^{T})\text{circ}((\mathcal{A}^{+})^{T})(F_{n_3}^{*} \otimes I) &= \text{diag}(D_i^{T}(D_i^{+})^{T}) \qquad (4.28) \\
&= \text{diag}((D_i^{+} D_i)^{T}) \\
&= \text{diag}(D_i^{+} D_i).
\end{aligned}
$$

Now from the left side of (4.28), we have

$$
\begin{aligned}
&[\text{circ}((\mathcal{A}^{+}))\text{circ}(\mathcal{A})]^{T} \\
&\quad = \text{circ}(\mathcal{A}^{T})\text{circ}((\mathcal{A}^{+})^{T}) \\
&\quad = (F_{n_3}^{*} \otimes I)\text{diag}(D_i^{+} D_i)(F_{n_3} \otimes I) \\
&\quad = (F_{n_3}^{*} \otimes I)\text{diag}(D_i^{+})\text{diag}(D_i)(F_{n_3} \otimes I)
\end{aligned}
$$

$$
\begin{aligned}
&= (F_{n_3}^* \otimes I)\mathrm{diag}(D_i^+)(F_{n_3} \otimes I)(F_{n_3}^* \otimes I)\mathrm{diag}(D_i)(F_{n_3} \otimes I) \\
&= \mathrm{circ}(\mathcal{A}^+)\mathrm{circ}(\mathcal{A}).
\end{aligned}
$$

Now, it follows that

$$
\begin{aligned}
(\mathcal{A}^+ * \mathcal{A})^T &= (\mathcal{A})^T * (\mathcal{A}^+)^T \\
&= \mathrm{fold}[\mathrm{circ}(\mathcal{A}^T)\mathrm{MatVec}((\mathcal{A}^+)^T)] \\
&= \mathrm{fold}[\mathrm{circ}(\mathcal{A}^T)\mathrm{circ}((\mathcal{A}^+)^T)(e_1 \otimes I)] \\
&= \mathrm{fold}[(\mathrm{circ}(\mathcal{A}^+)\mathrm{circ}(\mathcal{A}))^T(e_1 \otimes I)] \\
&= \mathrm{fold}[\mathrm{circ}(\mathcal{A}^+)\mathrm{circ}(\mathcal{A})(e_1 \otimes I)] \\
&= \mathrm{fold}[\mathrm{circ}(\mathcal{A}^+)\mathrm{MatVec}(\mathcal{A})] \\
&= \mathcal{A}^+ * \mathcal{A}.
\end{aligned}
$$

4. Similarly we will show that $(\mathcal{A} * \mathcal{A}^+)^T = \mathcal{A} * \mathcal{A}^+$

$$
(F_{n_3} \otimes I)\mathrm{circ}((\mathcal{A}^+)^T)\mathrm{circ}(\mathcal{A}^T)(F_{n_3}^* \otimes I) = \mathrm{diag}(D_i D_i^+)
$$

and,

$$
\begin{aligned}
&[\mathrm{circ}(\mathcal{A})\mathrm{circ}(\mathcal{A}^+)]^T \\
&\quad = \mathrm{circ}(\mathcal{A}^+)^T\mathrm{circ}(\mathcal{A})^T \\
&\quad = \mathrm{circ}((\mathcal{A}^+)^T)\mathrm{circ}(\mathcal{A}^T) \\
&\quad = (F_{n_3}^* \otimes I)\mathrm{diag}(D_i D_i^+)(F_{n_3} \otimes I) \\
&\quad = (F_{n_3}^* \otimes I)\mathrm{diag}(D_i)(F_{n_3} \otimes I)(F_{n_3}^* \otimes I)\mathrm{diag}(D_i^+)(F_{n_3} \otimes I) \\
&\quad = \mathrm{circ}(\mathcal{A})\mathrm{circ}(\mathcal{A}^+).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
(\mathcal{A} * \mathcal{A}^+)^T &= (\mathcal{A}^+)^T * \mathcal{A}^T \\
&= \mathrm{fold}[\mathrm{circ}((\mathcal{A}^+)^T)\mathrm{MatVec}(\mathcal{A}^T)] \\
&= \mathrm{fold}[\mathrm{circ}((\mathcal{A}^+)^T)\mathrm{circ}(\mathcal{A}^T)(e_1 \otimes I)] \\
&= \mathrm{fold}[\mathrm{circ}(\mathcal{A}^+)^T\mathrm{circ}(\mathcal{A})^T(e_1 \otimes I)] \\
&= \mathrm{fold}[(\mathrm{circ}(\mathcal{A})\mathrm{circ}(\mathcal{A}^+))^T(e_1 \otimes I)] \\
&= \mathrm{fold}[\mathrm{circ}(\mathcal{A})\mathrm{circ}(\mathcal{A}^+)(e_1 \otimes I)] \\
&= \mathrm{fold}[\mathrm{circ}(\mathcal{A})\mathrm{MatVec}(\mathcal{A}^+)] \\
&= \mathcal{A} * \mathcal{A}^+.
\end{aligned}
$$

For uniqueness, let $\mathcal{A}^+$ and $\mathcal{B}^+$ be the Moore-Penrose inverse of $\mathcal{A}$, then

$$\begin{aligned}
\mathcal{A}^+ &= \mathcal{A}^+ * \mathcal{A} * \mathcal{A}^+ \\
&= \mathcal{A}^+ * \mathcal{A} * \mathcal{B}^+ * \mathcal{A} * \mathcal{A}^+ \\
&= \mathcal{A}^+ * (\mathcal{A} * \mathcal{B}^+)^T * (\mathcal{A} * \mathcal{A}^+)^T \\
&= \mathcal{A}^+ * (\mathcal{B}^+)^T * \mathcal{A}^T * (\mathcal{A}^+)^T * \mathcal{A}^T \\
&= \mathcal{A}^+ * (\mathcal{B}^+)^T * (\mathcal{A} * \mathcal{A}^+ * \mathcal{A})^T \\
&= \mathcal{A}^+ * (\mathcal{B}^+)^T * \mathcal{A}^T \\
&= \mathcal{A}^+ * (\mathcal{A} * \mathcal{B}^+)^T \\
&= \mathcal{A}^+ * \mathcal{A} * \mathcal{B}^+ \\
&= (\mathcal{A}^+ * \mathcal{A})^T * \mathcal{B}^+ \\
&= (\mathcal{A})^T * (\mathcal{A}^+)^T * \mathcal{B}^+ \\
&= (\mathcal{A} * \mathcal{B}^+ * \mathcal{A})^T * (\mathcal{A}^+)^T * \mathcal{B}^+ \\
&= (\mathcal{B}^+ * \mathcal{A})^T * \mathcal{A}^T * (\mathcal{A}^+)^T * \mathcal{B}^+ \\
&= (\mathcal{B}^+ * \mathcal{A})^T * (\mathcal{A}^+ * \mathcal{A})^T * \mathcal{B}^+ \\
&= \mathcal{B}^+ * \mathcal{A} * \mathcal{A}^+ * \mathcal{A} * \mathcal{B}^+ \\
&= \mathcal{B}^+ * \mathcal{A} * \mathcal{B}^+ \\
&= \mathcal{B}^+.
\end{aligned}$$

□

Corollary 4.2.5. *Let $\mathcal{A}$ be a third order tensor with singular value decomposition $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$, then the Moore-Penrose pseudo inverse is*

$$\mathcal{A}^+ = \mathcal{V} * \mathcal{S}^+ * \mathcal{U}^T, \tag{4.29}$$

where $\mathcal{U}$ and $\mathcal{V}$ are orthogonal tensors and $\mathcal{S}^+$ is f*-diagonal.*

Proof. The proof is similar to the proof of Theorem 4.2.3. □

Chapter 5

Tensor-tensor product

In this chapter, we will introduce the general tensor-tensor product, the product is the generalization of the matrix product. We will study this product for higher-order tensors and compare its algebraic properties with the linear algebraic properties. The general tensor-tensor product as defined in [16, 2] is similar to Liqun Qi's product [14]. We approach this product using the composition of multilinear functions.

5.1 General tensor product

Definition 5.1.1. Let $\mathcal{A}$ be a $m \geq 2$ order tensor and $\mathcal{B}$ be a $k \geq 1$ order tensor. The product $\mathcal{C}$, between $\mathcal{A}$ and $\mathcal{B}$ is an order $(m-1)(k-1)+1$ n-dimension tensor with entries

$$c_{i\alpha_1\cdots\alpha_{m-1}} = \sum_{i_2,\ldots,i_m=1} a_{ii_2\cdots i_m} b_{i_2\alpha_1} \cdots b_{i_m\alpha_{m-1}}, \tag{5.1}$$

where $\alpha_j := (\alpha_{j2}, \alpha_{j3}, \ldots, \alpha_{jk}) \in \mathbb{N}^{k-1}$ each denote $k-1$ indices.

We will show that the product in Definition 5.1.1 is compositional. Suppose that $\phi_{\mathcal{A}}$ and $\phi_{\mathcal{B}}$ are multilinear functions given by

$$\phi_{\mathcal{A}}(e_{i_2}, \ldots, e_{i_m}) = \sum_{i_1} a_{i_1 i_2 i_3\cdots i_m} e_{i_1}, \quad \phi_{\mathcal{B}}(e_{i_2}, \ldots, e_{i_k}) = \sum_{i_1} b_{i_1 i_2 i_3\cdots i_m} e_{i_1}. \tag{5.2}$$

$\mathcal{A}$ is a representation tensor of $\phi_{\mathcal{A}}$ and $\mathcal{B}$ is a representation tensor of $\phi_{\mathcal{B}}$. Let $\phi_{\mathcal{B}}^{m-1} : \mathbb{F}^{(m-1)(k-1)} \to \mathbb{F}^{m-1}$ be a function, such that for $x_{11}, \ldots, x_{(m-1)(k-1)}$

$$\phi_{\mathcal{B}}^{m-1}(x_{11}, \ldots, x_{(m-1)(k-1)}) = (\phi_{\mathcal{B}}(x_{11}, \ldots, x_{1(k-1)}), \ldots, \phi_{\mathcal{B}}(x_{(m-1)1}, \ldots, x_{(m-1)(k-1)})).$$

The composition $\phi_{\mathcal{A}} \circ \phi_{\mathcal{B}}^{m-1}$ is defined by its action on the standard basis,

$$\begin{aligned}
&\phi_{\mathcal{A}} \circ \phi_{\mathcal{B}}^{m-1}(e_{\alpha_{22}}, \ldots, e_{\alpha_{m(k-1)}}) \\
&\quad = \phi_{\mathcal{A}}(\phi_{\mathcal{B}}(e_{\alpha_{22}}, \ldots, e_{\alpha_{2(k-1)}}), \ldots, \phi_{\mathcal{B}}(e_{\alpha_{m2}}, \ldots, e_{\alpha_{m(k-1)}})) \\
&\quad = \phi_{\mathcal{A}}\left(\sum_{i_2} b_{i_2\alpha_2} e_{i_2}, \ldots, \sum_{i_m} b_{i_m\alpha_m} e_{i_m}\right) \\
&\quad = \sum_{i_2,\ldots,i_m} b_{i_2\alpha_2} b_{i_3\alpha_3} \cdots b_{i_m\alpha_m} \phi_{\mathcal{A}}(e_{i_2}, e_{i_3}, \ldots, e_{i_m}) \\
&\quad = \sum_{i_2,\ldots,i_m=1}^{n} a_{ii_2\cdots i_m} b_{i_2\alpha_2} \cdots b_{i_m\alpha_m} e_i.
\end{aligned}$$

The general tensor-tensor product may be illustrated by the following diagram. Let $\phi_{\mathcal{A}} : B^m \to A$ and $\phi_{\mathcal{B}} : C^k \to B$, then

$$\begin{tikzcd}
 & & & (C^k)^m \arrow[dlll, "\pi_1"'] \arrow[dll, "\pi_2"] \arrow[d, "\pi_i"] \arrow[drr, "\pi_m"] & & \\
C^k \arrow[d, "\phi_B"] & C^k \arrow[d, "\phi_B"] & \cdots & C^k \arrow[d, "\phi_B"] & \cdots & C^k \arrow[d, "\phi_B"] \\
B & B & \cdots & B & \cdots & B \\
 & & & B^m \arrow[ulll, "\pi_1"] \arrow[ull, "\pi_2"'] \arrow[u, "\pi_i"'] \arrow[urr, "\pi_m"'] \arrow[d, "\phi_A"] & & \\
 & & & A & &
\end{tikzcd}$$

Figure 5.1

where $\pi_j(c_1, \ldots, c_m) = c_j$ is the usual projection. Here, unlike the usual diagrammatic representation, $B^m \xrightarrow{\pi_1} B \xleftarrow{\phi_B} C^k$ denotes the equality $\pi_1(b) = \phi_B(c)$ which is sensible in this diagram. We compose $\phi_{\mathcal{A}}$ with all the $\phi_{\mathcal{B}}$'s multilinearly. Let us consider the product between a third-order tensor and a second-order tensor.

