Higher-Order Singular Value Decomposition (HOSVD) for structured tensors

Definition and applications

Rémy Boyer

Laboratoire des Signaux et Système (L2S)
Université Paris-Sud XI

GDR ISIS, January 16, 2012

Joint work and articles

Collaborations:

- Roland Badeau (ENST-Paris, LTSI lab./COD team)
- Gérard Favier (CNRS, I3S lab.)

Publications:

- SIAM J. on Matrix Analysis and Applications: ”Fast multilinear singular value decomposition for structured tensors”, 2008, 30, 3.
- ICASSP’06: ”Adaptive multilinear SVD for structured tensors”
- ICASSP’11: ”Fast orthogonal decomposition of Volterra cubic kernels using oblique unfolding”

Context of this work

- The subject of multilinear decompositions is relatively new in the SP community.

- There are essentially two families:

  I CP (CANonical DECOMPosition/PARAllel FACtors) model [Carroll][Harshman]. Relied to the tensor rank.

  II.a Tucker model [Tucker].
  Related to the n-mode ranks. The CP is a special case of the Tucker model.

  II.b If orthogonality constraints are imposed, one can refer to the Higher-Order Singular Value Decomposition (HOSVD) [DeLathauwer] or multilinear SVD.
  $\Rightarrow$ The HOSVD is the subject of this work.

Where can we find HOSVD ?

This decomposition plays an important role in various domains, such as:

- ⋆ Spectral analysis,
- ⋆ Non-linear modeling,
- ⋆ Communication and Radar processing,
- ⋆ blind source separation,
- ⋆ image processing,
- ⋆ biomedical applications (magnetic resonance imaging and electrocardiography),
- ⋆ web search,
- ⋆ computer facial recognition,
- ⋆ handwriting analysis,
- ⋆ ...

Definition of the HOSVD

Every $I_1 \times I_2 \times I_3$ tensor $\boldsymbol{\mathcal{A}}$ can be written as:

$$\boldsymbol{\mathcal{A}} = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \sum_{i_3=1}^{I_3} \sigma_{i_1 i_2 i_3} \left(\mathbf{u}_{i_1}^{(1)} \circ \mathbf{u}_{i_2}^{(2)} \circ \mathbf{u}_{i_3}^{(3)} \right)$$

where $\circ$ is the outer product and

$$\mathbf{U}^{(1)} = \begin{bmatrix} \mathbf{u}_1^{(1)} & \ldots & \mathbf{u}_{I_1}^{(1)} \end{bmatrix}, \ \mathbf{U}^{(2)} = \begin{bmatrix} \mathbf{u}_1^{(2)} & \ldots & \mathbf{u}_{I_2}^{(2)} \end{bmatrix}, \ \mathbf{U}^{(3)} = \begin{bmatrix} \mathbf{u}_1^{(3)} & \ldots & \mathbf{u}_{I_3}^{(3)} \end{bmatrix}$$

are three $I_s \times I_s$ unitary matrices. An other expression of the HOSVD based on the Tucker product is

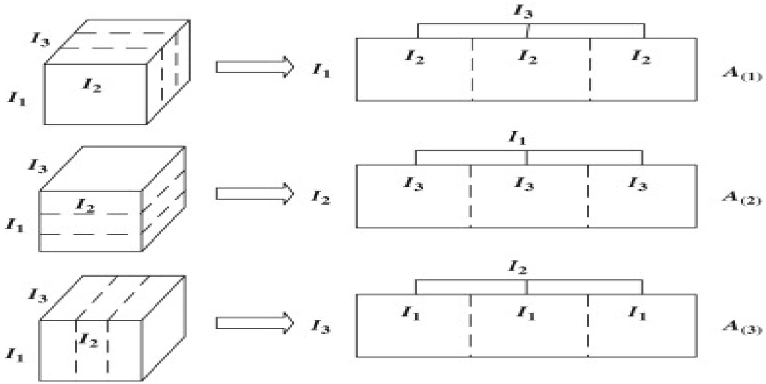
$$\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{S}} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)}$$

where

$$[\boldsymbol{\mathcal{S}}]_{i_1 i_2 i_3} = \sigma_{i_1 i_2 i_3}$$

is a all-orthogonal core tensor.

