We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Chapter

Bilinear Applications and Tensors

Rodrigo Garcia Eustaquio

Abstract

In this chapter, a theoretical approach to the vector space of tensor of order 3 and the vector space of bilinear applications will be presented in order to present an isomorphism between these spaces and several properties about tensor and bilinear applications. With this well-defined isomorphism, we will present how to calculate the product between tensor of second derivatives and a vector, where such a product is used in several numerical methods such as Chebyshev-Halley class and others mentioned in the introduction. In addition, concepts on differentiability are presented, allowing a better understanding for the reader about second-order derivatives seen as a tensor.

Keywords: tensor, bilinear application, isomorphism, second derivative, inexact tensor-free Chebyshev-Halley class

1. Introduction

Frequently, discretization of mathematical models demands solving a system of equations, which is generally nonlinear. Such mathematical problems might be written as

$$\text{find } x^* \in \mathbb{R}^n \text{ such that } F(x^*) = 0 \tag{1}$$

where $F : \text{—} \mathbb{R}^n \to \mathbb{R}^n$.

There exist iterative methods for solving (1) that have cubic convergence rate, for instance, the methods belonging to the following class of methods named Chebyshev-Halley class, which was introduced by Hernández and Gutiérrez in [1]:

$$x^{k+1} = x^k - \left[I + \frac{1}{2}\mathcal{L}(x^k)\left(I - \alpha\mathcal{L}(x^k)\right)^{-1} \right] J_F(x^k)^{-1} F(x^k), \tag{2}$$

for all $k \in \mathbb{N}$, where

$$\mathcal{L}(x) = J_F(x)^{-1} \mathcal{T}_F(x)\left(J_F(x)^{-1} F(x) \right), \tag{3}$$

and $J_F(x)$ and $\mathcal{T}_F(x)$ denote the first and second derivatives of F evaluated at x, respectively. The parameter α is a real number and I is the identity matrix in $\mathbb{R}^{n \times n}$.

Discretized versions of Chebyshev-Halley class have already been considered in [2] in such a way that the tensor of second derivatives of the function F was approximated by bilinear operators. A tensor is a multi-way array or multidimensional matrix. A generalization of the Chebyshev-Halley class (2) where no second-order derivative information is required but that also has cubic

convergence rate, named inexact tensor-free Chebyshev-Halley class, was introduced by Eustaquio, Ribeiro, and Dumett [3]. Other families of iterative methods with cubic convergence rate were extensively described in Traub's book [4].

Several alternatives exist for the product of the tensor of second derivatives of F by vectors [5–8], and this needs to be elucidated.

The aim of this chapter is to present concepts and relationships between tensors of order 3 and bilinear applications, in order to relate them to the second derivative of a two-differentiable application. We will see later that given the vectors $u, v \in \mathbb{R}^n$, the i-th row of the matrix $\mathcal{T}_F(x)v$ is defined by $v^T \nabla^2 f_i(x)$, where $\nabla^2 f_i(x)$ is the Hessian of the i-th component of F evaluated at x. The i-th component of the vector $\mathcal{T}_F(x)vu$ is defined by $v^T \nabla^2 f_i(x)u$.

2. Tensors

Tensors naturally arise in some applications, such as chemometry [9], signal processing [10], and others. According to [8], for many applications involving high-order tensors, the known results of matrix algebra seemed to be insufficient in the twentieth century. There were some workshops and congresses on the study of tensors, such as:

- Workshop on Tensor Decomposition at the American Institute of Mathematics which took place at the Palo Alto, California, 2004, organized by Golub, Kolda, Nagy, and Van Loan. Details in [11];

- Workshop on Tensor Decompositions and Applications, 2005, organized by Comon and De Lathauwer. Details in [12]; and

- Minisymposium on Numerical Multilinear Algebra: A New Beginning, 2007, organized by Golub, Comon, De Lathauwer, and Lim and which took place at the Zurich.

Readers interested in multilinear singular value decomposition, eigenvalues, and eigenvectors may consult as references [5–8, 13, 14]. In this text, we will focus our attention on tensors of order 3.

Let I_1, I_2, and I_3 be three positive integers. A tensor $\mathcal{T}$ of order 3 is an three-way array where its elements $t_{i_1 i_2}^{i_3}$ are indexed by $i_1 = 1, \ldots, I_1$, $i_2 = 1, \ldots, I_2$, and $i_3 = 1, \ldots, I_3$ and the n-th dimension of the tensor is denoted by I_n, for $n = 1, 2, 3$. For example, the first, second, and third dimensions of a tensor $\mathcal{T} \in \mathbb{R}^{2 \times 4 \times 3}$ are $2, 4, 3$, respectively.

Obviously, tensors are generalizations of matrices. A matrix can be viewed as a tensor of order 2, while a vector can be viewed as a tensor of order 1.

From an algebraic point of view, a tensor $\mathcal{T}$ of order 3 is an element of the vector space $\mathbb{R}^{I_1 \times I_2 \times I_3}$, whereas from the geometric point of view, a tensor $\mathcal{T}$ of order 3 can be seen as a parallelepiped [15], with I_1 rows, I_2 columns, and I_3 tubes. **Figure 1** illustrates a tensor $\mathcal{T} \in \mathbb{R}^{2 \times 4 \times 3}$.

In linear algebra, it is common to see a matrix through its columns. If $A \in \mathbb{R}^{m \times n}$, then A can be viewed as $A = [a_1 \ldots a_n]$, where $a_j \in \mathbb{R}^m$ denotes the j-th column of the matrix A. In the case of tensor of order 3, we can see them through fibers and slices. Hence follow the definitions.

