Lecture VI Tensor Spaces and Tensors

Lei Du
dulei@dlut.edu.cn

School of Mathematical Sciences
Dalian University of Technology

October, 2021

Contents

Introduction

- define the concepts of hypermatrix and outer product, as well as the operations of contraction and n-mode hypermatrix-matrix product
- establish the links between multilinear forms and multilinear maps, and homogeneous polynomials and hypermatrices
- consider the special cases of bilinear maps and bilinear forms
- introduce concept of symmetric hypermatrix, especially symmetric hypermatrices of third and fourth orders
- introduce Nth-order tensors based on the tensor product of N v.s., called tensor space
- highlight the hypermatrix of coordinates of a tensor with respect to a given basis, and present the canonical writing of tensors
- analyze the effect of a change of basis in the vector spaces of a tensor product, on the coordinate hypermatrix of a tensor
- introduce the notions of tensor rank and tensor decomposition through the presentation of the canonical polyadic decomposition (CPD)
- define the notions of eigenvalue and singular value of an hypermatrix
- finally, describe different examples of isomorphisms of tensor spaces

Hypermatrix vector spaces

A hypermatrix $\mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ of order N and of dimensions $I_1 \times \cdots \times I_N$, denoted $[a_{i_1,\cdots,i_N}]$ or $[a_{i_1 \cdots i_N}]$, is an array composed of elements depending on N indices $i_n \in \langle I_n \rangle$, with $n \in \langle N \rangle$. Indices are also called modes, dimensions or axes. The hypermatrix $\mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ can be defined as a map $f \colon \mathop{\times}\limits_{n=1}^{N} \langle I_n \rangle \to \mathbb{K}$ such that:

$$\mathop{\times}_{n=1}^{N} \langle I_n \rangle \ni (i_1, \cdots, i_N) \mapsto f(i_1, \cdots, i_N) = a_{i_1,\cdots,i_N} \in \mathbb{K}$$

where I_n is the upper bound of index i_n, corresponding to the n-mode dimension.

Hypermatrix vector spaces

Hypermatrices generalize vectors and matrices to orders higher than two, and the special cases of hypermatrices of orders $N = 1$ and $N = 2$ correspond to vectors and matrices, respectively.

- if $I_1 = \cdots = I_N = I$, the hypermatrix $\mathcal{A} \in \mathbb{K}^{I \times I \times \cdots \times I}$ of order N is said to be hypercubic.
- if $a_{i_1, \cdots, i_N} \neq 0$ only for $i_1 = i_2 = \cdots = i_N$, then $\mathcal{A}$ is called a diagonal hypermatrix. The set of elements $a_{i,i,\cdots,i}$, with $i \in \langle I \rangle$, form the diagonal of $\mathcal{A}$, and the sum of these elements is the trace of $\mathcal{A}$, denoted by $\mathrm{tr}(\mathcal{A})$: $\mathrm{tr}(\mathcal{A}) = \sum_{i=1}^{I} a_{i,i,\cdots,i}$.
- the identity hypermatrix of order N and of dimensions $I \times \cdots \times I$ is denoted $\mathcal{I}_{N,I} = [\delta_{i_1,\cdots,i_N}]$, with $i_n \in \langle I \rangle$ for $n \in \langle N \rangle$, or simply $\mathcal{I}$. This is a hypercubic hypermatrix whose elements are defined using the generalized Kronecker delta:

$$\delta_{i_1,\cdots,i_N} = \begin{cases} 1 & \text{if } i_1 = \cdots = i_N \\ 0 & \text{otherwise} \end{cases} .$$

Hypermatrix vector spaces

Let $\mathcal{A} = [a_{i_1}, \cdots, i_N], \mathcal{B} = [b_{i_1}, \cdots, i_N]$, and $\lambda \in \mathbb{K}$. By defining addition and scalar multiplication operations for hypermatrices of $\mathbb{K}^{I_1 \times \cdots \times I_N}$, such as:

$$\mathcal{A} + \mathcal{B} = \left[a_{i_1,\cdots,i_N} + b_{i_1,\cdots,i_N}\right],$$
$$\lambda\mathcal{A} = \left[\lambda a_{i_1,\cdots,i_N}\right],$$

the set $\mathbb{K}^{I_1 \times \cdots \times I_N}$ of hypermatrices, of order N, is a v.s. of dimension $\prod_{n=1}^{N} I_n$. This dimension expresses the fact that a hypermatrix of $\mathbb{K}^{I_1 \times \cdots \times I_N}$ can be vectorized as a vector of $\mathbb{K}^{\prod_{n=1}^{N} I_n}$.

Hypermatrix inner product and Frobenius norm

Given two complex hypermatrices $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, their Hermitian inner product is defined as:

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} a_{i_1,\cdots,i_N} b^*_{i_1,\cdots,i_N} = a_{i_1,\cdots,i_N} b^*_{i_1,\cdots,i_N}, \tag{1}$$

where the last equality results from the use of Einstein's summation convention which consists in summing over repeated indices.

For example, $\sum_{i=1}^{I} x_i \mathbf{e}_i^{(I)}$ is replaced by $x_i \mathbf{e}_i^{(I)}$. Similarly, $\sum_{j=1}^{J} a_{ij} b_{jk}$ will be replaced by $a_{ij} b_{jk}$. This convention, which makes it possible to simplify the writing of equations involving quantities defined using multiple indices, will be utilized later in this chapter.

Hypermatrix inner product and Frobenius norm

The Frobenius norm, also called Hilbert-Schmidt norm, of $\mathcal{A}$ is defined as:

$$\|\mathcal{A}\|_F^2 = \langle \mathcal{A}, \mathcal{A} \rangle = a_{i_1,\cdots,i_N} a^*_{i_1,\cdots,i_N} = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} \left| a_{i_1,\cdots,i_N} \right|^2 . \qquad (2)$$

In the matrix case, for $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{I \times J}$, equations (1) and (2) become:

$$\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{i=1}^{I} \sum_{j=1}^{J} a_{ij} b^*_{ij} = a_{ij} b^*_{ij}$$

$$\|\mathbf{A}\|_F^2 = \langle \mathbf{A}, \mathbf{A} \rangle = \sum_{i=1}^{I} \sum_{j=1}^{J} \left| a_{ij} \right|^2 = a_{ij} a^*_{ij}$$

NOTE: In the case of two real hypermatrices $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$, their Euclidean inner product and the Frobenius norm are obtained by removing the conjugation in definitions (1) and (2).

Two important operations are defined here.

- the contraction of two hypermatrices sharing common modes
- the n-mode product of a hypermatrix with a matrix having a common mode, which can be viewed as a special case of contraction

Contraction operation

Let $\mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ and $\mathcal{B} \in \mathbb{K}^{J_1 \times \cdots \times J_P}$ be two hypermatrices of orders N and P, respectively, with M common modes. Contraction corresponds to summations over common modes. The result of this contraction is a hypermatrix $\mathcal{C}$ of order $N + P - 2M$.

Contraction operation

Let $\mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ and $\mathcal{B} \in \mathbb{K}^{J_1 \times \cdots \times J_P}$ be two hypermatrices of orders N and P, respectively, with M common modes. Contraction corresponds to summations over common modes. The result of this contraction is a hypermatrix $\mathcal{C}$ of order $N + P - 2M$.

EXAMPLE: For $\mathcal{A} \in \mathbb{K}^{I_1 \times I_2 \times I_3}$ and $\mathcal{B} \in \mathbb{K}^{I_2 \times I_3 \times J_1 \times J_2}$, the contraction over common indices i_2 and i_3 gives the contracted hypermatrix $\mathcal{C} \in \mathbb{K}^{I_1 \times J_1 \times J_2}$ such that $c_{i_1, j_1, j_2} = \sum_{i_2=1}^{I_2} \sum_{i_3=1}^{I_3} a_{i_1, i_2, i_3} b_{i_2, i_3, j_1, j_2}$. We write $C = \mathcal{A} \underset{1,2}{\overset{2,3}{\times}} \mathcal{B}$, where numbers $\{1, 2\}$ are relative to the position of the common indices of tensor $\mathcal{B}$, while numbers $\{2, 3\}$ define the position of the common indices of $\mathcal{A}$.

Contraction operation

Let $\mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ and $\mathcal{B} \in \mathbb{K}^{J_1 \times \cdots \times J_P}$ be two hypermatrices of orders N and P, respectively, with M common modes. Contraction corresponds to summations over common modes. The result of this contraction is a hypermatrix $\mathcal{C}$ of order $N + P - 2M$.

EXAMPLE: For $\mathcal{A} \in \mathbb{K}^{I_1 \times I_2 \times I_3}$ and $\mathcal{B} \in \mathbb{K}^{I_2 \times I_3 \times J_1 \times J_2}$, the contraction over common indices i_2 and i_3 gives the contracted hypermatrix $\mathcal{C} \in \mathbb{K}^{I_1 \times J_1 \times J_2}$ such that $c_{i_1,j_1,j_2} = \sum_{i_2=1}^{I_2} \sum_{i_3=1}^{I_3} a_{i_1,i_2,i_3} b_{i_2,i_3,j_1,j_2}$. We write $C = \mathcal{A} \underset{1,2}{\overset{2,3}{\times}} \mathcal{B}$, where numbers $\{1, 2\}$ are relative to the position of the common indices of tensor $\mathcal{B}$, while numbers $\{2, 3\}$ define the position of the common indices of $\mathcal{A}$.

FACT: For $\mathbb{K} = \mathbb{R}$ and $M = N = P$, the contraction of $\mathcal{A}$ and $\mathcal{B} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ corresponds to their inner product:

$$\mathcal{A} \underset{1,2,\cdots,N}{\overset{1,2,\cdots,N}{\times}} \mathcal{B} = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} a_{i_1,\cdots,i_N} b_{i_1,\cdots,i_N} = \langle \mathcal{A}, \mathcal{B} \rangle.$$

n-mode hypermatrix-matrix product

The n-mode product of the hypermatrix $\mathcal{X} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ with the matrix $\mathbf{A} \in \mathbb{K}^{J_n \times I_n}$, denoted by $\mathcal{X} \times_n \mathbf{A}$, is equivalent to a contraction of the hypermatrix $\mathcal{X}$ with matrix $\mathbf{A}$ over their common mode i_n. This contraction gives the N-order hypermatrix $\mathcal{Y}$ of dimensions $I_1 \times \cdots \times I_{n-1} \times J_n \times I_{n+1} \times \cdots \times I_N$, such that:

$$y_{i_1,\ldots,i_{n-1},j_n,i_{n+1},\ldots,i_N} = \sum_{i_n=1}^{I_N} a_{j_n,i_n} x_{i_1,\ldots,i_{n-1},i_n,i_{n+1},\ldots,i_N}. \quad (3)$$

Thus, for $\mathcal{X} \in \mathbb{K}^{I \times J \times K}, \mathbf{A} \in \mathbb{K}^{P \times I}, \mathbf{B} \in \mathbb{K}^{Q \times J}, \mathbf{C} \in \mathbb{K}^{R \times K}$, we have:

$$(\mathcal{X} \times_1 \mathbf{A})_{p,j,k} = \sum_{i=1}^{I} a_{pi} x_{i,j,k}$$

$$(\mathcal{X} \times_2 \mathbf{B})_{i,q,k} = \sum_{j=1}^{J} b_{qj} x_{i,j,k}$$

$$(\mathcal{X} \times_3 \mathbf{C})_{i,j,r} = \sum^{K} c_{rk} x_{i,j,k}$$

n-mode hypermatrix-matrix product

and therefore, we deduce that:

$$\mathcal{Y} = \mathcal{X} \times {}_1\mathbf{A} \times {}_2\mathbf{B} \times {}_3\mathbf{C} \Leftrightarrow y_{p,q,r} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} a_{pi} b_{qj} c_{rk} x_{i,j,k}.$$

More generally, by considering a permutation $\pi(.)$ of the first N natural numbers $n \in \langle N \rangle$ such as $p_n = \pi(n)$, a series of p_n-mode products of $\mathcal{X} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ with $\mathbf{A}^{(p_n)} \in \mathbb{K}^{J_{p_n} \times I_{p_n}}$, $n \in \langle N \rangle$, will be concisely denoted as:

$$\mathcal{X} \times_{p_1} \mathbf{A}^{(p_1)} \cdots \times_{p_N} \mathbf{A}^{(p_N)} = \mathcal{X} \times_{p=p_1}^{p_N} \mathbf{A}^{(p)} \in \mathbb{K}^{J_{p_1} \times \cdots \times J_{p_N}}. \tag{4}$$

n-mode hypermatrix-matrix product

In the case where the hypermatrix-matrix product is performed for all modes, another concise notation of (4) was proposed by de Silva and Lim (2008):

$$\mathcal{X} \times_{p_1} \mathbf{A}^{(p_1)} \cdots \times_{p_N} \mathbf{A}^{(p_N)} = \left(\mathbf{A}^{(p_1)}, \cdots, \mathbf{A}^{(p_N)}\right) \cdot \mathcal{X} \tag{5}$$

where the order of the matrices $\mathbf{A}^{(p_1)}, \cdots, \mathbf{A}^{(p_N)}$ is identical to the order of the p_n-mode products, that is, $\times_{p_1}, \cdots, \times_{p_N}$.
Similarly, it is possible to define n-mode product of $\mathcal{X} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ with vector $\mathbf{u}^{(n)} \in \mathbb{K}^{I_n}$, for $n \in \langle N \rangle$. This gives a scalar:

$$y = \mathcal{X} \times_{n=1}^{N} \mathbf{u}^{(n)} = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} u_i^{(1)} \cdots u_i^{(N)} x_{i_1, \cdots, i_N}.$$

Properties of hypermatrix-matrix product

The n-mode product plays a very important role in tensor calculus, so we describe its main properties hereafter.