HOSVD and standard unfoldings



Using the fast SVD implementation based on orthogonal iteration algorithm:

Operation	Cost per iteration
SVD of $\mathbf{A}_1$	$O(M_1 I_1 I_2 I_3)$
SVD of $\mathbf{A}_2$	$O(M_2 I_1 I_2 I_3)$
SVD of $\mathbf{A}_3$	$O(M_3 I_1 I_2 I_3)$
Final cost	$O(MI^3)$

<u>Very high computational cost !</u>

The structure, the heart of many signal processing applications

- In many signal processing applications, structured tensors are involved. For instance,
  - Higher Principal component analysis (PCA)
  - Kernel decomposition in Volterra series
  - 1-D Harmonic retrieval problem with subspace methods
- Standard modes do not present a particularly noticeable structure even in the case of structured tensors.
- **BUT**... the modes (up to column-permutations) are IN FACT strongly structured.

  $\Longrightarrow$ Need a new way to unfold a structured tensor.
- To the best of our knowledge, there are no specific HOSVD algorithms proposed in the literature for exploiting tensors structures.

Higher PCA for real moment and cumulant

- The HOSVD can be viewed as a higher Principal Component Analysis (PCA) for data dimensional reduction.

- Third-order moment and cumulant tensors are defined according to

$$[\mathcal{M}]_{t_1 t_2 t_3} = E\{x(t_1)x(t_2)x(t_3)\}, \tag{1}$$

$$\begin{aligned}[\mathcal{C}]_{t_1 t_2 t_3} = & \; E\{x(t_1)x(t_2)x(t_3)\} + 2E\{x(t_1)\}E\{x(t_2)\}E\{x(t_3)\} \\ - & \; E\{x(t_1)\}E\{x(t_2)x(t_3)\} - E\{x(t_2)\}E\{x(t_1)x(t_3)\} - E\{x(t_3)\}E\{x(t_1)x(t_2)\}.\end{aligned} \tag{2}$$

  where $t_1, t_2, t_3 \in \{0 \ldots I-1\}$, and $x(t)$ is a real random process.

- Moment and cumulant are symmetric tensors.

- If $x(t)$ is a third-order stationary process, the moment and cumulant tensors are third-order Toeplitz tensors (if $x(t)$ is a stationary process, its probability distribution is invariant to temporal translations: $[\mathcal{C}]_{t+i_1,t+i_2,t+i_3} = [\mathcal{C}]_{i_1 i_2 i_3}$).

Kernel decomposition in Volterra series [ICASSP'11]

The input/output of a multidimensional convolution of the input signal is

$$y[n] = \sum_{m=1}^{M} \sum_{k_1,\ldots,k_m=0}^{I-1} [\boldsymbol{\mathcal{H}}_m]_{k_1\ldots k_m} x[n-k_1]\ldots x[n-k_m] \tag{3}$$

where

- The aim is to factorize the cubic Volterra kernel $\boldsymbol{\mathcal{H}}_m$.
- If there is no symmetric relations between the entries, there always exists an associated symmetrized kernel according to

$$[\boldsymbol{\mathcal{S}}_m]_{k_1\ldots k_m} = \frac{1}{m!} \sum_{\pi\in\mathcal{P}} [\boldsymbol{\mathcal{H}}_m]_{k_{\pi(1)}\ldots k_{\pi(m)}} \tag{4}$$

  where $\mathcal{P}$ is the set of cardinal $m!$ of all the permutations of $\{1,\ldots,m\}$.

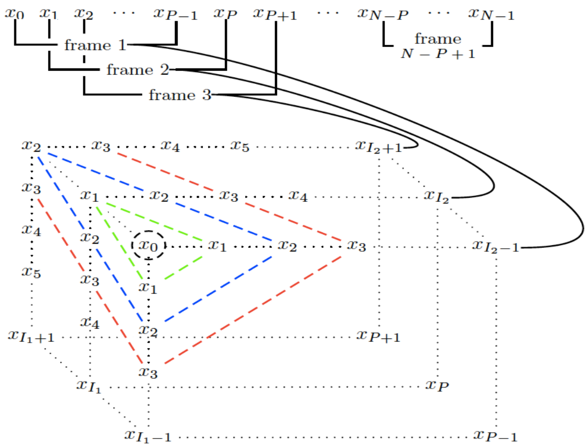
1-D harmonic retrieval problem

- The complex harmonic model is

$$x_n = \sum_{m=1}^{M} \alpha_m e^{(\delta_m + i\phi_m)n}, \quad \text{for } \; n \in [0 : N-1] \tag{5}$$