Example 5.1.2. Let $\mathcal{A} \in \mathbb{F}^{2\times2\times2}$ be a third-order tensor and $B \in \mathbb{F}^{2\times2}$ be a second-order tensor, where

$$\mathrm{frnt}(\mathcal{A}) = \left(\begin{array}{cc|cc} 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{array}\right) \text{ and } B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.$$

The product $\mathcal{A}B$ defined in Definition 5.1.1 is

$$\begin{aligned} c_{i\alpha_1\alpha_2} &= \sum_{i_2,i_3=1}^{2} a_{ii_2i_3} b_{i_2\alpha_1} b_{i_3\alpha_2} \\ &= a_{i11}b_{1\alpha_1}b_{1\alpha_2} + a_{i12}b_{1\alpha_1}b_{2\alpha_2} + a_{i21}b_{2\alpha_1}b_{1\alpha_2} + a_{i22}b_{2\alpha_1}b_{2\alpha_2}. \end{aligned}$$

Hence,

$$\operatorname{frnt}(\mathcal{C}) = \left(\begin{array}{cc|cc} 11 & 4 & 5 & 2 \\ 4 & 2 & 2 & 1 \end{array}\right).$$

Theorem 5.1.3. *The tensor-tensor product has the following properties.*

1. $(\mathcal{A}+\mathcal{B})\mathcal{C} = \mathcal{A}\mathcal{C} + \mathcal{B}\mathcal{C}$, *(where $\mathcal{A}$ and $\mathcal{B}$ have the same order)*
2. $\mathcal{A}(\mathcal{B}+\mathcal{C}) = \mathcal{A}\mathcal{B} + \mathcal{A}\mathcal{C}$,*(where $\mathcal{A}$ is a matrix)*
3. $(\lambda\mathcal{A})\mathcal{B} = \lambda(\mathcal{A}\mathcal{B})$,
4. $\mathcal{A}(\lambda\mathcal{B}) = \lambda^{m-1}\mathcal{A}\mathcal{B}$.*(where $m \geq 2$ and $k \geq 1$)*

Proof. **1.** By the pointwise definition of the addition of tensors

$$\begin{aligned} &(\phi_{\mathcal{A}} + \phi_{\mathcal{B}}) \circ \phi_{\mathcal{C}}^{m-1}(e_{\alpha_{22}}, \ldots, e_{\alpha_{m(k-1)}}) \\ &= (\phi_{\mathcal{A}} + \phi_{\mathcal{B}})(\phi_{\mathcal{C}}(e_{\alpha_{22}}, \ldots, e_{\alpha_{2(k-1)}}), \ldots, \phi_{\mathcal{C}}(e_{\alpha_{m2}}, \ldots, e_{\alpha_{m(k-1)}})) \\ &= \phi_{\mathcal{A}}(\phi_{\mathcal{C}}(e_{\alpha_{22}}, \ldots, e_{\alpha_{2(k-1)}}), \ldots, \phi_{\mathcal{C}}(e_{\alpha_{m2}}, \ldots, e_{\alpha_{m(k-1)}})) \\ &+ \phi_{\mathcal{B}}(\phi_{\mathcal{C}}(e_{\alpha_{22}}, \ldots, e_{\alpha_{2(k-1)}}), \ldots, \phi_{\mathcal{C}}(e_{\alpha_{m2}}, \ldots, e_{\alpha_{m(k-1)}})) \\ &= \phi_{\mathcal{A}} \circ \phi_{\mathcal{C}}^{m-1}(e_{\alpha_{22}}, \ldots, e_{\alpha_{m(k-1)}}) + \phi_{\mathcal{B}} \circ \phi_{\mathcal{C}}^{m-1}(e_{\alpha_{22}}, \ldots, e_{\alpha_{m(k-1)}}). \end{aligned}$$

2. Let A be a matrix and $\mathcal{B}$ a tensor then

$$\phi_A \circ \phi_{\mathcal{B}}^{m-1}(e_{\alpha_{12}}, \ldots, e_{\alpha_{1(k-1)}}) = \phi_A(\phi_{\mathcal{B}}^{m-1}(e_{\alpha_{12}}, e_{\alpha_{13}}, \ldots, e_{\alpha_{1(k-1)}}))$$

We also know that from (5.2),

$$\phi_{\mathcal{B}}(e_{i_2}, \ldots, e_{i_m}) = \sum_{i_1} b_{i_1i_2i_3\cdots i_m} e_{i_1}, \ \phi_{\mathcal{C}}(e_{j_2}, \ldots, e_{j_k}) = \sum_{j_1} c_{j_1j_2j_3\cdots j_m} e_{j_1}.$$

Thus,

$$\begin{aligned} \phi_{\mathcal{B}+\mathcal{C}}(e_{i_2}, \ldots, e_{i_m}) &= \sum_{i_1} (b_{i_1i_2\cdots i_m} + c_{i_1i_2\cdots i_m}) e_{i_1} \\ &= \sum_{i_1} b_{i_1i_2\cdots i_m} e_{i_1} + \sum_{i_1} c_{i_1i_2\cdots i_m} e_{i_1} \end{aligned}$$

$$= \phi_{\mathcal{B}}(e_{i_2}, \ldots, e_{i_m}) + \phi_{\mathcal{C}}(e_{i_2}, \ldots, e_{i_m}).$$

Therefore by the linearity of ϕ_A (where A is a matrix) it follows that

$$\begin{aligned}
&\phi_A \circ \phi_{\mathcal{B}+\mathcal{C}}(e_{\alpha_{12}}, \ldots, e_{\alpha_{1(k-1)}}) \\
&= \phi_A(\phi_{\mathcal{B}}(e_{\alpha_{12}}, \ldots, e_{\alpha_{1(k-1)}}) + \phi_{\mathcal{C}}(e_{\alpha_{12}}, \ldots, e_{\alpha_{1(k-1)}})) \\
&= \phi_A(\phi_{\mathcal{B}}(e_{\alpha_{12}}, \ldots, e_{\alpha_{1(k-1)}})) + \phi_A(\phi_{\mathcal{C}}(e_{\alpha_{12}}, \ldots, e_{\alpha_{1(k-1)}})).
\end{aligned}$$

3. By pointwise scalar multiplication of tensors

$$\begin{aligned}
&(\lambda\phi_A) \circ \phi_B^{m-1}(e_{\alpha_{22}}, \ldots, e_{\alpha_{m(k-1)}}) \\
&\quad = \lambda\phi_A(\phi_B(e_{\alpha_{22}}, \ldots, e_{\alpha_{2(k-1)}}), \ldots, \phi_B(e_{\alpha_{m2}}, \ldots, e_{\alpha_{m(k-1)}})). \\
&\quad = \lambda(\phi_A \circ \phi_B^{m-1}(e_{\alpha_{22}}, \ldots, e_{\alpha_{m(k-1)}}).
\end{aligned}$$

4. By multilinearity of ϕ_A,

$$\begin{aligned}
&\phi_A \circ (\lambda\phi_B^{m-1})(e_{\alpha_{22}}, \ldots, e_{\alpha_{m(k-1)}}) \\
&\quad = \phi_A(\lambda\phi_B)(e_{\alpha_{22}}, \ldots, e_{\alpha_{m(k-1)}}) \\
&\quad = \phi_A(\lambda\phi_B(e_{\alpha_{22}}, \ldots, e_{\alpha_{2(k-1)}}), \ldots, \lambda\phi_B(e_{\alpha_{m2}}, \ldots, e_{\alpha_{m(k-1)}})) \\
&\quad = \lambda^{m-1}\phi_A(\phi_B(e_{\alpha_{22}}, \ldots, e_{\alpha_{2(k-1)}}), \ldots, \phi_B(e_{\alpha_{m2}}, \ldots, e_{\alpha_{m(k-1)}})) \\
&\quad = \lambda^{m-1}(\phi_A \circ \phi_B^{m-1}(e_{\alpha_{22}}, \ldots, e_{\alpha_{m(k-1)}})).
\end{aligned}$$

□

Theorem 5.1.4. *Let $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ be an order $m+1$, an order $k+1$ and an order $r+1$, dimension n tensor, respectively. Then*

$$\mathcal{A}(\mathcal{B}\mathcal{C}) = (\mathcal{A}\mathcal{B})\mathcal{C}.$$

Proof. We will illustrate the proof by constructing the compositional diagram. Associativity follows from the associativity of the Cartesian product and the fact that the diagrams are composed of paths from $(D^r)^{mk}$ to A. Each path in the first diagram is equivalent to a distinct path in the second diagram and vice-versa. We use the fact that compositions of projections are projections again due to the associativity of the Cartesian product, i.e., the isomorphism $(C^k)^m \cong C^{mk}$ in the first diagram expands to the second diagram,

$$(\phi_A \circ \phi_B^{m-1}) \circ \phi_C^{mk-1} \equiv$$

$$\begin{tikzcd}
 & & & (D^r)^{mk} \arrow[dlll, "\pi_1"'] \arrow[dll, "\pi_2"] \arrow[d, "\pi_i"] \arrow[drr, "\pi_{mk}"] & & \\
D^r \arrow[d, "\phi_C"] & D^r \arrow[d, "\phi_C"] & \cdots & D^r \arrow[d, "\phi_C"] & \cdots & D^r \arrow[d, "\phi_C"] \\
C & C & \cdots & C & \cdots & C \\
 & & & C^{mk} \arrow[ulll, "\pi_1"] \arrow[ull, "\pi_2"'] \arrow[u, "\pi_i"] \arrow[urr, "\pi_m"'] \arrow[d, shift left] & & \\
 & & & (C^k)^m \arrow[u, shift left] \arrow[dlll, "\pi_1"'] \arrow[dll, "\pi_2"] \arrow[d, "\pi_i"] \arrow[drr, "\pi_m"] & & \\
C^k \arrow[d, "\phi_B"] & C^k \arrow[d, "\phi_B"] & \cdots & C^k \arrow[d, "\phi_B"] & \cdots & C^k \arrow[d, "\phi_B"] \\
B & B & \cdots & B & \cdots & B \\
 & & & B^m \arrow[ulll, "\pi_1"] \arrow[ull, "\pi_2"'] \arrow[u, "\pi_i"] \arrow[urr, "\pi_m"'] \arrow[d, "\phi_A"] & & \\
 & & & A & &
\end{tikzcd}$$

Figure 5.2

$$\phi_A \circ (\phi_B^{m-1} \circ \phi_C^{k-1})^{m-1} \equiv$$

$$\begin{tikzcd}
 & & & & & & (D^{rk})^m \arrow[dllll, "\pi_1"'] \arrow[d, "\pi_2"] \arrow[drrrrr, "\pi_m"] & & & & & & \\
 & & (D^r)^k \arrow[dll, "\pi_1"'] \arrow[dl, "\pi_2"] \arrow[dr, "\pi_k"] & & & & (D^r)^k \arrow[dll, "\pi_1"'] \arrow[dl, "\pi_2"] \arrow[dr, "\pi_k"] & & \cdots & & & (D^r)^k \arrow[dll, "\pi_1"'] \arrow[dl, "\pi_2"] \arrow[dr, "\pi_k"] & \\
D^r \arrow[d, "\phi_C"] & D^r \arrow[d, "\phi_C"] & \cdots & D^r \arrow[d, "\phi_C"] & D^r \arrow[d, "\phi_C"] & D^r \arrow[d, "\phi_C"] & \cdots & D^r \arrow[d, "\phi_C"] & \cdots & D^r \arrow[d, "\phi_C"] & D^r \arrow[d, "\phi_C"] & \cdots & D^r \arrow[d, "\phi_C"] \\
C & C & \cdots & C & C & C & \cdots & C & \cdots & C & C & \cdots & C \\
 & & C^k \arrow[ull, "\pi_1"] \arrow[ul, "\pi_2"'] \arrow[ur, "\pi_k"'] \arrow[d, "\phi_B"] & & & & C^k \arrow[ull, "\pi_1"] \arrow[ul, "\pi_2"'] \arrow[ur, "\pi_k"'] \arrow[d, "\phi_B"] & & \cdots & & & C^k \arrow[ull, "\pi_1"] \arrow[ul, "\pi_2"'] \arrow[ur, "\pi_k"'] \arrow[d, "\phi_B"] & \\
 & & B & & & & B & & \cdots & & & B & \\
 & & & & & & B^m \arrow[ullll, "\pi_1"] \arrow[u, "\pi_2"'] \arrow[urrrrr, "\pi_m"'] \arrow[d, "\phi_A"] & & & & & & \\
 & & & & & & A & & & & & &
\end{tikzcd}$$