Definition 1.1. A tensor fiber of a tensor of order 3 is a one-dimensional fragment obtained by fixing only two indices.
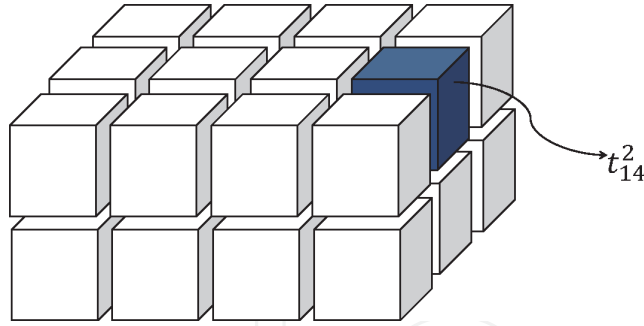
Figure 1.
A tensor $\mathcal{T} \in \mathbb{R}^{2 \times 4 \times 3}$.

Definition 1.2. A tensor slice of a tensor of order 3 is a two-dimensional section (fragment), obtained by fixing only one index.

Generally in tensors of order 3, a fiber is a vector and a slice is a matrix. We have three types of fibers:

- column fibers (or mode-1 fiber), where the indices i_2 and i_3 are fixed;

- row fibers (or mode-2 fiber), where the indices i_1 and i_3 are fixed; and

- tube fibers (or mode-3 fiber), where the indices i_1 and i_2 are fixed.

We also have three types of slices:

- horizontal slice, where the index i_1 is fixed;

- lateral slice, where the index i_2 is fixed; and

- frontal slice, where the index i_3 is fixed.

For example, consider a tensor $\mathcal{T} \in \mathbb{R}^{2 \times 4 \times 3}$ with $i = 1, 2, j = 1, 2, 3, 4$, and $k = 1, 2, 3$. The i-th horizontal slice, denoted by $\mathcal{T}^{i::}$, is the matrix

$$\mathcal{T}^{i::} = \begin{pmatrix} t_{i1}^1 & t_{i1}^2 & t_{i1}^3 \\ t_{i2}^1 & t_{i2}^2 & t_{i2}^3 \\ t_{i3}^1 & t_{i3}^2 & t_{i3}^3 \\ t_{i4}^1 & t_{i4}^2 & t_{i4}^3 \end{pmatrix},$$

the j-th lateral slice, denoted by $\mathcal{T}^{:j:}$, is the matrix

$$\mathcal{T}^{:j:} = \begin{pmatrix} t_{1j}^1 & t_{1j}^2 & t_{1j}^3 \\ t_{2j}^1 & t_{2j}^2 & t_{2j}^3 \end{pmatrix}$$

and the k-th frontal slice, denoted by $\mathcal{T}^{::k}$, is the matrix

$$\mathcal{T}^{::k} = \begin{pmatrix} t_{11}^k & t_{12}^k & t_{13}^k & t_{14}^k \\ t_{21}^k & t_{22}^k & t_{23}^k & t_{24}^k \end{pmatrix}. \tag{4}$$

Figures 2 and **3** illustrate the three types of fibers and slices, respectively, of a tensor $\mathcal{T} \in \mathbb{R}^{2 \times 4 \times 3}$.
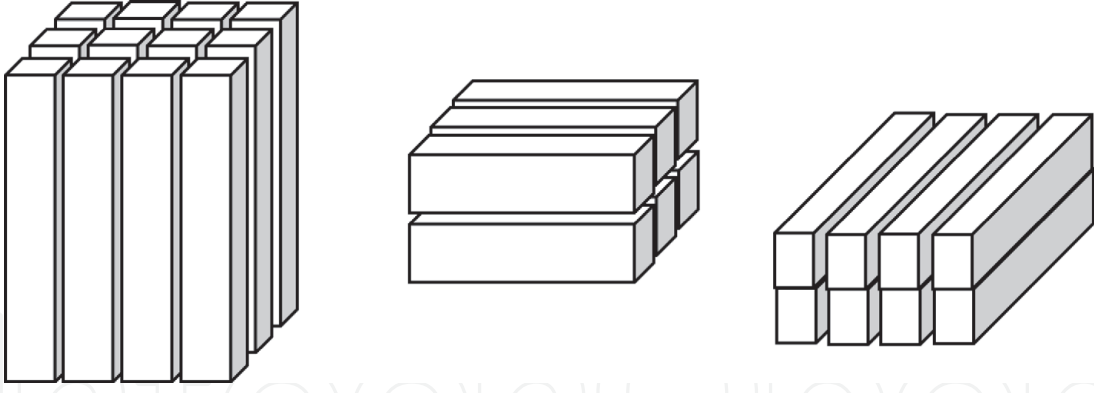
Figure 2.
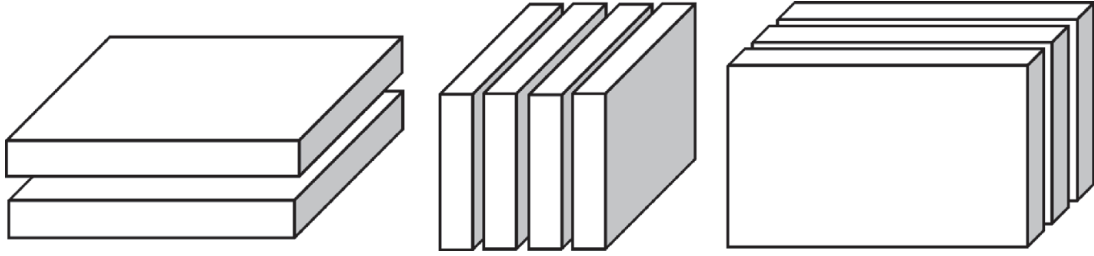*Columns, rows, and tube fibers, respectively.*

Figure 3.
Horizontal, lateral, and frontal slices, respectively.

2.1 Tensor operations

The first issue to consider in this subsection is how to calculate the product between tensors and matrices. It is well known from elementary algebra that given matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{R \times m}$, it is possible to calculate the product BA, because the first dimension (number of rows) of matrix A agrees with the second dimension (number of columns) of matrix B, and each product element is the result of the inner product between rows of matrix B and columns of matrix A.