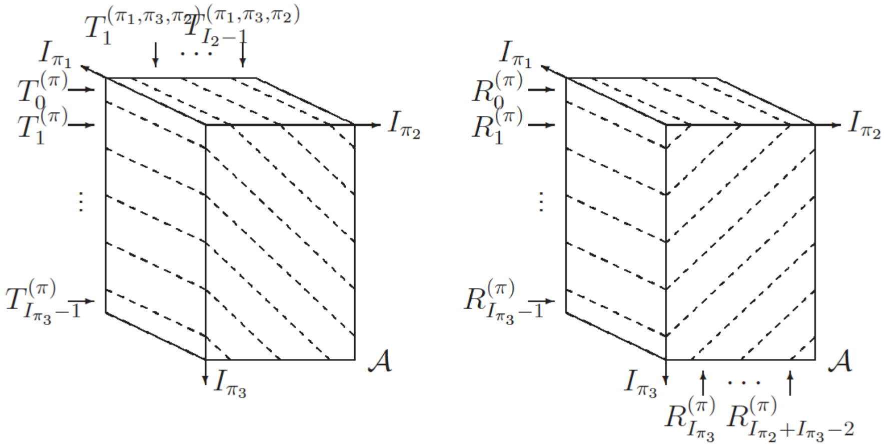
  where the aim is to estimate ϕ_m, the angular-frequency and δ_m, the damping factor.

- The Hankel tensor $[\boldsymbol{\mathcal{A}}]_{i_1 i_2 i_3} = x_{(i_1+i_2+i_3)}$ is diagonalizable in three Vandermonde basis [Papy, Delathauwer and VanHuffel]



New way to see a structured tensor

- Type-1 and type-2 oblique unfolding of a tensor:



Property 1 *Type-1 unfolding of symmetric tensors $\rightarrow$ reordered tensor modes has an axial blockwise symmetry w.r.t. its central oblique sub-matrix.*

Property 2 *Type-1 unfolding of Toeplitz tensors $\rightarrow$ reordered tensor modes are block-Toeplitz.*

Property 3 *Type-2 unfolding of Hankel tensors $\rightarrow$ reordered tensor modes are composed by rank-1 oblique sub-matrices.*

Oblique unfoldings to decrease the complexity

Structured modes by oblique unfolding

$\Downarrow$

fast techniques from numerical linear algebra

Thanks to the oblique type-1 and type-2 unfoldings, we have two improvements:

1. Exploit the column-redundancy resulting from oblique unfoldings
   $\rightarrow$ symmetric or Hankel tensors.

2. Exploit fast matrix-vector products
   $\rightarrow$ specific to Toeplitz and Hankel matrices.

Hankel tensor: an example

Let $[\mathcal{A}]_{ijk} = i + j + k$ be a Hankel tensor of size $4 \times 4 \times 4$.

- The 1-mode is

$$\mathbf{A}_1 = \left[\begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{bmatrix} \begin{bmatrix} 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \\ 5 & 6 & 7 & 8 \\ 6 & 7 & 8 & 9 \end{bmatrix} \right].$$

- A type-2 reordered tensor mode is formed of 7 rank-1 matrices $\mathbf{R}_0 \ldots \mathbf{R}_6$:

$$\mathbf{A}'_1 = \left[\underbrace{\begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}}_{\mathbf{R}_0} \underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \\ 4 & 4 \end{bmatrix}}_{\mathbf{R}_1} \underbrace{\begin{bmatrix} 2 & 2 & 2 \\ 3 & 3 & 3 \\ 4 & 4 & 4 \\ 5 & 5 & 5 \end{bmatrix}}_{\mathbf{R}_2} \underbrace{\begin{bmatrix} 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \\ 5 & 5 & 5 & 5 \\ 6 & 6 & 6 & 6 \end{bmatrix}}_{\mathbf{R}_3} \underbrace{\begin{bmatrix} 4 & 4 & 4 \\ 5 & 5 & 5 \\ 6 & 6 & 6 \\ 7 & 7 & 7 \end{bmatrix}}_{\mathbf{R}_4} \underbrace{\begin{bmatrix} 5 & 5 \\ 6 & 6 \\ 7 & 7 \\ 8 & 8 \end{bmatrix}}_{\mathbf{R}_5} \underbrace{\begin{bmatrix} 6 \\ 7 \\ 8 \\ 9 \end{bmatrix}}_{\mathbf{R}_6} \right].$$

- The first column is repeated 1 time, the second column is repeated 2 times...

