$(L_r, L_r, 1)$-decompositions, Sparse Component Analysis, and the Blind Separation of Sums of Exponentials

Nithin Govindarajan (KU Leuven)

Joint work with Ethan Epperly (Caltech), Lieven De Lathauwer (KU Leuven)

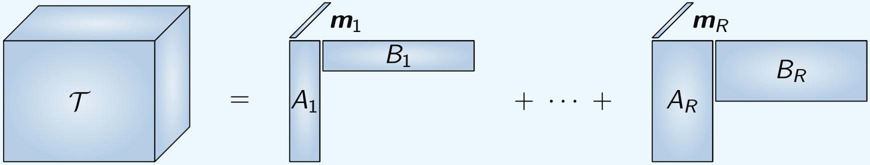
SeLMA Meeting - 12 October 2021

The $(L_r, L_r, 1)$-decomposition

Definition

A third-order tensor $\mathcal{T} \in \mathbb{C}^{I\times J\times K}$ is expanded in the form

$$\mathcal{T} = \sum_{r=1}^{R} H_r \otimes \boldsymbol{m}_r, \quad \operatorname{rank}(H_r) = \operatorname{rank}(A_r B_r^\top) = L_r > 0.$$

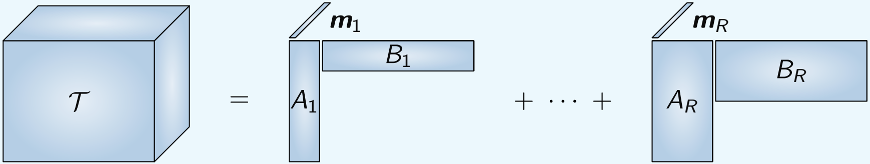


The $(L_r, L_r, 1)$-decomposition

Definition

A third-order tensor $\mathcal{T} \in \mathbb{C}^{I \times J \times K}$ is expanded in the form

$$\mathcal{T} = \sum_{r=1}^{R} H_r \otimes \boldsymbol{m}_r, \quad \operatorname{rank}(H_r) = \operatorname{rank}(A_r B_r^\top) = L_r > 0.$$



Special case: If $L_r = 1$ for $1 \leq r \leq R$, we have a polyadic decomposition (PD).

Uniqueness based on rank profile

Definition

A $(L_r, L_r, 1)$-decomposition of $\mathcal{T} \in \mathbb{C}^{I \times J \times K}$ with rank profile $(L_r = \operatorname{rank} H_r)_{1 \leq r \leq R}$ is *essentially unique* if every other $(L_r, L_r, 1)$-decomposition

$$\mathcal{T} = \sum_{r=1}^{R} H'_r \otimes \boldsymbol{m}'_r$$

satisfying the rank profile constraints:

$$\operatorname{rank} H'_r = L'_r \leq L_r, \qquad 1 \leq r \leq R$$

is the same up to *scaling and permutation ambiguity*.

Uniqueness based on rank profile

Definition

A $(L_r, L_r, 1)$-decomposition of $\mathcal{T} \in \mathbb{C}^{I \times J \times K}$ with rank profile $(L_r = \operatorname{rank} H_r)_{1 \leq r \leq R}$ is *essentially unique* if every other $(L_r, L_r, 1)$-decomposition

$$\mathcal{T} = \sum_{r=1}^{R} H'_r \otimes \boldsymbol{m}'_r$$

satisfying the rank profile constraints:

$$\operatorname{rank} H'_r = L'_r \leq L_r, \qquad 1 \leq r \leq R$$

is the same up to *scaling and permutation ambiguity*.

Special case: Rank profile $(L_r = 1)_{1 \leq r \leq R}$ returns uniqueness definition of the canonical polyadic decomposition (CPD).