定理

The n-mode product satisfies the following properties:

- *For any permutation $\pi(.)$ of the first N natural numbers $n \in \langle N \rangle$ such as $p_n = \pi(n)$, we have:*

$$\mathcal{X} \times_{p=p_1}^{p_N} \mathbf{A}^{(p)} = \mathcal{X} \times_{n=1}^{N} \mathbf{A}^{(n)},$$

  which means that the order of the n-mode products is irrelevant when modes n are all distinct.

- *For two products of $\mathcal{X} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ along n-mode, with $\mathbf{A} \in \mathbb{K}^{J_n \times I_n}$ and $\mathbf{B} \in \mathbb{K}^{K_n \times J_n}$, we have:*

$$\mathcal{Y} = \mathcal{X} \times_n \mathbf{A} \times {}_n\mathbf{B} = \mathcal{X} \times_n (\mathbf{BA}) \in \mathbb{K}^{I_1 \times \cdots \times I_{n-1} \times K_n \times I_{n+1} \times \cdots \times I_N}. \quad (6)$$

Properties of hypermatrix-matrix product

证明.

Defining $\mathcal{Z} = \mathcal{X} \times {}_n\mathbf{A}$, it is deduced that $\mathcal{Y} = \mathcal{Z} \times {}_n\mathbf{B}$. Then, using definition (3) of the n-mode product, we can write:

$$z_{i_1,\ldots,i_{n-1},j_n,i_{n+1},\ldots,i_N} = \sum_{i_n=1}^{I_n} a_{j_n,i_n} x_{i_1,\ldots,i_{n-1},i_n,i_{n+1},\ldots,i_N}, \tag{7}$$

$$y_{i_1,\ldots,i_{n-1},k_n,i_{n+1},\ldots,i_N} = \sum_{j_n=1}^{J_n} b_{k_n,j_n} z_{i_1,\ldots,i_{n-1},j_n,i_{n+1},\ldots,i_N}. \tag{8}$$

Using expression (7) in (8) and reversing the order of summations, we obtain:

$$y_{i_1,\ldots,i_{n-1},k_n,i_{n+1},\ldots,i_N} = \sum_{i_n=1}^{I_n} \left[\sum_{j_n=1}^{J_n} b_{k_n,j_n} a_{j_n,i_n} \right] x_{i_1,\ldots,i_{n-1},i_n,i_{n+1},\ldots,i_N}.$$

Properties of hypermatrix-matrix product

定理

The n-mode product satisfies the other two following properties:

- *For* $\mathcal{X} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$, $\mathbf{A}^{(n)} \in \mathbb{K}^{J_n \times I_n}$ *and* $\mathbf{B}^{(n)} \in \mathbb{K}^{K_n \times J_n}$, $n \in \langle N \rangle$, *property (6) can be generalized as follows:*

$$\mathcal{Y} = \mathcal{X} \times_{n=1}^{N} \mathbf{A}^{(n)} \times_{n=1}^{N} \mathbf{B}^{(n)} = \mathcal{X} \times_{n=1}^{N} \left(\mathbf{B}^{(n)} \mathbf{A}^{(n)}\right) \in \mathbb{K}^{K_1 \times \cdots \times K_N}.$$

- *If the factors* $\mathbf{A}^{(n)}$ *are full column rank, we have*

$$\mathcal{Y} = \mathcal{X} \times_{n=1}^{N} \mathbf{A}^{(n)} \Leftrightarrow \mathcal{X} = \mathcal{Y} \times_{n=1}^{N} \mathbf{A}^{(n)\dagger}$$

where $\mathbf{A}^{(n)\dagger}$ *denotes the Moore-Penrose pseudo-inverse of* $\mathbf{A}^{(n)}$.

NOTE: Considering the rows and the columns of a matrix as its 1-mode and 2-mode, respectively, the *n*-mode product can be used for matrix multiplication in such a way that: $\mathbf{A} \times_1 \mathbf{B} = \mathbf{B}\mathbf{A}$ and $\mathbf{A} \times_2 \mathbf{C} = \mathbf{A}\mathbf{C}^T$, which gives:

$$\mathbf{A} \times_1 \mathbf{B} \times_2 \mathbf{C} = \mathbf{B}\mathbf{A}\mathbf{C}^T \in \mathbb{K}^{K \times L}$$

with $\mathbf{A} \in \mathbb{K}^{I \times J}, \mathbf{B} \in \mathbb{K}^{K \times I}, \mathbf{C} \in \mathbb{K}^{L \times J}$.

Outer products

The outer product of two vectors $\mathbf{u} \in \mathbb{K}^I$ and $\mathbf{v} \in \mathbb{K}^J$, denoted $\mathbf{u} \circ \mathbf{v}$ and also called dyadic product of $\mathbf{u}$ with $\mathbf{v}$, gives a matrix $\mathbf{A} \in \mathbb{K}^{I \times J}$ of rank one, such that $a_{ij} = (\mathbf{u} \circ \mathbf{v})_{ij} = u_i v_j$, and therefore, $\mathbf{u} \circ \mathbf{v} = \mathbf{u}\mathbf{v}^T = [u_i v_j]$, with $i \in \langle I \rangle, j \in \langle J \rangle$. The outer product is to be compared with the inner product of $\mathbf{u}, \mathbf{v} \in \mathbb{K}^I$ which orovides a scalar: $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T\mathbf{v} = \sum_{i=1}^{I} u_i v_i$.

Outer products

The outer product of two vectors $\mathbf{u} \in \mathbb{K}^I$ and $\mathbf{v} \in \mathbb{K}^J$, denoted $\mathbf{u} \circ \mathbf{v}$ and also called dyadic product of $\mathbf{u}$ with $\mathbf{v}$, gives a matrix $\mathbf{A} \in \mathbb{K}^{I \times J}$ of rank one, such that $a_{ij} = (\mathbf{u} \circ \mathbf{v})_{ij} = u_i v_j$, and therefore, $\mathbf{u} \circ \mathbf{v} = \mathbf{u}\mathbf{v}^T = [u_i v_j]$, with $i \in \langle I \rangle, j \in \langle J \rangle$. The outer product is to be compared with the inner product of $\mathbf{u}, \mathbf{v} \in \mathbb{K}^I$ which orovides a scalar: $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T \mathbf{v} = \sum_{i=1}^{I} u_i v_i$.

EXAMPLE: For $I = 2, J = 3$, we have:

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \circ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & u_1 v_3 \\ u_2 v_1 & u_2 v_2 & u_2 v_3 \end{bmatrix}.$$

Outer products

- The outer product of three vectors $\mathbf{u} \in \mathbb{K}^{I}, \mathbf{v} \in \mathbb{K}^{J}$, and $\mathbf{w} \in \mathbb{K}^{K}$ gives a rank one third-order hypermatrix $\mathcal{A} \in \mathbb{K}^{I \times J \times K}$, such that $a_{ijk} = (\mathbf{u} \circ \mathbf{v} \circ \mathbf{w})_{ijk} = u_i v_j w_k$, and therefore, $\mathbf{u} \circ \mathbf{v} \circ \mathbf{w} = [u_i v_j w_k]$, with $i \in \langle I \rangle, j \in \langle J \rangle, k \in \langle K \rangle$.
- For the outer product of N vectors $\mathbf{u}^{(n)} \in \mathbb{K}^{I_n}$, denoted by $\underset{n=1}{\overset{N}{\circ}} \mathbf{u}^{(n)}$, a rank one hypermatrix $\mathcal{U} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$, of order N, is obtained such that:

$$\mathcal{U} = [u_{i_1, \cdots, i_N}] = \left[\prod_{n=1}^{N} u_{i_n}^{(n)} \right], i_n \in \langle I_n \rangle, n \in \langle N \rangle.$$

- When the vectors $\mathbf{u}^{(n)}$ are identical, that is, $\mathbf{u}^{(n)} = \mathbf{u}$, for $n \in \langle N \rangle$, their outer product $\underset{n=1}{\overset{N}{\circ}} \mathbf{u}^{(n)}$ is written as $\mathbf{u}^{\circ N}$.

Outer products

EXAMPLE: Consider the outer product defined as:

$$\mathcal{B} = \begin{bmatrix} i \\ 1 \end{bmatrix} \circ \begin{bmatrix} i \\ 1 \end{bmatrix} \circ \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} i \\ 1 \end{bmatrix}^{\circ 3} \in \mathbb{C}^{2\times 2\times 2},$$

with $i^2 = -1$. Ranging the elements b_{ijk} of the hypermatrix $\mathcal{B}$, with $i, j, k \in \langle 2 \rangle$, in a matrix defined as:

$$\mathbf{B} = \left[\begin{array}{cc|cc} b_{111} & b_{112} & b_{121} & b_{122} \\ b_{211} & b_{212} & b_{221} & b_{222} \end{array}\right] \in \mathbb{C}^{2\times 4}, \tag{9}$$

we have:

$$\mathbf{B} = \left[\begin{array}{cc|cc} -i & -1 & -1 & i \\ -1 & i & i & 1 \end{array}\right] \in \mathbb{C}^{2\times 4}.$$

Similarly, for the outer product:

$$\mathcal{C} = \begin{bmatrix} -i \\ 1 \end{bmatrix}^{\circ 3} \in \mathbb{C}^{2\times 2\times 2},$$

the matrix $\mathbf{C}$ corresponding to the arrangement (9) is given by:

$$\mathbf{C} = \left[\begin{array}{cc|cc} i & -1 & -1 & -i \\ -1 & -i & -i & 1 \end{array}\right] \in \mathbb{C}^{2\times 4}.$$

Matrices $\mathbf{B}$ and $\mathbf{C}$ above are called matrix unfoldings of the hypermatrices $\mathcal{B}$ and $\mathcal{C}$, respectively.

Outer products

The identity hypermatrix $\mathcal{I}_{N,I} = [\delta_{i_1,\cdots,i_N}]$ can be written using outer products of canonical basis vectors as:

$$\mathcal{I}_{N,I} = \sum_{i=1}^{I} \underbrace{\mathbf{e}_i^{(I)} \circ \cdots \circ \mathbf{e}_i^{(I)}}_{N \text{ terms}}$$

where $\mathbf{e}_i^{(I)} \circ \cdots \circ \mathbf{e}_i^{(I)}$ is the N th-order hypermatrix composed of 0's except one 1 on the diagonal, at position $i_1 = \cdots = i_N = i$, with $i \in \langle I \rangle$.

More generally, given $\mathbf{u} \in \mathbb{K}^K, \mathbf{A} \in \mathbb{K}^{I \times J}, \mathbf{B} \in \mathbb{K}^{K \times N}, \mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_M}$, and $\mathcal{B} \in \mathbb{K}^{J_1 \times \cdots \times J_N}$, we define the following outer products:

$$\begin{aligned}
\mathbf{A} \circ \mathbf{u} = \mathcal{C} \in \mathbb{K}^{I \times J \times K}, &\quad c_{ijk} = a_{ij} u_k \\
\mathbf{A} \circ \mathbf{B} = \mathcal{C} \in \mathbb{K}^{I \times J \times K \times N}, &\quad c_{ijkn} = a_{ij} b_{kn} \\
\mathcal{A} \circ \mathcal{B} = \mathcal{C} \in \mathbb{K}^{I_1 \times \cdots \times I_M \times J_1 \times \cdots \times J_N}, &\quad c_{i_1 \cdots i_M j_1 \cdots j_N} = a_{i_1 \cdots i_M} b_{j_1 \cdots j_N}.
\end{aligned}$$

Outer products

The vector spaces $\mathbb{K}^{I \times J}$ of matrices and $\mathbb{K}^{I_1 \times \cdots \times I_N}$ of hypermatrices of order N have the following respective canonical bases:

$$\mathbf{E}_{ij}^{(I \times J)} = \mathbf{e}_i^{(I)} \circ \mathbf{e}_j^{(J)}, \tag{10}$$

$$\mathcal{E}_{i_1 \cdots i_N}^{(I_1 \times \cdots \times I_N)} = \mathbf{e}_{i_1}^{(I_1)} \circ \cdots \circ \mathbf{e}_{i_N}^{(I_N)} = \mathop{\circ}_{n=1}^{N} \mathbf{e}_{i_n}^{(I_n)}, \tag{11}$$

with $i \in \langle I \rangle, j \in \langle J \rangle$, and $i_n \in \langle I_n \rangle$ for $n \in \langle N \rangle$. Matrix $\mathbf{E}_{ij}^{(I \times J)}$ and hypermatrix $\mathcal{E}_{i_1 \cdots i_N}^{(I_1 \times \cdots \times I_N)}$ contain one 1 at positions (i, j) and $(i_1, \cdots, i_N)$, respectively, and 0's elsewhere.