Figure 5.3

Definition 5.1.5. The m-th order, n dimensional tensor $\mathcal{I}$ is called the unit tensor if

$$\mathcal{I} = \delta_{i_1 i_2 \dots i_m} = \begin{cases} 1 & \text{if } i_1 = i_2 = \cdots = i_m, \\ 0 & \text{otherwise.} \end{cases}$$

The unit tensor behaves like a right identity in the following way,

$$\begin{aligned} (\mathcal{A}\mathcal{I})_{i\alpha_2\cdots\alpha_m} &= \sum_{i_1, i_2, \dots, i_m} a_{i_1 i_2 \cdots i_m} \delta_{i_2\alpha_2} \delta_{i_3\alpha_3} \cdots \delta_{i_m\alpha_m} \\ &= \begin{cases} a_{i_1 i_2 \cdots i_m} & \alpha_j = (i_j, i_j, \dots, i_j) \text{ for all } j \in \{2, \dots, m\}, \\ 0 & \alpha_j \neq (i_j, i_j, \dots, i_j) \text{ for some } j \in \{2, \dots, m\}. \end{cases} \end{aligned}$$

Thus $\mathcal{A}$, is embedded as a tensor $\mathcal{A}\mathcal{I}$ in a higher order tensor product space. On the other hand $\mathcal{I}$ is not a left identity,

$$\begin{aligned} (\mathcal{I}\mathcal{A})_{i\alpha_2\cdots\alpha_m} &= \sum_{i_2, \dots, i_m} \delta_{i i_2 \cdots i_m} a_{i_2\alpha_2} a_{i_3\alpha_3} \cdots a_{i_m\alpha_m} \\ &= a_{i\alpha_1} a_{i\alpha_2} \cdots a_{i\alpha_m} \\ &\neq (\mathcal{A})_{i\alpha_2\cdots\alpha_m}. \end{aligned}$$

In general, by nonlinearity in $\mathcal{B}$ (in Definition 5.1.1) there is no left identity. However, if $m = 2$ then $\mathcal{I} = I$ (the identity matrix) is a left identity.

Example 5.1.6. Suppose that $\mathcal{A} \in \mathbb{F}^{2\times2\times2}$ and $I \in \mathbb{F}^{2\times2}$, that is

$$\operatorname{frnt}(\mathcal{A}) = \left(\begin{array}{cc|cc} 1 & 0 & 1 & 2 \\ 2 & 5 & 0 & 1 \end{array} \right) \text{ and } I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The product $\mathcal{A}I = \mathcal{C}$ is given entrywise by

$$\begin{aligned} &c_{111} = 1, \quad c_{121} = 0, \quad c_{211} = 1, \quad c_{221} = 2, \\ &c_{112} = 2, \quad c_{122} = 5, \quad c_{212} = 0, \quad c_{222} = 1. \end{aligned}$$

Hence,

$$\operatorname{frnt}(\mathcal{C}) = \left(\begin{array}{cc|cc} 1 & 0 & 1 & 2 \\ 2 & 5 & 0 & 1 \end{array} \right) = \operatorname{frnt}(\mathcal{A}).$$

Chapter 6

Ternary products

In this chapter, we will briefly study a ternary product of tensors arising from an association scheme [13]. Association schemes arise in combinatorics and form a unitary commutative subalgebra (Bose-Mesner algebra) of matrices. Here we describe the association scheme on triples and the corresponding ternary product. The product can be generalised to an n-ary product which includes the matrix product as a special case. Throughout this chapter, concepts will be developed in terms of the ternary product in order to describe the spectral decomposition (described in [7] section 6.2.5).

6.1 Association scheme on triples

Definition 6.1.1. An association scheme on triples (AST) on a set Ω is a partition of $\Omega \times \Omega \times \Omega$ into relations R_i, where $i \in \{0, \ldots, m\}$ and $m \geq 4$, such that the following conditions are satisfied:

(1) For any two distinct elements $x, y \in \Omega$, the pair (x, y) is in a uniform number of triples $(x, y, z) \in R_i$, i.e, for all $x, x', y, y' \in \Omega$ with $x \neq y$ and $x' \neq y'$

$$|\{z : (x, y, z) \in R_i\}| = |\{z : (x', y', z) \in R_i\}|.$$

(2) For any $i, j, k, l \in \{0, \ldots, m\}$ and for a given triple $(x, y, z) \in R_l$, there is a uniform number of elements $w \in \Omega$ such that

$$(w, y, z) \in R_i,\ (x, w, z) \in R_j,\ \text{and}\ (x, y, w) \in R_k,$$

where w is independent of (x, y, z) and only depends on i, j, k, l and is denoted by P^l_{ijk}.

(3) For any relation R_i, where $i \in \{0, \dots, m\}$ and any permutation σ of $\{1, 2, 3\}$, the set

$$\{(x_{\sigma(1)}, x_{\sigma(2)}, x_{\sigma(3)}) : (x_1, x_2, x_3) \in R_i\}$$

equals some relation R_j where $j \in \{0, \dots, m\}$.

(4) The relations R_i, where $i \in \{0, \dots, 3\}$ are trivial relations such that, for distinct elements $x, y \in \Omega$

$$(x, x, x) \in R_0,\ (x, y, y) \in R_1,\ (x, y, x) \in R_2, \text{ and } (x, x, y) \in R_3.$$

We will work with the frontal slices of a third-order tensor. The addition and scalar multiplication of these third-order tensors is defined entrywise in the usual way. If the cardinality of the set Ω is v (finite), that is $|\Omega| = v$, then each relation R_i is determined by its $v \times v \times v$ adjacency tensor $\mathcal{A}$ with entries (x, y, z) denoted $(\mathcal{A}_i)_{xyz}$, where

$$(\mathcal{A}_i)_{xyz} = \begin{cases} 1 & \text{if } (x, y, z) \in R_i, \\ 0 & \text{otherwise.} \end{cases} \tag{6.1}$$

It is convenient to express these adjacency tensors in terms of their frontal slices.

Example 6.1.2. Let Ω be a set with $x, y \in \Omega$ such that x and y are distinct and $|\Omega| = 3$, then the trivial relations of the $3 \times 3 \times 3$ adjacency tensors $\mathcal{A}_0$, $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ (as defined in (6.1)) are

$$\text{frnt}((\mathcal{A}_0)_{xxx}) = \left(\begin{array}{ccc|ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right),$$

$$\text{frnt}((\mathcal{A}_1)_{xyy}) = \left(\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{array}\right),$$

$$\text{frnt}((\mathcal{A}_2)_{xyx}) = \left(\begin{array}{ccc|ccc|ccc} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right),$$

$$\text{frnt}((\mathcal{A}_3)_{xxy}) = \left(\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right).$$

The adjacency tensor for the relation, R_4, where x, y, z are all distinct is the following,

$$\operatorname{frnt}((\mathcal{A}_4)_{xyz}) = \left(\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Then $R_0, \ldots, R_4$ forms an AST.

We define in Definition 6.1.3 the ternary product. The connection between this ternary product and AST is given in Theorem 6.1.4.

Definition 6.1.3. For any third-order tensors $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ of size $v \times v \times v$ indexed by the set Ω, the ternary product $\mathcal{D}$ is the third-order tensor of size $v \times v \times v$ given by

$$(\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C})_{xyz} = \sum_{w \in \Omega} a_{wyz} b_{xwz} c_{xyw}, \tag{6.2}$$

where a_{wyz} denotes the (w, y, z) entry of $\mathcal{A}$, b_{xwz} denotes the (x, w, z) entry of $\mathcal{B}$ and c_{xyw} denotes the (x, y, w) entry of $\mathcal{C}$.

From the definition, it is clear that the ternary product preserves the order of the tensors.

Theorem 6.1.4. *For an AST defined on the set Ω with the relations $R_0, \ldots, R_m$ and the corresponding adjacency tensors $\mathcal{A}_0, \ldots, \mathcal{A}_m$, respectively, then for $i, j, k \in \{0, 1, \ldots, m\}$ we have that*

$$\mathcal{A}_i \cdot \mathcal{A}_j \cdot \mathcal{A}_k = \sum_{l=0}^{m} P_{ijk}^{l} \mathcal{A}_l. \tag{6.3}$$

We omit the proof of the theorem, in order to emphasize the algebraic and spectral properties (presented in [13]).

In [7], an equivalent product $\circ(\mathcal{A}, \mathcal{B}, \mathcal{C})$ is presented, where $\circ(\mathcal{A}, \mathcal{B}, \mathcal{C}) = \mathcal{C} \cdot \mathcal{A} \cdot \mathcal{B}$ in terms of the product in Definition 6.1.3. Throughout the remaining of the chapter we will adopt the product of [13] as given in Definition 6.1.3 and also present the results from [7] in terms of this product.

6.2 Properties of the ternary product

First we provide a connection between the matrix product and the product in Definition 6.1.3. Matrix products can be written as a summation of outer products, that is, for matrices A and B

$$(AB)_{ij} = \sum_{k=1}^{n} (A)_{ik}(B)_{kj},$$

so that

$$\begin{aligned} AB &= \sum_{k=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ik}b_{kj} \otimes e_i e_j^T \\ &= \sum_{k=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ik}e_i \otimes b_{kj} e_j^T \\ &= \sum_{k=1}^{n}\left(\sum_{i=1}^{n} a_{ik}e_i\right) \otimes \left(\sum_{j=1}^{n} b_{kj}e_j^T\right). \end{aligned}$$

The ternary product of tensors can be written as a sum of third-order outer products. The outer product generalizes in the following way. Let $a, b \in \mathbb{R}^n$ be vectors considered as $1 \times n$ and $n \times 1$ matrices respectively, then the outer product is given entrywise by the products

$$\mathrm{out}(a, b)_{ij} = b_{j1}a_{1i}, \tag{6.4}$$

indexed by the coordinates. For matrices, a ternary outer product is given entrywise by

$$\mathrm{out}(A, B, C)_{ijk} = (C)_{ij}(B)_{ik}(A)_{jk}. \tag{6.5}$$

For third order tensors $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$

$$(\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C})_{xyz} = \sum_{w=1}^{n} a_{wyz}b_{xwz}c_{xyw},$$

so that

$$\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C} = \sum_{w=1}^{n} \mathrm{out}(\mathcal{A}_{w\cdot\cdot}, \mathcal{B}_{\cdot w\cdot}, \mathcal{C}_{\cdot\cdot w}) \tag{6.6}$$

where,

$$\mathrm{out}(\mathcal{A}_{w\cdot\cdot}, \mathcal{B}_{\cdot w\cdot}, \mathcal{C}_{\cdot\cdot w}) := \sum_{x=1}^{n}\sum_{y=1}^{n}\sum_{z=1}^{n} a_{wyz}b_{xwz}c_{xyw}(e_x \otimes e_y \otimes e_z)$$

is the ternary outer product of tensors as defined in [7].