Uniqueness properties of $(L_r, L_r, 1)$-decomposition have direct application in the blind separation of sums of exponentials (BSSE) problem

Consider the linear observations

$$y_k(t) = m_{k1}s_1(t) + m_{k2}s_2(t) + \ldots + m_{kR}s_R(t), \qquad k = 1, \ldots, K$$

where:

$$s_r(t) = \sum_{j=1}^{L_r} \alpha_{r,j} z_{r,j}^t, \qquad 0 \leq t < T,$$

Uniqueness properties of $(L_r, L_r, 1)$-decomposition have direct application in the blind separation of sums of exponentials (BSSE) problem

Problem

Given $Y \in \mathbb{C}^{K\times T}$, find the factorization:

$$Y = MS,$$

where:

$$Y := \begin{bmatrix} y_1(0) & \cdots & y_1(T-1) \\ \vdots & & \vdots \\ y_K(0) & \cdots & y_K(T-1) \end{bmatrix}, \qquad S := \begin{bmatrix} s_1(0) & \cdots & s_1(T-1) \\ \vdots & & \vdots \\ s_R(0) & \cdots & s_R(T-1) \end{bmatrix}.$$

Uniqueness properties of $(L_r, L_r, 1)$-decomposition have direct application in the blind separation of sums of exponentials (BSSE) problem

Define the tensor

$$\mathcal{H}[Y](:,:,k) = H[y_k] := \begin{bmatrix} y_k(0) & y_k(1) & \cdots & y_k(T_2-1) \\ y_k(1) & y_k(2) & \cdots & y_k(T_2) \\ \vdots & \vdots & \ddots & \vdots \\ y_k(T_1-1) & y_k(T_1) & \cdots & y_k(T_1+T_2-2) \end{bmatrix}.$$

Problem is *solved* by recovering the $(L_r, L_r, 1)$-decomposition

$$\mathcal{H}[Y] = \sum_{r=1}^{R} H[s_r] \otimes \boldsymbol{m}_r,$$

which is *unique* under mild assumptions.

A tensor may admit multiple "unique" $(L_r, L_r, 1)$-decompositions

Let

$$A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 & \boldsymbol{a}_4 \end{bmatrix}, \quad B = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \boldsymbol{b}_3 & \boldsymbol{b}_4 \end{bmatrix}, \quad M = \begin{bmatrix} \boldsymbol{m}_1 & \boldsymbol{m}_2 \end{bmatrix}$$

be full column rank matrices.

$$\overbrace{(\boldsymbol{a}_1\boldsymbol{b}_1^\top + \boldsymbol{a}_2\boldsymbol{b}_2^\top) \otimes \boldsymbol{m}_1 + (\boldsymbol{a}_2\boldsymbol{b}_2^\top + \boldsymbol{a}_3\boldsymbol{b}_3^\top + \boldsymbol{a}_4\boldsymbol{b}_4^\top) \otimes \boldsymbol{m}_2}^{\text{unique under rank profile } (L_1, L_2) = (2,3)}$$

$$=$$

$$\underbrace{(\boldsymbol{a}_1\boldsymbol{b}_1^\top) \otimes \boldsymbol{m}_1 + (\boldsymbol{a}_2\boldsymbol{b}_2^\top) \otimes (\boldsymbol{m}_1 + \boldsymbol{m}_2) + (\boldsymbol{a}_3\boldsymbol{b}_3^\top + \boldsymbol{a}_4\boldsymbol{b}_4^\top) \otimes \boldsymbol{m}_2}_{\text{unique under rank profile } (L_1, L_2, L_3) = (1,1,2)}$$

Constraints on the rank profile are unknowns in the BSSE problem

The Hankel tensor of the BSSE problem:

$$\begin{aligned} s_1(t) &= z_1^t + z_2^t \\ s_2(t) &= z_2^t + z_3^t + z_4^t \end{aligned} , \quad M = \begin{bmatrix} \boldsymbol{m}_1 & \boldsymbol{m}_2 \end{bmatrix} \text{ full rank}$$

has *two* essentially unique $(L_r, L_r, 1)$-decompositions with rank profiles:

$$(2, 3) \text{ and } (1, 1, 2).$$

Constraints on the rank profile are unknowns in the BSSE problem

The Hankel tensor of the BSSE problem:

$$\begin{array}{rcl} s_1(t) & = & z_1^t + z_2^t \\ s_2(t) & = & z_2^t + z_3^t + z_4^t \end{array}, \quad M = \begin{bmatrix} \boldsymbol{m}_1 & \boldsymbol{m}_2 \end{bmatrix} \text{ full rank}$$

has *two* essentially unique $(L_r, L_r, 1)$-decompositions with rank profiles:

$$(2, 3) \text{ and } (1, 1, 2).$$

So... how should one *know* which rank profile must be used?