Hypermatrix associated to a multilinear form

Consider a multilinear form $\psi : \underset{n=1}{\overset{N}{\times}} E_n \to \mathbb{K}$, with $\dim(E_n) = I_n$, such that

$$\underset{n=1}{\overset{N}{\times}} E_n \ni \left(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}\right) \mapsto \psi\left(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}\right) \in \mathbb{K},$$

with the basis $\mathcal{B}^{(I_n)} = \left\{\mathbf{b}_1^{(I_n)}, \cdots, \mathbf{b}_{I_n}^{(I_n)}\right\}$ in E_n, $n \in \langle N \rangle$. Let us define the scalar coefficients equal to the images of these basis vectors by ψ :

$$b_{i_1,\cdots,i_N} = \psi\left(\mathbf{b}_{i_1}^{(I_1)}, \cdots, \mathbf{b}_{i_N}^{(I_N)}\right) \in \mathbb{K}, i_n \in \langle I_n \rangle, n \in \langle N \rangle,$$

and express $\psi\left(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}\right)$ in terms of the coordinates of the vectors $\mathbf{x}^{(n)}$, in the bases $\mathcal{B}^{(I_n)}$, that is, $\mathbf{x}^{(n)} = \sum_{i_n=1}^{I_n} c_{i_n}^{(n)} \mathbf{b}_{i_n}^{(I_n)}$.

Hypermatrix associated to a multilinear form

Using the multilinearity property of ψ and Einstein's convention, we get:

$$\begin{aligned}
\psi\left(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}\right) &= \psi\left(\sum_{i_1=1}^{I_1} c_{i_1}^{(1)} \mathbf{b}_{i_1}^{(I_1)}, \cdots, \sum_{i_N=1}^{I_N} c_{i_N}^{(N)} \mathbf{b}_{i_N}^{(I_N)}\right) \\
&= \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} c_{i_1}^{(1)} \cdots c_{i_N}^{(N)} \psi\left(\mathbf{b}_{i_1}^{(I_1)}, \cdots, \mathbf{b}_{i_N}^{(I_N)}\right) \\
&= c_{i_1}^{(1)} \cdots c_{i_N}^{(N)} \psi\left(\mathbf{b}_{i_1}^{(I_1)}, \cdots, \mathbf{b}_{i_N}^{(I_N)}\right) && (12) \\
&= b_{i_1,\cdots,i_N} c_{i_1}^{(1)} \cdots c_{i_N}^{(N)}. && (13)
\end{aligned}$$

In this equation, $\mathcal{B} = [b_{i_1,\cdots,i_N}] \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ is called the hypermatrix associated to the multilinear form ψ, and its elements $b_{i_1,\cdots,i_N}$ represent the components of ψ with respect to the N-linear terms $c_{i_1}^{(1)} \cdots c_{i_N}^{(N)}$. This equation can be interpreted in terms of homogeneous polynomial of degree N in the components $c_{i_n}^{(n)}$, $i_n \in \langle I_n \rangle$, of vectors $\mathbf{x}^{(n)}$. It is linear in every component of these vectors. The terms $b_{i_1}, \cdots, i_N$ are the coefficients of the polynomial.

Hypermatrix associated to a multilinear form

The N-linear terms are the elements of the rank-one hypermatrix $\mathcal{C} = \underset{n=1}{\overset{N}{\circ}} \mathbf{c}^{(n)} = [c_{i_1,\cdots,i_N}] = \left[c_{i_1}^{(1)} \cdots c_{i_N}^{(N)}\right]$, where $\mathbf{c}^{(n)} = \left[c_1^{(n)}, \cdots, c_{I_n}^{(n)}\right]^T$ is the coordinate vector of $\mathbf{x}^{(n)}$ in the basis $\mathcal{B}^{(I_n)}$. Therefore, (13) can also be viewed as the contraction of order N of the hypermatrices $\mathcal{B}$ and $\mathcal{C}$; in other words, their inner product:

$$\psi\left(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}\right) = \langle \mathcal{B}, \mathcal{C} \rangle.$$

FACT: When $E_n = \mathbb{K}^{I_n}$, for $n \in \langle N \rangle$, with the canonical bases $\mathbf{b}_{i_n}^{(I_n)} = \mathbf{e}_{i_n}^{(I_n)}, i_n \in \langle I_n \rangle$, the coordinates $c_{i_n}^{(n)}$ are the components of the vector $\mathbf{x}^{(n)}$ in the canonical basis of $\mathbb{K}^{I_n}$, that is, $\mathbf{x}^{(n)} = \sum_{i_n=1}^{I_n} x_{i_n}^{(n)} \mathbf{e}_{i_n}^{(I_n)}$, and equation (12) then becomes:

$$\psi\left(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}\right) = x_{i_1}^{(1)} \cdots x_{i_N}^{(N)} \psi\left(\mathbf{e}_{i_1}^{(I_1)}, \cdots, \mathbf{e}_{i_N}^{(I_N)}\right). \tag{14}$$

The term $\prod_{n=1}^{N} x_{i_n}^{(n)}$ can be interpreted as an element of the rank-one hypermatrix $\mathcal{X} = \underset{n=1}{\overset{N}{\circ}} \mathbf{x}^{(n)}$. Equation (14) expresses the multilinear form ψ in terms of N-tuples of the canonical bases transformed by ψ.

When $E_n = E, \forall n \in \langle N \rangle$, with $\dim(E) = J$, the multilinear form $\psi \in \mathcal{ML}_N(E, \mathbb{K})$ is symmetric if and only if for any permutation $\pi \in S_{\mathcal{N}}$, we have:

$$\psi\left(\mathbf{x}^{(\pi(1))}, \cdots, \mathbf{x}^{(\pi(N))}\right) = \psi\left(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}\right). \tag{15}$$

From (13), it is inferred that:

$$\psi\left(\mathbf{x}^{(\pi(1))}, \cdots, \mathbf{x}^{(\pi(N))}\right) = b_{i_{\pi(1)}, \cdots, i_{\pi(N)}} c_{i_{\pi(1)}}^{(\pi(1))} \cdots c_{i_{\pi(N)}}^{(\pi(N))}$$

and consequently, equation (15) implies that:

$$b_{i_{\pi(1)}, \cdots, i_{\pi(N)}} = b_{i_1, \cdots, i_N},$$

that is, the symmetry of the hypermatrix $\mathcal{B}$. This means that the elements of $\mathcal{B}$ are invariant with respect to $N!$ permutations $\pi(.)$ of indices $i_n \in \langle J \rangle, n \in \langle N \rangle$.

For example, for a bilinear form $\psi : E^2 \to \mathbb{K}$, with $\dim(E) = J$, such that:

$$E^2 \ni (\mathbf{x}, \mathbf{y}) \mapsto \psi(\mathbf{x}, \mathbf{y}) \in \mathbb{K},$$

the expansion of vectors $\mathbf{x} = \sum_{i=1}^{J} x_i \mathbf{u}_i$ and $\mathbf{y} = \sum_{j=1}^{J} y_j \mathbf{u}_j$ in the basis $\{\mathbf{u}\} = \{\mathbf{u}_1, \cdots, \mathbf{u}_J\}$ of E leads to:

$$\psi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{J} \sum_{j=1}^{J} x_i y_j \psi(\mathbf{u}_i, \mathbf{u}_j) = \sum_{i=1}^{J} \sum_{j=1}^{J} b_{ij} x_i y_j = \mathbf{x}_{\mathbf{u}}^T \mathbf{B}_{\mathbf{u}} \mathbf{y}_{\mathbf{u}}, \tag{16}$$

where $\mathbf{x}_{\mathbf{u}}$ and $\mathbf{y}_{\mathbf{u}}$ are the coordinate vectors of $\mathbf{x}$ and $\mathbf{y}$ in the basis $\{\mathbf{u}\}$, and the matrix $\mathbf{B}_{\mathbf{u}} = (b_{ij}) \in \mathbb{K}^{J \times J}$ associated with the bilinear form.

Equation (16) is a special case of equation (13) corresponding to $N = 2$. From equation (16), we can conclude that the symmetry assumption for ψ implies that $\mathbf{B}_{\mathbf{u}}$ is itself symmetric:

$$b_{ij} = b_{ji}, \forall i, j \in \langle J \rangle$$

$$\mathbf{B}_{\mathbf{u}}^T = \mathbf{B}_{\mathbf{u}}.$$

For a trilinear form $\psi : \mathbb{R}^J \times \mathbb{R}^J \times \mathbb{R}^J \to \mathbb{R}$, equation (13) becomes:

$$\mathbb{R}^J \times \mathbb{R}^J \times \mathbb{R}^J \ni (\mathbf{x}, \mathbf{y}, \mathbf{t}) \mapsto \psi(\mathbf{x}, \mathbf{y}, \mathbf{t}) = b_{ijk} x_i y_j t_k,$$

with:

$$\mathbf{x} = \sum_{i=1}^{J} x_i \mathbf{u}_i, \mathbf{y} = \sum_{j=1}^{J} y_j \mathbf{u}_j, \mathbf{t} = \sum_{k=1}^{J} t_k \mathbf{u}_k$$
$$b_{ijk} = \psi\left(\mathbf{u}_i, \mathbf{u}_j, \mathbf{u}_k\right).$$

The symmetry of the trilinear form ψ implies that of the hypermatrix $\mathcal{B}$, which gives the following equalities:

$$\psi(\mathbf{x}, \mathbf{y}, \mathbf{t}) = \psi(\mathbf{x}, \mathbf{t}, \mathbf{y}) = \psi(\mathbf{y}, \mathbf{x}, \mathbf{t}) = \psi(\mathbf{y}, \mathbf{t}, \mathbf{x}) = \psi(\mathbf{t}, \mathbf{x}, \mathbf{y}) = \psi(\mathbf{t}, \mathbf{y}, \mathbf{x})$$
$$b_{ijk} = b_{ikj} = b_{jik} = b_{jki} = b_{kij} = b_{kji}, \quad \forall i, j, k \in \langle J \rangle.$$

Similarly, for a quadrilinear form ψ such that:

$$\mathbb{R}^J \times \mathbb{R}^J \times \mathbb{R}^J \times \mathbb{R}^J \ni (\mathbf{x}, \mathbf{y}, \mathbf{t}, \mathbf{w}) \mapsto \psi(\mathbf{x}, \mathbf{y}, \mathbf{t}, \mathbf{w}) = b_{ijkl} x_i y_j t_k w_l,$$

the symmetry of ψ means that:

$$\psi(\mathbf{x}, \mathbf{y}, \mathbf{t}, \mathbf{w}) = \psi(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{t}) = \psi(\mathbf{x}, \mathbf{t}, \mathbf{y}, \mathbf{w}) = \cdots = \psi(\mathbf{w}, \mathbf{t}, \mathbf{y}, \mathbf{x}),$$

and the symmetry of the hypermatrix $\mathcal{B}$ then implies that the coefficient b_{ijkl} is not altered by the 24 permutations of the indices i, j, k and l, that is:

$$b_{ijkl} = b_{ijlk} = b_{ikjl} = b_{iklj} = \cdots = b_{lkji}, \quad \forall i, j, k, l \in \langle J \rangle.$$

FACT: The following observations can be made:

- The set of symmetric hypermatrices of order N form a subspace of the v.s. $\mathbb{K}^{I_1 \times \cdots \times I_N}$ of the hypermatrices.
- A rank-one N th-order hypermatrix $\mathcal{B} \in \mathbb{K}^{J \times J \times \cdots \times J}$ is symmetric if it can be written as the outer product of N identical vectors, that is:

$$\mathcal{B} = \underbrace{\mathbf{u} \circ \cdots \circ \mathbf{u}}_{N \text{ terms}} = \mathbf{u}^{\circ N}$$

- Partial symmetries can be defined with respect to a subset of indices. For example, the cubic hypermatrix $\mathcal{B} = \left[b_{ijk}\right]$ is said to be partially symmetric with respect to its first two indices if $b_{ijk} = b_{jik}$, $\forall i, j, k \in \langle J \rangle$. An hypermatrix can always be transformed into a symmetric hypermatrix with respect to a subset of indices. For example, $\mathcal{B} = \left[b_{ijk}\right]$ can be transformed into a symmetric hypermatrix $\mathcal{C}$ with respect to its first two indices, by defining $c_{ijk} = \frac{1}{2}\left(b_{ijk} + b_{jik}\right)$.