The product between tensors of order 3 and matrices or vectors is a bit more complicated. In order to obtain an element of the product between a tensor and a matrix, it is necessary to specify what dimension of the tensor will be chosen to agree with the number of columns of the matrix, and each resulting element will be a result of the inner product between the mode-n fibers (column, row, or tube) and the columns of the matrix. We will use the solution adopted by [8], which defines the product mode-n between tensors and matrices and the solution adopted by [5] that defines the contracted product mode-n between tensors and vectors.

The mode-n product is useful when one wants to decompose into singular values a high-order tensor in order to avoid the use of the generalized transpose concept. We refer to [5, 7, 8, 13] for details.

Definition 1.3. (mode-n tensor matrix product) The mode-1 product between a tensor $\mathcal{T} \in \mathbb{R}^{m \times n \times p}$ and a matrix $A \in \mathbb{R}^{R \times m}$ is a tensor

$$\mathcal{Y} = \mathcal{T} \times_1 A \in \mathbb{R}^{R \times n \times p}$$

where its elements are defined by

$$y_{rj}^k = \sum_{i=1}^{m} t_{ij}^k a_{ri} \quad \text{where} \quad r = 1, \dots, R, j = 1, \dots, n, \text{ and } k = 1, \dots, p.$$

The mode-2 product between a tensor $T \in \mathbb{R}^{m \times n \times p}$ and a matrix $A \in \mathbb{R}^{R \times n}$ is a tensor

$$\mathcal{Y} = T \times_2 A \in \mathbb{R}^{m \times R \times p}$$

where its elements are defined by

$$y_{ir}^k = \sum_{j=1}^{n} t_{ij}^k a_{rj} \quad \text{where} \quad i = 1, \dots, m, r = 1, \dots, R \text{ and } k = 1, \dots, p.$$

The mode-3 product between a tensor $T \in \mathbb{R}^{m \times n \times p}$ and a matrix $A \in \mathbb{R}^{R \times p}$ is a tensor

$$\mathcal{Y} = T \times_3 A \in \mathbb{R}^{m \times n \times R}$$

where its elements are defined by

$$y_{ij}^r = \sum_{k=1}^{p} t_{ij}^k a_{rk} \quad \text{where} \quad i = 1, \dots, m, j = 1, \dots, n \text{ and } r = 1, \dots, R.$$

To understand the mode-n product in terms of matrix, consider matrices $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{k \times m}$, and $C \in \mathbb{R}^{q \times n}$. By Definition 1.3 we have

$$A \times_1 B = BA \in \mathbb{R}^{k \times n} \text{ and } A \times_2 C = AC^T \in \mathbb{R}^{m \times q}.$$

Thus, the singular value decomposition of matrix A can be written as

$$U\Sigma V^T = (\Sigma \times_1 U) \times_2 V = (\Sigma \times_2 V) \times_1 U.$$

The mode-n product satisfies the following property [8]:
Property 1. Let T be a tensor of order 3 and matrices A and B of convenient sizes. We have for all $r, s = 1, 2, 3$

$$(T \times_r A) \times_s B = (T \times_s B) \times_r A = T \times_r A \times_s B \quad \text{for } r \neq s \text{ and} \qquad (5)$$

$$(T \times_r A) \times_r B = T \times_r (BA) \qquad (6)$$

The idea of Bader and Kolda [5] to calculate the product between tensor and vector is to calculate the inner product of each mode-n fiber (column, row, or tube) with the vector. It is not advantageous to treat an m-dimensional vector as a matrix $m \times 1$. For example, if we consider a tensor $T \in \mathbb{R}^{m \times n \times p}$ and a vector $v \in \mathbb{R}^{m \times 1}$, with $m, n, p \neq 1$, by Definition 1.3, the product between T and v is not well defined, but it is possible to calculate $T \times_1 v^T$.
Definition 1.4. (Contracted product mode-n between tensors and vectors) The contracted product mode-1 between a tensor $T \in \mathbb{R}^{m \times n \times p}$ and a vector $v \in \mathbb{R}^m$ is the matrix

$$A = T \overline{\times}_1 v \in \mathbb{R}^{n \times p}$$

where its elements are defined by

$$a_{jk} = \sum_{i=1}^{m} t_{ij}^k v_i \quad \text{where} \quad j = 1, \dots, n \text{ and } k = 1, \dots, p$$

where v_i is the i-th component of the vector v.

The contracted product mode-2 between a tensor $\mathcal{T} \in \mathbb{R}^{m \times n \times p}$ and a vector $v \in \mathbb{R}^n$ is the matrix

$$A = \mathcal{T} \overline{\times}_2 v \in \mathbb{R}^{m \times p}$$

where its elements are defined by

$$a_{ik} = \sum_{j=1}^{n} t_{ij}^k v_j \quad \text{where} \quad i = 1, \ldots, m \text{ and } k = 1, \ldots, p$$

where v_j is the j-th component of the vector v.

The contracted product mode-3 between a tensor $\mathcal{T} \in \mathbb{R}^{m \times n \times p}$ and a vector $v \in \mathbb{R}^p$ is the matrix

$$A = \mathcal{T} \overline{\times}_3 v \in \mathbb{R}^{m \times n}$$

where its elements are defined by

$$a_{ij} = \sum_{k=1}^{p} t_{ij}^k v_k \quad \text{where} \quad i = 1, \ldots, m \text{ and } j = 1, \ldots, n$$

where v_k is the k-th component of the vector v.