Dictionary representations: towards an alternative definition of uniqueness

Consider the $(L_r, L_r, 1)$-decomposition

$$\mathcal{T} = H_1 \otimes \boldsymbol{m}_1 + H_2 \otimes \boldsymbol{m}_2, \qquad H_1 = \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \boldsymbol{u}_2 \boldsymbol{v}_2^\top, \quad H_2 = \boldsymbol{u}_2 \boldsymbol{v}_2^\top + \boldsymbol{u}_3 \boldsymbol{v}_3^\top.$$

"Original representation" in terms of the triple (U, V, M), where:

$$U = \left[\begin{array}{cc|cc} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{array}\right], \quad V = \left[\begin{array}{cc|cc} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{array}\right], \quad M = \begin{bmatrix} \boldsymbol{m}_1 & \boldsymbol{m}_2 \end{bmatrix}$$

Dictionary representations: towards an alternative definition of uniqueness

Consider the $(L_r, L_r, 1)$-decomposition

$$\mathcal{T} = H_1 \otimes \boldsymbol{m}_1 + H_2 \otimes \boldsymbol{m}_2, \qquad H_1 = \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \boldsymbol{u}_2 \boldsymbol{v}_2^\top, \quad H_2 = \boldsymbol{u}_2 \boldsymbol{v}_2^\top + \boldsymbol{u}_3 \boldsymbol{v}_3^\top.$$

Alternatively, we may express

$$H_1 = \underbrace{\xi_{11}}_{=1} \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \underbrace{\xi_{12}}_{=1} \boldsymbol{u}_2 \boldsymbol{v}_2^\top + \underbrace{\xi_{13}}_{=0} \boldsymbol{u}_3 \boldsymbol{v}_3^\top, \qquad H_2 = \underbrace{\xi_{21}}_{=0} \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \underbrace{\xi_{22}}_{=1} \boldsymbol{u}_2 \boldsymbol{v}_2^\top + \underbrace{\xi_{23}}_{=1} \boldsymbol{u}_3 \boldsymbol{v}_3^\top,$$

$$\boldsymbol{\Xi} = \begin{bmatrix} \xi_{11} & \xi_{12} & \xi_{13} \\ \xi_{21} & \xi_{22} & \xi_{23} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Dictionary representations: towards an alternative definition of uniqueness

Consider the $(L_r, L_r, 1)$-decomposition

$$\mathcal{T} = H_1 \otimes \boldsymbol{m}_1 + H_2 \otimes \boldsymbol{m}_2, \qquad H_1 = \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \boldsymbol{u}_2 \boldsymbol{v}_2^\top, \quad H_2 = \boldsymbol{u}_2 \boldsymbol{v}_2^\top + \boldsymbol{u}_3 \boldsymbol{v}_3^\top.$$

This leads to the "Dictionary representation" in terms of the tuple (A, B, M, Ξ), where:

$$A = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{bmatrix}, \quad B = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{bmatrix}, \quad M = \begin{bmatrix} \boldsymbol{m}_1 & \boldsymbol{m}_2 \end{bmatrix}, \quad \Xi = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

A uniqueness definition based on dictionary representation

Let (1)-(P) denote a list of properties satisfied by the matrices A, B, M, and Ξ.

$$\mathscr{C} := \begin{array}{c}\text{Collection of all } (L_r, L_r, 1)\text{-decompositions that admit a} \\ \text{dictionary representation satisfying properties (1)-}(P).\end{array}$$

Definition

A tensor $\mathcal{T} \in \mathbb{C}^{I \times J \times K}$ has a *unique* $(L_r, L_r, 1)$*-decomposition satisfying properties (1)-*(P) if any other $(L_r, L_r, 1)$-decomposition <u>in $\mathscr{C}$</u> that describes $\mathcal{T}$ is the same *up to scaling and permutation ambiguity*.