Multilinear maps and homogeneous polynomials

Consider a multilinear map $\psi : \underset{n=1}{\overset{N}{\times}} E_n \to F$, with $\dim(E_n) = I_n$ for $n \in \langle N \rangle$ and $\dim(F) = P$. Let $\mathcal{B}^{(I_n)} = \left\{\mathbf{b}_1^{(I_n)}, \cdots, \mathbf{b}_{I_n}^{(I_n)}\right\}$ be a basis in E_n, $n \in \langle N \rangle$, and $\mathcal{B}^{(P)} = \left\{\mathbf{b}_p^{(P)}, p \in \langle P \rangle\right\}$ a basis in F. Expanding $\psi(\mathbf{b}_{i_1}^{(I_1)}, \cdots, \mathbf{b}_{i_N}^{(I_N)})$ in the basis $\mathcal{B}^{(P)}$ of F gives:

$$\psi(\mathbf{b}_{i_1}^{(I_1)}, \cdots, \mathbf{b}_{i_N}^{(I_N)}) = \sum_{p=1}^{P} a_{i_1,\cdots,i_N,p}\mathbf{b}_p^{(P)} = a_{i_1,\cdots,i_N,p}\mathbf{b}_p^{(P)}$$

and (12) becomes:

$$\psi(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}) = c_{i_1}^{(1)} \cdots c_{i_N}^{(N)} a_{i_1,\cdots,i_N,p}\mathbf{b}_p^{(P)} = d_p\mathbf{b}_p^{(P)}.$$

This expansion is now expressed using $P\prod_{n=1}^{N} I_n$ components. The P coordinates $\{d_p, p \in \langle P \rangle\}$ of $\psi\left(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}\right)$ in the basis of F are P homogeneous polynomials of degree N in the components of vectors $\mathbf{x}^{(n)}$, and linear in the components of each vector. The coefficients of these polynomials are the coordinates $a_{i_1,\cdots,i_N,p}$ of $\psi(\mathbf{b}_{i_1}^{(I_1)}, \cdots, \mathbf{b}_{i_N}^{(I_N)})$ in the basis of F.
Defining the hypermatrix $\mathcal{A} = \left[a_{i_1,\cdots,i_N,p}\right]$ of order $N+1$ and dimensions $I_1 \times \cdots \times I_N \times P$, the coordinate d_p can be written as the contraction:

$$d_p = c_{i_1}^{(1)} \cdots c_{i_N}^{(N)} a_{i_1,\cdots,i_N,p} = \mathcal{A} \underset{1,2,\cdots,N}{\overset{1,2,\cdots,N}{\times}} \mathcal{C},$$

where $\mathcal{C} = \left[c_{i_1,\cdots,i_N}\right] = \left[c_{i_1}^{(1)} \cdots c_{i_N}^{(N)}\right]$ is defined as for a multilinear form.

Definitions

Just as a vector is defined as an element of a v.s., also called linear space, a tensor of order N can be defined as an element of a tensor space, that is, a tensor product of N v.s. The linearity of each v.s. induces the multilinearity (N-linearity) property of the tensor space. There are several ways to define the concept of tensor product. Here, we choose an approach based on the use of a multilinear map and the choice of a basis.

Definitions

Just as a vector is defined as an element of a v.s., also called linear space, a tensor of order N can be defined as an element of a tensor space, that is, a tensor product of N v.s. The linearity of each v.s. induces the multilinearity (N-linearity) property of the tensor space. There are several ways to define the concept of tensor product. Here, we choose an approach based on the use of a multilinear map and the choice of a basis.

The tensor product of N $\mathbb{K}$-vector spaces E_n of dimension I_n, with $n \in \langle N \rangle$, denoted $\underset{n=1}{\overset{N}{\otimes}} E_n$, can be defined as the image of a multilinear map ψ :

$$\underset{n=1}{\overset{N}{\times}} E_n \xrightarrow{\psi} \underset{n=1}{\overset{N}{\otimes}} E_n$$

satisfying the following condition: considering a basis $\mathcal{B}^{(I_n)} = \left\{ \mathbf{b}_1^{(I_n)}, \cdots, \mathbf{b}_{I_n}^{(I_n)} \right\}$ for E_n, with $n \in \langle N \rangle$, then the set:

$$\mathcal{B} = \left\{ \psi(\mathbf{b}_{i_1}^{(I_1)}, \cdots, \mathbf{b}_{i_N}^{(I_N)}) = \mathbf{b}_{i_1}^{(I_1)} \otimes \cdots \otimes \mathbf{b}_{i_N}^{(I_N)}, i_n \in \langle I_n \rangle, n \in \langle N \rangle \right\} \tag{17}$$

constitutes a basis for the tensor product space $\underset{n=1}{\overset{N}{\otimes}} E_n$.

An example

EXAMPLE: Given two $\mathbb{K}$-vector spaces E and F, of respective dimensions I and J, with the respective bases $\left\{\mathbf{b}_i^{(I)}, i \in \langle I\rangle\right\}$ and $\left\{\mathbf{b}_j^{(J)}, j \in \langle J\rangle\right\}$, then the IJ vectors $\mathbf{b}_i^{(I)} \otimes \mathbf{b}_j^{(J)}$ constitute a basis of $E \otimes F$.

Definitions

The tensor product $\overset{N}{\underset{n=1}{\otimes}} \mathbf{u}^{(n)}$ of N vectors $\mathbf{u}^{(n)} \in E_n$ represents an N-order tensor, of rank one, called elementary tensor, or pure tensor, or still decomposable tensor. The tensor product space $\overset{N}{\underset{n=1}{\otimes}} E_n$, also referred to as tensor space, consists of the set of linear combinations of N-order elementary tensors:

$$\overset{N}{\underset{n=1}{\otimes}} E_n = \mathrm{lc}\left\{ \overset{N}{\underset{n=1}{\otimes}} \mathbf{u}^{(n)}, \mathbf{u}^{(n)} \in E_n, n \in \langle N \rangle \right\}.$$

The tensor space $\overset{N}{\underset{n=1}{\otimes}} E_n$ is a v.s. of dimension:

$$\dim\left(\overset{N}{\underset{n=1}{\otimes}} E_n \right) = \prod_{n=1}^{N} \dim\left(E_n\right) = \prod_{n=1}^{N} I_n.$$

Multilinearity and associativity

The tensor product satisfies the multilinearity property, that is, for all $\mathbf{x}^{(n)}, \mathbf{y}^{(n)} \in E_n, n \in \langle N \rangle$, and all $\alpha \in \mathbb{K}$, we have:

$$\begin{aligned}\mathbf{x}^{(1)} \otimes \cdots \otimes \left(\alpha \mathbf{x}^{(n)} + \mathbf{y}^{(n)}\right) \otimes \cdots \otimes \mathbf{x}^{(N)} =& \alpha \left(\mathbf{x}^{(1)} \otimes \cdots \otimes x^{(n)} \otimes \cdots \otimes \mathbf{x}^{(N)}\right) + \\ & \left(\mathbf{x}^{(1)} \otimes \cdots \otimes \mathbf{y}^{(n)} \otimes \cdots \otimes \mathbf{x}^{(N)}\right).\end{aligned}$$

Multilinearity and associativity

The tensor product satisfies the multilinearity property, that is, for all $\mathbf{x}^{(n)}, \mathbf{y}^{(n)} \in E_n, n \in \langle N \rangle$, and all $\alpha \in \mathbb{K}$, we have:

$$\mathbf{x}^{(1)} \otimes \cdots \otimes \left(\alpha \mathbf{x}^{(n)} + \mathbf{y}^{(n)}\right) \otimes \cdots \otimes \mathbf{x}^{(N)} = \alpha \left(\mathbf{x}^{(1)} \otimes \cdots \otimes x^{(n)} \otimes \cdots \otimes \mathbf{x}^{(N)}\right) + \left(\mathbf{x}^{(1)} \otimes \cdots \otimes \mathbf{y}^{(n)} \otimes \cdots \otimes \mathbf{x}^{(N)}\right).$$

EXAMPLE: Consider the case $N = 2$ corresponding to bilinear maps. Let E and F be two $\mathbb{K}$-v.s. of respective dimensions I and J. Their tensor product $E \otimes F$ consists of elements of the form $\mathbf{u} \otimes \mathbf{v}$, with $\mathbf{u} \in E$ and $\mathbf{v} \in F$, and we have:

$$\forall \mathbf{u} \in E, \forall \left(\mathbf{v}_1, \mathbf{v}_2\right) \in F^2, \mathbf{u} \otimes \left(\mathbf{v}_1 + \mathbf{v}_2\right) = \mathbf{u} \otimes \mathbf{v}_1 + \mathbf{u} \otimes \mathbf{v}_2$$

$$\forall \left(\mathbf{u}_1, \mathbf{u}_2\right) \in E^2, \forall \mathbf{v} \in F, \left(\mathbf{u}_1 + \mathbf{u}_2\right) \otimes \mathbf{v} = \mathbf{u}_1 \otimes \mathbf{v} + \mathbf{u}_2 \otimes \mathbf{v}$$

$$\forall (\mathbf{u}, \mathbf{v}) \in E \times F, \forall \lambda \in \mathbb{K}, \lambda(\mathbf{u} \otimes \mathbf{v}) = \lambda \mathbf{u} \otimes \mathbf{v} = \mathbf{u} \otimes \lambda \mathbf{v}.$$

Multilinearity and associativity

The tensor product satisfies the multilinearity property, that is, for all $\mathbf{x}^{(n)}, \mathbf{y}^{(n)} \in E_n, n \in \langle N \rangle$, and all $\alpha \in \mathbb{K}$, we have:

$$\mathbf{x}^{(1)} \otimes \cdots \otimes \left(\alpha \mathbf{x}^{(n)} + \mathbf{y}^{(n)}\right) \otimes \cdots \otimes \mathbf{x}^{(N)} = \alpha \left(\mathbf{x}^{(1)} \otimes \cdots \otimes x^{(n)} \otimes \cdots \otimes \mathbf{x}^{(N)}\right) + \left(\mathbf{x}^{(1)} \otimes \cdots \otimes \mathbf{y}^{(n)} \otimes \cdots \otimes \mathbf{x}^{(N)}\right).$$

EXAMPLE: Consider the case $N = 2$ corresponding to bilinear maps. Let E and F be two $\mathbb{K}$-v.s. of respective dimensions I and J. Their tensor product $E \otimes F$ consists of elements of the form $\mathbf{u} \otimes \mathbf{v}$, with $\mathbf{u} \in E$ and $\mathbf{v} \in F$, and we have:

$$\forall \mathbf{u} \in E, \forall (\mathbf{v}_1, \mathbf{v}_2) \in F^2, \mathbf{u} \otimes (\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{u} \otimes \mathbf{v}_1 + \mathbf{u} \otimes \mathbf{v}_2$$
$$\forall (\mathbf{u}_1, \mathbf{u}_2) \in E^2, \forall \mathbf{v} \in F, (\mathbf{u}_1 + \mathbf{u}_2) \otimes \mathbf{v} = \mathbf{u}_1 \otimes \mathbf{v} + \mathbf{u}_2 \otimes \mathbf{v}$$
$$\forall (\mathbf{u}, \mathbf{v}) \in E \times F, \forall \lambda \in \mathbb{K}, \lambda(\mathbf{u} \otimes \mathbf{v}) = \lambda \mathbf{u} \otimes \mathbf{v} = \mathbf{u} \otimes \lambda \mathbf{v}.$$

The tensor product is associative. Thus, for three $\mathbb{K}$-v.s. E, F, G, we have:

$$(E \otimes F) \otimes G = E \otimes (F \otimes G) = E \otimes F \otimes G.$$

This property directly results from the definition of tensor product and from the property related to the basis of a Cartesian product of v.s..

Tensors and coordinate hypermatrices

A tensor $\mathcal{X} \in \bigotimes_{n=1}^{N} E_n$ is written in the bases $\mathcal{B}^{(I_n)} = \left\{ \mathbf{b}_1^{(I_n)}, \cdots, \mathbf{b}_{I_n}^{(I_n)} \right\}$ of E_n with $n \in \langle N \rangle$, as:

$$
\begin{aligned}
\mathcal{X} &= \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} a_{i_1,\cdots,i_N} \mathbf{b}_{i_1}^{(I_1)} \otimes \cdots \otimes \mathbf{b}_{i_N}^{(I_N)} \\
&= a_{i_1,\cdots,i_N} \bigotimes_{n=1}^{N} \mathbf{b}_{i_n}^{(I_n)} \text{(with Einstein's convention)}. \qquad (18)
\end{aligned}
$$

The coefficients $a_{i_1,\cdots,i_N}$ are the coordinates of $\mathcal{X}$ in the basis $\mathcal{B}$ defined in (17). These coordinates can be viewed as the elements of a hypermatrix $\mathcal{A} = [a_{i_1,\cdots,i_N}]$ of the space $\mathbb{K}^{I_1 \times \cdots \times I_N}$, that is, a multidimensional array of numbers accessible by way of N indices, each index i_n being associated with a mode of the tensor.

Tensors and coordinate hypermatrices

It should be noted that for every set of bases $\left\{\mathcal{B}^{(I_1)}, \cdots, \mathcal{B}^{(I_N)}\right\}$ corresponds a different coordinate hypermatrix.

It is important to point out the distinction between tensors and hypermatrices. The first ones express mathematical objects in a given basis of the tensor space under consideration, while the second ones correspond to arrays of numbers resulting in practice of measurements with units set during acquisition.

In summary, formally speaking, a tensor $\mathcal{X}$ can be associated with a set $\left\{\mathcal{B}^{(I_n)}, n \in \langle N \rangle, \mathcal{A}\right\}$ comprising the bases of the v.s. E_n and the coordinate hypermatrix in these bases.

Canonical writing of tensors

If $E_n = \mathbb{K}^{I_n}$, by choosing the canonical basis $\{\mathbf{e}_1^{(I_n)}, \cdots, \mathbf{e}_{I_n}^{(I_n)}\}$, then the tensor $\overset{N}{\underset{n=1}{\otimes}} \mathbf{e}_{i_n}^{(I_n)}$ is an element of the canonical basis of the tensor space $\overset{N}{\underset{n=1}{\otimes}} \mathbb{K}^{I_n}$. It should be noted that in the tensor product $\overset{N}{\underset{n=1}{\otimes}} \mathbf{e}_{i_n}^{(I_n)}$, the symbol $\otimes$ can be replaced by the outer product, as defined by (11).

