



HAL
open science

An algebraic algorithm for the ParaTuck-2 decomposition

Konstantin Usevich

► **To cite this version:**

Konstantin Usevich. An algebraic algorithm for the ParaTuck-2 decomposition. 2024. hal-03966869v2

HAL Id: hal-03966869

<https://hal.science/hal-03966869v2>

Preprint submitted on 26 Feb 2024

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.

An algebraic algorithm for the ParaTuck-2 decomposition

Konstantin Usevich^a

^a*Université de Lorraine, CNRS, CRAN, Nancy F-54000, France.*

Abstract

In this note we consider the problem of ParaTuck-2 decomposition of a three-way tensor. We provide an algebraic algorithm for finding the ParaTuck-2 decomposition for the case when the ParaTuck-2 ranks are smaller than the frontal dimensions of the tensors. Our approach relies only on linear algebra operations and is based on finding the kernel of a structured matrix constructed from the tensor.

1. Introduction

ParaTuck-2 decomposition of 3-rd order tensors (multi-way arrays), proposed in [HL96], can be viewed as a 2-level extension of the well-known CP (canonical polyadic) decomposition. It is relevant in several applications, such as chemometrics [Bro98] and telecommunications [FdA14]. A tensor (3-way array) $\mathcal{T} \in \mathbb{F}^{N_1 \times N_2 \times N_3}$ over a field \mathbb{F} ($\mathbb{F} = \mathbb{R}$ or \mathbb{C}) admits a ParaTuck-2 decomposition with rank (R, S) if its frontal slices can be written as

$$\mathcal{T}_{:, :, k} = \mathbf{A} \mathbf{D}_k(\mathbf{G}) \mathbf{F} \mathbf{D}_k(\mathbf{H}) \mathbf{B}^\top, \quad (1)$$

$\mathbf{A} \in \mathbb{F}^{N_1 \times R}$, $\mathbf{B} \in \mathbb{F}^{N_2 \times S}$, $\mathbf{F} \in \mathbb{F}^{R \times S}$, $\mathbf{G} \in \mathbb{F}^{R \times N_3}$, $\mathbf{H} \in \mathbb{F}^{S \times N_3}$, and $\mathbf{D}_k(\mathbf{G})$ denotes the diagonal matrix built from a k -th column of the matrix \mathbf{G} . In a very special case, the Paratuck-2 decomposition reduces to the CP decomposition (when $R = S$ and $\mathbf{F} = \mathbf{D}_k(\mathbf{H}) = \mathbf{I}_R$ for all k in (1)). Generalizations of Paratuck-2 were also proposed for higher orders [FdA14], but we focus in this paper only on the third-order case ((1)). In this paper, we also assume that $R \leq N_1, S \leq N_1$.

Similarly to the CP decomposition, ParaTuck-2 enjoys strong uniqueness properties [HL96]. Under some conditions, the factors \mathbf{A} , \mathbf{B} , \mathbf{F} , $\mathbf{D}_k(\mathbf{G})$, $\mathbf{D}_k(\mathbf{H})$ can be recovered uniquely, subject to scaling and permutation ambiguities (similarly to the CP model). Despite the usefulness and the nice features of ParaTuck-2, there are very few effective algorithms to find a factorization (1). Many standard optimization tools, such as alternating least squares [Bro98, pp. 68–71] suffer from slow convergence and local minima. In many applications and algorithms, however, it is often considered that at least one of the factors (for example \mathbf{A}) is known [dOFFB19], [Bro98, p. 213], which simplifies the problem considerably. In [Nas20] it was suggested that double contractions may lead to another way of computing updates in the alternating least squares strategy. There are also reformulations of the ParaTuck-2 decomposition as a structured polyadic decomposition [FdA14], but these approaches do not currently yield a reliable way to compute the decomposition.

In this paper, we focus on the case when $R \leq N_1, S \geq N_2$ and N_3 sufficiently large. We provide an algebraic algorithm to compute the ParaTuck-2 decomposition under mild assumptions that hold in the generic case. Our approach relies only on the linear algebraic operations and make use of nonlinearly structured matrices constructed from a tensor. To our knowledge, it is the first algebraic algorithm for ParaTuck-2 decomposition.