A caution must be added when calculating the product between matrices and vectors by considering the definitions 1.3 and 1.4. For example, note that if $A \in \mathbb{R}^{m \times n}$, $u \in \mathbb{R}^n$, and $v \in \mathbb{R}^m$, then $A \overline{\times}_2 u$ and $A \times_2 u^T$ have the same elements, but

$$A \overline{\times}_2 u \neq A \times_2 u^T,$$

because $A \overline{\times}_2 u \in \mathbb{R}^m$ (column vector) and $A \times_2 u^T \in \mathbb{R}^{1 \times m}$ (row vector). Note that, in relation to the matrix product of elementary algebra, we have

$$Au = A \overline{\times}_2 u \tag{7}$$

$$v^T A = A \times_1 v^T \neq A \overline{\times}_1 v. \tag{8}$$

In particular, given a tensor $\mathcal{T} \in \mathbb{R}^{n \times m \times m}$ and a vector $v \in \mathbb{R}^m$, by Definition 1.4 together with (8), it follows that $\mathcal{T} \overline{\times}_2 v \in \mathbb{R}^{n \times m}$ and

$$(\mathcal{T} \overline{\times}_2 v) \overline{\times}_2 v = (\mathcal{T} \overline{\times}_2 v) v \in \mathbb{R}^n.$$

The contracted product mode-n satisfies the following property [5]:

Property 2. Given a tensor $\mathcal{T}$ of order 3 and vectors u and v of convenient sizes, we have for all $r = 1, 2, 3$ and $s = 2, 3$ that

$$(\mathcal{T} \overline{\times}_r u) \overline{\times}_{s-1} v = (\mathcal{T} \overline{\times}_s v) \overline{\times}_r u \quad \text{for } r < s.$$

For example, consider a tensor $\mathcal{T} \in \mathbb{R}^{2 \times 4 \times 3}$, and denote the k-th column and the q-th row of matrix A by $\mathrm{col}_k(A)$ and $\mathrm{row}_q(A)$, respectively. Note that if:

1. $x \in \mathbb{R}^2$, then $\mathcal{T} \overline{\times}_1 x \in \mathbb{R}^{4 \times 3}$ and

$$\text{col}_k(\mathcal{T}\,\overline{\times}_1 x) = \begin{pmatrix} a_{1k} \\ a_{2k} \\ a_{3k} \\ a_{4k} \end{pmatrix} = \begin{pmatrix} t_{11}^k & t_{21}^k \\ t_{12}^k & t_{22}^k \\ t_{13}^k & t_{23}^k \\ t_{14}^k & t_{24}^k \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \left(\mathcal{T}^{::k}\right)^T x \text{ and}$$

$$\text{row}_j(\mathcal{T}\,\overline{\times}_1 x) = \begin{pmatrix} a_{j1} & a_{j2} & a_{j3} \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} t_{1j}^1 & t_{1j}^2 & t_{11}^3 \\ t_{2j}^1 & t_{2j}^2 & t_{21}^3 \end{pmatrix} = x^T \mathcal{T}^{:j:}$$

2. $x \in \mathbb{R}^4$, then $\mathcal{T}\,\overline{\times}_2 x \in \mathbb{R}^{2 \times 3}$ and

$$\text{col}_k(\mathcal{T}\,\overline{\times}_2 x) = \begin{pmatrix} a_{1k} \\ a_{2k} \end{pmatrix} = \begin{pmatrix} t_{11}^k & t_{12}^k & t_{13}^k & t_{14}^k \\ t_{21}^k & t_{22}^k & t_{23}^k & t_{24}^k \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \left(\mathcal{T}^{::k}\right) x \text{ and}$$

$$\text{row}_i(\mathcal{T}\,\overline{\times}_2 x) = \begin{pmatrix} a_{i1} & a_{i2} & a_{i3} \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \end{pmatrix} \begin{pmatrix} t_{i1}^1 & t_{i1}^2 & t_{i1}^3 \\ t_{i2}^1 & t_{i2}^2 & t_{i2}^3 \\ t_{i3}^1 & t_{i3}^2 & t_{i3}^3 \\ t_{i4}^1 & t_{i4}^2 & t_{i4}^3 \end{pmatrix} = x^T \mathcal{T}^{i::}$$

3. $x \in \mathbb{R}^3$, then $\mathcal{T}\,\overline{\times}_3 x \in \mathbb{R}^{2 \times 4}$ and

$$\text{col}_j(\mathcal{T}\,\overline{\times}_3 x) = \begin{pmatrix} a_{1j} \\ a_{2j} \end{pmatrix} = \begin{pmatrix} t_{1j}^1 & t_{1j}^2 & t_{1j}^3 \\ t_{2j}^1 & t_{2j}^2 & t_{2j}^3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \left(\mathcal{T}^{:j:}\right) x \text{ and}$$

$$\text{row}_i(\mathcal{T}\,\overline{\times}_3 x) = \begin{pmatrix} a_{i1} & a_{i2} & a_{i3} \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} t_{i1}^1 & t_{i2}^1 & t_{i3}^1 & t_{i4}^1 \\ t_{i1}^2 & t_{i2}^2 & t_{i3}^2 & t_{i4}^2 \\ t_{i1}^3 & t_{i2}^3 & t_{i3}^3 & t_{i4}^3 \end{pmatrix} = x^T \left(\mathcal{T}^{i::}\right)^T$$

This example can be easily generalized to arbitrary dimensions. In particular, for a tensor $\mathcal{T} \in \mathbb{R}^{m \times n \times n}$ and a vector $x \in \mathbb{R}^n$, we have

$$\text{row}_i(\mathcal{T}\,\overline{\times}_2 x) = x^T \mathcal{T}^{i::} \tag{9}$$

$$\text{row}_i(\mathcal{T}\,\overline{\times}_3 x) = x^T \left(\mathcal{T}^{i::}\right)^T \tag{10}$$

Lemma 1.5. Let $\mathcal{T} \in \mathbb{R}^{n \times n \times n}$ be a tensor. If $\mathcal{T}^{i::}$ is a symmetric matrix for all $i = 1, \ldots, n$, then

$$(\mathcal{T}\,\overline{\times}_2 u)v = (\mathcal{T}\,\overline{\times}_2 v)u$$

for all $u, v \in \mathbb{R}^n$.