Theorem 6.2.1. *The ternary product is trilinear.*

Proof. The ternary product is distributive in the first entry, that is

$$(\mathcal{A}+\mathcal{A}')\cdot\mathcal{C}\cdot\mathcal{D}=\mathcal{A}\cdot\mathcal{C}\cdot\mathcal{D}+\mathcal{A}'\cdot\mathcal{C}\cdot\mathcal{D}. \tag{6.7}$$

To show this, let us consider third-order tensors $\mathcal{A}, \mathcal{A}', \mathcal{C}$, $\mathcal{D}$ and scalars $\alpha, \beta \in \mathbb{F}$. It now follows that

$$\begin{aligned}[(\alpha\mathcal{A}+\beta\mathcal{A}')\cdot\mathcal{C}\cdot\mathcal{D}]_{xyz} &= \sum_{w=1}^{n}(\alpha a_{wyz}+\beta a'_{wyz})c_{xwz}d_{xyw} \\ &= \alpha\sum_{w=1}^{n} a_{wyz}c_{xwz}d_{xyw}+\beta\sum_{w=1}^{n} a'_{wyz}c_{xwz}d_{xyw} \\ &= \alpha(\mathcal{A}\cdot\mathcal{C}\cdot\mathcal{D})_{xyz}+\beta(\mathcal{A}'\cdot\mathcal{C}\cdot\mathcal{D})_{xyz}.\end{aligned}$$

Similarly the distributivity in the second and third argument is as follow

$$\mathcal{A}\cdot(\mathcal{B}+\mathcal{B}')\cdot\mathcal{C}=\mathcal{A}\cdot\mathcal{B}\cdot\mathcal{C}+\mathcal{A}\cdot\mathcal{B}'\cdot\mathcal{C}, \tag{6.8}$$

$$\mathcal{A}\cdot\mathcal{B}\cdot(\mathcal{C}+\mathcal{C}')=\mathcal{A}\cdot\mathcal{B}\cdot\mathcal{C}+\mathcal{A}\cdot\mathcal{B}\cdot\mathcal{C}'. \tag{6.9}$$

□

6.2.1 Tensor transpose

Definition 6.2.2. Let $\mathcal{A}$, be a third-order tensor with entries, $(\mathcal{A})_{xyz} = a_{xyz}$. Then the transpose of $\mathcal{A}$, denoted $\mathcal{A}^T$, has the entries

$$(\mathcal{A}^T)_{xyz} = a_{yzx}. \tag{6.10}$$

Consequently $((\mathcal{A}^T)^T)_{xyz} = a_{zxy}$ and $(((\mathcal{A}^T)^T)^T)_{xyz} = a_{xyz}$. Notice that in (6.10) the transpose is a cyclic permutation of the indices of $\mathcal{A}$. We will use the notation

$$\mathcal{A}^{T^n} = \underbrace{((((\mathcal{A}^T)^T)^T)\cdots)^T}_{n\text{-times}}.$$

Definition 6.2.3. The tensor $\mathcal{A}$ is symmetric if

$$\mathcal{A}=\mathcal{A}^T=\mathcal{A}^{T^2}, \tag{6.11}$$

where $a_{xyz} = a_{yzx} = a_{zxy}$.

Proposition 6.2.4. *The transpose of the ternary product is given by*

$$(\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C})^T = \mathcal{C}^T \cdot \mathcal{A}^T \cdot \mathcal{B}^T. \tag{6.12}$$

Proof.

$$\begin{aligned}
(\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C})^T_{xyz} &= (\mathcal{C} \cdot \mathcal{A} \cdot \mathcal{B})_{yzx} \\
&= \sum_{w=1}^{n} c_{wyz} a_{xwz} b_{xyw} \\
&= (\mathcal{C}^T \cdot \mathcal{A}^T \cdot \mathcal{B}^T)_{xyz}.
\end{aligned}$$

Thus, $(\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C})^T = \mathcal{C}^T \cdot \mathcal{A}^T \cdot \mathcal{B}^T$. □

Corollary 6.2.5. $(\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C})^{T^2} = \mathcal{B}^{T^2} \cdot \mathcal{C}^{T^2} \cdot \mathcal{A}^{T^2}$.

The proof of the Corollary is similar to the proof of Proposition 6.2.4.

Proposition 6.2.6. *If $\mathcal{A}$ is symmetric, then $\mathcal{A} \cdot \mathcal{A}^T \cdot \mathcal{A}^{T^2}$, $\mathcal{A}^{T^2} \cdot \mathcal{A} \cdot \mathcal{A}^T$ and $\mathcal{A}^T \cdot \mathcal{A}^{T^2} \cdot \mathcal{A}$ are symmetric ternary products.*

Proof. Let us consider $\mathcal{A} \cdot \mathcal{A}^T \cdot \mathcal{A}^{T^2}$, then using Proposition 6.2.4 and Definition 6.2.2, it follows that

$$\begin{aligned}
(\mathcal{A} \cdot \mathcal{A}^T \cdot \mathcal{A}^{T^2})^T &= (\mathcal{A}^{T^2})^T \cdot (\mathcal{A})^T \cdot (\mathcal{A}^T)^T \\
&= \mathcal{A}^{T^3} \cdot \mathcal{A}^T \cdot \mathcal{A}^{T^2} \\
&= \mathcal{A} \cdot \mathcal{A}^T \cdot \mathcal{A}^{T^2}.
\end{aligned}$$

For the second part of the proof, it follows that

$$\begin{aligned}
(\mathcal{A} \cdot \mathcal{A}^T \cdot \mathcal{A}^{T^2})^{T^2} &= (\mathcal{A}^T)^{T^2} \cdot (\mathcal{A}^{T^2})^{T^2} \cdot (\mathcal{A})^{T^2} \\
&= \mathcal{A}^{T^3} \cdot (\mathcal{A}^{T^3})^T \cdot \mathcal{A}^{T^2} \\
&= \mathcal{A} \cdot \mathcal{A}^T \cdot \mathcal{A}^{T^2}.
\end{aligned}$$

Therefore,

$$(\mathcal{A} \cdot \mathcal{A}^T \cdot \mathcal{A}^{T^2})^{T^2} = (\mathcal{A} \cdot \mathcal{A}^T \cdot \mathcal{A}^{T^2})^T = \mathcal{A} \cdot \mathcal{A}^T \cdot \mathcal{A}^{T^2},$$

so $\mathcal{A} \cdot \mathcal{A}^T \cdot \mathcal{A}^{T^2}$ is symmetric. □

Theorem 6.2.7. *Let $\mathcal{A}$ be a third-order tensor, there exists a third-order ten-*

sor, $\mathcal{I}$, such that

$$\mathcal{A} \cdot \mathcal{I} \cdot \mathcal{I}^{T^2} = \mathcal{A}, \tag{6.13}$$

for all tensors $\mathcal{A}$.

Proof. We will show that (6.13) holds for $(I)_{ijk} = \delta_{jk}$, where δ_{jk} denotes the Kronecker delta. We first let $\mathcal{A} = \mathcal{E}_{uvw}$, the standard basis element with a 1 in row u, column v and fiber w and 0 elsewhere, thus

$$\mathcal{A} = (a)_{ijk} = \begin{cases} 1 & \delta_{iu}\delta_{jv}\delta_{kw}, \textit{ when } i = u,\ j = v,\ k = w, \\ 0 & \textit{otherwise.} \end{cases} \tag{6.14}$$

Using equation (6.13) with the above, we have

$$\begin{aligned} \sum_{t=1}^{n} a_{tjk}(\mathcal{I})_{itk}(\mathcal{I}^{T^2})_{ijt} = a_{ijk} &\Rightarrow \sum_{t=1}^{n} \delta_{tu}\delta_{jv}\delta_{kw}(\mathcal{I})_{itk}(\mathcal{I}^{T^2})_{ijt} = \delta_{iu}\delta_{jv}\delta_{kw} \\ &\Rightarrow \delta_{jv}\delta_{kw}\sum_{t=1}^{n} \delta_{tu}(\mathcal{I})_{itk}(\mathcal{I}^{T^2})_{ijt} = \delta_{iu}\delta_{jv}\delta_{kw} \\ &\Rightarrow (\mathcal{I})_{iuk}(\mathcal{I}^{T^2})_{iju} = \delta_{iu} \\ &\Rightarrow (\mathcal{I})_{iuk}(\mathcal{I})_{uij} = \delta_{iu}, \end{aligned}$$

where the last two equalities hold when $v = j$ and $w = k$. In particular, when $j = k$ we have $(\mathcal{I})^2_{iuk} = \delta_{iu}$. Moreover, if the underlying field is $\mathbb{R}$ or $\mathbb{C}$, then there is a unique tensor $\mathcal{I}$ with non-negative entries obeying (6.14). Therefore,

$$\mathcal{I}_{ijk} = \delta_{ij}, \tag{6.15}$$

satisfies (6.13). Now we need to show that $\mathcal{A} \cdot \mathcal{I} \cdot \mathcal{I}^{T^2} = \mathcal{A}$,

$$\begin{aligned} (\mathcal{A} \cdot \mathcal{I} \cdot \mathcal{I}^{T^2})_{ijk} &= \sum_{t=1}^{n} a_{tjk}(\mathcal{I})_{itk}(\mathcal{I}^2)_{ijt} \\ &= \sum_{t=1}^{n} a_{tjk}\delta_{it}\delta_{ti} \\ &= a_{ijk} \\ &= \mathcal{A}_{ijk}. \end{aligned}$$

So, $\mathcal{A} \cdot \mathcal{I} \cdot \mathcal{I}^{T^2} = \mathcal{A}$. □

If the field is $\mathbb{R}$ or $\mathbb{C}$, then $\mathcal{I}_{ijk} = \delta_{ij}$ is the unique non-negative square root of δ_{ij}. Consequently, from Theorem 6.2.7 it follows that $(\mathcal{I}^T)_{ijk} = \delta_{jk}$, and $(\mathcal{I}^{T^2})_{ijk} = \delta_{ki}$.

Lemma 6.2.8. *Let $\mathcal{A}$, be a third-order tensor. Then,*

$$\mathcal{I} \cdot \mathcal{A} \cdot \mathcal{I}^T = \mathcal{A}, \tag{6.16}$$

$$\mathcal{I}^{T^2} \cdot \mathcal{I}^T \cdot \mathcal{A} = \mathcal{A}. \tag{6.17}$$

Proof. Let the entries of $\mathcal{A}$ be $(\mathcal{A})_{ijk} = a_{ijk}$. Starting from the left side of (6.16), we have

$$\begin{aligned}(\mathcal{I} \cdot \mathcal{A} \cdot \mathcal{I}^T)_{ijk} &= \sum_{t=1}^{n} (\mathcal{I})_{tjk} a_{itk} (\mathcal{I}^T)_{ijt} \\ &= \sum_{t=1}^{n} \delta_{tj} a_{itk} \delta_{jt} \\ &= a_{ijk}.\end{aligned}$$

Thus $\mathcal{I} \cdot \mathcal{A} \cdot \mathcal{I}^T = \mathcal{A}$. For (6.17), one have

$$\begin{aligned}(\mathcal{I}^{T^2} \cdot \mathcal{I}^T \cdot \mathcal{A})_{ijk} &= \sum_{t=1}^{n} (\mathcal{I}^{T^2})_{tjk} (\mathcal{I}^T)_{itk} a_{ijt} \\ &= \sum_{t=1}^{n} \delta_{kt} \delta_{tk} a_{ijt} \\ &= a_{ijk}.\end{aligned}$$

Thus, $\mathcal{I}^{T^2} \cdot \mathcal{I}^T \cdot \mathcal{A} = \mathcal{A}$. □

6.2.2 Associativity of the ternary product

Associativity of a ternary product may be defined in multiple ways, since "adjacency" in the terms does not provide a canonical ordering for evaluation. We consider first the obvious generalization of binary associativity to the ternary product. At the end of this section we will show that a non-standard form of associativity does hold for the ternary product.

Counter examples of the following Lemma 6.2.9 will be provided in 6.2.10 and 6.2.11.

Lemma 6.2.9. *Let $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D},$ and $\mathcal{E}$ be third-order tensors and*

$$(\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C})_{ijk} = \sum_{t=1}^{n} a_{tjk} b_{itk} c_{ijt}.$$

We want to show that in general

$$(\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C}) \cdot \mathcal{D} \cdot \mathcal{E} \neq \mathcal{A} \cdot \mathcal{B} \cdot (\mathcal{C} \cdot \mathcal{D} \cdot \mathcal{E}) \neq \mathcal{A} \cdot (\mathcal{B} \cdot \mathcal{C} \cdot \mathcal{D}) \cdot \mathcal{E}. \tag{6.18}$$

In total we will have four cases.