Notation. Tensor arrays, matrices, and vectors, will be respectively denoted by bold calligraphic letters, e.g. \mathcal{A} , with bold uppercase letters, e.g. \mathbf{M} , and with bold lowercase letters, e.g. \mathbf{u} ; corresponding entries will be denoted by \mathcal{A}_{ijk} , M_{ij} , and u_i . Operator \bullet_p denotes contraction on the p th index of a tensor; when contracted with a matrix, it is understood that summation is always performed on the second index of the matrix. For instance, $(\mathcal{A} \bullet_1 \mathbf{M})_{ijk} = \sum_\ell \mathcal{A}_{\ell jk} M_{i\ell}$. We denote by $\mathbf{T}^{(1)} \in \mathbb{F}^{N_1 \times (N_2 N_3)}$ and $\mathbf{T}^{(2)} \in \mathbb{F}^{N_2 \times (N_1 N_3)}$, $\mathbf{T}^{(3)} \in \mathbb{F}^{N_3 \times (N_1 N_2)}$ the first, second and third unfoldings of a tensor \mathcal{T} , respectively. Let $\text{vec}\{\mathcal{T}\} \in \mathbb{F}^{N_1 N_2 N_3}$

Email address: `firstname.lastname@cnrs.fr` (Konstantin Usevich)

denote the vectorization of a tensor, and by $\text{matr}_{N_1, N_2}\{\mathbf{v}\}$ we denote for a vector $\mathbf{v} \in \mathbb{F}^{N_1 N_2}$ the matrix \mathbf{V} such that $\mathbf{v} = \text{vec}\{\mathbf{V}\}$.

2. Core tensor and its properties

2.1. Core tensor and its decomposition

In what follows, for simplicity, we assume $\mathbb{F} = \mathbb{R}$ (however, the case $\mathbb{F} = \mathbb{C}$ can be treated similarly). We first note that the Paratuck-2 decomposition can be compactly written using contractions

$$\mathcal{T} = \mathcal{C} \underset{1}{\bullet} \mathbf{A} \underset{2}{\bullet} \mathbf{B}, \quad (2)$$

where $\mathcal{C} \in \mathbb{F}^{R \times S \times N_3}$ is the core tensor with slices

$$\mathcal{C}_{:, :, k} = \mathbf{D}_k(\mathbf{G}) \mathbf{F} \mathbf{D}_k(\mathbf{H}). \quad (3)$$

It is easy to see that (3) can be equivalently written as:

$$\mathcal{C}_{ijk} = F_{ij} G_{ik} H_{jk}, \quad (4)$$

where (4) can be viewed as a generalization of the tensor product of three vectors to a product of three matrices.

If the core tensor is known, it is easy to find its factors under a simple assumption.

Assumption 1. *There exists one fixed index r such that $\mathcal{C}_{ijr} \neq 0$ for any i, j .*

Under Assumption 1, the factors can be obtained from the ratios of the slices. Indeed, for any k

$$\frac{\mathcal{C}_{ijk}}{\mathcal{C}_{ijr}} = \frac{G_{ik}}{G_{ir}} \frac{H_{jk}}{H_{jr}}, \quad (5)$$

that is, for each k , the matrix $\mathbf{X}^{(k)} \in \mathbb{F}^{N_1 \times N_2}$ obtained by elementwise division of $\mathcal{C}_{:, :, k}$ by $\mathcal{C}_{:, :, r}$ must be rank-one. Therefore, the matrices \mathbf{G} and \mathbf{H} can be recovered, for example, from the SVD of $\mathbf{X}^{(k)}$.

2.2. Implicit equations for the core tensor

In this subsection, we provide equations that must be satisfied by a core tensor of the form (4). Note that the condition (5) requires the matrices $\mathbf{X}^{(k)}$ defined in the same section to be rank one. It is well known that the defining equations for the set of matrices of rank ≤ 1 is given by 2×2 minors. Therefore, the elements of the matrix $\mathbf{X}^{(k)}$ must satisfy:

$$X_{ij}^{(k)} X_{pq}^{(k)} - X_{iq}^{(k)} X_{pj}^{(k)} = \frac{\mathcal{C}_{ijk} \mathcal{C}_{pqk}}{\mathcal{C}_{ijr} \mathcal{C}_{pqr}} - \frac{\mathcal{C}_{iqk} \mathcal{C}_{pjk}}{\mathcal{C}_{iqr} \mathcal{C}_{pjr}} = 0$$

Therefore, by getting rid of the denominators, we get the equations

$$\mathcal{C}_{ijk} \mathcal{C}_{pqk} \mathcal{C}_{iqr} \mathcal{C}_{pjr} - \mathcal{C}_{iqk} \mathcal{C}_{pjk} \mathcal{C}_{ijr} \mathcal{C}_{pqr} = 0. \quad (6)$$

Any core tensor of the form (4) satisfies (6) (see also Figure 1). As we will show later, under some assumptions, equations (6) are the only nontrivial equations for the core tensor.

