



HAL
open science

Tenseurs en sciences de données

Pierre Comon

► **To cite this version:**

Pierre Comon. Tenseurs en sciences de données. Techniques de l'ingénieur. Mathématiques pour l'ingénieur, , 2021, 10.51257/a-v1-af114 . hal-03266497v2

HAL Id: hal-03266497

<https://hal.science/hal-03266497v2>

Submitted on 15 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Tenseurs

en sciences des données

Tensors in Data Sciences

par **Pierre COMON**

Directeur de Recherche

Univ. Grenoble Alpes, CNRS, Grenoble INP, Gipsa-Lab, 38000 Grenoble

Résumé

Le principal argument souvent avancé en physique pour utiliser les tenseurs est leur définition intrinsèque permettant l'invariance de leurs propriétés vis à vis du système de coordonnées. Dans cet article, un autre intérêt des tenseurs est mis en avant, à savoir l'unicité de leur décomposition en somme de tenseurs simples. Cette unicité permet d'identifier ces tenseurs simples à des grandeurs ayant un sens physique. Cette propriété unique, décrite en détail dans cet article, a inspiré de nombreux travaux ces dernières années dans des domaines applicatifs très variés, notamment en science des données, dont il est donné un aperçu.

Abstract

The main argument often put forward in physics to use tensors is their intrinsic definition allowing the invariance of their properties with respect to the coordinate system. In this paper, another interest of tensors is put forward, namely the uniqueness of their decomposition into a sum of simple tensors. This uniqueness allows to identify these simple tensors to quantities having a physical meaning. This unique property, described in detail in this article, has inspired numerous works in recent years in a wide variety of application domains, particularly in data science, which is outlined here.

Mots-clés

rang tensoriel, décomposition Canonique Polyadique (CP), Tucker, MLSVD, séparation aveugle de sources, traitement d'antenne, spectrométrie fluorescente, complexité arithmétique, polynômes multi-variés, compression de données

Keywords

tensor rank, Canonical Polyadic (CP) decomposition, Tucker, MLSVD, blind source separation, antenna array processing, fluorescence spectroscopy, arithmetic complexity, multivariate polynomials, data compression

Table des matières

1 Définitions et propriétés	3
1.1 Applications multilinéaires	3
1.2 Exemples	5
1.3 Notations et transformations	6
1.3.1 Contraction	6
1.3.2 Changement de base	7
1.3.3 Produit tensoriel	9
1.3.4 Norme euclidienne	9
2 Décompositions tensorielles exactes	10
2.1 Rang tensoriel, décomposition Canonique Polyadique (CP)	10
2.2 Lien avec les polynômes multivariés	13
2.3 Rang multilinéaire (Tucker, MLSVD)	14
2.4 Autres rangs (non-négatif, symétrique, orthogonal)	15
2.5 Trains tensoriels	16
3 Décompositions tensorielles approchées	17
3.1 Diagonalisation orthogonale approximative	17
3.2 Approximation de faible rang multilinéaire	17
3.3 Approximation de rang 1	18
3.4 Rang frontière	19
3.5 Approximation de faible rang sous contrainte	19
4 Applications	20
4.1 Analyse en Composantes Indépendantes	20
4.1.1 Identification de mélanges linéaires	21
4.1.2 Maximisation d'un contraste quand $P \leq N$	22
4.1.3 Séparation aveugle de sources indépendantes	23
4.1.4 Fouille documentaire	24
4.2 Diversités physiques	25
4.2.1 Spectrométrie fluorescente	25
4.2.2 Imagerie planétaire	26
4.2.3 Traitement d'antenne	27
5 Conclusion	29
6 Annexes	30
6.1 Notations supplémentaires	30
6.2 Démonstration de la propriété 5 et de (12)	31
6.3 Démonstration de la propriété 10 et de (23)	31
7 Glossaire	31

Introduction

Les tenseurs sont des objets utilisés depuis longtemps en physique, car ils jouissent de propriétés d'invariance vis à vis des systèmes de coordonnées utilisés. Ils apparaissent aussi lorsqu'on veut évaluer la complexité arithmétique de certains problèmes. Enfin, on les rencontre en statistiques, avec les moments et les cumulants de variables aléatoires multivariées. Mais plus récemment, les tenseurs ont fait leur entrée dans d'autres secteurs du monde de l'ingénieur, tels que les télécommunications, l'ingénierie biomédicale, la chimométrie, le traitement du signal, et bien d'autres. Pourtant ce n'est pas l'indépendance du système de coordonnées dans leur définition qui a permis aux tenseurs de revenir au coeur des outils algébriques utilisés par les ingénieurs. Alors que s'est-il passé ?