Proof. By Property 2, it follows that $(\mathcal{T} \overline{\times}_2 u)v = (\mathcal{T} \overline{\times}_3 v)u$. By (10), (11), and the symmetry of $\mathcal{T}^{i::}$, we have $\mathcal{T} \overline{\times}_3 v = \mathcal{T} \overline{\times}_2 v$. $\qquad\square$

3. Space of bilinear applications

In this section, we define bilinear applications on finite dimensional vector spaces, in order to relate them to the second derivative of a two-differentiable application, as well as a tensor of order 3.

Definition 1.6. Let U, V, W be vector spaces. An application $f : U \times V \rightarrow W$ is a bilinear application if:

i. $f(\lambda u_1 + u_2, v) = \lambda f(u_1, v) + f(u_2, v)$ for all $\lambda \in \mathbb{R}, u_1, u_2 \in U$, and $v \in V$.

ii. $f(u, \lambda v_1 + v_2) = \lambda f(u, v_1) + f(u, v_2)$ for all $\lambda \in \mathbb{R}, u \in U$, and $v_1, v_2 \in V$.

In other words, an application $f : U \times V \rightarrow W$ is a bilinear application if it is linear in each of the variables when the other variable is fixed. We denote by $\mathcal{B}(U \times V, W)$ the set of all bilinear applications of $U \times V$ in W. In particular, if $U = V$ and $W = \mathbb{R}$ in Definition 1.6, then $f : U \times U \rightarrow \mathbb{R}$ is a bilinear form in which we are used to quadratic forms, for example.

A simple example of bilinear application is the function $f : U \times V \rightarrow \mathbb{R}$ defined by

$$f(u, v) = h(u)g(v), \tag{11}$$

with $h \in U^*$ and $g \in V^*$, where U^* denotes the dual space to U. In fact, we have for all $\lambda \in \mathbb{R}, u_1, u_2 \in U$ and $v \in V$ such that

$$f(\lambda u_1 + u_2, v) = h(\lambda u_1 + u_2)g(v) = (\lambda h(u_1) + h(u_2))g(v) = \lambda f(u_1, v) + f(u_2, v).$$

Similarly, it is easy to see that $f(u, \lambda v_1 + v_2) = \lambda f(u, v_1) + f(u, v_2)$ for all $\lambda \in \mathbb{R}, u \in U$ and $v_1, v_2 \in V$.

The next theorem ensures that a bilinear application $f : U \times V \rightarrow W$ is well defined when the image of f applied in the bases elements of U and V is known.

Theorem 1.7. Let U, V, and W be vector spaces; $\{u_1, \ldots, u_m\}$ and $\{v_1, \ldots, v_n\}$ bases of the U and V, respectively; and $\{w_{ij} | i = 1, \ldots, m$ and $j = 1, \ldots, n\}$ a subset of W. Then, there exists an only bilinear application $f : U \times V \rightarrow W$ such that $f(u_i, v_j) = w_{ij}$.

Proof. Let $u = \sum_{i=1}^{m} \alpha_i u_i$ and $v = \sum_{j=1}^{n} \beta_j v_j$ be arbitrary elements of U and V, respectively. We defined an application $f : U \times V \rightarrow W$ by

$$f(u, v) = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i \beta_j w_{ij}.$$

It is easy to see that f is a bilinear application and $f(u_i, v_j) = w_{ij}$. Such an application is unique because if g is another bilinear application satisfying $g(u_i, v_j) = w_{ij}$, then

$$g(u, v) = g\left(\sum_{i=1}^{m} \alpha_i u_i, \sum_{j=1}^{n} \beta_j v_j\right) = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i \beta_j g(u_i, v_j) = \tag{12}$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i \beta_j w_{ij} = f(u, v). \tag{13}$$

Therefore $g = f$. $\qquad\square$

The following theorem guarantees the isomorphism between space of bilinear applications and space of tensor of order 3.

Theorem 1.8. Let U, V, and W be vector spaces with dimensions n, p, and m, respectively. Then, the space $\mathcal{B}(U \times V, W)$ has dimension mnp.

Proof. The idea of the proof is to exhibit a basis for space $\mathcal{B}(U \times V, W)$. For this, let $\{w_1, \ldots, w_m\}$, $\{u_1, \ldots, u_n\}$, and $\{v_1, \ldots, v_p\}$ be bases of the W, U, and V, respectively. For each triple (i, j, k), with $i = 1, \ldots, m, j = 1, \ldots, n$, and $k = 1, \ldots, p$, we define a bilinear application $f_{ij}^k : U \times V \to W$ such that

$$f_{ij}^k(u_r, v_s) = \begin{cases} w_i & \text{if } r = j \text{ and } s = k \\ 0 & \text{if } r \neq j \text{ or } s \neq k. \end{cases} \tag{14}$$

Theorem 1.7 ensures the existence of the f_{ij}^k. We will then show that the set

$$A = \left\{ f_{ij}^k | i = 1, \ldots, m, j = 1, \ldots, n \text{ and } k = 1, \ldots, p \right\}$$

is a basis of the space $\mathcal{B}(U \times V, W)$. Let $f \in \mathcal{B}(U \times V, W)$. We note in passing that

$$f(u_r, v_s) = \sum_{i=1}^{m} a_{ir}^s w_i \tag{15}$$

for all $r = 1, \ldots, n$ and $s = 1, \ldots, p$. Consider the bilinear application

$$g = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} a_{ij}^k f_{ij}^k.$$

Our goal is to show that $g = f$. In particular, we have

$$g(u_r, v_s) = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} a_{ij}^k f_{ij}^k(u_r, v_s) = \sum_{i=1}^{m} a_{ir}^s w_i = f(u_r, v_s)$$

for all $r = 1, \ldots, n$ and $s = 1, \ldots, p$. Therefore $g = f$. The set A is linearly independent, because if