Exploiting column-redundancy (1)

- Define the $I_s \times J_s$ matrix $\mathbf{H}_s$ as the matrix obtained by removing the repeated columns in the s-mode with $J_s = \sum_{s \neq s'} I_s - 1$.

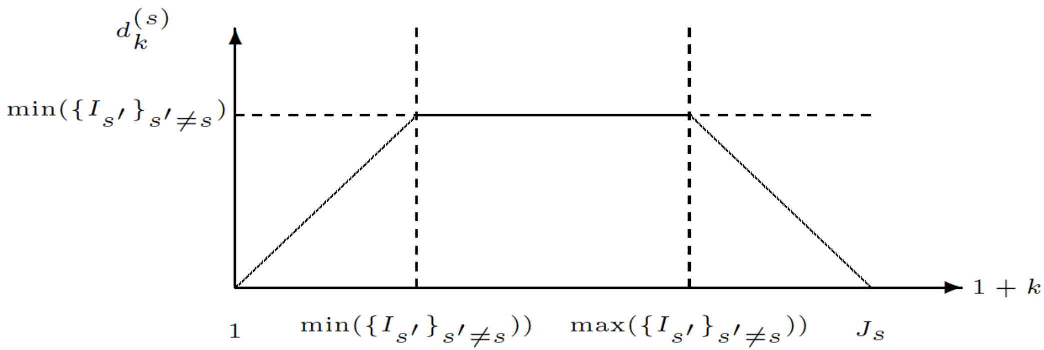
  In the proposed example ($J_1 = 7$): $\mathbf{H}_1 = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix}$.

- Let $d_k^{(s)}$ be the number of occurrences of the k^{th} column of $\mathbf{H}_s$ in the s-mode.

  In the proposed example:

$$d_1^{(1)} = d_7^{(1)} = 1, d_2^{(1)} = d_6^{(1)} = 2, d_3^{(1)} = d_5^{(1)} = 3, d_4^{(1)} = 4.$$

- The general form of the occurrence is given by:



Exploiting column-redundancy (2)

- Let $\mathbf{a}_k^{(s)}$ be the k-th column of mode $\mathbf{A}_s$, using the column-redundancy property, we have

$$\mathbf{A}_s\mathbf{A}_s^H = \sum_{k=1}^{I_s} \mathbf{a}_k^{(s)}\mathbf{a}_k^{(s)H} = \sum_{k\in\Omega_s} d_k^{(s)}\mathbf{a}_k^{(s)}\mathbf{a}_k^{(s)H}$$

  with $|\Omega_s| = J_s$.

- Same dominant left singular space:

$$\mathcal{R}\{\mathbf{A}_s\} = \mathcal{R}\left\{\mathbf{H}_s\begin{bmatrix}\sqrt{d_0^{(s)}} & & \\ & \ddots & \\ & & \sqrt{d_{J_s-1}^{(s)}}\end{bmatrix}\right\}.$$

- The column dimensions of $\mathbf{H}_s\,\mathbf{D}_s$ and $\mathbf{A}_s$ satisfiy

$$J_s < \prod_{s\neq s'} I_s$$

- SVD of modes with a smaller number of columns !

$\rightarrow O(MI^2)$: Gain of one order of magnitude w.r.t. the standard algorithm.

Complexities of the HOSVD algorithms

Computation of the (M, M, M)-HOSVD by the orthogonal iteration method for a cubic tensor.

Structure	Global cost per iteration
unstructured	$12MI^3$
symmetric	$6MI^3$
Toeplitz (fast)	$540MI^2\log_2(I)$
symmetric Toeplitz (fast)	$90MI^2\log_2(I)$
Hankel (fast)	$24MI^2$
cubic Hankel (fast)	$8MI^2$
Hankel (ultra-fast)	$270MI\log_2(I)$
cubic Hankel (ultra-fast)	$60MI\log_2(I)$

Conclusion and perspectives

- Structured tensors imply strongly structured modes if oblique unfoldings are used. Not true for standard unfoldings !

- Increasing the structure of the modes allows to exploit fast techniques from numerical linear algebra based on

  - the column-redundancy property

  - fast products vector/matrix for Toeplitz or Hankel matrices.

- Fastest implementation of the rank-truncated HOSVD (dedicated to Hankel tensors) has a quasilinear complexity w.r.t. the tensor dimensions.

- Generalize to tensors of order > 3.

- Extend to other HOSVD (constrained HOSVD, cross-HOSVD,...)?