2.3. Determinantal representation

Note that the equation (6) can be rewritten

$$\det \begin{bmatrix} \mathcal{C}_{ijk} \mathcal{C}_{pqk} & \mathcal{C}_{ijr} \mathcal{C}_{pqr} \\ \mathcal{C}_{iqk} \mathcal{C}_{pjk} & \mathcal{C}_{iqr} \mathcal{C}_{pjr} \end{bmatrix} = 0 \quad (7)$$

Then the system of equations (7) for all $k, r \in \{1, \dots, N_3\}$ is equivalent to vanishing of all 2×2 minors of

$$\mathbf{\Psi}^{(ijpq)} := \mathbf{\Psi} = [\Psi_{\ell k}]_{\ell, k=1}^{2, N_3} \quad (8)$$

$$\begin{aligned} & \mathcal{C}_{ijk}\mathcal{C}_{pqk}\mathcal{C}_{iqr}\mathcal{C}_{pjr} - \\ & \mathcal{C}_{iqk}\mathcal{C}_{pj k}\mathcal{C}_{ijr}\mathcal{C}_{pqr} = 0 \end{aligned}$$

Figure 1: Equations for core tensor via in terms of the elements of the $2 \times 2 \times 2$ subtensor $\mathcal{C}_{(i,p),(j,q),(k,r)}$.

where

$$\begin{bmatrix} \Psi_{1k} \\ \Psi_{2k} \end{bmatrix} = \begin{bmatrix} \mathcal{C}_{ijk}\mathcal{C}_{pqk} \\ \mathcal{C}_{iqk}\mathcal{C}_{pj k} \end{bmatrix};$$

hence, for fixed (i, j) and (p, q) , the set of equations (6) combined for all k, r is equivalent to

$$\text{rank}\{\Psi^{(ijpq)}\} \leq 1.$$

Remark 2.1. Under Assumption 1, we have that, for some index r , $\Psi_{1r}, \Psi_{2r} \neq 0$. Therefore, the rank of $\Psi^{(ijpq)}$ is exactly 1, and there is only one (up to scaling) vector in the left kernel:

$$\begin{bmatrix} \theta_1 & \theta_2 \end{bmatrix} \Psi^{(ijpq)} = 0. \quad (9)$$

Note that, in terms of the factor matrices,

$$\begin{bmatrix} \Psi_{1k} \\ \Psi_{2k} \end{bmatrix} = G_{ik}G_{pk}H_{jk}H_{qk} \begin{bmatrix} F_{ij}F_{pq} \\ F_{iq}F_{pj} \end{bmatrix},$$

hence the vector in the left kernel can be chosen as

$$\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} \theta_1^{(ijpq)} \\ \theta_2^{(ijpq)} \end{bmatrix} := \begin{bmatrix} -F_{iq}F_{pj} \\ F_{ij}F_{pq} \end{bmatrix}. \quad (10)$$

3. An algebraic approach to Paratuck-2 decomposition

The difficulty in finding the ParaTuck decomposition lies in finding the factor matrices \mathbf{A} and \mathbf{B} , and we will show in this subsection how to find these matrix from a kernel of a structured matrix constructed from the tensor. We will use the following following (lifting) matrix

$$\Phi(\mathcal{T}) = (\mathbf{T}^{(3)})^\top \odot (\mathbf{T}^{(3)})^\top = [\text{vec}\{\mathcal{T}_{::,1}\} \boxtimes \text{vec}\{\mathcal{T}_{::,1}\} \cdots \text{vec}\{\mathcal{T}_{::,N_3}\} \boxtimes \text{vec}\{\mathcal{T}_{::,N_3}\}] \in \mathbb{R}^{(N_1 N_2)^2 \times N_3},$$

Remark 3.1. The columns of $\Phi(\mathcal{T})$, in fact, correspond 4D tensors $\{\mathcal{T}_{::,k} \otimes \mathcal{T}_{::,k}\}_{k=1}^{N_3}$ reshaped in a special way.

3.1. Key assumptions

In this paper, we use the following main assumptions.

Assumption 2. The matrix \mathbf{F} does not have zero elements.

Assumption 3. The matrix $H \odot H \odot G \odot G$ has the maximal possible rank

$$\text{rank}\{H \odot H \odot G \odot G\} = \frac{R(R+1)}{2} \frac{S(S+1)}{2}.$$

Remark 3.2. The assumption (3) holds for general matrices H and G when $N_3 \geq \frac{R(R+1)}{2} \frac{S(S+1)}{2}$. Indeed, the column space of $H \odot H \odot G \odot G$ corresponds to the space of tensors $(\mathbf{v} \otimes \mathbf{v}) \otimes (\mathbf{w} \otimes \mathbf{w})$, whose dimension does not exceed $\frac{R(R+1)}{2} \frac{S(S+1)}{2}$ and is equal to this value for vectors in a general position.

Note that the Assumption 2 also holds for general matrices **CoreFront**.

We will also use the following simplifying assumption.

Assumption 4. *The matrices **A** and **B** are square (i.e., $N_1 = R$, $N_2 = S$) and invertible.*

Remark 3.3. *Assumption 4 can be relaxed to the following two conditions:*

- **A** and **B** are full column rank,
- and the first and second unfoldings of the core tensor $\mathbf{C}^{(1)}$ and $\mathbf{C}^{(2)}$ are full row rank.

