JEVD and CPD: algorithms and application to ICA and fluorescence spectroscopy. PART I

Laurent Albera [1,2,3], Xavier Luciani [4,5]

[1] INSERM U1099, Rennes F-35000, France.
[2] Université de Rennes 1, LTSI, Rennes F-35000, France.
[3] INRIA, Centre Inria Rennes - Bretagne Atlantique, 35042 Rennes Cedex, France.

[4] Aix Marseille Université, CNRS, ENSAM, LSIS, UMR 7296, 13397 Marseille, France.
[5] Université de Toulon, CNRS, LSIS, UMR 7296, 83957 La Garde, France.

TRICAP : ThRee-way methods In Chemistry And Psychology
7th edition, June 1st-5th, 2015, Belluno

Outline

OUTLINE

1. HOW TO MODIFY BIOME TO COMPUTE THE CP DECOMPOSITION OF SYMMETRIC EVEN HIGHER ORDER ARRAYS?
   - Introduction
   - BIOME and its modified version

TENSORS WITH SOME STRUCTURE

Definition (cubical array)

A Q-way ($Q \geq 2$) array $\mathcal{T} \in \mathbb{C}^{N_1 \times \cdots \times N_Q}$ is "cubical" if all its Q dimensions are identical, i.e. $N_1 = \cdots = N_Q = N$

Definition (symmetric array)

Let $\mathfrak{S}_Q$ be the symmetric group of permutations on $[\![1, Q]\!]$. A cubical Q-way ($Q \geq 2$) array $\mathcal{T} \in \mathbb{C}^{N \times \cdots \times N}$ will be called "symmetric" if:

$$\forall (n_1, \cdots, n_Q) \in [1; Q]_{\mathbb{N}}^{Q}, \qquad \mathcal{T}_{\sigma(n_1), \cdots, \sigma(n_Q)} = \mathcal{T}_{n_1, \cdots, n_Q}$$

for all permutations $\sigma \in \mathfrak{S}_Q$.

We define the subset $[m; n]_{\mathbb{N}}$ of integers by $[\![m, n]\!] = \{m, m+1, \ldots, n\} = [m, n] \bigcap \mathbb{N}$.

CANONICAL POLYADIC (CP) DECOMPOSITION

Definition (Sets)

Let $C^Q(\mathbb{C}^N) = \mathbb{C}^N \otimes \cdots \otimes \mathbb{C}^N$ (Q copies) be the set of all order-Q dimension-N cubical tensors. Then the set of symmetric tensors in $C^Q(\mathbb{C}^N)$ will be denoted by $S^Q(\mathbb{C}^N)$.

Lemma (Symmetric CP model)

Let $\mathcal{T} \in S^Q(\mathbb{C}^N)$. Then there exists $\mathbf{H} = (H_{n,p}) \in \mathbb{C}^{N\times P}$ and a diagonal array $\mathcal{S} \in S^Q(\mathbb{R}^P)$ such that [Comon et al., 2008]:

$$\forall(n_1, \cdots, n_Q) \in [1;Q]_{\mathbb{N}}^Q, \qquad \mathcal{T}_{n_1,\cdots,n_Q} = \sum_{p=1}^{P} \mathcal{S}_{p,\cdots,p} H_{n_1,p} \cdots H_{n_Q,p} \tag{1}$$

Definition (Positive semi-definiteness)

An array $\mathcal{T} \in S^Q(\mathbb{C}^N)$ is called Positive Semi-Definite (PSD) if the P components $\mathcal{S}_{p,\cdots,p}$ of the diagonal core array $\mathcal{S} \in S^Q(\mathbb{R}^P)$ are positive.

CP DECOMPOSITION ALGORITHMS

Iterative algorithms

- ALS [Harshman, 1970] [Bro, 1998] [Smilde, Bro and Geladi, 2004];
- Gauss-Newton, LM, CG [Tomasi and Bro, 2006] [Acar et al., 2011];
- for semi-symmetric 3rd order tensors: [Carroll and Chang, 1970] [Yeredor, 2002] [Maurandi and Moreau, 2014]
- for semi-nonnegative semi-symmetric 3rd order tensors: [Wang et al., 2013] [Coloigner et al., 2014] [Wang et al., 2014]

Semi-algebraic approaches

Idea: rewrite the CP decomposition as a matrix decomposition problem

- By congruence [De Lathauwer, 2004];
- By similarity [Roemer and Haardt, 2008] [Luciani and Albera 2011];
- For PSD symmetric even order tensors (BIOME): [Albera et al., 2004].