Canonical writing of tensors

If $E_n = \mathbb{K}^{I_n}$, by choosing the canonical basis $\{\mathbf{e}_1^{(I_n)}, \cdots, \mathbf{e}_{I_n}^{(I_n)}\}$, then the tensor $\underset{n=1}{\overset{N}{\otimes}} \mathbf{e}_{i_n}^{(I_n)}$ is an element of the canonical basis of the tensor space $\underset{n=1}{\overset{N}{\otimes}} \mathbb{K}^{I_n}$. It should be noted that in the tensor product $\underset{n=1}{\overset{N}{\otimes}} \mathbf{e}_{i_n}^{(I_n)}$, the symbol $\otimes$ can be replaced by the outer product, as defined by (11).

Any tensor $\mathcal{X}$ of the tensor space $\underset{n=1}{\overset{N}{\otimes}} \mathbb{K}^{I_n}$ is then written as a linear combination of these basis tensors, namely:

$$\begin{aligned}\mathcal{X} &= \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} x_{i_1,\cdots,i_N} \underset{n=1}{\overset{N}{\circ}} \mathbf{e}_{i_n}^{(I_n)} \\ &= x_{i_1,\cdots,i_N} \mathcal{E}_{i_1\cdots i_N}^{(I_1\times\cdots\times I_N)},\end{aligned}$$

where the coefficients $x_{i_1,\cdots,i_N}$ are the coordinates of $\mathcal{X}$ in the canonical basis, the last equality resulting from the use of Einstein's convention.

Canonical writing of tensors

It should be noted that the coordinates define a hypermatrix $\mathcal{X}$ which will be generally denoted as the tensor itself. The following table summarizes the expansions of tensors of orders $1, 2, 3$, and N, in their respective canonical basis, with Einstein's convention.

Tensors	Spaces	Canonical expansions	Dimensions
Vectors	$\mathbf{x} \in \mathbb{K}^{I}$	$\mathbf{x} = x_i \mathbf{e}_i^{(I)}$	I
Matrices	$\mathbf{X} \in \mathbb{K}^{I \times J}$	$\mathbf{X} = x_{ij} \mathbf{e}_i^{(I)} \circ \mathbf{e}_j^{(J)}$	$I \times J$
Third-order tensors	$\mathcal{X} \in \mathbb{K}^{I \times J \times K}$	$\mathcal{X} = x_{ijk} \mathbf{e}_i^{(I)} \circ \mathbf{e}_j^{(J)} \circ \mathbf{e}_k^{(K)}$	$I \times J \times K$
N-order tensors	$\mathcal{X} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$	$\mathcal{X} = x_{i_1}, \cdots, i_N \overset{N}{\underset{n=1}{\circ}} \mathbf{e}_{i_n}^{(I_n)}$	$I_1 \times \cdots \times I_N$

Table: Canonical expansions of tensors

Canonical writing of tensors

For example, in the particular case $N = 2$, the tensor space $\bigotimes_{n=1}^{2} \mathbb{K}^{I_n} = \mathbb{K}^{I_1} \otimes \mathbb{K}^{I_2} = \mathbb{K}^{I_1 \times I_2}$ is the vector space of matrices of dimensions $I_1 \times I_2$, and in the canonical basis (10), $\mathbf{X} \in \mathbb{K}^{I_1 \times I_2}$ is written as:

$$\mathbf{X} = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} x_{i_1 i_2} \mathbf{e}_{i_1}^{(I_1)} \circ \mathbf{e}_{i_2}^{(I_2)} = x_{i_1 i_2} \mathbf{E}_{i_1 i_2}^{(I_1 \times I_2)}.$$

Canonical writing of tensors

For example, in the particular case $N = 2$, the tensor space $\bigotimes_{n=1}^{2} \mathbb{K}^{I_n} = \mathbb{K}^{I_1} \otimes \mathbb{K}^{I_2} = \mathbb{K}^{I_1 \times I_2}$ is the vector space of matrices of dimensions $I_1 \times I_2$, and in the canonical basis (10), $\mathbf{X} \in \mathbb{K}^{I_1 \times I_2}$ is written as:

$$\mathbf{X} = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} x_{i_1 i_2} \mathbf{e}_{i_1}^{(I_1)} \circ \mathbf{e}_{i_2}^{(I_2)} = x_{i_1 i_2} \mathbf{E}_{i_1 i_2}^{(I_1 \times I_2)}.$$

EXAMPLE: For $I_1 = I_2 = 2$, we have:

$$\begin{aligned} \mathbf{X} &= x_{11} \mathbf{E}_{11}^{(2\times 2)} + x_{12} \mathbf{E}_{12}^{(2\times 2)} + x_{21} \mathbf{E}_{21}^{(2\times 2)} + x_{22} \mathbf{E}_{22}^{(2\times 2)} \\ &= x_{11} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + x_{12} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + x_{21} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + x_{22} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \end{aligned}$$

with the following coordinate matrix in the canonical basis:

$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$. As already mentioned, the same notation $(\mathbf{X})$ is employed to designate the element of the matrix space $\mathbb{K}^{2\times 2}$ and its coordinate matrix in the canonical basis.

Canonical writing of tensors

EXAMPLE: Let a tensor $\mathcal{X} \in \mathbb{K}^{I \times J \times K}$ with $I = J = K = 2$, such that:

$$\begin{aligned}\mathcal{X} &= x_{111}\mathbf{e}_1^{(I)} \otimes \mathbf{e}_1^{(J)} \otimes \mathbf{e}_1^{(K)} + x_{212}\mathbf{e}_2^{(I)} \otimes \mathbf{e}_1^{(J)} \otimes \mathbf{e}_2^{(K)} \\ &\quad + x_{221}\mathbf{e}_2^{(I)} \otimes \mathbf{e}_2^{(J)} \otimes \mathbf{e}_1^{(K)} + x_{222}\mathbf{e}_2^{(I)} \otimes \mathbf{e}_2^{(J)} \otimes \mathbf{e}_2^{(K)} \\ &= x_{111}\mathcal{E}_{111}^{(2\times2\times2)} + x_{212}\mathcal{E}_{212}^{(2\times2\times2)} + x_{221}\mathcal{E}_{221}^{(2\times2\times2)} + x_{222}\mathcal{E}_{222}^{(2\times2\times2)}. \end{aligned} \quad (19)$$

The coordinate hypermatrix in the canonical basis is given by:

$$\begin{aligned}\mathcal{X} = \begin{bmatrix}1\\0\end{bmatrix} \circ \begin{bmatrix}1\\0\end{bmatrix} \circ \begin{bmatrix}x_{111}\\0\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix} \circ \begin{bmatrix}1\\0\end{bmatrix} \circ \begin{bmatrix}0\\x_{212}\end{bmatrix} \\ + \begin{bmatrix}0\\1\end{bmatrix} \circ \begin{bmatrix}0\\1\end{bmatrix} \circ \begin{bmatrix}x_{221}\\x_{222}\end{bmatrix}\end{aligned}$$

It should be noted that, according to (19), four coordinates of the tensor $\mathcal{X}$ are zero.

Expansion of the tensor product of N vectors

Consider a rank-one tensor $\mathcal{X}$ resulting from the tensor product of N vectors $\mathbf{x}^{(n)} \in E_n$, expanded in the bases $\mathcal{B}^{(I_n)} = \{\mathbf{b}_1^{(I_n)}, \cdots, \mathbf{b}_{I_n}^{(I_n)}\}$, for $n \in \langle N \rangle$, that is, $\mathbf{x}^{(n)} = \sum_{i_n=1}^{I_n} c_{i_n}^{(n)} \mathbf{b}_{i_n}^{(I_n)}$. Using the multilinearity property of the tensor product and Einstein's convention, we get:

$$\begin{aligned} \bigotimes_{n=1}^{N} \mathbf{x}^{(n)} &= \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} c_{i_1}^{(1)} \cdots c_{i_N}^{(N)} \bigotimes_{n=1}^{N} \mathbf{b}_{i_n}^{(I_n)} \\ &= c_{i_1}^{(1)} \cdots c_{i_N}^{(N)} \bigotimes_{n=1}^{N} \mathbf{b}_{i_n}^{(I_n)}. \end{aligned} \tag{20}$$

The coordinate hypermatrix $\mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ is then defined as $[a_{i_1, \cdots, i_N}] = [c_{i_1}^{(1)} \cdots c_{i_N}^{(N)}]$, with $i_n \in \langle I_n \rangle$, $n \in \langle N \rangle$.

Expansion of the tensor product of N vectors

Consider a rank-one tensor $\mathcal{X}$ resulting from the tensor product of N vectors $\mathbf{x}^{(n)} \in E_n$, expanded in the bases $\mathcal{B}^{(I_n)} = \{\mathbf{b}_1^{(I_n)}, \cdots, \mathbf{b}_{I_n}^{(I_n)}\}$, for $n \in \langle N \rangle$, that is, $\mathbf{x}^{(n)} = \sum_{i_n=1}^{I_n} c_{i_n}^{(n)} \mathbf{b}_{i_n}^{(I_n)}$. Using the multilinearity property of the tensor product and Einstein's convention, we get:

$$\begin{aligned} \underset{n=1}{\overset{N}{\otimes}} \mathbf{x}^{(n)} &= \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} c_{i_1}^{(1)} \cdots c_{i_N}^{(N)} \underset{n=1}{\overset{N}{\otimes}} \mathbf{b}_{i_n}^{(I_n)} \\ &= c_{i_1}^{(1)} \cdots c_{i_N}^{(N)} \underset{n=1}{\overset{N}{\otimes}} \mathbf{b}_{i_n}^{(I_n)}. \end{aligned} \tag{20}$$

The coordinate hypermatrix $\mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ is then defined as $[a_{i_1, \cdots, i_N}] = [c_{i_1}^{(1)} \cdots c_{i_N}^{(N)}]$, with $i_n \in \langle I_n \rangle, n \in \langle N \rangle$. By comparing (20) with (18), we can conclude that if any N th-order tensor is of the form (18), the decomposition of the coefficients $a_{i_1, \cdots, i_N}$ of the coordinate hypermatrix $\mathcal{A}$ in the form of products $c_{i_1}^{(1)} \cdots c_{i_N}^{(N)}$ is characteristic of a tensor of rank one. This hypermatrix is equal to the outer product of the N coordinate vectors $\mathbf{c}^{(n)} = [c_1^{(n)}, \cdots, c_{I_n}^{(n)}]^T$, that is, $\mathcal{A} = [c_{i_1}^{(1)} \cdots c_{i_N}^{(N)}] = \underset{n=1}{\overset{N}{\circ}} \mathbf{c}^{(n)}$, is a rank-one hypermatrix.

Expansion of the tensor product of N vectors

FACT: In the case where $E_n = \mathbb{K}^{I_n}$ with $\mathbf{x}^{(n)}$ expressed in the canonical basis of E_n, namely, $\mathbf{x}^{(n)} = \sum_{i_n=1}^{I_n} x_{i_n}^{(n)} \mathbf{e}_{i_n}^{(I_n)}$, equation (20) becomes:

$$\underset{n=1}{\overset{N}{\otimes}} \mathbf{x}^{(n)} = x_{i_1}^{(1)} \cdots x_{i_N}^{(N)} \underset{n=1}{\overset{N}{\circ}} \mathbf{e}_{i_n}^{(I_n)} \tag{21}$$

$$= x_{i_1}^{(1)} \cdots x_{i_N}^{(N)} \mathcal{E}_{i_1 \cdots i_N}^{(I_1 \times \cdots \times I_N)}, \tag{22}$$

which represents the expansion of the rank-one tensor $\underset{n=1}{\overset{N}{\otimes}} \mathbf{x}^{(n)} = \underset{n=1}{\overset{N}{\circ}} \mathbf{x}^{(n)}$ in the canonical basis $\{\mathcal{E}_{i_1 \cdots i_N}^{(I_1 \times \cdots \times I_N)}\}$.

Expansion of the tensor product of N vectors

FACT: Consider the case $N = 2$, with $E_1 = E_2 = \mathbb{K}^3$. Let $\mathbf{x} = \sum_{i=1}^{3} x_i \mathbf{e}_i^{(3)}$ and $\mathbf{y} = \sum_{j=1}^{3} y_j \mathbf{e}_j^{(3)}$. The tensor product $\mathbf{x} \otimes \mathbf{y}$ defines a matrix of rank one:

$$\mathbf{x} \otimes \mathbf{y} = \sum_{i,j=1}^{3} x_i y_j \mathbf{e}_i^{(3)} \otimes \mathbf{e}_j^{(3)} = x_i y_j \mathbf{E}_{ij}^{(3\times 3)},$$

and the coordinate matrix in the canonical basis is given by:

$$\begin{bmatrix} x_1 y_1 & x_1 y_2 & x_1 y_3 \\ x_2 y_1 & x_2 y_2 & x_2 y_3 \\ x_3 y_1 & x_3 y_2 & x_3 y_3 \end{bmatrix} \in \mathbb{K}^{3\times 3}.$$

Symmetry property of the tensor product

When the N v.s. are identical $(E_n = E)$, for $n \in \langle N \rangle$, with $\dim(E) = I$, the elements of the tensor product $\underbrace{E \otimes E \otimes \cdots \otimes E}_{N \text{ terms}}$, which is written as $E^{\otimes N}$, are hypercubic tensors of dimensions $I \times \cdots \times I$. Since for a symmetric multilinear form, a hypercubic tensor is symmetric if its coordinate hypermatrix $\mathcal{A}$ defined in (18) is symmetric, which means that:

$$a_{i_{\pi(1)},\cdots,i_{\pi(N)}} = a_{i_1,\cdots,i_N}, \tag{23}$$

for any permutation $\pi \in \mathcal{S}_{\mathcal{N}}$.