Indeed, under such conditions, we can transform the problem to Assumption 4, by computing the Tucker decomposition of \mathcal{T} first.

Remark 3.4. *The relaxed assumption Remark 3.3 holds for general (random) matrices **A** and **B**, **F**, **G**, **H**, as long as $N_1 \geq R$, $N_2 \geq R$ and $N_3 \geq \max(R, S)$. Indeed, if $N_1 \geq R$, $N_2 \geq R$, general matrices **A**, **B** are full column rank. The condition on N_3 is needed to obtain the correct ranks of the unfolding For example $\mathbf{C}^{(1)}$ is full row rank (for general factors **F**, **G**, **H**) if $N_3 \geq R$ because*

$$\mathbf{C}^{(1)} = [\text{Diag}\{\mathbf{G}_{:1}\} \mathbf{F} \text{Diag}\{\mathbf{H}_{:1}\} \cdots \text{Diag}\{\mathbf{G}_{:N_3}\} \mathbf{F} \text{Diag}\{\mathbf{H}_{:N_3}\}].$$

Then, under these assumptions, we get the following result on the rank of $\Phi(\mathcal{T})$.

Lemma 3.5. *Under Assumptions 2–4, the matrix $\Phi(\mathcal{T})$ has the maximal possible rank:*

$$\text{rank}\{\Phi(\mathcal{T})\} = \frac{R(R+1)}{2} \frac{S(S+1)}{2}.$$

Proof. Note that

$$\Phi(\mathcal{T}) = (\mathbf{B} \boxtimes \mathbf{A} \boxtimes \mathbf{B} \boxtimes \mathbf{A}) \text{Diag}\{\text{vec}\{\mathbf{F}\} \boxtimes \text{vec}\{\mathbf{F}\}\} (\mathbf{H} \odot \mathbf{G} \odot \mathbf{H} \odot \mathbf{G}).$$

The matrices $(\mathbf{B} \boxtimes \mathbf{A} \boxtimes \mathbf{B} \boxtimes \mathbf{A})$ and $\text{Diag}\{\text{vec}\{\mathbf{F} \boxtimes \mathbf{F}\}\}$ are invertible by Assumption 4 and Assumption 3, which completes the proof. \square

3.2. Basis of the left kernel

Thanks to the assumptions from the previous subsection, and using the equations (6) of the core tensor, we can get the basis of the kernel of the structured matrix $\Phi(\mathcal{T})$. Let $\tilde{\mathbf{A}} = \mathbf{A}^{-1}$ and $\tilde{\mathbf{B}} = \mathbf{B}^{-1}$ and

$$\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}_1 \cdots \tilde{\mathbf{a}}_R], \quad \tilde{\mathbf{B}} = [\tilde{\mathbf{b}}_1 \cdots \tilde{\mathbf{b}}_S].$$

We also define the linear symmetrization map $\sigma : \mathbb{R}^{SRSR} \rightarrow \mathbb{R}^{SRSR}$ that maps

$$\sigma : \mathbf{y} \boxtimes \mathbf{z} \mapsto \frac{1}{2}(\mathbf{y} \boxtimes \mathbf{z} + \mathbf{z} \boxtimes \mathbf{y}), \quad \text{for all } \mathbf{y}, \mathbf{z} \in \mathbb{R}^{RS},$$

i.e. it is a symmetrized matricization of the vector. The space $\sigma(\mathbb{R}^{SRSR})$ is thus the space of symmetric vectors of dimension

$$\frac{RS(RS+1)}{2}.$$

Then we have the following proposition.

Proposition 3.6. *For each pair $1 \leq i < p \leq R$, $1 \leq j < q \leq S$, define the following vector $\mathbf{v}^{(ijpq)} \in \sigma(\mathbb{R}^{SRSR})$:*

$$\mathbf{v}^{(ijpq)} = \theta_1^{(ijpq)} \sigma(\tilde{\mathbf{b}}_j \boxtimes \tilde{\mathbf{a}}_i \boxtimes \tilde{\mathbf{b}}_q \boxtimes \tilde{\mathbf{a}}_p) + \theta_2^{(ijpq)} \sigma(\tilde{\mathbf{b}}_q \boxtimes \tilde{\mathbf{a}}_i \boxtimes \tilde{\mathbf{b}}_j \boxtimes \tilde{\mathbf{a}}_p),$$

where $\theta_1^{(ijpq)}, \theta_2^{(ijpq)}$ are defined in (10). Then we have that:

1. Each vector $\mathbf{v}^{(ijpq)}$ belongs to the left kernel of $\Phi(\mathcal{T})$;

2. Under Assumptions 1, 2 and 4, these vectors are linearly independent and the vector space

$$\mathcal{V} = \{\mathbf{v}^{(ijpq)} : 1 \leq i < p \leq R, 1 \leq j < q \leq S\} \quad (11)$$

has dimension

$$\dim \mathcal{V} = \frac{R(R-1)}{2} \frac{S(S-1)}{2}; \quad (12)$$

moreover, if, in addition, Assumption 3 is satisfied, they form a basis of the symmetrized left kernel

$$\mathcal{V} = \sigma(\text{lker}\{\Phi(\mathcal{T})\}).$$

Proof. 1. We can rewrite (9) as

$$(\theta_1 \mathbf{e}_j \boxtimes \mathbf{e}_i \boxtimes \mathbf{e}_q \boxtimes \mathbf{e}_p + \theta_2 \mathbf{e}_q \boxtimes \mathbf{e}_i \boxtimes \mathbf{e}_j \boxtimes \mathbf{e}_p) \Phi(\mathbf{C}) = 0$$

Note that the columns of $\Phi(\mathbf{C})$ are spanned by the vectors $\mathbf{w} \boxtimes \mathbf{w}$ with $\mathbf{w} \in \mathbb{R}^{RS}$, therefore the corresponding symmetrization

$$(\theta_1 \sigma(\mathbf{e}_j \boxtimes \mathbf{e}_i \boxtimes \mathbf{e}_q \boxtimes \mathbf{e}_p) + \theta_2 \sigma(\mathbf{e}_q \boxtimes \mathbf{e}_i \boxtimes \mathbf{e}_j \boxtimes \mathbf{e}_p)) \quad (13)$$

lies also in the left kernel of $\Phi(\mathbf{C})$. Now define $\tilde{\mathbf{A}} = \mathbf{A}^{-1}$ and $\tilde{\mathbf{B}} = \mathbf{B}^{-1}$. Then we have

$$\Phi(\mathbf{C}) = (\tilde{\mathbf{B}} \boxtimes \tilde{\mathbf{A}} \boxtimes \tilde{\mathbf{B}} \boxtimes \tilde{\mathbf{A}}) \Phi(\mathcal{T}),$$

hence $\mathbf{v}^{(ijpq)}$ (which is a multilinear transformation applied to (13)) is in the left kernel of $\Phi(\mathcal{T})$.

2. Thanks to Assumptions 1 and 2, the vectors in (13) are nonzero and they are linearly independent because $1 \leq i < p \leq R, 1 \leq j < q \leq S$, hence the set of \mathcal{V} is also linearly independent.
3. Note that from the previous points of the proof, $\mathcal{V} \subseteq \sigma(\text{lker}\{\Phi(\mathcal{T})\})$. To prove the equality, we count the dimensions

$$\begin{aligned} \dim \sigma(\text{lker}\{\Phi(\mathcal{T})\}) &= \dim \sigma(\mathbb{R}^{SRSR}) - \text{rank}\{\Phi(\mathcal{T})\} = \frac{RS(RS+1)}{2} - \frac{R(R+1)}{2} \frac{S(S+1)}{2} \\ &= \frac{R(R-1)}{2} \frac{S(S-1)}{2} = \dim(\mathcal{V}), \end{aligned}$$

which completes the proof. □

3.3. Decoupling the factors from the left kernel

In this subsection, we show how to recover the factors from \mathcal{V} defined in (11). Consider another symmetrization combined with permutation $\pi : \mathbb{R}^{SRSR} \rightarrow \mathbb{R}^{SSRR}$

$$\pi : \mathbf{z} \boxtimes \mathbf{y} \boxtimes \mathbf{b} \boxtimes \mathbf{a} \mapsto \frac{1}{2}(\mathbf{b} \boxtimes \mathbf{z} + \mathbf{z} \boxtimes \mathbf{b}) + \frac{1}{2}(\mathbf{a} \boxtimes \mathbf{y} + \mathbf{y} \boxtimes \mathbf{a}), \quad \text{for all } \mathbf{y}, \mathbf{a} \in \mathbb{R}^R, \mathbf{z}, \mathbf{b} \in \mathbb{R}^S.$$

Note that $\pi(\mathbb{R}^{SRSR})$ is spanned by the vectors of form $(\mathbf{b} \boxtimes \mathbf{b}) \boxtimes (\mathbf{a} \boxtimes \mathbf{a})$ and has dimension

$$\frac{R(R+1)}{2} \frac{S(S+1)}{2}.$$

Then, obviously, we have

$$\pi(\mathbf{v}^{(ijpq)}) = (\theta_1^{(ijpq)} + \theta_2^{(ijpq)}) \sigma(\tilde{\mathbf{b}}_j \boxtimes \tilde{\mathbf{b}}_q) \boxtimes \sigma(\tilde{\mathbf{a}}_i \boxtimes \tilde{\mathbf{a}}_p). \quad (14)$$

Note that

$$(\theta_1^{(ijpq)} + \theta_2^{(ijpq)}) = \det \mathbf{F}_{\{i,p\},\{j,q\}}, \quad \text{where } \mathbf{F}_{\{i,p\},\{j,q\}} = \begin{bmatrix} F_{ij} & F_{pj} \\ F_{iq} & F_{pq} \end{bmatrix}$$

Then if all the 2×2 minors of \mathbf{F} are nonzero, then the dimension of $\pi(\mathcal{V})$ does not drop (is equal to (12)). However, we can find the factors \mathbf{A} and \mathbf{B} under the following weaker conditions.