Une des propriétés fondamentales des tenseurs est qu'ils peuvent se décomposer de manière unique en une somme de tenseurs plus simples, sous des hypothèses assez faibles. Et dans bien des situations, ces termes plus simples revêtent une signification physique intéressante. C'est cette propriété d'unicité qui leur a valu ce regain d'intérêt depuis une douzaine d'années. Nous décrivons cette propriété fondamentale dans la section 2. Toutefois, en présence d'erreurs de mesure (e.g. bruit), le tenseur mesuré n'a pas le rang escompté de sorte qu'il faudrait calculer la meilleure approximation de rang faible. Or cette approximation n'existe pas toujours comme nous le montrons dans la section 3, ce qui complique l'implantation des algorithmes. Enfin plusieurs applications phare sont décrites dans la section 4.

Nota : le lecteur trouvera en fin d'article un glossaire des termes et expressions importants, ainsi qu'un tableau des sigles, notations et symboles utilisés tout au long de l'article.

1 Définitions et propriétés

On notera les vecteurs avec des lettres grasses minuscules, e.g. \mathbf{x} , les matrices avec des majuscules grasses, e.g. \mathbf{M} , et les tenseurs par des lettres calligraphiques grasses, e.g. \mathcal{T} . L'opérateur de transposition d'un vecteur ou d'une matrice sera noté $(^T)$; dans \mathbb{C} l'opérateur de transposition hermitienne sera noté $(^H)$. Les composantes des vecteurs, matrices et tenseurs dans un système de coordonnées prédéfini seront des nombres scalaires, représentés par des lettres non grasses indexées, par exemple $x_i, M_{ij}, \mathcal{T}_{ijk}$. Dans la mesure du possible, un indice n variera entre 1 et N , i.e. la même lettre sera choisie pour la borne supérieure de l'intervalle, mais en majuscule.

1.1 Applications multilinéaires

Dimension. La dimension d'un espace vectoriel (de dimension finie) est le nombre de vecteurs linéairement indépendants nécessaires pour en constituer une base. Le corps de construction est de dimension 1, car dans la base un seul nombre non nul suffit.

Linéarité. Soient E et F deux espaces vectoriels, de dimension finie M et N respectivement. Le corps de construction \mathbb{K} des espaces désignera \mathbb{R} ou \mathbb{C} , suivant le besoin. Une application g de E vers F est linéaire si elle vérifie la propriété suivante :

$$g(\alpha \mathbf{x} + \alpha' \mathbf{x}') = \alpha g(\mathbf{x}) + \alpha' g(\mathbf{x}'), \quad \forall \mathbf{x}, \mathbf{x}' \in E, \forall \alpha, \alpha' \in \mathbb{K}.$$

Si on fixe des bases pour les espaces E et F , nous savons que tout vecteur de E (resp. F) peut être représenté par un tableau de M (resp. N) coordonnées, et que toute application linéaire de E vers F peut être représentée par une matrice de dimensions $M \times N$.

Bilinéarité. Si E_1 et E_2 sont deux espaces vectoriels, une application $g(\mathbf{x}_1, \mathbf{x}_2)$ de $E_1 \times E_2$ dans F est bilinéaire si elle est linéaire par rapport à \mathbf{x}_1 et \mathbf{x}_2 . Autrement dit, $g(\alpha \mathbf{x}_1 + \alpha' \mathbf{x}'_1, \mathbf{x}_2) = \alpha g(\mathbf{x}_1, \mathbf{x}_2) + \alpha' g(\mathbf{x}'_1, \mathbf{x}_2)$ et $g(\mathbf{x}_1, \beta \mathbf{x}_2 + \beta' \mathbf{x}'_2) = \beta g(\mathbf{x}_1, \mathbf{x}_2) + \beta' g(\mathbf{x}_1, \mathbf{x}'_2)$. Un exemple typique est la forme bilinéaire réelle que nous connaissons bien, qui à tout couple de vecteurs réels $(\mathbf{x}_1, \mathbf{x}_2)$ associe le nombre $g(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1^T \mathbf{A} \mathbf{x}_2 \in \mathbb{K}$, où $(^T)$ désigne la transposition.