A uniqueness definition based on dictionary representation

Let (1)-(P) denote a list of properties satisfied by the matrices A, B, M, and Ξ.

$\mathscr{C} :=$ Collection of all $(L_r, L_r, 1)$-decompositions that admit a dictionary representation satisfying properties (1)-(P).

Definition

A tensor $\mathcal{T} \in \mathbb{C}^{I\times J\times K}$ has a *unique* $(L_r, L_r, 1)$*-decomposition satisfying properties (1)-(P)* if any other $(L_r, L_r, 1)$-decomposition in $\mathscr{C}$ that describes $\mathcal{T}$ is the same *up to scaling and permutation ambiguity*.

How can we derive uniqueness results with this new definition?

Introducing sparse component analysis (SCA)

- Let $\mathcal{D}$ denote a set of *"admissible"* pairs $(M, \Xi) \in (\mathbb{C}^{K\times R}, \mathbb{C}^{R\times N})$
- In SCA:

  $\mathcal{D}$ is defined through *rank* and *sparsity constraints* on M and Ξ, respectively.
- Suppose that $C = M\Xi$ for some $(M, \Xi) \in \mathcal{D}$.

Problem:

Given $C \in \mathbb{C}^{K\times R}$, *recover* $C = M\Xi$ up to a *scaling and permutation ambiguity*

A mechanism to derive uniqueness results for $(L_r, L_r, 1)$-decompositions

Technical assumptions on admissible set $\mathscr{D}$:

(a) every $(M, \Xi) \in \mathscr{D}$ is proportionality-revealing,

(b) $\mathscr{D}$ is scaled permutation invariant.

Lemma

(A, B, M, Ξ) is the unique $(L_r, L_r, 1)$-decomposition satisfying properties:

(1) A and B have full column rank,

(2) $(M, \Xi) \in \mathscr{D}$.

$$\Longleftrightarrow \underbrace{C = M\Xi \text{ is unique in } \mathscr{D}}_{\text{Hence, proving this is enough!}}.$$

SCA: the intuition behind proving uniqueness of $C = M\Xi$

A 2-step *constructive* approach to unique recovery:

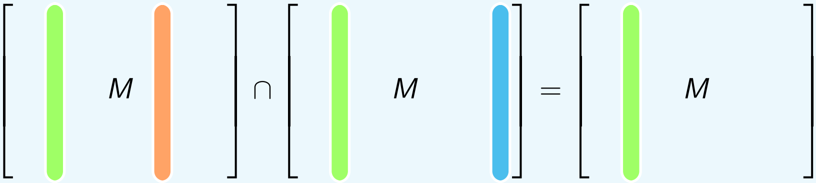
1. Recover the columns of $M \in \mathbb{C}^{K\times R}$ up to permutation and scaling ambiguity.
2. Recover $C \in \mathbb{C}^{R\times N}$ by solving

$$M\boldsymbol{\xi}_n = \boldsymbol{c}_n, \qquad n = 1, \dots N,$$

   with $\boldsymbol{c}_k$ subjected to "sparsity constraints"

SCA: the intuition behind proving uniqueness of $C = M\Xi$

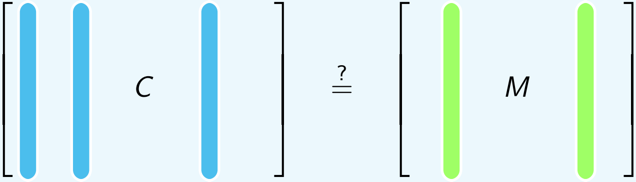
Given a collection of subspaces of columns of M, when is it possible to *all* retrieve individual columns of M?