Lemma 3.7. Let $\mathbf{p}_1, \dots, \mathbf{p}_L$ be the basis of $\pi(\mathcal{V})$.

1. Define the matrix

$$\mathbf{S}_A = [\text{matr}_{R^2, S^2}\{\mathbf{p}_1\} \quad \cdots \quad \text{matr}_{R^2, S^2}\{\mathbf{p}_L\}] \in \mathbb{R}^{R^2 \times LS^2}.$$

Then if the row Kruskal rank of \mathbf{F} is at least 2:

$$\text{krank}\{\mathbf{F}^\top\} \geq 2,$$

then we have that

$$\text{rank}\{\mathbf{S}_A\} = \frac{R(R-1)}{2},$$

and the column space of \mathbf{S}_A is spanned by $\{\sigma(\tilde{\mathbf{a}}_i \boxtimes \tilde{\mathbf{a}}_p)\}_{1 \leq i < p \leq R}$

2. Similarly, for the matrix Define the matrix

$$\mathbf{S}_B = [(\text{matr}_{R^2, S^2}\{\mathbf{p}_1\})^\top \quad \cdots \quad (\text{matr}_{R^2, S^2}\{\mathbf{p}_L\})^\top] \in \mathbb{R}^{S^2 \times LR^2}.$$

Then if the (column) Kruskal rank of \mathbf{F} is at least 2:

$$\text{krank}\{\mathbf{F}\} \geq 2,$$

then we have that

$$\text{rank}\{\mathbf{S}_B\} = \frac{R(R-1)}{2},$$

and the column space of \mathbf{S}_B is spanned by $\{\sigma(\tilde{\mathbf{b}}_j \boxtimes \tilde{\mathbf{b}}_q)\}_{1 \leq j < q \leq S}$.

Proof. We will prove only the first statement, because the second is analogous. Indeed, we have under the row Kruskal rank condition, for any $1 \leq i < p \leq R$ there exist $1 \leq j < q \leq R$ such that $\det \mathbf{F}_{\{i,p\}, \{j,q\}} \neq 0$. Therefore the space $\pi(\mathcal{V})$ includes as a subspace the space of vectors

$$\text{span}\{\sigma(\tilde{\mathbf{a}}_i \boxtimes \tilde{\mathbf{a}}_p) \boxtimes \mathbf{w}_{(i,p)}\}_{1 \leq i < p \leq R},$$

where $\mathbf{w}_{(i,p)} \in \mathbb{R}^{R \times R}$. Therefore, after the matricization, the column space of \mathbf{S}_A is spanned by

$$\text{span}\{\sigma(\tilde{\mathbf{a}}_i \boxtimes \tilde{\mathbf{a}}_p)\}_{1 \leq i < p \leq R},$$

which completes the proof. \square

3.4. Extracting the spaces from the symmetrizations

Finally, we show how to extract the matrix \mathbf{A} from

$$\mathcal{V}_A = \text{span}\{\sigma(\tilde{\mathbf{a}}_i \boxtimes \tilde{\mathbf{a}}_p)\}_{1 \leq i < p \leq R}.$$

Under Assumption 4, this space has dimension $R(R-1)/2$ as all the vectors are linearly independent. Let us denote by $\mathbf{q}_1, \dots, \mathbf{q}_{R(R-1)/2} \in \mathbb{R}^{R^2}$ the basis of this vector space \mathcal{V}_A , and by $\mathbf{Q}_1, \dots, \mathbf{Q}_{R(R-1)/2} \in \mathbb{R}^{R \times R}$ their matricizations

$$\mathbf{Q}_k = \text{matr}_{R,R}\{\mathbf{q}_k\}.$$

We make use of the following lemma.

Lemma 3.8. The system of equations

$$\begin{aligned} \mathbf{a}^\top \mathbf{Q}_1 \mathbf{a} &= 0 \\ &\vdots \\ \mathbf{a}^\top \mathbf{Q}_{R(R-1)/2} \mathbf{a} &= 0 \end{aligned} \tag{15}$$

has exactly R nonzero solutions (up to multiplications by scalar), which are $\{\mathbf{a}_1, \dots, \mathbf{a}_R\}_{i=1}^R$

Proof. Note that the system of equations (15) is equivalent to

$$\mathbf{a}^\top \sigma(\tilde{\mathbf{a}}_i \boxtimes \tilde{\mathbf{a}}_p) \mathbf{a}, \quad \text{for all } 1 \leq i < p \leq R,$$

where each individual equation is equivalent to

$$\langle \tilde{\mathbf{a}}_i, \mathbf{a} \rangle \langle \tilde{\mathbf{a}}_p, \mathbf{a} \rangle = 0.$$

The statement can be proved by induction on R . For $R = 2$ (the base), we have just one equation

$$\langle \tilde{\mathbf{a}}_1, \mathbf{a} \rangle \langle \tilde{\mathbf{a}}_2, \mathbf{a} \rangle = 0.$$

which gives the two possible solutions \mathbf{a}_1 or \mathbf{a}_2 .