BIOME ALGORITHM (AT ORDER 4)

In the following we describe the real version of the BIOME (Blind Identification of Over-complete MixturEs of sources) algorithm but the complex version is almost identical.

- Let $\mathcal{T} \in S^4(\mathbb{R}^N)$ have the symmetric CP decomposition (1);
- Let $\mathbf{T} \in S^2(\mathbb{R}^{N^2})$ be the unfolding matrix of $\mathcal{T}$ such that:
$$\mathbf{T} = (\mathbf{H} \odot \mathbf{H})\mathbf{S}(\mathbf{H} \odot \mathbf{H})^\mathsf{T} \tag{2}$$
with $\mathbf{S} \in S^2(\mathbb{R}^{P^2})$ the corresponding diagonal unfolding matrix of $\mathcal{S}$ (1);
- Let $\mathbf{V}\Sigma\mathbf{V}^\mathsf{T}$ be its rank-P truncated EigenValue Decomposition (EVD).

Then it exists a non-singular matrix $\mathbf{W} \in \mathbb{R}^{P \times P}$ such that:

$$(\mathbf{H} \odot \mathbf{H})\,\mathbf{S}^{1/2} = \mathbf{V}\,\Sigma^{1/2}\,\mathbf{W} \quad \text{and} \quad \mathbf{S}^{1/2}\,(\mathbf{H} \odot \mathbf{H})^\mathsf{T} = \mathbf{W}^{-1}\,\Sigma^{1/2}\,\mathbf{V}^\mathsf{T} \tag{3}$$

Let $\phi^{(1)}, \cdots, \phi^{(N)}$ be the diagonal matrices built from the rows of matrix $\mathbf{H}$.

$$\mathbf{S}^{1/2}(\mathbf{H} \odot \mathbf{H})^\mathsf{T} = \left[\Phi^{(1)}\mathbf{H}^\mathsf{T}, \cdots, \Phi^{(N)}\mathbf{H}^\mathsf{T}\right] \tag{4}$$

$$\Sigma^{1/2}\mathbf{V}^\mathsf{T} = \left[\underbrace{\mathbf{W}\Phi^{(1)}\mathbf{H}^\mathsf{T}}_{\Gamma^{(1)\mathsf{T}}}, \cdots, \underbrace{\mathbf{W}\Phi^{(N)}\mathbf{H}^\mathsf{T}}_{\Gamma^{(N)\mathsf{T}}}\right] \tag{5}$$

Defining $\mathbf{M}^{(k_1,k_2)} = \Gamma^{(k_1)\sharp}\Gamma^{(k_2)}$ and $\Lambda^{(k_1,k_2)} = \Phi^{(k_1)}{}^{-1}\Phi^{(k_2)}$, we have:

$$\boxed{\forall (k_1, k_2) \in [1; N]^2_{\mathbb{N}},\ k_2 > k_1, \quad \mathbf{M}^{(k_1,k_2)} = \mathbf{W}^{-\mathsf{T}}\Lambda^{(k_1,k_2)}\mathbf{W}^\mathsf{T}} \tag{6}$$

Joint EVD (JEVD) problem

- $\mathcal{T}$ is PSD $\Rightarrow \mathbf{W} \in O(P)$;
- Orthogonal JEVD computed using the JAD (Joint Approximate Diagonalization) algorithm [Cardoso and Souloumiac, 1996].

$O(P)$ and $\sharp$ denote the group of orthogonal $P \times P$ matrices and the Moore-Penrose matrix inverse operator, respectively.

Let $\phi^{(1)}, \cdots, \phi^{(N)}$ be the diagonal matrices built from the rows of matrix $\mathbf{H}$.

$$\mathbf{S}^{1/2}(\mathbf{H} \odot \mathbf{H})^{\mathsf{T}} = \left[\mathbf{S}^{1/2}\Phi^{(1)}\mathbf{H}^{\mathsf{T}}, \cdots, \mathbf{S}^{1/2}\Phi^{(N)}\mathbf{H}^{\mathsf{T}}\right] \tag{7}$$

$$\Sigma^{1/2}\mathbf{V}^{\mathsf{T}} = \left[\underbrace{\mathbf{W}\Phi^{(1)}\mathbf{H}^{\mathsf{T}}}_{\Gamma^{(1)\mathsf{T}}}, \cdots, \underbrace{\mathbf{W}\Phi^{(K)}\mathbf{H}^{\mathsf{T}}}_{\Gamma^{(N)\mathsf{T}}}\right] \tag{8}$$