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} a_{ij}^k f_{ij}^k = 0,$$

then

$$0 = \sum_{k=1}^{p} \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^k f_{ij}^k(u_r, v_s) = \sum_{i=1}^{m} a_{ir}^s w_i.$$

Since $\{w_1, \ldots, w_m\}$ is a basis of W, it follows that $a_{ir}^s = 0$ for all $i = 1, \ldots, m, r = 1, \ldots, n$, and $k = 1, \ldots, p$. $\square$

In particular, if the dimensions of the vector spaces U and V are m and n, respectively, then the vector space $\mathcal{B}(U \times V, \mathbb{R})$ has dimension mn. Now, as two vector spaces of the same finite dimension are isomorphic [16], there exists a matrix

$m \times n$ associated with each $f \in \mathcal{B}(U \times V, \mathbb{R})$. By considering $B = \{u_1, \dots, u_m\}$ and $C = \{v_1, \dots, v_n\}$ bases of U and V, respectively, and if $u = \sum_{i=1}^{m} \alpha_i u_i$ and $v = \sum_{j=1}^{n} \beta_j v_j$, then by doing $f(u_i, v_j) = a_{ij}$ for all $i = 1, \dots, m$ and $j = 1, \dots, n$, we have

$$f(u, v) = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i a_{ij} \beta_j,$$

which in matrix form is $f(u, v) = [u]_B^T A [v]_C$, where $A = (a_{ij})$ and $[v]_C$ denote the vector components v in the basis C. Hence follows the next definition:

Definition 1.9. Let U and V be vector spaces of finite dimension and ordered bases $B = \{u_1, \dots, u_m\} \subset U$ and $C = \{v_1, \dots, v_n\} \subset V$. We define, for each $f \in \mathcal{B}(U \times V, \mathbb{R})$, the matrix $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ of the f relative to the ordered bases B and C, whose elements are given by $a_{ij} = f(u_i, v_j)$ with $i = 1, \dots, m$ and $j = 1, \dots, n$.

Consider now the space $\mathcal{B}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^p)$ and the canonical bases $\{e_1, \dots, e_m\}$, $\{\bar{e}_1, \dots, \bar{e}_n\}$, $\{\hat{e}_1, \dots, \hat{e}_p\}$ of the $\mathbb{R}^m$, $\mathbb{R}^n$, and $\mathbb{R}^p$, respectively. Consider $f \in \mathcal{B}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^p)$. For all $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$, we have

$$f(u, v) = \sum_{j=1}^{m} \sum_{k=1}^{n} u_j v_k f(e_j, \bar{e}_k)$$

where u_j and v_k are the components of the u and v in the canonical bases of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. Denote the i-th component of the f by f_i. Note that $f_i \in \mathcal{B}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R})$. So for each $i = 1, \dots, p$, we have

$$f_i(u, v) = \sum_{j=1}^{m} \sum_{k=1}^{n} u_j v_k f_i(e_j, \bar{e}_k).$$

By Definition 1.9, the matrix of the f_i in relation to the canonical bases is the matrix

$$A_i = (t_{ij}^k) \in \mathbb{R}^{m \times n},$$

where $t_{ij}^k = f_i(e_j, \bar{e}_k)$. So, we can write

$$f_i(u, v) = u^T A_i v.$$

In general, we can define p matrices $m \times n$ as a tensor $\mathcal{T} \in \mathbb{R}^{p \times m \times n}$; this means that the p matrices can be seen as the horizontal slices of the tensor $\mathcal{T}$. We note in passing that we can write $f(u, v)$ as a product between tensor $\mathcal{T}$ and vectors u and v, that is,

$$f(u, v) = \begin{pmatrix} u^T A_1 v \\ u^T A_2 v \\ \vdots \\ u^T A_p v \end{pmatrix} = (\mathcal{T} \overline{\times}_2 u) v. \tag{16}$$

Thus, we can generalize Definition 1.9 as follows:

Definition 1.10. Let U and V be finite dimension vector spaces. For fixed bases $B = \{u_1, \ldots, u_m\}$ and $C = \{v_1, \ldots, v_n\}$ of the U and V, respectively, we define, for each $f \in \mathcal{B}(U \times V, \mathbb{R}^p)$, the tensor $\mathcal{T} = \left(t_{ij}^k\right) \in \mathbb{R}^{p \times m \times n}$ of the f relative to the ordered bases B and C, whose elements are given by $t_{ij}^k = f_i(u_j, v_k)$ where f_i is the i-th component of the f, that is, $f_i \in \mathcal{B}(U \times V, \mathbb{R})$, with $i = 1, \ldots, p$, $j = 1, \ldots, m$, and $k = 1, \ldots, n$.

4. Differentiability

Let U be an open subset of $\mathbb{R}^m$ and $F : U \subset \mathbb{R}^m \to \mathbb{R}^n$ a differentiable application throughout U and $a \in U$. Denote $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ the set of all linear applications of $\mathbb{R}^m$ in $\mathbb{R}^n$. When $F' : U \subset \mathbb{R}^m \to \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ is differentiable in $a \in U$, we say that the application F is twice differentiable in $a \in U$ and then the linear transformation $F''(a) \in \mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$ is the second derivative of F in $a \in U$.

The norm of $F''(a)$ is naturally defined. For any $h \in \mathbb{R}^m$, it follows that

$$\left\|F''(a)h\right\| = \sup_{\|k\|=1} \left\{\left\|F''(a)hk\right\| \operatorname{com} k \in \mathbb{R}^m\right\}$$

and then

$$\left\|F''(a)\right\| = \sup_{\|h\|=1} \left\|F''(a)h\right\| = \sup_{\|h\|=1} \sup_{\|k\|=1} \left\|F''(a)hk\right\|.$$

An important observation with respect to Theorem 1.8 is that the spaces $\mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$ and $\mathcal{B}(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$ are isomorphic. This means that $F''(a)$ is a bilinear application belonging to space $\mathcal{B}(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$. Such isomorphism can be found in classical analysis books [17, 18]. On the other hand, by the same theorem, the space of bilinear applications $\mathcal{B}(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$ and space of tensor $\mathbb{R}^{n \times m \times m}$ are also isomorphic.