If $R \geq 3$, then we there exists at least one index s such that $\langle \mathbf{a}, \tilde{\mathbf{a}}_s \rangle = 0$. We can remove all the equations with $i = s$ or $p = s$. Therefore $\mathbf{a} = \mathbf{A}_{:, \neq s} \mathbf{a}'$, where $\mathbf{A}_{:, \neq s} \in \mathbb{R}^{R \times (R-1)}$ is the matrix obtained by removing the s -th column of \mathbf{A} . Then if we define $\tilde{\mathbf{a}}' = \mathbf{A}_{:, \neq s}^\top$ then

$$\langle \mathbf{A}_{:, \neq s} \mathbf{a}', \tilde{\mathbf{a}}_j \rangle = \langle \mathbf{a}', \tilde{\mathbf{a}}_j \rangle,$$

and note that

$$\mathbf{A}_{:, \neq s}^\top [\tilde{\mathbf{a}}_1 \quad \cdots \quad \tilde{\mathbf{a}}_s \quad \cdots \quad \tilde{\mathbf{a}}_R] = \mathbf{I}_{R-1};$$

hence, by the induction assumption, $\mathbf{a}' = \mathbf{e}_j$ for some $j \in \{1, \dots, R-1\}$, which completes the proof. \square

Finally, we show how to solve the system (15). By Lemma 3.8, the symmetrization of the left kernel of the reshapings of \mathbf{Q} is spanned by

$$\sigma(\text{lker}\{\text{vec}\{\mathbf{Q}_1\} \quad \cdots \quad \text{vec}\{\mathbf{Q}_{R(R-1)/2}\}\}) = \text{span}\{\mathbf{a}_i \boxtimes \mathbf{a}_i\}_{i=1}^N.$$

Let $\tilde{\mathbf{Q}}_1, \dots, \tilde{\mathbf{Q}}_R$ be the set of linearly independent symmetric matrices orthogonal to

$$\text{span}\{\mathbf{Q}_k\}_{k=1}^{R(R-1)/2}.$$

Then we have that there exists a matrix \mathbf{W} such that

$$\tilde{\mathbf{Q}}_k = \sum_{j=1}^R W_{kj} \mathbf{a}_j \mathbf{a}_j^\top.$$

Thus the tensor $\tilde{\mathbf{Q}} \in \mathbb{R}^{R \times R \times R}$, whose frontal slices are $\tilde{\mathbf{Q}}_{:, :, k} = \tilde{\mathbf{Q}}_k$, must have the CPD

$$\tilde{\mathbf{Q}} = \llbracket \mathbf{A}, \mathbf{A}, \mathbf{W} \rrbracket,$$

and therefore \mathbf{A} can be obtained from the unique CPD of $\tilde{\mathbf{Q}}$.

4. The algorithm and the implementation details

4.1. Overall algorithm

Here we summarize the complete algorithm. Note that the algorithm is able to compute the decomposition under Assumptions 1–4 and also when both \mathbf{F} and \mathbf{F}^\top have Kruskal rank at least 2.

Algorithm 1: Algebraic decomposition algorithm for Paratuck**input :** $\mathcal{T} \in \mathbb{R}^{N_1, N_2, N_3}$, R, S **output:** $\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{F}}, \widehat{\mathbf{G}}, \widehat{\mathbf{H}}$ 1. Compute the HOSVD of rank (R, S, N_3) : $\mathcal{T} = \llbracket \mathcal{T}_c; \mathbf{A}_c, \mathbf{B}_c, \mathbf{I}_{I_1} \rrbracket$, $\mathbf{A}_c \in \mathbb{F}^{I_1 \times R}$, $\mathbf{B}_c \in \mathbb{F}^{I_2 \times S}$;2. Construct $\Phi(\mathcal{T}_c)$ 3. Compute (using the SVD) the matrix containing the basis $\mathbf{V} \in \mathbb{R}^{R^2 S^2 \times L}$ of the symmetric part $\Phi(\mathcal{T}_c)$ of the left nullspace of SVD (i.e., take the last $(R^2 S^2 - RS(R+1)(S+1)/4)$ left singular vectors and symmetrize them);4. Build the matrices \mathbf{S}_A and \mathbf{S}_B respectively. For each of the matrices, compute the approximate symmetric nullspaces and the corresponding bases