Symmetry property of the tensor product

When the N v.s. are identical ($E_n = E$), for $n \in \langle N \rangle$, with $\dim(E) = I$, the elements of the tensor product $\underbrace{E \otimes E \otimes \cdots \otimes E}_{N \text{ terms}}$, which is written as $E^{\otimes N}$, are hypercubic tensors of dimensions $I \times \cdots \times I$. Since for a symmetric multilinear form, a hypercubic tensor is symmetric if its coordinate hypermatrix $\mathcal{A}$ defined in (18) is symmetric, which means that:

$$a_{i_{\pi(1)},\cdots,i_{\pi(N)}} = a_{i_1,\cdots,i_N}, \tag{23}$$

for any permutation $\pi \in \mathcal{S}_{\mathcal{N}}$.

EXAMPLE: Consider the tensor $\mathcal{X} \in \mathbb{K}^{I \times J \times K}$ with $I = J = K = 2$, expanded in the canonical basis as:

$$\mathcal{X} = \mathcal{E}_{221}^{(2\times2\times2)} + \mathcal{E}_{122}^{(2\times2\times2)} + \mathcal{E}_{212}^{(2\times2\times2)} - \mathcal{E}_{111}^{(2\times2\times2)}$$

The coordinate hypermatrix being such that:

$$x_{121} = x_{211} = x_{112} = x_{222} = 0, x_{221} = x_{122} = x_{212} = 1 \text{ and } x_{111} = -1,$$

it satisfies constraints (23), which means that $\mathcal{X}$ is a symmetric tensor.

定理

For any multilinear map $\varphi \in \mathcal{ML}(E_1, \cdots, E_N; F)$: $\mathop{\times}\limits_{n=1}^{N} E_n \xrightarrow{\varphi} F$, *there exists a unique linear map* $f \in \mathcal{L}(\bigotimes\limits_{n=1}^{N} E_{n=1}; F) : \bigotimes\limits_{n=1}^{N} E_n \to F$, *such that the multilinear map* φ *can be broken down as:*

$$\mathop{\times}\limits_{n=1}^{N} E_n \xrightarrow{\psi} \bigotimes\limits_{n=1}^{N} E_{n=1} \xrightarrow{f} F$$

that is, $\varphi = f \circ \psi$, *which is equivalent to the following diagram, called commutative diagram:*

$$\begin{tikzcd}
\mathop{\times}\limits_{n=1}^{N} E_n \arrow[rr, "\varphi"] \arrow[dr, "\psi"'] & & F \\
& \bigotimes\limits_{n=1}^{N} E_n \arrow[ur, "f"'] &
\end{tikzcd}$$

So, for $\mathbf{u}_n \in E_n$, *with* $n \in \langle N \rangle$, *we have:*

$$\varphi(\mathbf{u}_1, \cdots, \mathbf{u}_N) = f \circ \psi(\mathbf{u}_1, \cdots, \mathbf{u}_N) = f(\psi(\mathbf{u}_1, \cdots, \mathbf{u}_N)) = f(\bigotimes\limits_{n=1}^{N} \mathbf{u}_n).$$

Universal property of the tensor product

It should be noted that since a vector in the tensor space $\bigotimes_{n=1}^{N} E_n$ is expressed as a linear combination of elementary tensors, the linear function f can be determined uniquely by transformation of the vectors of a basis $\mathcal{B}$ such as defined in (17), leading to the equation

$\varphi\left(\mathbf{b}_{i_1}^{(I_1)}, \cdots, \mathbf{b}_{i_N}^{(I_N)}\right) = f\left(\mathbf{b}_{i_1}^{(I_1)} \otimes \cdots \otimes \mathbf{b}_{i_N}^{(I_N)}\right)$, for all $\mathbf{b}_{i_n}^{(I_n)} \in \mathcal{B}^{(I_n)}, n \in \langle N \rangle$.

Universal property of the tensor product

It should be noted that since a vector in the tensor space $\underset{n=1}{\overset{N}{\otimes}} E_n$ is expressed as a linear combination of elementary tensors, the linear function f can be determined uniquely by transformation of the vectors of a basis $\mathcal{B}$ such as defined in (17), leading to the equation
$\varphi\left(\mathbf{b}_{i_1}^{(I_1)}, \cdots, \mathbf{b}_{i_N}^{(I_N)}\right) = f\left(\mathbf{b}_{i_1}^{(I_1)} \otimes \cdots \otimes \mathbf{b}_{i_N}^{(I_N)}\right)$, for all $\mathbf{b}_{i_n}^{(I_n)} \in \mathcal{B}^{(I_n)}, n \in \langle N \rangle$.

This universal property reflects the fact that there exists an isomorphism between the v.s. $\mathcal{ML}(E_1, \cdots, E_N; F)$ of multilinear maps over the Cartesian product $\underset{n=1}{\overset{N}{\times}} E_n$ and the v.s. $\mathcal{L}\left(\underset{n=1}{\overset{N}{\otimes}} E_n; F\right)$ of the linear maps built over the tensor product $\underset{n=1}{\overset{N}{\otimes}} E_n$. This property is often used to define the tensor product of two v.s., namely, when $N = 2$ corresponding to bilinear maps. The tensor product of N v.s., with $N > 2$, then constitutes a simple extension to N-linear maps.

Universal property of the tensor product

FACT (Universal property when $E_n = \mathbb{K}^{I_n}$): When $E_n = \mathbb{K}^{I_n}$, for $n \in \langle N \rangle$, the commutative diagram becomes:

$$
\begin{tikzcd}
\bigtimes_{n=1}^{N} \mathbb{K}^{I_n} \arrow[rr, "\varphi"] \arrow[dr, "\psi"'] & & \mathbb{K}^{I_1 \times \cdots \times I_n} \\
& \bigotimes_{n=1}^{N} \mathbb{K}^{I_n} \arrow[ur, "f"'] &
\end{tikzcd}
$$

Universal property of the tensor product

In this diagram, the multilinear map φ transforms the N-tuple $(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}) \in \bigtimes_{n=1}^{N} \mathbb{K}^{I_n}$ into the hypermatrix $\underset{n=1}{\overset{N}{\circ}} \mathbf{x}^{(n)} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$, while the map ψ transforms this N-tuple into an elementary tensor $\bigotimes_{n=1}^{N} \mathbf{x}^{(n)}$ of the tensor space $\bigotimes_{n=1}^{N} \mathbb{K}^{I_n}$, and f assigns to this tensor its coordinate hypermatrix:

$$\varphi(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}) = \underset{n=1}{\overset{N}{\circ}} \mathbf{x}^{(n)} = [x_{i_1}^{(1)} \cdots x_{i_N}^{(N)}],$$

$$\psi(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}) = \bigotimes_{n=1}^{N} \mathbf{x}^{(n)},$$

$$f(\bigotimes_{n=1}^{N} \mathbf{x}^{(n)}) = [x_{i_1}^{(1)} \cdots x_{i_N}^{(N)}].$$

Universal property of the tensor product

Expanding each vector $\mathbf{x}^{(n)}$ in the canonical basis of $\mathbb{K}^{I_n}$, that is, $\mathbf{x}^{(n)} = \sum_{i_n=1}^{I_n} x_{i_n}^{(n)} \mathbf{e}_{i_n}^{(I_n)}$, the map φ breaks down into:

$$\begin{aligned}
\varphi\left(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}\right) &= f \circ \psi\left(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}\right) = f\left(\psi\left(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(N)}\right)\right) \\
&= f\left(\overset{N}{\otimes} \mathbf{x}^{(n)}\right) = f\left(x_{i_1}^{(1)} \cdots x_{i_N}^{(N)} \mathcal{E}_{i_1 \cdots i_N}^{(I_1 \times \cdots \times I_N)}\right) \text{ according to } (22) \\
&= \left[x_{i_1}^{(1)} \cdots x_{i_N}^{(N)}\right] \text{ by definition of } f,
\end{aligned}$$

that is, the coordinate hypermatrix of the tensor $\overset{N}{\underset{n=1}{\otimes}} \mathbf{x}^{(n)}$ in the canonical basis. This coordinate hypermatrix being also given by $\overset{N}{\underset{n=1}{\circ}} \mathbf{x}^{(n)}$, the notations $\overset{N}{\underset{n=1}{\otimes}} \mathbf{x}^{(n)}$ and $\overset{N}{\underset{n=1}{\otimes}} \mathbf{x}^{(n)}$ will be used indistinctly, which amount to interchanging the tensor and its coordinate hypermatrix, as previously discussed for vectors and matrices. The tensor product of vectors is then equivalent to the outer product of these vectors. We have:

$$\dim\left(\overset{N}{\underset{n=1}{\times}} \mathbb{K}^{I_n}\right) = \dim\left(\overset{N}{\underset{n=1}{\otimes}} \mathbb{K}^{I_n}\right) = \dim\left(\mathbb{K}^{I_1 \times \cdots \times I_N}\right) = \prod_{n=1}^{N} I_n.$$

Change of basis formula

Let us consider the tensor $\mathcal{X} \in \bigotimes_{n=1}^{N} E_n$ defined in (18), and changes of basis in the v.s. E_n, for $n \in \langle N \rangle$, such that:

$$\mathbf{b}_{i_n}^{(I_n)} = \sum_{j_n=1}^{I_n} p_{i_n,j_n}^{(n)} \mathbf{b}_{j_n}^{(I_n)}, \, i_n \in \langle I_n \rangle. \tag{24}$$

By replacing the basis vectors $\mathrm{b}_{i_\alpha}^{(I_n)}$ by their expressions (24) in (18), and using the multilinearity property of the tensor product, we get:

$$\begin{aligned}
\mathcal{X} &= \sum_{i_1,\cdots,i_N=1}^{I_1,\cdots,I_N} a_{i_1,\cdots,i_N} \bigotimes_{n=1}^{N} \mathbf{b}_{i_n}^{(I_n)} = a_{i_1,\cdots,i_N} \bigotimes_{n=1}^{N} \mathbf{b}_{i_n}^{(I_n)} \\
&= \sum_{i_1,\cdots,i_N=1}^{I_1,\cdots,I_N} a_{i_1,\cdots,i_N}, \bigotimes_{n=1}^{N} \Big(\sum_{j_n=1}^{I_n} p_{i_n,j_n}^{(n)} \mathbf{b}_{j_n}^{\prime(I_n)} \Big) \\
&= \sum_{i_1,\cdots,i_N=1}^{I_1,\cdots,I_N} a_{i_1,\cdots,i_N} \sum_{j_1,\cdots,j_N=1}^{I_1,\cdots,I_N} p_{i_1,j_1}^{(1)} \cdots p_{i_N,j_N}^{(N)} \bigotimes_{n=1}^{N} \mathbf{b}_{j_n}^{\prime(I_n)} \\
&= \sum_{j_1,\cdots,j_N=1}^{I_1,\cdots,I_N} a'_{j_2,\cdots,j_N} \bigotimes_{n=1}^{N} \mathbf{b}_{j_n}^{\prime(I_n)} = a'_{j_1,\cdots,j_N} \bigotimes_{n=1}^{N} \mathbf{b}_{n=1}^{\prime(I_n)},
\end{aligned} \tag{25}$$

Change of basis formula

with:

$$a'_{j_1,\cdots,j_N} = \sum_{i_1,\ldots,i_N=1}^{I_1,\cdots,I_N} a_{i_1,\cdots,i_N} p^{(1)}_{i_1,j_1} \cdots p^{(N)}_{i_N,j_N}$$
$$= a_{i_1,\cdots,i_N} p^{(1)}_{i_1,j_1} \ldots p^{(N)}_{i_N,j_N} \text{ (with Einstein's convention).} \tag{26}$$

The hypermatrix $\mathcal{A}' = [a_{j_1,\cdots,j_N}] \in \mathbf{K}^{I_1\times\cdots\times I_N}$ contains the coordinates of the tensor $\mathcal{X}$ in the new bases $\mathcal{B}'^{(I_n)} = \{\mathbf{b}_1^{(I_n)}, \cdots, \mathbf{b}_{I_n}'^{(I_n)}\}$, for $n \in \langle N\rangle$. Let us define the matrices of changes of basis $\mathbf{P}^{(n)} = [p^{(n)}_{i_n,j_n}] \in \mathbb{K}^{I_n\times I_n}$. By comparison with (4) and (5), the equation of definition (26) of the hypermatrix $\mathcal{A}'$ can be written in the two following compact equivalent forms:

$$\mathcal{A}' = \mathcal{A} \times_1 \mathbf{P}^{(1)} \times_2 \cdots \times_N \mathbf{P}^{(N)} = \mathcal{A} \mathop{\times}_{n=1}^{N} \mathbf{P}^{(n)}$$
$$= (\mathbf{P}^{(1)}, \cdots, \mathbf{P}^{(N)}) \cdot \mathcal{A}$$