Defining $\mathbf{M}^{(k_1,k_2)} = \Gamma^{(k_1)\sharp}\Gamma^{(k_2)}$ and $\Lambda^{(k_1,k_2)} = {\Phi^{(k_1)}}^{-1}\Phi^{(k_2)}$, we have:

$$\boxed{\forall (k_1, k_2) \in [1; N]^2_{\mathbb{N}},\ k_2 > k_1, \qquad \mathbf{M}^{(k_1,k_2)} = \mathbf{W}^{-\mathsf{T}}\Lambda^{(k_1,k_2)}\mathbf{W}^{\mathsf{T}}} \tag{9}$$

Joint EVD (JEVD) problem

- $\mathcal{T}$ is not PSD $\Rightarrow \mathbf{W} \in \mathbb{R}^{P\times P}$;
- Need for an efficient non-orthogonal JEVD solver!

OUTLINE

2 JET: A SOLUTION TO THE JEVD PROBLEM
- The JEVD problem
- Two algorithms based on the **LU** decomposition

Problem formulation

Find a non-singular matrix $\mathbf{A} \in \mathbb{R}^{P \times P}$ from a set of non-defective matrices $\mathbf{M}^{(k)}$ so that:

$$\forall k \in [1; K]_{\mathbb{N}}, \qquad \mathbf{M}^{(k)} = \mathbf{A}\mathbf{D}^{(k)}\mathbf{A}^{-1} \tag{10}$$

where the K matrices $\mathbf{D}^{(k)} \in \mathbb{R}^{P \times P}$ are diagonal and unknown.

State of the art

All these algorithm resort to a Jacobi-like iterative procedure.

- sh-rt [Fu, 2006] based on the polar decomposition of $\mathbf{A}$.
- JUST [Iferroudjene, 2009] based on the polar decomposition of $\mathbf{A}$.
- JDTM [Luciani and Albera, 2010] based on the polar decomposition of $\mathbf{A}$.
- JET (JET-U and JET-O) [Luciani and Albera, 2011 and 2015] based on the **LU** decomposition of $\mathbf{A}$.
- JDTE [André, 2015] global estimation of $\mathbf{A}$ at each iteration.

LU decomposition

Due to the indeterminacies of the JEVD problem, the matrix $\mathbf{A}$ (12) can be chosen of the form $\mathbf{A} = \mathbf{LU}$ (without any loss of generality) with:

- $\mathbf{L}$: unit lower triangular matrix (1 on the diagonal)
- $\mathbf{U}$: unit upper triangular matrix (1 on the diagonal)

Joint triangularization

Let $\mathbf{R}^{(k)}$ be given by $\mathbf{R}^{(k)} = \mathbf{U}\mathbf{D}^{(k)}\mathbf{U}^{-1}$ for any $k \in [1;K]_{\mathbb{N}}$.

- Joint triangularization of the K matrices $\mathbf{M}^{(k)}$ by $\mathbf{L}$:

$$\forall k \in [1;K]_{\mathbb{N}}, \qquad \mathbf{M}^{(k)} = \mathbf{L}\,\mathbf{R}^{(k)}\,\mathbf{L}^{-1}$$

- Direct computation of the unit upper triangular matrix $\mathbf{U}$ from the set of matrices $\mathbf{R}^{(k)}$ (component by component).

Definition (elementary lower triangular matrix)

An elementary lower triangular matrix $\mathbf{L}^{(i,j)}(a)$ *is a unit lower triangular matrix with only one non-null off-diagonal component a located at the i-th row and the j-th column.*

Lemma (**LU** factorization)

Any unit lower triangular matrix $\mathbf{L}$ *of size* $(P \times P)$ *can be factorized as a product of* $M = P(P-1)/2$ *elementary lower triangular matrices:*

$$\mathbf{L} = \prod_{j=1}^{P-1} \prod_{i=j+1}^{P} \mathbf{L}^{(i,j)}(\ell_{i,j})$$

Algorithm (Jacobi-like procedure)

Repeat several times a series (called "sweep") of M sequential optimizations with respect to only one parameter.