Answer: when the *"richness"* property is satisfied + and M has sufficiently high Kruskal rank

SCA: the intuition behind proving uniqueness of $C = M\Xi$

Given sparsity constraints on Ξ, under what *assumption* can one decisively recover subspaces spanned by columns of M from subspaces spanned by columns of C?



Answer: if the so-called *"non-degeneracy"* assumption is satisfied

This leads to the following new uniqueness result

Theorem

Given $A \in \mathbb{C}^{I \times N}$, $B \in \mathbb{C}^{I \times N}$, $M \in \mathbb{C}^{K \times R}$, *and* $\Xi \in \mathbb{C}^{R \times N}$, *fix:*

$$2 \leq p \leq K, \qquad m := \lfloor \tfrac{p}{2} \rfloor.$$

(A, B, M, Ξ) *is the unique* $(L_r, L_r, 1)$*-decomposition satisfying properties:*

(1) A *and* B *have full column rank,*

(2) rank constraint: k-rank $M \geq p$,

(3) sparsity constraint: Ξ *has no zero rows and every column of* Ξ *has at least one and at most* m *nonzero entries,*

$\Longleftarrow$ Ξ *is sufficiently rich with parameter* m.

(4) (M, Ξ) *is non-degenerate up to parameter* m.

What does this mean for BSSE problem?
Interpretation of sparsity pattern in Ξ

Let $\{z_n\}_{n=1}^{N}$ be an enumeration of all the dinstinct poles in

$$\{z_{r,j} : 1 \leq r \leq R,\ 1 \leq j \leq L_r\}.$$

$$s_r(t) = \sum_{j=1}^{L_r} \alpha_{r,j} z_{r,j}^{t}, \qquad 1 \leq r \leq R,$$

What does this mean for BSSE problem?
Interpretation of sparsity pattern in Ξ

Let $\{z_n\}_{n=1}^N$ be an enumeration of all the dinstinct poles in

$$\{z_{r,j} : 1 \leq r \leq R,\ 1 \leq j \leq L_r\}.$$

$$s_r(t) = \sum_{n=1}^{N} \xi_{rn} z_n^t, \qquad 1 \leq r \leq R,$$

What does this mean for BSSE problem?
Interpretation of sparsity pattern in Ξ

Let $\{z_n\}_{n=1}^{N}$ be an enumeration of all the dinstinct poles in

$$\{z_{r,j} : 1 \leq r \leq R,\ 1 \leq j \leq L_r\}.$$

$$s_r(t) = \sum_{n=1}^{N} \xi_{rn} z_n^t, \qquad 1 \leq r \leq R,$$

nonzero entries in:

$$\begin{aligned} \textit{rows} \text{ of } \Xi \quad &= \quad \text{which poles are present in which source signal} \\ \textit{columns} \text{ of } \Xi \quad &= \quad \text{which poles are ``shared'' amongst which source signals} \end{aligned}$$

Under a mild technical assumption, a larger family of BSSE problems can be solved

K = number of observations, R = number of source signals

$K < R$	$K \geq R$
poles must be distinct	poles can be shared

Under a mild technical assumption, a larger family of BSSE problems can be solved

$K =$ number of observations, $R =$ number of source signals

$K < R$

poles can be shared
(provided non-degeneracy assumption)

$K \geq R$

poles can be shared

Contributions

- A new uniqueness result for $(L_r, L_r, 1)$-decompositions inspired from SCA principles.
- Direct application to BSSE, allowing for a richer set of problems to be solved.
- **(not discussed)** Numerical implications:

$$\text{LL1} = \text{CPD} + \text{SCA on third factor matrix.}$$

Open question

So far, we have *not* been able to answer the following question:

If the richness property is not satisfied,
can we conclude that the $(L_r, L_r, 1)$-decomposition is *not* unique?

Further details

Manuscript awaiting review in SIMAX:

$(L_r, L_r, 1)$*-decompositions, Sparse Component Analysis, and the blind separation of sums of exponentials*

N. Govindarajan, E. Epperly, L. De Lathauwer.