These two equations express the fact that a linear transformation of matrix $\mathbf{P}^{(n)}$ is applied to each v.s. E_n, for $n \in \langle N\rangle$. They involve N hypermatrixmatrix multiplications, that is, the n-mode products of the hypermatrix $\mathcal{A}$ with $\mathbf{P}^{(\mathrm{n})} \in \mathbb{K}^{I_n\times I_0}$, $n \in \langle N\rangle$ or, equivalently, an N-linear (or multilinear) multiplication. The term multilinear multiplication comes from the fact that, for all matrices $P^{(n)}, Q^{(n)} \in \mathbb{K}^{I_n\times I_n}$, and for all $\alpha, \beta \in \mathrm{K}$, we have:

$$(\mathbf{P}^{(1)}, \cdots, \alpha\mathbf{P}^{(n)} + \beta\mathbf{Q}^{(n)}, \cdots, \mathbf{P}^{(N)}) \cdot \mathcal{A} = \alpha(\mathbf{P}^{(1)}, \cdots, \mathbf{P}^{(n)}, \cdots, \mathbf{P}^{(N)}) \cdot \mathcal{A}$$
$$+ \beta(\mathbf{P}^{(1)}, \cdots, \mathbf{Q}^{(n)}, \cdots, \mathbf{P}^{(N)}) \cdot \mathcal{A}.$$

Matrix rank

In the matrix case, a matrix $\mathbf{A} \in \mathbb{K}^{I \times J}$ is of rank one if and only if it can be written as the outer product of two non-zero vectors:

$$\mathbf{A} = \mathbf{u} \circ \mathbf{v}, \mathbf{u} \in \mathbb{K}^{I}, \mathbf{v} \in \mathbb{K}^{J} \Leftrightarrow a_{ij} = u_i v_j.$$

The rank of a matrix $\mathbf{A} \in \mathbb{K}^{I \times J}$, denoted $\mathrm{r}_{\mathbf{A}}$ or $\mathrm{r}(\mathbf{A})$, is defined as the smallest integer R such that $\mathbf{A}$ be written as the sum of R matrices of rank one:

$$\mathrm{r}(\mathbf{A}) = \min \left\{ R : \mathbf{A} = \sum_{r=1}^{R} \mathbf{u}_r \circ \mathbf{v}_r, \mathbf{u}_r \in \mathbb{K}^{I}, \mathbf{v}_r \in \mathbb{K}^{J} \right\}.$$

Matrix rank

In the matrix case, a matrix $\mathbf{A} \in \mathbb{K}^{I \times J}$ is of rank one if and only if it can be written as the outer product of two non-zero vectors:

$$\mathbf{A} = \mathbf{u} \circ \mathbf{v}, \mathbf{u} \in \mathbb{K}^{I}, \mathbf{v} \in \mathbb{K}^{J} \Leftrightarrow a_{ij} = u_i v_j.$$

The rank of a matrix $\mathbf{A} \in \mathbb{K}^{I \times J}$, denoted $\mathrm{r}_{\mathbf{A}}$ or $\mathrm{r}(\mathbf{A})$, is defined as the smallest integer R such that $\mathbf{A}$ be written as the sum of R matrices of rank one:

$$\mathrm{r}(\mathbf{A}) = \min \left\{ R : \mathbf{A} = \sum_{r=1}^{R} \mathbf{u}_r \circ \mathbf{v}_r, \mathbf{u}_r \in \mathbb{K}^{I}, \mathbf{v}_r \in \mathbb{K}^{J} \right\}.$$

The dyadic decomposition of $\mathbf{A}$ can be written as following:

$$\mathbf{A} = \mathbf{U}\mathbf{V}^{T},$$

where

$$\mathbf{U} = \left[\mathbf{u}_1 \mathbf{u}_2 \cdots \mathbf{u}_R \right], \mathbf{V} = \left[\mathbf{v}_1 \mathbf{v}_2 \cdots \mathbf{v}_R \right].$$

Hypermatrix rank

In the case of a tensor of order higher than two, there are several ways to define the rank. We only consider the definition of tensor rank from its decomposition into a linear combination of elementary tensors, that is, a sum of rank-one tensors, which constitutes a generalization of the matrix rank to orders higher than two. The multilinear rank of an N-order tensor, related to the Tucker decomposition and to the N modal matricizations of the tensor, will be not introduced here.

Consider a hypermatrix $\mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ of order N. This hypermatrix is of rank one if and only if there exists N non-zero vectors $\mathbf{u}^{(n)} \in \mathbb{K}^{I_n}$, $n \in \langle N \rangle$, such that $\mathcal{A} = \mathop{\circ}\limits_{n=1}^{N} \mathbf{u}^{(n)}$, that is, $\mathcal{A}$ is written as the outer product of N vectors.

Hypermatrix rank

The rank of $\mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$, denoted by $\mathrm{r}_{\mathcal{A}}$ or $\mathrm{r}(\mathcal{A})$, is defined as the smallest integer R such that $\mathcal{A}$ is written as a sum of R rank-one hypermatrices:

$$\mathbf{r}(\mathcal{A}) = \min\{R : \mathcal{A} = \sum_{r=1}^{R} \mathbf{u}_r^{(1)} \circ \cdots \circ \mathbf{u}_r^{(N)}, \mathbf{u}_r^{(n)} \in \mathbb{K}^{I_n}, n \in \langle N \rangle\}. \qquad (27)$$

This equation defines a canonical polyadic decomposition (CPD) of the hypermatrix $\mathcal{A}$. This decomposition is also called PARAFAC [for parallel factors] or CANDECOMP [for canonical decomposition]. The CPD introduced in (27) involves N factor matrices $\mathbf{U}^{(n)} = \left[\mathbf{u}_1^{(n)} \cdots \mathbf{u}_R^{(n)}\right] \in \mathbb{K}^{I_n \times R}, n \in \langle N \rangle$.

Symmetric rank of a hypermatrix

In the case of a symmetric hypermatrix $\mathcal{A} \in \mathbb{K}^{I \times \cdots \times I}$ of order N, the symmetric rank of $\mathcal{A}$, denoted by $\mathbf{r}_S(\mathcal{A})$, is defined as the smallest integer R_S such that $\mathcal{A}$ can be written as a sum of R_S symmetric rank-one hypermatrices:

$$\mathbf{r}_S(\mathcal{A}) = \min \left\{ R_S : \mathcal{A} = \sum_{r=1}^{R_S} \mathbf{u}_r \circ \cdots \circ \mathbf{u}_r = \sum_{r=1}^{R_S} \mathbf{u}_r^{\circ N}, \mathbf{u}_r \in \mathbb{K}^I \right\}.$$

Symmetric rank of a hypermatrix

In the case of a symmetric hypermatrix $\mathcal{A} \in \mathbb{K}^{I \times \cdots \times I}$ of order N, the symmetric rank of $\mathcal{A}$, denoted by $\mathbf{r}_S(\mathcal{A})$, is defined as the smallest integer R_S such that $\mathcal{A}$ can be written as a sum of R_S symmetric rank-one hypermatrices:

$$\mathbf{r}_S(\mathcal{A}) = \min \left\{ R_S : \mathcal{A} = \sum_{r=1}^{R_S} \mathbf{u}_r \circ \cdots \circ \mathbf{u}_r = \sum_{r=1}^{R_S} \mathbf{u}_r^{\circ N}, \mathbf{u}_r \in \mathbb{K}^I \right\}.$$

For a complex symmetric hypermatrix, the symmetric rank R_S and the tensor rank R, as defined in (27), are such that:

$$R \leq R_S.$$

In Comon et al. (2008), it is shown that $R = R_S$ generically (i.e. with probability one), when $R_s \leq I$ and N is large enough compared to I.

Comparative properties of hypermatrices and matrices

Several properties differentiate hypermatrices of order $N \geq 3$ from matrices.

- The rank of a hypermatrix $\mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ of order higher than two depends on the field $\mathbb{K}$ over which its elements are defined. Indeed, since $\mathbb{R}^{I_1 \times \cdots \times I_N} \subseteq \mathbb{C}^{I_1 \times \cdots \times I_N}$ for $N \geq 3$, the tensor ranks over $\mathbb{K} = \mathbb{R}$ and over $\mathbb{K} = \mathbb{C}$, respectively, denoted by $\mathrm{r}_{\mathbb{R}}$ and $\mathrm{r}_{\mathbb{C}}$, are such that $\mathrm{r}_{\mathbb{C}}(\mathcal{A}) \leq \mathrm{r}_{\mathbb{R}}(\mathcal{A})$. This is not the case for matrices for which we have $\mathrm{r}_{\mathbb{C}}(\mathbf{A}) = \mathrm{r}_{\mathrm{R}}(\mathbf{A})$.

Several properties differentiate hypermatrices of order $N \geq 3$ from matrices.

- The rank of a hypermatrix $\mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ of order higher than two depends on the field $\mathbb{K}$ over which its elements are defined. Indeed, since $\mathbb{R}^{I_1 \times \cdots \times I_N} \subseteq \mathbb{C}^{I_1 \times \cdots \times I_N}$ for $N \geq 3$, the tensor ranks over $\mathbb{K} = \mathbb{R}$ and over $\mathbb{K} = \mathbb{C}$, respectively, denoted by $\mathrm{r}_{\mathbb{R}}$ and $\mathrm{r}_{\mathbb{C}}$, are such that $\mathrm{r}_{\mathbb{C}}(\mathcal{A}) \leq \mathrm{r}_{\mathbb{R}}(\mathcal{A})$. This is not the case for matrices for which we have $\mathrm{r}_{\mathbb{C}}(\mathbf{A}) = \mathrm{r}_{\mathrm{R}}(\mathbf{A})$.

EXAMPLE:

$$\begin{aligned}
\mathcal{A} = a_{121} = a_{211} = a_{112} = 0, a_{221} = a_{122} = a_{212} = 1, a_{111} = -1, a_{222} = 0 \\
= \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}^{\circ 3} + \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}^{\circ 3} - 2\begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\circ 3} \\
= \frac{i}{2}\begin{bmatrix} -i \\ 1 \end{bmatrix}^{\circ 3} - \frac{i}{2}\begin{bmatrix} i \\ 1 \end{bmatrix}^{\circ 3}, \text{ where } i^2 = -1
\end{aligned}$$

- The rank of a real hypermatrix $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ may be greater than the largest of its dimensions, that is, one can have $R > \max\{I_1, \cdots, I_N\}$, whereas for a matrix $\mathbf{A} \in \mathbb{K}^{I \times J}$, the rank is at most equal to the smallest of its dimensions, that is, $R \leq \min(I, J)$.
- While a random matrix is of full rank with probability one, a real random hypermatrix of order higher than two can have several ranks called typical ranks. So, Kruskal (1989) showed that a third-order real hypermatrix of dimensions $2 \times 2 \times 2$, has two typical ranks two and three. For complex hypermatrices, there is only one typical rank, called a generic rank, with probability one.
- A fundamental question concerns the determination of an upper bound for the rank of a hypermatrix. Thus, for a third-order hypermatrix $\mathcal{A} \in \mathbb{K}^{I \times J \times K}$, the rank is bounded by (Kruskal 1989):

$$R \leq \min\{IJ, JK, KI\}.$$

As we have just seen, there are several notions of rank for a hypermatrix, and therefore for a tensor. One can define the non-negative rank of a non-negative real hypermatrix $\mathcal{A}$, that is, a hypermatrix whose all elements are positive or zero. This rank, denoted $\mathrm{r}_+(\mathcal{A})$, is the minimal number of rank-one non-negative terms that form the CPD. Similarly to the matrix case, we have $\mathrm{r}(\mathcal{A}) \leq \mathrm{r}_+(\mathcal{A})$. For a review of the main results on the different notions of rank of a hypermatrix, refer to Sidiropoulos et al. (2017).

N.D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E.E. Papalexakis, C. Faloutsos
Tensor Decomposition for Signal Processing and Machine Learning, IEEE Transactions on Signal Processing, 65(13):3551-3582, 2017.

CPD and dimensionality reduction

In the era of big data, complexity reduction is a major issue for storage, visualization, representation, analysis and classification of massive data, or megadata such as, for example, in recommendation systems, or for processing medical, astronomical, climatic observation, or social networks databases. This complexity reduction can be obtained in two different ways, either at the level of the data representation model or through the use of numerically efficient processing methods.

The previously introduced CPD plays a fundamental role in reducing the dimensionality of a hypermatrix. Thus, for a hypermatrix $\mathcal{A} \in \mathbb{K}^{I \times \cdots \times I}$, of order N and of the same dimension for each mode $(I_n = I, \forall n \in \langle N \rangle)$, the number of elements and, therefore, the memory needed to store all the data contained in $\mathcal{A}$, is I^N, while for CPD (27), the amount of data to be stored, that is, the elements of the N factor matrices $\mathbf{U}^{(n)} \in \mathbb{K}^{I \times R}, n \in \langle N \rangle$, is equal to NRI. This constitutes a very significant dimensionality reduction for large values of N and I, that is, for very large hypermatrices.