Case (I) $(\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C}) \cdot \mathcal{D} \cdot \mathcal{E} \neq \mathcal{A} \cdot \mathcal{B} \cdot (\mathcal{C} \cdot \mathcal{D} \cdot \mathcal{E})$,

The left side equals:

$$\begin{aligned} ((\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C}) \cdot \mathcal{D} \cdot \mathcal{E})_{ijk} &= \sum_{s=1}^{n} (\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C})_{sjk} d_{isk} e_{ijs} \\ &= \sum_{s=1}^{n} \sum_{t=1}^{n} a_{tjk} b_{stk} c_{sjt} d_{isk} e_{ijs}. \end{aligned} \tag{6.19}$$

On the right side we have

$$\begin{aligned} (\mathcal{A} \cdot \mathcal{B} \cdot (\mathcal{C} \cdot \mathcal{D} \cdot \mathcal{E}))_{ijk} &= \sum_{t=1}^{n} a_{tjk} b_{itk} (\mathcal{C} \cdot \mathcal{D} \cdot \mathcal{E})_{ijt} \\ &= \sum_{t=1}^{n} \sum_{s=1}^{n} a_{tjk} b_{itk} c_{sjt} d_{ist} e_{ijs}. \end{aligned} \tag{6.20}$$

Comparing (6.19) and (6.20) we can see that in general,

$$a_{tjk} b_{stk} c_{sjt} d_{isk} e_{ijs} \neq a_{tjk} b_{itk} c_{sjt} d_{ist} e_{ijs}, \tag{6.21}$$

because generally, $b_{stk} d_{isk} \neq b_{itk} d_{ist}$. The following example illustrates case (I).

Example 6.2.10. Let $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$, and $\mathcal{E}$ be third-order tensors, it follows that if $a_{ijk} = c_{ijk} = e_{ijk} = 1$,

$$b_{stk} = \begin{cases} 1 & i = j = k = 1, \\ 0 & otherwise, \end{cases} \quad \text{and} \quad d_{ijk} = \begin{cases} 1 & i = 2, j = k = 1, \\ 0 & otherwise, \end{cases}$$

then, $(\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C}) \cdot \mathcal{D} \cdot \mathcal{E} \neq \mathcal{A} \cdot \mathcal{B} \cdot (\mathcal{C} \cdot \mathcal{D} \cdot \mathcal{E})$.

Case (II) $\mathcal{A} \cdot \mathcal{B} \cdot (\mathcal{C} \cdot \mathcal{D} \cdot \mathcal{E}) \neq \mathcal{A} \cdot (\mathcal{B} \cdot \mathcal{C} \cdot \mathcal{D}) \cdot \mathcal{E}$.

$$\begin{aligned} (\mathcal{A} \cdot (\mathcal{B} \cdot \mathcal{C} \cdot \mathcal{D}) \cdot \mathcal{E})_{ijk} &= \sum_{s=1}^{n} a_{sjk} (\mathcal{B} \cdot \mathcal{C} \cdot \mathcal{D})_{isk} e_{ijs} \\ &= \sum_{s=1}^{n} \sum_{t=1}^{n} a_{sjk} b_{tsk} c_{itk} d_{ist} e_{ijs}. \end{aligned} \tag{6.22}$$

Using (6.20) and (6.22) we can see that, in general

$$a_{tjk} b_{itk} c_{sjt} d_{ist} e_{ijs} \neq a_{sjk} b_{tsk} c_{itk} d_{ist} e_{ijs}, \tag{6.23}$$

in general, $a_{tjk} b_{itk} c_{sjt} \neq a_{sjk} b_{tsk} c_{itk}$. We provide a counter example where

(6.23) does not hold to prove case (II).

Example 6.2.11. Let $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$, and $\mathcal{E}$ be third-order tensors, with $c_{ijk} = d_{ijk} = 1$, where

$$a_{ijk} = \begin{cases} 1 & i = j = k = 1, \\ 0 & \textit{otherwise}, \end{cases}, \qquad b_{ijk} = \begin{cases} 1 & i = 2, j = k = 1, \\ 0 & \textit{otherwise}, \end{cases}$$

and,

$$e_{ijk} = \begin{cases} 1 & i = 2, j = k = 1, \\ 0 & \textit{otherwise}. \end{cases}$$

Thus, $\mathcal{A} \cdot \mathcal{B} \cdot (\mathcal{C} \cdot \mathcal{D} \cdot \mathcal{E}) \neq \mathcal{A} \cdot (\mathcal{B} \cdot \mathcal{C} \cdot \mathcal{D}) \cdot \mathcal{E}$.

The remaining cases:

$$\mathcal{A} \cdot \mathcal{B} \cdot (\mathcal{C} \cdot \mathcal{D} \cdot \mathcal{E}) \neq (\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C}) \cdot \mathcal{D} \cdot \mathcal{E},$$
$$\mathcal{A} \cdot (\mathcal{B} \cdot \mathcal{C} \cdot \mathcal{D}) \cdot \mathcal{E} \neq (\mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C}) \cdot \mathcal{D} \cdot \mathcal{E},$$

follow a similar reasoning and so in general associativity does not hold for the ternary product (6.2).

For ternary products, the notion of associativity generalizes in different ways from the notion of associativity for binary products. Let us first look at the results for matrices. A novel approach to introduce associativity follows from a result by Eckmann and Hilton (the so called Eckmann-Hilton argument [5]), which states that the internal unital magmas in the category of unital magmas are commutative monoids. The argument can be adopted in the following way.

Definition 6.2.12. Let $* : M \times M \to M$, be a binary operation acting on the set M. A magma is a set $(M, *)$ equipped with a binary operation. A unital magma is a triple $(M, *, 1)$ where $1 \in M$, such that for $x \in M$, $x*1 = 1*x = x$.

Associativity in unital magmas may be described as a special kind of interchange law. Two binary products, $\cdot$ and $*$, are said to obey the interchange law if

$$(A \cdot B) * (C \cdot D) = (A * C) \cdot (B * D). \tag{6.24}$$

Now if $* = \cdot$, then a restriction of the interchange law describes associativity. Suppose that A and B are matrices and I be the identity matrix. Using the special case of the interchange law (6.24), it follows that

$$(A \cdot I) \cdot (B \cdot C) = (A \cdot B) \cdot (I \cdot C) \Longleftrightarrow A(BC) = (AB)C. \tag{6.25}$$

Similarly we want to investigate a notion of associativity for the ternary product. We will present a form of associativity involving the identity tensor. Consider the following product

$$
\begin{aligned}
[(a\alpha A)(b\beta B)(r\gamma R)]_{ijk} &= \sum_{t=1}^{n} (a\alpha A)_{tjk}(b\beta B)_{itk}(r\gamma R)_{ijt} \\
&= \sum_{t=1}^{n} \left[\sum_{l=1}^{n} a_{ljk}\alpha_{tlk}A_{tjl} \sum_{u=1}^{n} b_{utk}\beta_{iuk}B_{itu} \sum_{v=1}^{n} r_{vjt}\gamma_{ivt}R_{ijv} \right] \\
&= \sum_{l=1}^{n}\sum_{u=1}^{n}\sum_{v=1}^{n} a_{ljk}\beta_{iuk}R_{ijv} \sum_{t=1}^{n} \left[\alpha_{tlk}A_{tjl}b_{utk}B_{itu}r_{vjt}\gamma_{ivt}\right].
\end{aligned}
\tag{6.26}
$$

The next theorem shows that, for matrix-like tensors, a form of associativity holds, where each term in a product is associated with a second and third term of a product.

Theorem 6.2.13. *Let A', b' and γ' be $n \times n$ matrices and let A, b and γ denote the tensors, given by*

$$
A = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \otimes A', \quad b = \left[\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \otimes b' \right]^{T}, \quad \textit{and} \quad \gamma = \left[\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \otimes \gamma' \right]^{T^2}.
$$

Furthermore, let α, B and r be arbitrary third-order n-dimensional tensors, then

$$
(\mathcal{I}\alpha A)(b\mathcal{I}^T B)(r\gamma\mathcal{I}^{T^2}) = \mathcal{I}(b\mathcal{I}^T(\alpha B r))(A\gamma\mathcal{I}^{T^2}). \tag{6.27}
$$

Proof.

$$
[(\mathcal{I}\alpha A)(b\mathcal{I}^T B)(r\gamma\mathcal{I}^{T^2})]_{ijk} = \sum_{l=1}^{n}\sum_{u=1}^{n}\sum_{v=1}^{n} \delta_{lj}\delta_{uk}\delta_{vi} \sum_{t=1}^{n} \left[\alpha_{tjk}A_{tjl}b_{utk}B_{itk}r_{ijt}\gamma_{ivt}\right].
$$

Since $A_{tjl} = A_{vjl}$, $b_{utk} = b_{ulk}$ and $\gamma_{ivt} = \gamma_{ivl}$, we have

$$
\begin{aligned}
&(\mathcal{I}\alpha A)(b\mathcal{I}^T B)(r\gamma\mathcal{I}^{T^2})_{ijk} \\
&= \sum_{l=1}^{n}\sum_{u=1}^{n}\sum_{v=1}^{n} \mathcal{I}_{ljk}(\mathcal{I}^T)_{iuk}(\mathcal{I}^{T^2})_{ijv}A_{vjl}b_{ulk}\gamma_{ivl}(\alpha B r)_{ijk} \\
&= \sum_{l=1}^{n} \mathcal{I}_{ljk} \left(\sum_{u=1}^{n} b_{ulk}(\mathcal{I}^T)_{iuk}(\alpha B r)_{ilu} \right) \left(\sum_{v=1}^{n} A_{vjl}\gamma_{ivl}(\mathcal{I}^{T^2})_{ijv} \right) \\
&= \sum_{l=1}^{n} \mathcal{I}_{ljk} \left(b(\mathcal{I}^T)(\alpha B r)\right)_{ilk} \left(A\gamma(\mathcal{I}^{T^2})\right)_{ijl}
\end{aligned}
$$

$$= [\mathcal{I}(b\mathcal{I}^T(\alpha Br))(A\gamma\mathcal{I}^{T^2})]_{ijk}. \qquad \square$$

Remark 6.2.14. Here r, α and B are first, second and third terms in a product and associate on the right (in different order). The tensors A, γ and $\mathcal{I}^{T^2}$ associate in a similar manner (but are not arbitrary tensors).

The proof of the following result of associativity, appears in appendix A.1.

$$[(\mathcal{I}\alpha A)(b\mathcal{I}B)(r\gamma\mathcal{I}^T)] = [\mathcal{I}\left(\alpha\mathcal{I}B\right)\left(r(Ab\gamma)\mathcal{I}^T\right)]. \tag{6.28}$$

The remaining sections define analogues to the elementary operations and culminate in a ternary version of the spectral decomposition. The spectral decomposition described in terms of orthogonal (or unitary) matrices, scaling matrices and orthogonal matrices. The tensor analogues follow in Section 6.2.4.