In many practical applications, such as algorithm implementations, the second derivative $F''(a)$ may be implemented as a tensor belonging to space $\mathbb{R}^{n \times m \times m}$. The question now is how the tensor elements are formed. For this, consider the application $A : \mathbb{R} \to \mathbb{R}^{n \times m}$ and $\alpha \in \mathbb{R}$. We have $A(\alpha)$ as a matrix with n rows and m columns. Its elements are denoted by $a_{ij}(\alpha)$ where a_{ij} are components functions of A with $i = 1, \ldots, n$ and $j = 1, \ldots, m$. Case $a_{ij} : \mathbb{R} \to \mathbb{R}$ is differentiable in α for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$; the derivative of A in α is the matrix

$$A'(\alpha) = \left(a'_{ij}(\alpha)\right) \in \mathbb{R}^{n \times m}. \tag{17}$$

The definition of the derivative of $A(\alpha)$ (17) is a classical definition. We refer to [19] for details.

In the sense of generalizing (17), consider now $A : U \subset \mathbb{R}^p \to \mathbb{R}^{n \times m}$ a differentiable application in $u \in U$ with component function $a_{ij} : \mathbb{R}^p \to \mathbb{R}$ with $i = 1, \ldots, n$ and $j = 1, \ldots, m$. When a_{ij} is differentiable in u for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$, we defined the derivative of A in u as the tensor

$$A'(u) = \left(\nabla a_{ij}(u)\right) \in \mathbb{R}^{n \times m \times p}. \tag{18}$$

Note that in fact (18) is a generalization of (17). With fixed i and j, $\nabla a_{ij}(u)$ is a tube fiber of the tensor $A'(u)$, whose elements are

$$A'(u)_{ij}^k = \frac{\partial a_{ij}}{\partial x_k}(u) \tag{19}$$

for all $k = 1, \ldots, p$.

For example, consider an application $F : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ twice differentiable in $a \in U$ where U is an open set. The Jacobian matrix of F in a is given by

$$J_F(a) = \begin{pmatrix} \nabla f_1(a)^T \\ \nabla f_2(a)^T \\ \nabla f_3(a)^T \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dfrac{\partial f_1}{\partial x_2}(a) \\[2ex] \dfrac{\partial f_2}{\partial x_1}(a) & \dfrac{\partial f_2}{\partial x_2}(a) \\[2ex] \dfrac{\partial f_3}{\partial x_1}(a) & \dfrac{\partial f_3}{\partial x_2}(a) \end{pmatrix}$$

and its derivative is, by (18), the tensor

$$J_{F'}(a) = \mathcal{T}_F(a) = \left(\nabla \frac{\partial f_i}{\partial x_j}(a) \right) \in \mathbb{R}^{3 \times 2 \times 2} \tag{20}$$

where, by (19), its elements are described as

$$t_{ij}^k = \frac{\partial^2 f_i}{\partial x_k \partial x_j}(a).$$

With fixed i, it is easy to see that the i-th horizontal slice of the $\mathcal{T}_F(a)$ is the Hessian matrix $\nabla^2 f_i(a)$ defined by

$$\nabla^2 f_i(a) = \mathcal{T}_F(a)^{i::} = \begin{pmatrix} \dfrac{\partial^2 f_i}{\partial x_1 \partial x_1}(a) & \dfrac{\partial^2 f_i}{\partial x_1 \partial x_2}(a) \\[2ex] \dfrac{\partial^2 f_i}{\partial x_2 \partial x_1}(a) & \dfrac{\partial^2 f_i}{\partial x_2 \partial x_2}(a) \end{pmatrix}. \tag{21}$$

We note in passing that any column of the matrix $\nabla^2 f_i(x)$ is a row fiber of the i-th horizontal slice.

As mentioned in the introduction, some numerical methods need to calculate the product between tensor $\mathcal{T}_F(a)$ and vectors in $\mathbb{R}^2$.

From Definition 1.4, it is possible to calculate the contracted product mode-2 and mode-3. As Hessian matrices are symmetrical, given $v \in \mathbb{R}^2$, by Lemma 1.5 together with (10) and (11), we have

$$\mathcal{T}_F(a)\overline{\times}_3 v = \mathcal{T}_F(a)\overline{\times}_2 v = \begin{pmatrix} \mathrm{row}_1(\mathcal{T}_F(a)\overline{\times}_2 v) \\ \mathrm{row}_2(\mathcal{T}_F(a)\overline{\times}_2 v) \\ \mathrm{row}_3(\mathcal{T}_F(a)\overline{\times}_2 v) \end{pmatrix} = \begin{pmatrix} v^T \nabla^2 f_1(a) \\ v^T \nabla^2 f_2(a) \\ v^T \nabla^2 f_3(a) \end{pmatrix} \in \mathbb{R}^{3 \times 2}$$

and consequently it follows that

$$(\mathcal{T}_F(a) \overline{\times}_2 v)u = \begin{pmatrix} v^T \nabla^2 f_1(a)u \\ v^T \nabla^2 f_2(a)u \\ v^T \nabla^2 f_3(a)u \end{pmatrix} \in \mathbb{R}^3 \qquad (22)$$

for all $u, v \in \mathbb{R}^2$.