Update of the matrices to be triangularized

$$\forall (i,j) \in [1;P]_{\mathbb{N}}^2,\ i > j,\ \forall k \in [1;K]_{\mathbb{N}},\ \mathbf{M}^{(k)} \leftarrow \left(\mathbf{L}^{(ij)}(x_{i,j})\right)^{-1} \mathbf{M}^{(k)} \mathbf{L}^{(ij)}(x_{i,j}) \quad (11)$$

- Each of these updates only depends on one parameter $x_{i,j}$;
- Each parameter $x_{i,j}$ is computed in order to sequentially improve the upper triangular structure of the $\mathbf{L}^{(i,j)}(x_{i,j})$-updated matrices.

Objective functions

$$\zeta_O(x_{i,j}) = \sum_{k=1}^{K} \sum_{q=1}^{P-1} \sum_{p=q+1}^{P} \left(M_{p,q}^{(k)}\right)^2 \quad (12)$$

$$\zeta_U(x_{i,j}) = \sum_{k=1}^{K} \left(M_{i,j}^{(k)}\right)^2 \quad (13)$$

OUTLINE

3. APPLICATION TO ICA
   - Cumulant-based ICA
   - Computer results

Definition (Independent Component Analysis (ICA))

ICA of a N-dimensional random vector $\mathbf{x}$ *of finite covariance matrix* $\mathbf{C}_x$ *is given by a pair of matrices* $(\mathbf{H}, \mathbf{C}_x)$ *such that:*

1. *the covariance matrix* $\mathbf{C}_x$ *factorizes into* $\mathbf{C}_x = \mathbf{H}\,\mathbf{C}_s\,\mathbf{H}^\top$ *where* $\mathbf{C}_s$ *is diagonal real positive and is full column rank P;*
2. *the random vector* $\mathbf{x}$ *can be written as* $\mathbf{x} = \mathbf{H}\,\mathbf{s}$ *where* $\mathbf{s}$ *is a P-dimensional random vector with covariance* $\mathbf{C}_s$ *and whose components are "the most independent possible".*

Interest

Solving Blind Source Separation (BSS) or blind source subspace identification problems as those encountered in biomedical engineering.

Definition (Cumulants)

Let $\Psi_{\mathbf{x}}$ be the second characteristic function of a real-valued random vector $\mathbf{x}$ defined by $\Psi_{\mathbf{x}}(\mathbf{u}) = \log(\mathsf{E}[\mathrm{e}^{\mathrm{i}\,\mathbf{u}^{\mathsf{T}}\mathbf{x}}])$, where $\mathsf{E}[z]$ denotes the mathematical expectation of z. The the Q-th order ($Q \geq 1$) cumulant array $\mathcal{C}_{Q,\mathbf{x}} = \left(\mathcal{C}_{n_1,\ldots,n_Q,\mathbf{x}}\right)$ of $\mathbf{x}$ is then defined by:

$$\left(\mathcal{C}_{n_1,\ldots,n_Q,\mathbf{x}}\right) = (-i)^Q \left.\frac{\partial^Q \Psi_{\mathbf{x}}(\mathbf{u})}{\partial u_{n_1} \cdots \partial u_{n_Q}}\right|_{\mathbf{u}=\mathbf{0}} \tag{14}$$

Definition (Leonov-Shirayev formula)

There is a link between Q-th order cumulants and moments of order less than Q. For instance, the components of the FO cumulant array of a zero-mean random vector $\mathbf{x}$ are given by:

$$\begin{aligned}\mathcal{C}_{n_1,n_2,n_3,n_4,\mathbf{x}} = {} & \mathsf{E}[x_{n_1}x_{n_2}x_{n_3}x_{n_4}] - \mathsf{E}[x_{n_1}x_{n_2}]\mathsf{E}[x_{n_3}x_{n_4}] \\ & - \mathsf{E}[x_{n_1}x_{n_3}]\mathsf{E}[x_{n_2}x_{n_4}] - \mathsf{E}[x_{n_1}x_{n_4}]\mathsf{E}[x_{n_2}x_{n_3}]\end{aligned}$$

Some properties...