6.2.3 Permutation tensors

Definition 6.2.15. Let σ be an element of the permutation group S_n, then for all $\sigma \in S_n$, the σ-permutation tensor is given by

$$(\mathcal{P}_\sigma)_{xyz} = \delta_{z\sigma(y)}. \tag{6.29}$$

Proposition 6.2.16. *Let $1_{n\times n\times n}$ be a third-order tensor with all entries equal to one, then*

$$\mathcal{P}_\sigma = \sum_{t=1}^{n}\left(\sum_{k=1}^{n} e_k \otimes e_k \otimes e_{\sigma(k)}\right)\cdot 1_{n\times n\times n}\cdot 1_{n\times n\times n}. \tag{6.30}$$

Proof. Since $(\mathcal{P}_\sigma)_{xyz} = \delta_{z\sigma(y)}$, we have

$$\begin{aligned}
(P_\sigma)_{xyz} &= \sum_{t=1}^{n}\left(\sum_{k=1}^{n} e_k \otimes e_k \otimes e_{\sigma(k)}\right)_{tyz}\cdot(1)_{xtz}\cdot(1)_{xyt}\\
&= \sum_{t=1}^{n}\left(\sum_{k=1}^{n} e_k \otimes e_k \otimes e_{\sigma(k)}\right)_{tyz}\\
&= \sum_{t=1}^{n}\sum_{k=1}^{n}\delta_{tk}\delta_{yk}\delta_{z\sigma(k)}\\
&= \sum_{k=1}^{n}\delta_{yk}\delta_{yk}\delta_{z\sigma(k)}\\
&= \delta_{z\sigma(y)}.
\end{aligned} \qquad \square$$

Consequently,

$$(P_\sigma)^T_{xyz} = (P_\sigma)_{yzx} = \delta_{x\sigma(z)},$$

and

$$(P_\sigma)^{T^2}_{xyz} = (P_\sigma)_{zxy} = \delta_{y\sigma(x)}.$$

Lemma 6.2.17. *The permutation tensor P_σ performs the permutation σ on the depth slices of a tensor $\mathcal{A}$, through the product $\mathcal{P}_\sigma^T \cdot \mathcal{P}_{\sigma^{-1}} \cdot \mathcal{A}$.*

Proof. Straightforward computation yields

$$\begin{aligned}
(\mathcal{P}_\sigma^T \cdot \mathcal{P}_{\sigma^{-1}} \cdot \mathcal{A})_{xyz} &= \sum_{t=1}^{n} (P_\sigma^T)_{tyz} (P_{\sigma^{-1}})_{xtz} a_{xyt} \\
&= \sum_{t=1}^{n} \delta_{t\sigma(z)} \delta_{z\sigma^{-1}(t)} a_{xyt} \\
&= \sum_{t=1}^{n} \delta_{t\sigma(z)} \delta_{t\sigma(z)} a_{xyt} \\
&= a_{xy\sigma(z)}.
\end{aligned}$$

□

Similarly, the permutation tensor $\mathcal{P}_\sigma$ performs the permutation of the vertical and horizontal slices through the products $\mathcal{P}^{T^2}_{\sigma^{-1}} \cdot \mathcal{A} \cdot \mathcal{P}_\sigma$ and $\mathcal{A} \cdot \mathcal{P}^{T^2}_\sigma \cdot \mathcal{P}^T_{\sigma^{-1}}$ respectively, that is

$$\begin{aligned}
(\mathcal{P}^{T^2}_{\sigma^{-1}} \cdot \mathcal{A} \cdot \mathcal{P}_\sigma)_{xyz} &= \sum_{t=1}^{n} (P^{T^2}_{\sigma^{-1}})_{tyz} a_{xtz} (P_\sigma)_{xyt} \\
&= \sum_{t=1}^{n} \delta_{y\sigma^{-1}(t)} a_{xtz} \delta_{t\sigma(y)} \\
&= \sum_{t=1}^{n} \delta_{t\sigma(y)} a_{xtz} \delta_{t\sigma(y)} \\
&= a_{x\sigma(y)z},
\end{aligned}$$

and similarly,

$$\begin{aligned}
(\mathcal{A} \cdot \mathcal{P}^{T^2}_\sigma \cdot \mathcal{P}^T_{\sigma^{-1}})_{xyz} &= \sum_{t=1}^{n} a_{tyz} (P^{T^2}_\sigma)_{xtz} (P^T_{\sigma^{-1}})_{xyt} \\
&= \sum_{t=1}^{n} a_{tyz} \delta_{t\sigma(x)} \delta_{x\sigma^{-1}(t)} \\
&= \sum_{t=1}^{n} a_{tyz} \delta_{t\sigma(x)} \delta_{t\sigma(x)}
\end{aligned}$$

$$= a_{\sigma(x)yz}.$$

6.2.4 Orthogonality and scaling tensor

We will first consider what orthogonality means for matrices.

Definition 6.2.18. Let A be a $n \times n$ matrix, then A is said to be orthogonal if

$$A^T A = AA^T = I, \tag{6.31}$$

where I is the $n \times n$ identity matrix.

Now we will extend orthogonality of a matrix to third-order tensors.

Definition 6.2.19. Let $\mathcal{A}$, be a third-order tensor, then $\mathcal{A}$ is orthogonal if

$$(\mathcal{A}^T \mathcal{A} \mathcal{A}^{T^2})_{xyz} = \delta_{xyz}. \tag{6.32}$$

where $\delta_{xyz} = 1$ when $x = y = z$, and zero otherwise.

In matrix algebra, multiplication of a diagonal matrix by another matrix is a row or column scaling operation. If D is a $m \times m$ diagonal matrix and A an $m \times n$ matrix then

$$\begin{aligned} DA &= \begin{pmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{mm} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \\ &= \begin{pmatrix} d_{11}a_{11} & d_{11}a_{12} & \cdots & d_{11}a_{1n} \\ d_{22}a_{21} & d_{22}a_{22} & \cdots & d_{22}a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{mm}a_{m1} & d_{mm}a_{m2} & \cdots & d_{mm}a_{mn} \end{pmatrix}, \end{aligned}$$

thus, D scales the rows of matrix A. Similarly, if the product AD is conformable then we can conclude that matrix D scales the columns of matrix A.

Proposition 6.2.20. *An $n \times n$ real matrix D is a diagonal if and only if*

$$(DD^T)_{ij} = D^2_{ij}$$

Proof. $(\Rightarrow)$ A diagonal matrix, D, is a square matrix, such that its entries obey $d_{ij} = d_{ii}\delta_{ij} = d_{jj}\delta_{ij}$. Hence,

$$\begin{aligned}
(DD^T)_{ij} &= \sum_{k=1}^{n} d_{ik}d^T_{kj} \\
&= \sum_{k=1}^{n} d_{ik}d_{jk} \\
&= \sum_{k=1}^{n} d_{ii}\delta_{ik}d_{jj}\delta_{jk} \\
&= (d_{ii})^2(\delta_{ij})^2 \\
&= D^2_{ij}.
\end{aligned}$$

($\Leftarrow$) Suppose that $(DD^T)_{ij} = D^2_{ij}$, we want to show that D is diagonal. First, set $i = j$, then we have

$$\begin{aligned}
(DD^T)_{ii} = D^2_{ii} &\Rightarrow \sum_{k=1}^{n} d_{ik}d^T_{ki} = d^2_{ii} \\
&\Rightarrow \sum_{k=1}^{n} d_{ik}d_{ik} = d^2_{ii} \\
&\Rightarrow \sum_{k=1}^{n} d^2_{ik} = d^2_{ii} \\
&\Rightarrow d^2_{i1} + d^2_{i2} + \cdots + d^2_{i(i-1)} + d^2_{i(i+1)} + \cdots + d^2_{in} = 0.
\end{aligned}$$

Since D is real, $d_{ik} = 0$ for all $k \neq i$, therefore D is diagonal. □

Proposition 6.2.20 provides the motivation for the following definition of a diagonal tensor [7].

Definition 6.2.21. Let $\mathcal{D}$ be a real third-order tensor, we say that the tensor $\mathcal{D}$ is diagonal if

$$(\mathcal{D} \cdot \mathcal{D}^T \cdot \mathcal{D}^{T^2})_{xyz} = d^3_{xyz}. \tag{6.33}$$

For tensor scaling, let $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ be tensors and let

$$b_{xyz} = \delta_{yz}w_{xz}, \quad c_{xyz} = \delta_{xy}w_{xz}, \tag{6.34}$$

then

$$\begin{aligned}
(\mathcal{C} \cdot \mathcal{A} \cdot \mathcal{B})_{xyz} &= \sum_{t=1}^{n} c_{tyz}a_{xtz}b_{xyt} \\
&= \sum_{t=1}^{n} \delta_{ty}w_{tz}a_{xtz}\delta_{yt}w_{xt} \\
&= w_{yz}a_{xyz}w_{xy}.
\end{aligned} \tag{6.35}$$

Thus, the entry a_{xyz} is scaled on the left and on the right by w_{yz} and w_{xy}, respectively.

In the following section, we will use the Frobenius norm (which is a p-norm, with $p = 2$).

Definition 6.2.22. Let $\| \cdot \|_p : \mathbb{R}^n \to \mathbb{R}^+$ be a norm operator such that for $\underline{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$,

$$\|\underline{x}\|_p := \left[\sum_{k=1}^{n} |x_k|^p \right]^{\frac{1}{p}} . \tag{6.36}$$

The Frobenius norm of a third-order tensor is defined as follows.

Definition 6.2.23. Let $\mathcal{A}$ be an $m \times n \times p$ dimensional, third-order tensor, then

$$\|\mathcal{A}\|_3 := \left[\sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} |a_{ijk}|^3 \right]^{\frac{1}{3}} \tag{6.37}$$

defines the Frobenius norm for tensor $\mathcal{A}$.

6.2.5 Spectral decomposition

To understand the spectral theorem for the ternary product, let us first consider the eigenvalue decomposition for a square matrix. Eigenvalue decomposition is a factorization of a matrix into a canonical form, that is, the matrix is factored in terms of its eigenvalues and eigenvectors.

In the following, we assume that the underlying field is $\mathbb{R}$. In [7], the same structure is also considered over $\mathbb{C}$. The spectral theorem proofs carry over in a straightforward way, but the notion of Hermitian/self-adjoint does not. We give an example of this in Appendix A.2.

Proposition 6.2.24. *Let A be a $n \times n$ symmetric matrix, then there exists an orthogonal matrix U such that*

$$A = U \Lambda U^T \tag{6.38}$$

where $UU^T = I$ and $\Lambda = diag(\lambda_1, \dots, \lambda_n)$. Here $\lambda_1, \dots, \lambda_n$ are the n real eigenvalues of A and U has the corresponding eigenvectors as its columns.

In general, there is no analogue for the spectral decomposition as defined in (6.38) for higher-order tensors. We will rewrite (6.38) in a form that will allow us to define the spectral decomposition for the ternary product, following the approach in [7].

There exist matrices D and E, a pair of diagonal matrices, Q and R, such that $\Lambda = D^T E$ and $QR^T = Q^T R = I$ with

$$A = (DQ)^T ER. \tag{6.39}$$

In particular, if $Q = U^T$ and $R = U^T$, then

$$\begin{aligned} U\Lambda U^T &= UD^T EU^T \\ &= (DU^T)^T EU^T \\ &= (DQ)^T ER. \end{aligned}$$

To prove that we can decompose A as in (6.39), let us consider the following.

Lemma 6.2.25. *Let $C_N(\mathbb{R})$ be the ring of real valued continuous functions (with pointwise addition and multiplication) with domain $\mathbb{R}^N$ and let $X \subseteq C_N(\mathbb{R})$, and let S be a set of all common zeros of the functions in X, i.e., $S = \{s \in \mathbb{R}^N : \forall P \in X : P(s) = 0\}$, then*

$$\langle X \rangle = \langle \{Y \in C_N(\mathbb{R}) : \forall s \in S : Y(s) = 0\} \rangle \,, \tag{6.40}$$

is an ideal in $C_N(\mathbb{R})$.

Proof. First we need to show that for $f, g \in \langle X \rangle$ we have $f + g \in \langle X \rangle$. Let $f, g \in \langle X \rangle$, then

$$f, g \in \langle X \rangle \Longrightarrow f(s) = 0 \text{ and } g(s) = 0, \quad \forall s \in S,$$

so

$$f(s) + g(s) = 0 + 0 = 0, \ \forall s \in S.$$

Consequently, $f + g \in \langle X \rangle$. Secondly, we need to show that for $f \in \langle X \rangle$ and $h \in C_N(\mathbb{R})$ then $h \cdot f \in \langle X \rangle$. Let $h \in C_N(\mathbb{R})$, then for all $s \in S$ it follows that

$$f \in \langle X \rangle \Rightarrow f(s) = 0,$$

hence

$$h(s) \cdot f(s) = h(s) \cdot 0 = 0, \ \forall s \in S.$$

Thus, $h \cdot f \in \langle X \rangle$. Therefore $\langle X \rangle$ is an ideal. $\square$

Lemma 6.2.26. *Let X and S be defined as in Lemma 6.2.25. If S is an empty* *set, then* $\langle X\rangle = C_N(\mathbb{R})$.

The proof of this lemma is trivial. However, the following corollary is essential for our proof that a decomposition exists. This proof (of the matrix spectral decomposition) provides the setting for the proof of the tensor spectral decomposition theorem.

Corollary 6.2.27. *Let* X *and* S *be defined as in Lemma 6.2.25. If* $\langle X\rangle \neq C_N(\mathbb{R})$ *then* S *is a non-empty set, i.e., there exists a common zero of all functions in* X.