CPD and dimensionality reduction

In the case of a third-order hypermatrix $\mathcal{X} \in \mathbb{K}^{I \times J \times K}$ of rank R, a CPD is written as:

$$x_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr}, \tag{28}$$

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{A}_{.r} \circ \mathbf{B}_{.r} \circ \mathbf{C}_{.r}, \tag{29}$$

$$= \mathcal{I}_{3,R} \times {}_1\mathbf{A} \times {}_2\mathbf{B} \times {}_3\mathbf{C}, \tag{30}$$

where $\mathbf{A}_{.r} \in \mathbb{K}^I, \mathbf{B}_{.r} \in \mathbb{K}^J, \mathbf{C}_{.r} \in \mathbb{K}^K$ are the r th column vectors of the factor matrices

$$\mathbf{A} = [\mathbf{A}_{.1}, \cdots, \mathbf{A}_{.R}], \mathbf{B} = [\mathbf{B}_{.1}, \cdots, \mathbf{B}_{.R}], \mathbf{C} = [\mathbf{C}_{.1}, \cdots, \mathbf{C}_{.R}].$$

This PARAFAC model, of rank R, is denoted by $\|\mathbf{A}, \mathbf{B}, \mathbf{C}; R\|$.

CPD and dimensionality reduction

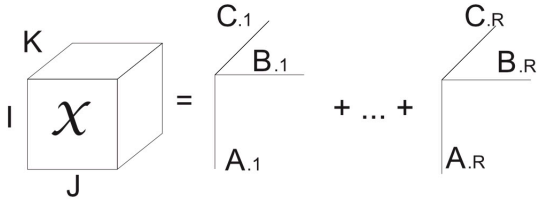

图: Third-order PARAFAC model

This figure illustrates the PARAFAC model $\|\mathbf{A}, \mathbf{B}, \mathbf{C}; R\|$ in the form (29) of a sum of R rank-one hypermatrices, each one being equal to the outer product of three column vectors $(\mathbf{A}_{.r}, \mathbf{B}_{.r}, \mathbf{C}_{.r})$, with $r \in \langle R \rangle$.

CPD and dimensionality reduction

A fundamental difference between matrices and hypermatrices concerns the uniqueness properties of their dyadic and polyadic decomposition, respectively.

It should be noted that when $\mathbf{A}$ and $\mathcal{A}$ are of rank one, that is, $\mathbf{A} = \mathbf{u} \circ \mathbf{v}$ and $\mathcal{A} = \mathbf{a} \circ \mathbf{b} \circ \mathbf{c}$, then ambiguities are scalar, and all matrices and hypermatrices equivalent to $\mathbf{A}$ and $\mathcal{A}$ can be written as: $(\alpha \mathbf{u}) \circ \left(\frac{1}{\alpha}\mathbf{v}\right)$ and $(\alpha \mathbf{a}) \circ (\beta \mathbf{b}) \circ \left(\frac{1}{\alpha\beta}\mathbf{c}\right)$, respectively, with $\alpha, \beta \in \mathbb{K}$.

Determining the parameters of a decomposition from a data hypermatrix is a very important problem for tensor-based applications.

In contrast to the matrix case, for tensors of order higher than two, the tensor approximation problem using a lower-rank tensor is an ill-posed problem due to the fact that the set of tensors of rank equal to R is not generally closed. This means that a sequence of tensors of rank R can converge to a tensor of higher rank, as illustrated by the example below.

CPD and dimensionality reduction

EXAMPLE: Let the sequence of rank-two tensors:

$$\mathcal{X}_n = n\left(\mathbf{u} + \frac{1}{n}\mathbf{v}\right) \circ \left(\mathbf{u} + \frac{1}{n}\mathbf{v}\right) \circ \left(\mathbf{u} + \frac{1}{n}\mathbf{v}\right) - n\mathbf{u} \circ \mathbf{u} \circ \mathbf{u}$$

with $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$. Using the associativity property of the outer product over addition, $\mathcal{X}_n$ can be developed as:

$$\begin{aligned}\mathcal{X}_n &= \mathbf{u} \circ \mathbf{u} \circ \mathbf{v} + \mathbf{u} \circ \mathbf{v} \circ \mathbf{u} + \mathbf{v} \circ \mathbf{u} \circ \mathbf{u} + \frac{1}{n}(\mathbf{v} \circ \mathbf{v} \circ \mathbf{u} \\ &+ \mathbf{v} \circ \mathbf{u} \circ \mathbf{v} + \mathbf{u} \circ \mathbf{v} \circ \mathbf{v}) + \frac{1}{n^2}\mathbf{v} \circ \mathbf{v} \circ \mathbf{v} \\ &= \mathcal{X} + \text{ terms in } \mathcal{O}\left(\frac{1}{n}\right)\end{aligned}$$

When $n \to \infty$, the sequence $\mathcal{X}_n$ tends toward the tensor $\mathcal{X}$ of higher rank than $\mathcal{X}_n$ equal to three, defined by:

$$\mathcal{X} = \mathbf{u} \circ \mathbf{u} \circ \mathbf{v} + \mathbf{u} \circ \mathbf{v} \circ \mathbf{u} + \mathbf{v} \circ \mathbf{u} \circ \mathbf{u}.$$

Tensor rank

In the case of a tensor $\mathcal{X} \in \bigotimes_{n=1}^{N} E_n$, the rank of $\mathcal{X}$, denoted by $\mathbf{r}(\mathcal{X})$, is defined as the minimum number R of elementary tensors whose sum can represent $\mathcal{X}$:

$$\mathbf{r}(\mathcal{X}) = \min \left\{ R : \mathcal{X} = \sum_{r=1}^{R} \mathbf{u}_r^{(1)} \otimes \cdots \otimes \mathbf{u}_r^{(N)}, \mathbf{u}_r^{(n)} \in E_n, n \in \langle N \rangle \right\},$$

where $\mathbf{u}_r^{(1)} \otimes \cdots \otimes \mathbf{u}_r^{(N)}$ is to be considered as a tensor product. It should be noted that when one chooses the basis $\mathcal{B}$ defined in (17) for the tensor space $\bigotimes_{n=1}^{N} E_n$, then $\mathrm{rank}(\mathcal{A}) = \mathrm{rank}(\mathcal{X})$, which means that the tensor has the same rank as the hypermatrix of its coordinates in the basis $\mathcal{B}$.

The determination of the rank of a tensor of order higher than two and, thus, the computation of an exact PARAFAC decomposition are difficult problems because they are ill-posed.

Eigenvalues and singular values of a hypermatrix

$\mathbf{A} \in \mathbb{R}^{I\times I}$, $\mathbf{u} \in \mathbb{R}^{I}, \mathbf{u} \neq \mathbf{0}_I$
Eigenvalues λ
$\sum_{j=1}^{I} a_{ij}u_j = \lambda u_i$, $i \in \langle I\rangle$ $\Leftrightarrow$ $\mathbf{Au} = \lambda\mathbf{u}$
$\mathbf{A} \in \mathbb{R}^{I\times J}$, $\mathbf{u} \in \mathbb{R}^{I}$, $\mathbf{v} \in \mathbb{R}^{J}$
Singular values σ
$\sum_{j=1}^{J} a_{ij}v_j = \sigma u_i$, $i \in \langle I\rangle$ $\Leftrightarrow$ $\mathbf{Av} = \sigma\mathbf{u}$ $\Leftrightarrow$ $\mathbf{A}^T\mathbf{Av} = \sigma^2\mathbf{v}$ $\sum_{i=1}^{I} a_{ij}u_i = \sigma v_j$, $j \in \langle J\rangle$ $\Leftrightarrow$ $\mathbf{A}^T\mathbf{u} = \sigma\mathbf{v}$ $\Leftrightarrow$ $\mathbf{AA}^T\mathbf{u} = \sigma^2\mathbf{u}$
$\mathcal{A} \in \mathbb{R}^{I\times I\times I}$, $\mathbf{u} \in \mathbb{R}^{I}$, $\mathbf{u} \neq \mathbf{0}_I$
Eigenvalues λ
$\sum_{i,j=1}^{I} a_{ijk}u_iu_j = \lambda u_k$, $k \in \langle I\rangle$
$\mathcal{A} \in \mathbb{R}^{I\times J\times K}$, $\mathbf{u} \in \mathbb{R}^{I}$, $\mathbf{v} \in \mathbb{R}^{J}$, $\mathbf{w} \in \mathbb{R}^{K}$
Singular values σ
$\sum_{j,k=1}^{J,K} a_{ijk}v_jw_k = \sigma u_i$, $i \in \langle I\rangle$ $\sum_{i,k=1}^{I,K} a_{ijk}u_iw_k = \sigma v_j$, $j \in \langle J\rangle$ $\sum_{i,j=1}^{I,J} a_{ijk}u_iv_j = \sigma w_k$, $k \in \langle K\rangle$

Table: Matrix and third-order tensor eigenvalues and singular values

Isomorphisms of tensor spaces

Given that two $\mathbb{K}$-v.s. are isomorphic if they have the same dimension, it is possible to decompose a tensor space in different ways. This means that the elements of a hypermatrix of the space $\mathbb{K}^{I_1 \times \cdots \times I_N}$ can be stored in different ways by combination of modes. Two standard unfolding methods are vectorization and matricization (also called matrix unfolding), which consist of storing the elements in a column vector and a matrix, respectively. More generally, one can store the elements of an Nth order hypermatrix in a hypermatrix of order $P < N$, the cases $P = 1$ and $P = 2$ corresponding to vectorization and matricization, respectively.

Isomorphisms of tensor spaces: examples

- A fourth-order hypermatrix $\mathcal{A} \in \mathbb{K}^{I_1 \times I_2 \times I_3 \times I_4}$ can be vectorized into a vector $\mathbf{v} \in \mathbb{K}^{I_1 I_2 I_3 I_4}$. The order of dimensions in the product $I_1 I_2 I_3 I_4$ is directly related to the order of variation of the indices (i_1, i_2, i_3, i_4). By choosing to vary i_1 the most slowly and i_4 the most rapidly, we have $v_i = a_{i_1 i_2 i_3 i_4}$ with $i = i_4 + (i_3 - 1)I_4 + (i_2 - 1)I_3 I_4 + (i_1 - 1)I_2 I_3 I_4$.
- The hypermatrix $\mathcal{A}$ can also be unfolded into a matrix as, $\mathbf{M} \in \mathbb{K}^{I_1 \times I_2 I_3 I_4}$ such that $m_{i_1, j} = a_{i_1 i_2 i_3 i_4}$ with $j = i_4 + (i_3 - 1)I_4 + (i_2 - 1)I_3 I_4$.
- One can also unfold $\mathcal{A}$ into a third-order hypermatrix by combining, the last two modes. The unfolded hypermatrix $\mathcal{B} \in \mathbb{K}^{I_1 \times I_2 \times I_3 I_4}$ is such that $b_{i_1, i_2, k} = a_{i_1 i_2 i_3 i_4}$ with $k = i_4 + (i_3 - 1)I_4$.

NOTE: To highlight how the modes are combined to build the unfoldings, we shall use their dimensions to distinguish them. So, for the three above examples, we shall write the unfoldings as: $\mathbf{a}_{I_1 I_2 I_3 I_4}$, $\mathbf{A}_{I_1 \times I_2 I_3 I_4}$, and $\mathcal{A}_{I_1 \times I_2 \times I_3 I_4}$, instead of $\mathbf{v}, \mathbf{M}$, and $\mathcal{B}$, respectively

Isomorphisms of tensor spaces: examples

EXAMPLE: Let a third-order hypermatrix $\mathcal{A} = \left[a_{ijk}\right] \in \mathbb{K}^{2\times 2\times 2}$. We have:

$$\mathbf{a}_{IJK} = \left[a_{111}\ a_{112}\ a_{121}\ a_{122}\ a_{211}\ a_{212}\ a_{221}\ a_{222}\right]^T \in \mathbb{K}^{IJK},$$
$$\mathbf{a}_{JKI} = \left[a_{111}\ a_{211}\ a_{112}\ a_{212}\ a_{121}\ a_{221}\ a_{122}\ a_{222}\right]^T \in \mathbb{K}^{JKI},$$
$$\mathbf{A}_{IJ\times K} = \begin{bmatrix} a_{111} & a_{112} \\ a_{121} & a_{122} \\ a_{211} & a_{212} \\ a_{221} & a_{222} \end{bmatrix} \in \mathbb{K}^{IJ\times K}, \mathbf{A}_{IK\times J} = \begin{bmatrix} a_{111} & a_{121} \\ a_{112} & a_{122} \\ a_{211} & a_{221} \\ a_{212} & a_{222} \end{bmatrix} \in \mathbb{K}^{IK\times J}.$$

The hypermatrix space $\mathbb{K}^{2\times 2\times 2}$ is isomorphic to the vector space $\mathbb{K}^8$ and to the matrix space $\mathbb{K}^{4\times 2}$. This means that the addition of two hypermatrices and the multiplication of a hypermatrix by a scalar can be indifferently performed using the vectorized or matricized forms of the hypermatrices.