Multi-linéarité. Dans cet article, nous allons étudier les applications multilinéaires. Soient E_i des espaces vectoriels de dimension M_i , et F_j des espaces vectoriels de dimension N_j , $1 \leq i \leq I$, $1 \leq j \leq J$. Si \mathbf{g} désigne une application multilinéaire de $E_1 \times \dots \times E_I$ dans $F_1 \times \dots \times F_J$, elle vérifie $\forall i, 1 \leq i \leq I$:

$$\mathbf{g}(\mathbf{x}_1, \dots, \alpha \mathbf{x}_i + \alpha' \mathbf{x}'_i, \dots, \mathbf{x}_I) = \alpha \mathbf{g}(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_I) + \alpha' \mathbf{g}(\mathbf{x}_1, \dots, \mathbf{x}'_i, \dots, \mathbf{x}_I),$$

pour tout $\alpha, \alpha' \in \mathbb{K}$, et pour tous les vecteurs $\mathbf{x}_i, \mathbf{x}'_i \in E_i$.

Tableau de coordonnées. La propriété essentielle avancée généralement pour argumenter l'usage des tenseurs en physique est leur indépendance du système de coordonnées, c'est à dire des bases choisies pour les différents espaces [1, 2] [A1210] [A125]. Dans cet article, ce n'est précisément *pas cette propriété* qui est mise en avant, de sorte que nous pouvons travailler avec les tableaux de coordonnées. C'est ce que nous allons faire pour alléger les notations. Assimiler les vecteurs, les matrices, et plus généralement tous les tenseurs, à leur *tableau de coordonnées*, est légitime en dimension finie si des bases ont été préalablement définies dans chaque espace, ce que nous supposerons. En revanche les coordonnées d'un tenseur changeront évidemment lorsque les bases changent, de manière à ce que l'application multilinéaire que représente ce tenseur reste la même.

Ordre et valence. On appelle *valence* le couple (I, J) , où I désigne le nombre d'espaces de départ et J celui des espaces d'arrivée (on ne compte ici que les espaces de dimension supérieure à 1). On appelle *ordre* le nombre total $D = I + J$ d'espaces vectoriels impliqués dans la définition d'une forme multilinéaire. Pour une application linéaire, $I = J = 1$. Pour une forme bilinéaire, telle qu'un produit scalaire, $I = 2$, et l'ensemble d'arrivée est le corps de construction lui-même de sorte que $J = 0$.

Exemple 1 : forme trilinéaire

Si on prend une forme trilinéaire $g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ sur \mathbb{R} , $I = 3$ et $J = 0$, de sorte que $D = 3$. A cause de la multi-linéarité de g , nous avons pour tout triplet $(\alpha, \beta, \gamma) \in \mathbb{K}^3$ la propriété suivante :

$$g(\alpha \mathbf{x}_1, \beta \mathbf{x}_2, \gamma \mathbf{x}_3) = \alpha \beta \gamma g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3). \quad (1)$$

Ainsi, les triplets de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ et $(\alpha \mathbf{x}_1, \mathbf{x}_2, \frac{1}{\alpha} \mathbf{x}_3)$ ont la même image par toute forme trilinéaire.

Tenseurs. Fort de cette propriété (1), on peut construire une **classe d'équivalence**, notée $\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3$ à laquelle tous les triplets de vecteurs de la forme (1) appartiennent [1, 3].

Définition D1. Le sous-espace vectoriel engendré par ces classes d'équivalences est noté $E_1 \otimes E_2 \otimes E_3$. Le symbole \otimes désigne le produit tensoriel, et les éléments de $E_1 \otimes E_2 \otimes E_3$ sont appelés *tenseurs*.

Exemple 2 : produit tensoriel de deux vecteurs

Soient $\mathbf{x}_1 \in E_1$ et $\mathbf{x}_2 \in E_2$. Les vecteurs $(5\mathbf{x}_1, 2\mathbf{x}_2)$ et $(10\mathbf{x}_1, \mathbf{x}_2)$ sont différents dans $E_1 \times E_2$. Mais les tenseurs $5\mathbf{x}_1 \otimes 2\mathbf{x}_2$ et $10\mathbf{x}_1 \otimes \mathbf{x}_2$ sont identiques dans $E_1 \otimes E_2$.