In [7], Theorem 6.2.28 was proved in a ring of polynomials. However, introducing the absolute value necessitates working in a larger ring and we do so here.

To prove that the decomposition (6.39) exists, let us consider a matrix A with $\|A\|_2 \neq 1$, we want to show that (6.39) has a non-trivial solution using an ideal under the ring of continuous functions.

Let $D = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$, $E = \text{diag}(\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$. Write $q_x = (q_{x1}, \ldots, q_{xn})^T$ and $r_y = (r_{y1}, r_{y2}, \ldots, r_{yn})$. We will rewrite $A = (DQ)^T ER$ and $I = Q^T R$ as a system of polynomial equations, that is, $A = (DQ)^T ER$ is equivalent to finding a zero of the system of polynomials

$$P_1(D, E, q_x, r_y) := a_{xy} - \sum_{\substack{k=1\\1\leq x\leq y}}^{n} (\sigma_k q_{xk})\epsilon_k r_{ky} \tag{6.41}$$

and $I = Q^T R$ is equivalent to finding a zero of the system of polynomials

$$P_2(q_x, r_y) := \delta_{xy} - \sum_{\substack{k=1\\1\leq x\leq y}}^{n} q_{xk} r_{ky}. \tag{6.42}$$

Let $\mathcal{X} = \left\{\sum_{k=1} \sigma_k q_{xk}\epsilon_k r_{ky} - a_{xy};\ \sum_{k=1} q_{xk} r_{kj} - \delta_{xy}\right\}$ and consider the ideal $\langle \mathcal{X}\rangle$ generated by the set $\mathcal{X}$, where $\mathcal{X} \subseteq \mathbb{R}\left[\bigcup_j \{\sigma_j, \epsilon_j, q_j, r_j\}\right] \subseteq C_N(\mathbb{R}^N)$ then

$$\langle \mathcal{X}\rangle = \left\langle \left\{\sum_{k=1} \sigma_k q_{xk}\epsilon_k r_{ky} - a_{xy};\ \sum_{k=1} q_{xk} r_{kj} - \delta_{xy}\right\}\right\rangle. \tag{6.43}$$

The following theorem and its proof follows the proof of the weak spectral theorem in [7]. However, we have expressed some concepts in different ways and adopted the proof according. In this proof, we postpone an explanation of the choice of $Q(D, E, q_x, r_y)$ below to Theorem 6.2.30 where the third-order case is handled in much the same way.

Theorem 6.2.28. *Let A be an $n \times n$ symmetric matrix, with $\|A\|_2 \neq 1$, then the spectral system of polynomial equations*

$$\begin{cases} P_1(D, E, q_x, r_y) = 0, \\ P_2(q_x, r_y) = 0, \end{cases} \tag{6.44}$$

has a solution, where $P_1(D, E, q_x, r_y)$ and $P_2(q_x, r_y)$ are defined in terms of A and I respectively in equations (6.41) *and* (6.42).

Proof. Let $Q(D, E, q_x, r_y) \in C_N(\mathbb{R}^N)$ be given by

$$Q(D, E, q_x, r_y) := \left(\sum_{x,y=1}^{z} |P_1(D, E, q_x, r_y)|^2 + |a_{xy}|^2 \right)^2 - \|A\|_2^2. \tag{6.45}$$

We will show that $Q(D, E, q_x, r_y) \notin \langle \mathcal{X} \rangle$ using a proof by contradiction.

Suppose that $Q(D, E, q_x, r_y) \in \langle \mathcal{X} \rangle$, then for every common zero (D, E, q_x, r_y) of the polynomials in $\mathcal{X}$, using (6.45), it follows that

$$Q(D, E, q_x, r_y) = 0 \Rightarrow \left(\sum_{x,y=1}^{z} |P_1(D, E, q_x, r_y)|^2 + |a_{xy}|^2 \right)^2 - \|A\|_2^2 = 0.$$

So, $\left(\sum_{x,y=1}^{z} |P_1(D, E, q_x, r_y)|^2 + |a_{xy}|^2 \right)^2 = \|A\|_2^2$. When $\|A\|_2 = 0$, it follows that $A = 0$ and (6.44) has a trivial solution. So assume $\|A\|_2 \neq 0$. Since, (D, E, q_x, r_y) is a common solution of the polynomials in $\mathcal{X}$, we must have $P_1(D, E, q_x, r_y) = 0$, hence $(\|A\|_2^2)^2 = \|A\|_2^2$. However, this contradicts $\|A\|_2 \neq 1$, thus from Corollary 6.2.27, and the fact that $Q(D, E, q_x, r_y) \notin \langle \mathcal{X} \rangle$ and $Q(D, E, q_x, r_y) \in C_N(\mathbb{R}^N)$, (6.44) has a solution. $\square$

Corollary 6.2.29. *Let A be a $n \times n$ symmetric matrix, with $\|A\|_2 = 1$, then the spectral system of polynomial equations*

$$\begin{cases} P_1(D, E, q_x, r_y) = 0, \\ P_2(q_x, r_y) = 0, \end{cases} \tag{6.46}$$

has a solution, where $P_1(D, E, q_x, r_y)$ and $P_2(q_x, r_y)$ are defined in terms of A and I in equations (6.41) *and* (6.42)*, respectively.*

Proof. Let $\|A\|_2 = 1$ and $B = kA$, where $|k| \neq 1$, then

$$\|B\|_2 = \|kA\|_2 = |k|\|A\|_2 = |k| \neq 1. \tag{6.47}$$

Since B is a symmetric matrix with $\|B\|_2 \neq 1$, it follows from Theorem 6.2.28

that (6.46) has a solution, i.e., $B = (DQ)^T ER$ and hence

$$A = \left(\frac{1}{k}DQ\right)^T ER, \tag{6.48}$$

is a spectral decomposition of A. $\square$

We can now extend the spectral theorem for matrices to third-order tensors. Let $\mathcal{A}$ be a third-order symmetric tensor, then Theorem 6.2.30 below shows that

$$\begin{cases} \mathcal{A} &= (\mathcal{F}^T \mathcal{S} \mathcal{F})^T (\mathcal{D}^T \mathcal{Q} \mathcal{D})(\mathcal{E}^T \mathcal{R} \mathcal{E})^{T^2}, \\ \Delta &= \mathcal{S}^T \mathcal{Q} \mathcal{R}^{T^2}, \end{cases} \tag{6.49}$$

where $\mathcal{F}$, $\mathcal{D}$ and $\mathcal{E}$ are scaling tensors. Using the definition of the ternary product (6.35), we can rewrite (6.49) as

$$\begin{cases} a_{xyz} = \displaystyle\sum_{t=1}^{n} (\omega_{ty} s_{zty} \omega_{zt})(\sigma_{tz} q_{xtz} \sigma_{xt})(\epsilon_{tx} r_{ytx} \epsilon_{yt}), \\ \delta_{xyz} = \displaystyle\sum_{t=1}^{n} s_{zty} q_{xtz} r_{ytx}, \end{cases} \tag{6.50}$$

where $1 \le x \le y \le z \le n$. Satisfying (6.49) is equivalent to finding zeros of the system of polynomials

$$a_{xyz} - \sum_{t=1}^{n} (\omega_{ty} s_{zty} \omega_{zt})(\sigma_{tz} q_{xtz} \sigma_{xt})(\epsilon_{tx} r_{ytx} \epsilon_{yt}) := P_1(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx}) \tag{6.51}$$

and

$$\delta_{xyz} - \sum_{t=1}^{n} s_{zty} q_{xtz} r_{ytx} := P_2(s_{zty}, q_{xtz}, r_{ytx}). \tag{6.52}$$

We will consider the following ideal, $\langle \mathcal{L} \rangle$, generated by the set $\mathcal{L}$, where

$$\mathcal{L} \subseteq \mathbb{R}\left[\bigcup_j \{\omega_j, s_j, q_j, \sigma_j, r_j, \epsilon_j\}\right],$$

is given by

$$\langle \mathcal{L} \rangle = \{P_1(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx}) - a_{xyz}; P_2(s_{zty}, q_{xtz}, r_{ytx}) - \delta_{xyz}\}.$$

The following theorem extends from the matrix spectral decomposition. We note that we have a squared term in the expression for H below (unlike [7], which cubes the terms). This is necessary to contradict $\|\mathcal{A}\|_3^3 \neq 1$. The proof in [7] follows similarly by noting that $x^9 = x^3$ has the solutions $x = 0$ and $x = 1$ when $x > 0$. By taking the square of H (below), we avoid this small technical detail.

Theorem 6.2.30. *Let $\mathcal{A}$ be a symmetric tensor, with $\|\mathcal{A}\|_3^3 \neq 1$, then*

$$\begin{cases} P_1(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx}) = 0, \\ P_2(s_{zty}, q_{xtz}, r_{ytx}) = 0, \end{cases} \tag{6.53}$$

has a solution, where $P_1(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx})$ and $P_2(s_{zty}, q_{xtz}, r_{ytx})$ are defined in terms of $\mathcal{A}$ and Δ in (6.51) *and* (6.52)*, respectively.*

Proof. The proof follow the same approach as the proof of Theorem 6.2.28. We will describe a continuous function $H(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx})$ such that $H \notin \langle \mathcal{L} \rangle$. Let H be expressed in terms of the Frobenius norm of $P_1(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx})$, that is

$$H := \left(\sum_{x,y,z=1}^{n} |P_1(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx})|^3 + |a_{xyz}|^3 \right)^2 - \|\mathcal{A}\|_3^3.$$

If we assume that $H \in \langle \mathcal{L} \rangle$, for every common zero

$$(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx})$$

of polynomials in $\mathcal{L}$, then

$$H = 0 \Rightarrow \left(\sum_{x,y,z=1}^{n} |P_1(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx})|^3 + |a_{xyz}|^3 \right)^2 - \|\mathcal{A}\|_3^3 = 0,$$

so that,

$$\left(\sum_{x,y,z=1}^{n} |P_1(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx})|^3 + |a_{xyz}|^3 \right)^2 = \|\mathcal{A}\|_3^3.$$

When $\|\mathcal{A}\| = 0$, then $\mathcal{A} = 0$ and (6.53) has a trivial solution. Therefore assume that $\|\mathcal{A}\|_3 \neq 0$, since $(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx})$ is a common solution of the polynomial in $\mathcal{L}$, we must have that $P_1(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx}) = 0$. Hence, $(\|\mathcal{A}\|_3^3)^2 = \|\mathcal{A}\|_3^3$. which contradicts the assumption that $\|\mathcal{A}\|_3^3 \neq 1$, thus (6.2.30) has a solution. □

Corollary 6.2.31. *Let $\mathcal{A}$ be a symmetric tensor, with $\|\mathcal{A}\|_3^3 = 1$, then*

$$\begin{cases} P_1(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx}) = 0, \\ P_2(s_{zty}, q_{xtz}, r_{ytx}) = 0, \end{cases} \tag{6.54}$$

has a solution, where $P_1(F_{zy}, D_{xz}, E_{yx}, s_{zty}, q_{xtz}, r_{ytx})$ and $P_2(s_{zty}, q_{xtz}, r_{ytx})$ are defined in terms of $\mathcal{A}$ and Δ in (6.51) *and* (6.52) *respectively.*

Proof. Let $\mathcal{A}$ be a symmetric tensor, with $\|\mathcal{A}\|_3 = 1$, then it follows that (6.53) has a solution, since $\mathcal{B} = k\mathcal{A}$, where $|k| \neq 1$. Now,

$$\|\mathcal{B}\|_3 = \|k\mathcal{A}\|_3 = |k|\|\mathcal{A}\|_3 = |k| \neq 1.$$

Hence

$$\mathcal{A} = \left(\frac{1}{k}\mathcal{F}^T\mathcal{S}\mathcal{F}\right)^T (\mathcal{D}^T\mathcal{Q}\mathcal{D})(\mathcal{E}^T\mathcal{R}\mathcal{E})^{T^2}. \tag{6.55}$$

Thus, by Theorem 6.2.30, it follows that (6.55) has a solution. □

Chapter 7

Conclusion

Unlike matrix multiplication, the higher-order tensor product is not unique. Consequently, there are multiple ways to define the product and thus investigate algebraic properties. The products we investigated in Chapter 2, 3 and 4 involves unfolding tensors to matrices and thus making it easy to study the algebraic properties from the matrix space.