$$\{\widetilde{\mathbf{Q}}_{A,1}, \dots, \widetilde{\mathbf{Q}}_{A,R}\},$$

and

$$\{\widetilde{\mathbf{Q}}_{B,1}, \dots, \widetilde{\mathbf{Q}}_{B,S}\}.$$

5. Stack these matrices in tensors $\widetilde{\mathcal{Q}}_A$ and $\widetilde{\mathcal{Q}}_B$ and compute their CPDs

$$\widetilde{\mathcal{Q}}_A = \llbracket \mathbf{A}', \mathbf{A}', \mathbf{W}_A \rrbracket \quad \text{and} \quad \widetilde{\mathcal{Q}}_B = \llbracket \mathbf{B}', \mathbf{B}', \mathbf{W}_B \rrbracket,$$

by exploiting performing joint diagonalization of matrices. 6. Set $\widehat{\mathbf{A}} = \mathbf{A}_c \mathbf{A}'$, $\widehat{\mathbf{B}} = \mathbf{B}_c \mathbf{B}'$.8. Find the core tensor $\widehat{\mathcal{C}} = \mathcal{T}_c \bullet_1 \mathbf{A}'^\dagger \bullet_2 \mathbf{B}'^\dagger$.9. Set $\widehat{\mathbf{F}} = \mathcal{C}_{:, :, k}$, where k is some index.10. Determine the factors $\widehat{\mathbf{G}}$ and $\widehat{\mathbf{H}}$ from rank-one approximations of elementwise divisions of $\widehat{\mathcal{C}}_{:, :, k}$ by $\widehat{\mathbf{F}}$.

4.2. Some numerical experiments

4.2.1. Deterministic example

We consider the following test example:

$$\mathbf{F} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & -3 \end{bmatrix}$$

and

$$\mathbf{G} = \begin{bmatrix} -5 & -4 & -3 & -2 & -1 & 0 & 1 & 2 & 3 & 4 \\ 1 & 0 & 2 & 1 & -3 & 2 & -2 & -1 & 0 & 1 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} -5 & -4 & -3 & -2 & -1 & 0 & 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Algorithm 1 is able to compute a Paratuck-2 approximation with approximation error (squared Frobenius norm of the difference: $\|\mathcal{T} - \widehat{\mathcal{T}}\|_F^2$) as $3.65 \cdot 10^{-26}$; after refinement by alternating least squares, the approximation error is reduced to $8.45 \cdot 10^{-27}$. The alternating least squares was run for one iteration.

We also ran alternating least squares [Bro98, pp. 68–71] with random initialization (maximum 5000 iterations), and it gave an approximation error $\|\mathcal{T} - \widehat{\mathcal{T}}\|_F^2$ below 10^{-10} only in 8% of the cases (number of Monte-Carlo runs was 100).

4.2.2. Random example

We next consider random example with $I_1 = I_2 = 10$, $I_3 = 15$ and $R = S = 2$. We generate 100 examples, with \mathbf{A} , \mathbf{B} , \mathbf{H} , \mathbf{F} , \mathbf{G} having i.i.d. entries drawn randomly from $N(0, 1)$. In 99 of 100 runs the algebraic algorithm gave the approximation error $\|\mathcal{T} - \widehat{\mathcal{T}}\|_F^2$ below 10^{-20} (in the remaining one example it was $\sim 10^{-18}$).

5. Acknowledgements

This research was supported by the ANR (Agence Nationale de Recherche) grant LeafleT (ANR-19-CE23-0021). I also would like to thank Yassine Zniyed, Mariya Ishteva, Philippe Dreesen and Yang Qi for stimulating discussions.

References

- [Bro98] Rasmus Bro. *Multi-way Analysis in the Food Industry*. PhD thesis, Vrije Universiteit Brussel (VUB), 1998.
- [dOFFB19] Pedro Marinho R. de Oliveira, C. Alexandre Rolim Fernandes, Gérard Favier, and Remy Boyer. PARATUCK Semi-Blind Receivers for Relaying Multi-Hop MIMO Systems. *Digital Signal Processing*, 92:127–138, 2019.
- [FdA14] Gérard Favier and André LF de Almeida. Overview of constrained parafac models. *EURASIP Journal on Advances in Signal Processing*, 2014(1):142, 2014.
- [HL96] Richard A Harshman and Margaret E Lundy. Uniqueness proof for a family of models sharing features of tucker’s three-mode factor analysis and parafac/candecomp. *Psychometrika*, 61(1):133–154, 1996.
- [Nas20] Kristina Naskovska. *Advanced tensor based signal processing techniques for wireless communication systems and biomedical signal processing*. PhD thesis, Ilmenau, Jan 2020. Dissertation, Technische Universität Ilmenau, 2019.