This means that the tensor $\mathcal{T}_F(a)$ defined by (20) is the associate tensor to bilinear application $F''(a)$, in relation to canonical basis of $\mathbb{R}^2$, by means of Definition 1.10. Without loss of generality, we have

$$\mathcal{T}_F(a) \overline{\times}_3 v = \mathcal{T}_F(a) \overline{\times}_2 v = \mathcal{T}_F(a)v$$

and by means of Lemma 1.5, it follows that

$$(\mathcal{T}_F(a)u)v = (\mathcal{T}_F(a)v)u = \mathcal{T}_F(a)vu.$$

To finish, we consider the following particular case. We know that the k-th column of Jacobian $J_F(x)$ is equal to product $J_F(x)e_k$, where e_k is the k-th canonical vector of $\mathbb{R}^n$. It is worth noting what the slice of the matrix $\mathcal{T}_F(x)e_k$ is. By definition, we have

$$\mathcal{T}_F(x)e_k = \begin{pmatrix} e_k^T \nabla^2 f_1(x) \\ e_k^T \nabla^2 f_2(x) \\ \vdots \\ e_k^T \nabla^2 f_n(x) \end{pmatrix} = \begin{pmatrix} \text{row}_k \nabla^2 f_1(x) \\ \text{row}_k \nabla^2 f_2(x) \\ \vdots \\ \text{row}_k \nabla^2 f_n(x) \end{pmatrix}$$

Given that $\text{row}_k \nabla^2 f_i(x)$ is the k-th tube fiber of i-th horizontal slice, we have $\mathcal{T}_F(x)e_k$ as the k-th lateral slice, or, by symmetry of Hessians, it is the transpose of k-th frontal slice. In short, for the twice differentiable application $F : U \subset \mathbb{R}^n \to \mathbb{R}^m$, we have $\mathcal{T}_F(x) \in \mathbb{R}^{m \times n \times n}$ where the m horizontal slices are the Hessians $\nabla^2 f_i(x)$, with $i = 1, \ldots, m$ and the n lateral and frontal slices obtained by the following product $\mathcal{T}_F(x)e_k$, with $k = 1, \ldots, n$.

5. Conclusions

In this text, we have shown some properties of tensors, in particular those of order 3. In addition, we have approached bilinear applications, and we have shown the isomorphism between space of bilinear applications and of tensor of order 3. As mentioned in the introduction, to solve a nonlinear system, some numerical methods use tensors, either in the iterative scheme or in the proof of theorems. For this reason, we have written a section on differentiability of applications by showing how to calculate the product between tensor of second derivatives and vectors.

Author details

Rodrigo Garcia Eustaquio
Department of Mathematics, Federal Technological University of Paraná, Curitiba, PR, Brazil

*Address all correspondence to: eustaquio@utfpr.edu.br;
rodrigogeustaquio@gmail.com

IntechOpen

References

[1] Hernández MA, Gutiérrez JM. A family of Chebyshev-Halley type methods in Banach spaces. Bulletin of the Australian Mathematical Society. 1997;**55**:113-130

[2] Ehle GP, Schwetlick H. Discretized Euler-Chebyshev multistep methods. SIAM Journal on Numerical Analysis. 1976;**13**(3):432-447

[3] Eustaquio RG, Ribeiro AA, Dumett MA. A new class of root-finding methods in $\mathbb{R}^n$: The inexact tensor-free Chebyshev-Halley class. Computational and Applied Mathematics. 2018;**37**: 6654-6675

[4] Traub JF. Iterative Methods for the Solution of Equations. Prentice-Hall Series in Automatic Computation. Englewood Cliffs, NJ: Prentice-Hall; 1964

[5] Bader BW, Kolda TG. Algorithm 862: MATLAB tensor classes for fast algorithm prototyping. ACM Transactions on Mathematical Software. 2006;**32**(4):635-653. DOI: 10.1145/1186785.1186794

[6] Cichocki A, Zdunek R, Phan AH, Amari S. Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multiway Data Analysis and Blind Source Separation. John Wiley Sons, Ltd; 2009

[7] Kolda TG, Bader BW. Tensor decompositions and applications. SIAM Review. 2009;**51**(3):455-500

[8] De Lathauwer L, De Moor B, VandeWalle J. A multilinear singular value decomposition. SIAM Journal on Matrix Analysis and Applications. 2000;**21**:1253-1278

[9] Smilde A, Bro R, Geladi P. Multi-Way Analysis: Applications in the Chemical Sciences. Wiley; 2004

[10] Chen B, Petropulu A, De Lathauwer L. Blind identification of convolutive MIMO systems with 3 sources and 2 sensors. Applied Signal Processing. 2002;**5**:487-496. Special Issue: Space-time Coding and Its Applications—Part II. Available from: http://publi-etis.ensea.fr/2002/CPD02a

[11] Golub GH, Kolda TG, Nagy JG, Van Loan CF. Workshop on Tensor Decompositions. Palo Alto, California: American Institute of Mathematics; 2004. Available from http://www.aimath.org/WWN/tensordecomp/

[12] De Lathauwer L, Comon P. Workshop on Tensor Decompositions and Applications. Marseille, France; 2005. Available from: http://www.etis.ensea.fr/wtda/

[13] Bader BW, Kolda TG. Efficient MATLAB Computations with Sparse and Factored Tensors. Albuquerque, NM/Livermore, CA: Sandia National Laboratories; 2006. SAND2006-7592. Available from: http://www.prod.sandia.gov/cgi-bin/techlib/access-control.pl/2006/067592.pdf

[14] Ishteva M. Numerical methods for the best low multilinear rank approximation of higher-order tensors [PhD thesis]. Belgium: Faculty of Engineering, Katholieke Universiteit Leuven; 2009

[15] Kiers HAL. Towards a standardized notation and terminology in multiway analysis. Journal of Chemometrics. 2000;**14**:105-122

[16] Coelho FU, Loureno ML. Um Curso de lgebra Linear. So Paulo, Brasil: Editora da Universidade de So Paulo; 2007

[17] Lima EL. Anlise no Espao $\mathbb{R}^n$. So Paulo: Editora Universidade de Braslia; 1970

[18] Lima EL. Curso de Anlise, Volume 2. Rio de Janeiro, Brasil: IMPA; 1981

[19] Golub GH, Van Loan CF. Matrix Computations. The Johns Hopkins University Press; 1996