On notera que les *tenseurs élémentaires* associés aux formes multilinéaires (1) sont les plus simples, et sont parfois appelés *tenseurs simples*, *purs*, *décomposables*, *élémentaires* ou *séparables*. Dans le présent article, nous adopterons la terminologie de "tenseurs simples" comme dans [4]. On verra plus loin, lorsque le rang tensoriel sera défini, que de tels tenseurs ont un rang égal à 1.

Exemple 3 : discrétisation d'une fonction à variables séparées.

Si on échantillonne une fonction à variables séparées, $f(x, y, z) = a(x)b(y)c(z)$, on obtient un tableau \mathcal{D}_{ijk} qui peut s'écrire comme un tenseur simple : $\mathcal{D}_{ijk} = a_i b_j c_k$.

Nous entrevoyons déjà l'écueil que représente la manipulation des tableaux de coordonnées : il faut choisir un représentant dans la classe d'équivalence, et garder à l'esprit qu'il est défini aux facteurs d'échelle (α, β, γ) près. Nous adopterons la définition suivante.

Définition D2. Soient E_i et F_j des espaces de dimension supérieure à 1, $1 \leq i \leq I$, $1 \leq j \leq J$. Une application multilinéaire de $E_1 \otimes \cdots \otimes E_I$ dans $F_1 \otimes \cdots \otimes F_J$, avec $D = I + J$, est un tenseur d'ordre D . Si les bases des espaces E_i et F_j sont fixées, alors de tels tenseurs sont définis par un tableau à D indices.

Parmi ces D indices, I indices sont dits "contravariants" et J indices sont dits "covariants" [A1210]. Sauf mention contraire, nous nous intéresserons essentiellement aux tenseurs entièrement contravariants, c'est à dire aux formes multilinéaires ($D = I$ et $J = 0$), comme dans l'exemple 1, ou dans l'exemple 4 ci-après.

Pour terminer, le terme de "rang" a été utilisé à une époque en physique pour désigner l'ordre des tenseurs, mais nous proscrivons cette terminologie car le rang désigne déjà un autre concept en mathématiques, bien connu pour les opérateurs linéaires (représentés par des matrices). Le *rang tensoriel* sera défini un peu plus loin dans cet article, dans la section 2.1.

1.2 Exemples

Exemple 4 : forme trilinéaire

Soient $\mathbf{x} \in E_1$, $\mathbf{y} \in E_2$ et $\mathbf{z} \in E_3$, trois vecteurs. L'application de $E_1 \otimes E_2 \otimes E_3$ dans \mathbb{K} définie par

$$g(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{i,j,k} \mathcal{T}_{ijk} x_i y_j z_k$$

est un tenseur d'ordre 3. Nous voyons que l'image peut être vue comme un polynôme homogène de degré global 3, mais linéaire par rapport à chacune des 3 variables. C'est cette propriété du polynôme qui permet de lui conférer l'invariance par changement d'échelle décrite dans l'équation (1).

De manière plus générale et comme nous le verrons dans la section 2.2, tous les polynômes homogènes de degré D sont associés bijectivement à des tenseurs d'ordre D [3]. Notons que certains tenseurs peuvent jouir de symétrie partielle, comme ceux associés à des polynômes du type $g(\mathbf{x}, \mathbf{x}, \mathbf{z})$, ou de symétrie totale comme ceux associés à $g(\mathbf{x}, \mathbf{x}, \mathbf{x})$. Un tenseur symétrique peut être représenté par un tableau de coordonnées symétrique [5], c'est à dire vérifiant $\mathcal{T}_{\sigma(i,j,k)} = \mathcal{T}_{ijk}$ pour toute permutation σ .