- Symmetry of a cumulant array;
- Diagonality of the cumulant array of a random vector for which components are mutually independent;
- Additivity: $\mathcal{C}_{Q,\mathbf{x}_1+\mathbf{x}_2} = \mathcal{C}_{Q,\mathbf{x}_1} + \mathcal{C}_{Q,\mathbf{x}_2}$ if $\mathbf{x}_1$ and $\mathbf{x}_2$ are independent;
- Cancellation of the Q-th ($Q > 2$) cumulant array of a Gaussian vector;
- Multi-linearity of the cumulant function $\mathcal{C}_{n_1,\ldots,n_Q,(\cdot)}$ such that:

$$\mathcal{C}_{n_1,\ldots,n_Q,\mathbf{x}} = \sum_{p_1=1}^{P} \ldots \sum_{p_Q=1}^{P} \mathcal{C}_{p_1,\ldots,p_Q,\mathbf{s}} H_{n_1,p_1} \ldots H_{n_Q,p_Q} \;\text{ if }\; \mathbf{x} = \mathbf{H}\,\mathbf{s}$$

Very useful for ICA...

ICA can be performed by computing the rank-P symmetric CP decomposition (with modified BIOME) of $\mathcal{C}_{Q,\mathbf{x}}$. Indeed we have:

$$\mathcal{C}_{n_1,\ldots,n_Q,\mathbf{x}} = \sum_{p=1}^{P} \mathcal{C}_{p,\ldots,p,\mathbf{s}} H_{n_1,p} \ldots H_{n_Q,p}$$

ICA of ECoG signals: a biomedical engineering BSS problem

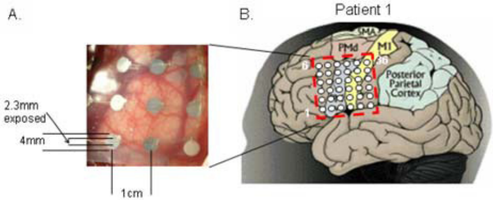
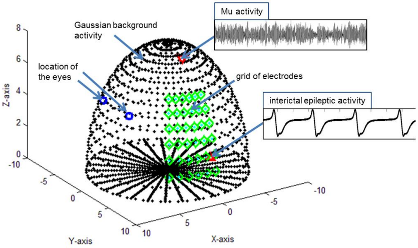
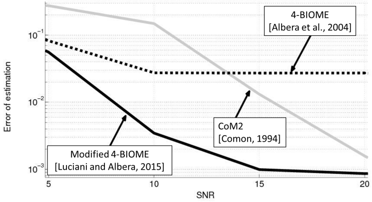


Figure: A. In vivo placement of the electrode grid in a 3x3cm area of cortex and the relative electrode, gyri, sulci, and vasculature relationships. B. Electrode placement. From [Gunduz et al., 2008] with permission.

- Design of neuroprosthesis from some brain electrical sources recorded by means of subdural electrodes, i.e. from ElectroCorticoGraphy (ECoG) signals;
- Coruption of the sources of interest, for instance the Mu rhythm, by epileptic activities.



(a) Simulated half-sphere-like cortex.



(b) Global estimation error of **H**.

$$\mathbf{x}[m] = \mathbf{H}\,\mathbf{s}[m] + \nu[m] \tag{15}$$

- Simulation of $N = 36$ electrodes recording ECoG data;
- Simulation of one epileptic activity in the left parietal lobe and one Mu activity in the somatosensory area (i.e. $P = 2$ sources);
- Gaussian additive noise with an unknown covariance matrix.

OUTLINE

4 CONCLUSION

Conclusion

- How to relax the positive semi-definiteness assumption of BIOME;
- Reformulating the symmetric CP decomposition as a non-orthogonal JEVD problem;
- Proposition of two JEVD algorithms based on the **LU** factorization for real- and complex-valued matrices;
- Application to ICA in order to process mixtures of sources involving a Gaussian noise with an unknown covariance matrix.

[Luciani and Albera, 2015] X. Luciani and L. Albera, "Joint eigenvalue decomposition of non-defective matrices based on the LU factorization with application to ICA," to appear in IEEE Transactions on Signal Processing.

Perspectives

- Extension of the modified BIOME algorithm to the case of non-symmetric arrays (presented in one minute);
- Reformulating the symmetric CP decomposition as a **J**-unitary JEVD problem.

A matrix $\mathbf{A}$ *is* $\mathbf{J}$*-unitary if* $\mathbf{A}\,\mathbf{J}\,\mathbf{A}^{\mathsf{H}} = \mathbf{J}$ *where* $\mathbf{J}$ *is a sign diagonal matrix.*