The products in chapter 5 and 6 are different in the sense that they do not have an underlying matrix product. For each product, we investigated the algebraic properties that arise in linear algebra. The products of tensors in chapter 2, 3 and 4 generalizes these algebraic properties (from matrices) and so, the products of tensors is a generalization of matrix multiplication.

For all the products in this dissertation (with the exception of Chapter 5), we studied the singular value decomposition and the Moore-Penrose pseudo inverse. We also note that every tensor that is decomposable (by the singular value decomposition) has a pseudo inverse given in terms of its singular value decomposition.

From our investigation of the products of tensors, the approach of the proof for the spectral decomposition of a tensor in chapter 6 can be applied to further investigate the decompositions for the n-ary product.

Appendix A

Additional comments on the ternary product

A.1 Associativity of the ternary product

Let α', B' and r' be $n \times n$ matrices and α, B and r be tensors such that

$$\alpha = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \otimes \alpha', \quad B = \left[\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \otimes B' \right]^T, \quad \text{and} \quad r = \left[\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \otimes r' \right]^{T^2}$$

and let A, b and γ be arbitrary third-order n-dimensional tensors, then

$$(\mathcal{I}\alpha A)(b\mathcal{I}B)(r\gamma\mathcal{I}^T) = \mathcal{I}\left(\alpha\mathcal{I}B\right)\left(r(Ab\gamma)\mathcal{I}^T\right). \tag{A.1}$$

Proof.

$$[(\mathcal{I}\alpha A)(b\mathcal{I}B)(r\gamma\mathcal{I}^T)]_{ijk} = \sum_{l=1}^{n}\sum_{u=1}^{n}\sum_{v=1}^{n} \delta_{lj}\delta_{iu}\delta_{vj} \sum_{t=1}^{n} \left[\alpha_{tlk}A_{tjk}b_{itk}B_{itu}r_{vjt}\gamma_{ijk}\right].$$

Since $\alpha_{tlk} = \alpha_{ulk}$, $B_{itu} = B_{ilu}$ and $r_{vjt} = r_{vju}$, then

$$\begin{aligned}
&[(\mathcal{I}\alpha A)(b\mathcal{I}B)(r\gamma\mathcal{I}^T)]_{ijk} \\
&\quad = \sum_{l=1}^{n}\sum_{u=1}^{n}\sum_{v=1}^{n} \delta_{lj}\delta_{iu}\delta_{vj} \sum_{t=1}^{n} \left[\alpha_{ulk}A_{tjk}b_{itk}B_{ilu}r_{vju}\gamma_{ijk}\right] \\
&\quad = \sum_{l=1}^{n}\sum_{u=1}^{n}\sum_{v=1}^{n} \mathcal{I}_{ljk}(\mathcal{I})_{iuk}(\mathcal{I}^T)_{ijv}\alpha_{ulk}B_{ilu}r_{vju}(Ab\gamma)_{ijk}
\end{aligned}$$

$$= \sum_{l=1}^{n} \mathcal{I}_{ljk} \left(\sum_{u=1}^{n} \alpha_{ulk} (\mathcal{I})_{iuk} B_{ilu} \right) \left(\sum_{v=1}^{n} r_{vjl} (Ab\gamma)_{ivl} (\mathcal{I}^T)_{ijv} \right)$$
$$= \sum_{l=1}^{n} \mathcal{I}_{ljk} \left(\alpha \mathcal{I} B\right)_{ilk} \left(r(Ab\gamma)\mathcal{I}^T\right)_{ijl}$$
$$= [\mathcal{I} \left(\alpha \mathcal{I} B\right) \left(r(Ab\gamma)\mathcal{I}^T\right)]_{ijk}. \qquad \square$$

A.2 Adjoint of the ternary product

To understand the adjoint operator for the ternary product as defined in [7], let us first define this for matrices. Let us first define the conjugate operation,

Definition A.2.1. Let $z = re^{i\theta}$, then the conjugate is the map $c_n : z \to \overline{z}$, where $\overline{z} = re^{-i\theta}$. The conjugate may be expressed as

$$c_n^k(z) = re^{i\theta e^{\frac{i2\pi k}{n}}}, \tag{A.2}$$

where $c_n^{n-1}(z) = \overline{z}$ and $c_n^n(z) = z$ and n is the order of the tensor with entries z.

Now the adjoint is the transpose conjugate, that is

Definition A.2.2. Let $A \in \mathbb{C}^{n\times n}$, then the adjoint is

$$\begin{cases} A^{\dagger} & = (c_2^1(a_{ij}))^T, \\ A^{\dagger^2} & = (c_2^2(a_{ij}))^{T^2}. \end{cases} \tag{A.3}$$

We can extend this definition to third order tensors.

Definition A.2.3. Let $\mathcal{A}$ be a third order tensor, then the adjoint of $\mathcal{A}$ is

$$\begin{cases} \mathcal{A}^{\dagger} & = (c_3^1(a_{ijk}))^T, \\ \mathcal{A}^{\dagger^2} & = (c_3^2(a_{ijk}))^{T^2}. \end{cases} \tag{A.4}$$

In [7] the authors claim that the product $\mathcal{A}^{\dagger}\mathcal{A}\mathcal{A}^{\dagger^2}$ is Hermitian. We will produce a counterexample to show that this is not true.

Example A.2.4. Let $\mathcal{A} \in \mathbb{C}^{2\times 2\times 2}$ be a third-order tensor with entries

$$a_{111} = 0, \quad a_{112} = 0, \quad a_{121} = 1, \quad a_{122} = i,$$
$$a_{211} = i, \quad a_{212} = 1, \quad a_{221} = 0, \quad a_{222} = 0.$$

Applying the conjugate to $\mathcal{A}$, gives us the tensor $c_3^1(\mathcal{A})$ such that

$$c_3^1(a_{111}) = 0, \quad c_3^1(a_{112}) = 0, \quad c_3^1(a_{121}) = 1, \quad c_3^1(a_{122}) = e^{-\frac{i\pi}{4}} e^{-\frac{\sqrt{3}\pi}{4}},$$
$$c_3^1(a_{211}) = e^{-\frac{i\pi}{4}} e^{-\frac{\sqrt{3}\pi}{4}}, \quad c_3^1(a_{212}) = 1, \quad c_3^1(a_{221}) = 0, \quad c_3^1(a_{222}) = 0.$$

Taking the transpose of $c_3^1(\mathcal{A})$, we get our adjoint $\mathcal{A}^\dagger$ with entries

$$a_{111} = 0, \quad a_{112} = 1, \quad a_{121} = e^{-\frac{i\pi}{4}} e^{-\frac{\sqrt{3}\pi}{4}}, \quad a_{122} = 0,$$
$$a_{211} = 0, \quad a_{212} = e^{-\frac{i\pi}{4}} e^{-\frac{\sqrt{3}\pi}{4}}, \quad a_{221} = 1, \quad a_{222} = 0.$$

For $\mathcal{A}^{\dagger^2}$, we first determine the entries of $C_3^2(\mathcal{A})$

$$c_3^2(a_{111}) = 0, \quad c_3^2(a_{112}) = 0, \quad c_3^2(a_{121}) = 1, \quad c_3^2(a_{122}) = e^{\frac{i\pi}{4}} e^{\frac{\sqrt{3}\pi}{4}},$$
$$c_3^2(a_{211}) = e^{-\frac{i\pi}{4}} e^{\frac{\sqrt{3}\pi}{4}}, \quad c_3^2(a_{212}) = 1, \quad c_3^2(a_{221}) = 0, \quad c_3^2(a_{222}) = 0.$$

Now, taking the transpose of $c_3^2(\mathcal{A})$ twice, we get $\mathcal{A}^{\dagger^2}$, with entries

$$a_{111} = 0, \quad a_{112} = e^{-\frac{i\pi}{4}} e^{-\frac{\sqrt{3}\pi}{4}}, \quad a_{121} = 0, \quad a_{122} = 1,$$
$$a_{211} = 0, \quad a_{212} = 0, \quad a_{221} = e^{-\frac{i\pi}{4}} e^{-\frac{\sqrt{3}\pi}{4}}, \quad a_{222} = 0.$$

Let $\mathcal{D} = \mathcal{A}^\dagger \mathcal{A} \mathcal{A}^{\dagger^2}$, to calculate the entries of $\mathcal{D}$, we use

$$d_{xyz} = \sum_{t=1}^{2} a^\dagger_{tyz} a_{xtz} a^{\dagger^2}_{xyt}. \tag{A.5}$$

Therefore, the conjugate of $\mathcal{D}$, $C_3^1(\mathcal{D})$ has the entries

$$d_{111} = 0, \quad d_{112} = 1, \quad d_{121} = 1, \quad d_{122} = 0,$$
$$d_{211} = 0, \quad d_{212} = 0, \quad d_{221} = 1, \quad d_{222} = 0.$$

The adjoint of $\mathcal{D}$ is $\mathcal{D}^\dagger = (C_3^1(\mathcal{D}))^T$ with entries

$$d_{111} = 0, \quad d_{112} = 1, \quad d_{121} = 0, \quad d_{122} = 1,$$
$$d_{211} = 1, \quad d_{212} = 0, \quad d_{221} = 0, \quad d_{222} = 0.$$

Thus $\mathcal{D}^\dagger \neq \mathcal{D}$, therefore $\mathcal{A}^\dagger \mathcal{A} \mathcal{A}^{\dagger^2}$ is not self-adjoint.

Bibliography

[1] Howard Anton and Chris Rorres, *Elementary linear algebra: applications version*, John Wiley & Sons, 2013.

[2] Changjiang Bu, Xu Zhang, Jiang Zhou, Wenzhe Wang, and Yimin Wei, *The inverse, rank and product of tensors*, Linear Algebra and Its Applications **446** (2014), 269–280.

[3] Pierre Comon, *Tensors: a brief introduction*, IEEE Signal Processing Magazine **31** (2014), no. 3, 44–53.

[4] Lieven De Lathauwer, Bart De Moor, and Joos Vandewalle, *A multilinear singular value decomposition*, SIAM journal on Matrix Analysis and Applications **21** (2000), no. 4, 1253–1278.

[5] Beno Eckmann and Peter John Hilton, *Group-like structures in general categories I multiplications and comultiplications.*, Mathematische Annalen **145** (1962), 227–255.

[6] Josiah Willard Gibbs, *Elements of vector analysis: arranged for the use of students in physics*, Tuttle, Morehouse & Taylor, 1884.

[7] Edinah Koffi Gnang, Ahmed Elgammal, and Vladimir Retakh, *A spectral theory for tensors*, Annales de la Faculté des Sciences de Toulouse: Mathématiques, vol. 20, 2011, pp. 801–841.

[8] Gene Howard Golub and Charles Francis Van Loan, *Matrix computations*, JHU press, 2013.

[9] Misha Elena Kilmer and Carla Dee Martin, *Factorization strategies for third-order tensors*, Linear Algebra and its Applications **435** (2011), no. 3, 641–658.

[10] Tamara Kolda and Brett Bader, *Tensor decompositions and applications*, SIAM review **51** (2009), no. 3, 455–500.

[11] Yueh-Cheng Kuo and Tsung-Lin Lee, *Computing the unique candecomp/parafac decomposition of unbalanced tensors by homotopy method*, Linear Algebra and its Applications **556** (2018), 238–264.

[12] Russell Merris, *Multilinear algebra*, Algebra, Logic and Applications, vol. 8, Gordon and Breach Science Publishers, Amsterdam, 1997.

[13] Dale Mesner and Prabir Bhattacharya, *Association schemes on triples and a ternary algebra*, Journal of Combinatorial Theory, Series A **55** (1990), no. 2, 204–234.

[14] Liqun Qi, *Eigenvalues of a real supersymmetric tensor*, Journal of Symbolic Computation **40** (2005), no. 6, 1302–1324.

[15] Kathrin Schäcke, *On the Kronecker product*, Master's thesis, University of Waterloo (2004).

[16] Jia-Yu Shao, *A general product of tensors with applications*, Linear Algebra and its Applications **439** (2013), no. 8, 2350–2366.

[17] Lizhu Sun, Baodong Zheng, Changjiang Bu, and Yimin Wei, *Moore–Penrose inverse of tensors via Einstein product*, Linear and Multilinear Algebra **64** (2016), no. 4, 686–698.