Exemple 5 : produit de deux nombres complexes

Le corps des complexes peut être vu comme un espace de dimension 2 sur \mathbb{R} . Ainsi, si on représente tout nombre complexe par le couple de ses parties réelle et imaginaire, alors le produit de deux nombres complexes (x_1, x_2) et (y_1, y_2) peut s'écrire dans \mathbb{R}^2 :

$$(z_1, z_2) = (x_1, x_2) \cdot (y_1, y_2) = (x_1 y_1 - x_2 y_2, x_1 y_2 + x_2 y_1).$$

L'application bilinéaire qui à deux nombres complexes associe leur produit est un tenseur d'ordre 3. Les bases canoniques étant fixées, elle est représentée par un tableau de coordonnées \mathcal{T} de dimension $2 \times 2 \times 2$ de sorte que les composantes de l'image peuvent s'écrire :

$$z_k = \sum_{i=1}^2 \sum_{j=1}^2 \mathcal{T}_{ijk} x_i y_j$$

En posant

$$\mathbf{T}_R = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathbf{T}_I = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

nous avons :

$$z_1 = (x_1, x_2) \mathbf{T}_R \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad \text{et} \quad z_2 = (x_1, x_2) \mathbf{T}_I \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Pour des raisons pratiques, il pourra être utile de ranger les valeurs du tableau \mathcal{T} dans une matrice unique, par exemple dans la matrice $\mathbf{T}_{(1)}$ suivante, de dimension 2×4 :

$$\mathbf{T}_{(1)} = [\mathbf{T}_R \mid \mathbf{T}_I].$$

Cette matrice $\mathbf{T}_{(1)}$ est la première “matrice dépliant” associée à \mathcal{T} .

Cet exemple a été utilisé dans notre cours à l'Univ. Grenoble Alpes en 2015-18 [6], et servira dans cet article à illustrer différents concepts de manière simple et concrète.

Exemple 6 : produit de deux matrices

Le produit de deux matrices \mathbf{A} et \mathbf{B} , de taille $M \times N$ et $N \times P$ respectivement, peut être vu comme une application bilinéaire $\mathbf{C} = g(\mathbf{A}, \mathbf{B})$. Or l'ensemble des matrices de taille $M \times N$ forme un espace vectoriel E_1 de dimension MN . Donc cette application bilinéaire g est définie de $E_1 \otimes E_2$ dans F , où E_2 est de dimension NP et F de dimension MP . Si les bases de ces espaces sont fixées (et il suffit de choisir comme base les matrices ayant un seul élément non nul, par exemple égal à 1), alors l'application bilinéaire g est définie par un tableau à trois indices, \mathcal{T}_{ijk} , avec $i \in \{1, \dots, MN\}$, $j \in \{1, \dots, NP\}$, et $k \in \{1, \dots, MP\}$. Nous ne donnons pas ici le tableau de coordonnées, car il est de dimensions trop grandes (même pour des matrices 2×2 , il est déjà de taille $4 \times 4 \times 4$).

Nota : Il existe un algorithme (proposé par Strassen en 1969) qui permet d'effectuer le produit de deux matrices 2×2 en seulement 7 multiplications [7, pp.10-12].

Le tenseur représentant le produit de deux matrices a fait l'objet de travaux de recherche très poussés, visant à réduire la complexité du calcul. En effet, le rang d'un tel tenseur représente le nombre minimal de multiplications nécessaires [7, ch.14]. Mais les algorithmes atteignant cette borne n'ont pas été découverts pour toutes les valeurs possibles des dimensions.

1.3 Notations et transformations

Comme nous allons travailler avec les tableaux de coordonnées, il est important de savoir comment ces derniers sont affectés par diverses transformations, dont les changements de base en particulier. Pour alléger les écritures, nous n'utiliserons pas la convention de sommation d'Einstein [2, 8, 9], mais plutôt des notations compactes.

1.3.1 Contraction

La première opération est la contraction sur un indice. Soient deux tenseurs \mathcal{A} et \mathcal{B} , d'ordre D_A et D_B respectivement. On peut effectuer un “produit” de ces deux tenseurs en sommant sur le p ème indice de \mathcal{A} et le q ème indice de \mathcal{B} , pourvu que les dimensions concernées soient compatibles. Dans ce cas, le tenseur obtenu, \mathcal{C} , est d'ordre $D_A + D_B - 2$. Un exemple suffit pour clarifier cette définition. Si $D_A = 3$, $D_B = 4$, $p = 2$ et $q = 3$, alors le tenseur \mathcal{C} peut être représenté par le tableau de coordonnées :

$$\mathcal{C}_{ijklm} = \sum_n \mathcal{A}_{inj} \mathcal{B}_{klnm}.$$