High order coupled matrix-tensor factorization

Journée GDR - Nouvelles méthodes tensorielles et applications
Ivry-sur-Seine, 17 juin 2019

Abdelhak BOUDEHANE
Laboratoire des signaux et système, Université de Paris-Sud

This is a joint work with

- Yassine Zniyed, Laboratoire des Signaux et Systèmes, Univ. Paris-Sud, France.
- Arthur Tenehaus, , Laboratoire des Signaux et Systèmes, Univ. Paris-Sud, France.
- Laurent Le Brusquet, , Laboratoire des Signaux et Systèmes, Univ. Paris-Sud, France.
- Rémy Boyer, Centre de Recherche en Informatique, Signal et Automatique de Lille, Univ. Lille, France.

References

- A. Boudehane, Y. Zniyed, A. Tenehaus, L. Le Brusquet and R. Boyer. *BREAKING THE CURSE OF DIMENSIONALITY FOR COUPLED TENSOR-MATRIX FACTORIZATION.* CAMSAP. 2019.
- Y. Zniyed, R. Boyer, A.L.F. de Almedia and G. Favier, *High-Order CPD Estimation with Dimensionality Reduction Using A Tensor Train Model*, EUSIPCO, 2018.
- Y. Zniyed, R. Boyer, A.L.F. de Almedia and G. Favier, *High-order tensor factorization via trains of coupled third-order CP and Tucker decompositions*, Linear Algebra and its Applications (LAA), 2018, submitted.
- Y. Zniyed, R. Boyer, A.L.F. de Almedia and G. Favier, *Multidimensional Harmonic Retrieval Based on Vandermonde Tensor Train*, Elsevier, Signal Processing, vol. 163, pp. 75-86, 2019.
- Y. Zniyed, R. Boyer, A.L.F. de Almedia and G. Favier, *A TT-Based Hierarchical Framework for Decomposing Big Data Tensors*, SIAM journal on Scientific Computing (SISC), 2018, submitted.
- Y. Zniyed, R. Boyer, A.L.F. de Almedia and G. Favier, *Tensor Train Representation of MIMO channels using the JIRAFE Method*, Elsevier, Signal Processing, 2019, submitted.
- Y. Zniyed, R. Boyer, A.L.F. de Almedia and G. Favier, *Tensor-Train modeling for MIMO-OFDM tensor coding-and-forwarding relay systems*, EUSIPCO, 2019.

Notations and tensor operations

Scalars, vectors, matrices, and tensors:

$$x, \mathbf{x}, \mathbf{X}, \boldsymbol{\mathcal{X}}$$

- Mode-n product

$$[\boldsymbol{\mathcal{X}} \times_n \mathbf{A}]_{i_1 \cdots i_{n-1} j i_{n+1} \cdots i_N} = \sum_{i_n=1}^{I_N} [\boldsymbol{\mathcal{X}}]_{i_1 i_2 \cdots i_N} [\mathbf{A}]_{j i_n}$$

- $\times_n^m$ product $[\boldsymbol{\mathcal{A}} \times_n^m \boldsymbol{\mathcal{B}}]_{i_1,\ldots,i_{n-1},i_{n+1},\ldots,i_N,j_1,\ldots,j_{m-1},j_{m+1},\ldots,j_M}$
$= \sum_{k=1}^{I_n} [\boldsymbol{\mathcal{A}}]_{i_1,\ldots,i_{n-1},k,i_{n+1},\ldots,i_N} [\boldsymbol{\mathcal{B}}]_{j_1,\ldots,j_{m-1},k,j_{m+1},\ldots,j_M}.$

Outline

1. Coupled Matrix/Tensor Factorization (CMTF)
2. All-At-Once Optimization
3. Curse of Dimensionality
4. dimensionality reduction : JIRAFE
5. Impact on CMTF
6. C-JIRAFE
7. Simulation

Coupled Matrix/Tensor Factorization (CMTF)

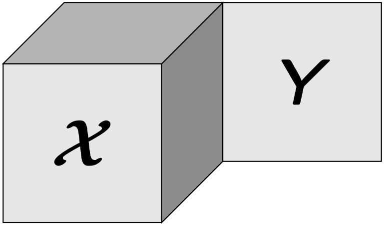


Figure: 3d-tensor $\boldsymbol{\mathcal{X}}$ coupled with matrix $\boldsymbol{Y}$ on the first mode

Canonical Polyadic (CP) model tensor

$$\boldsymbol{\mathcal{X}} = [\![\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_Q]\!]$$

Rank Factorization Matrix

$$\boldsymbol{Y} = \boldsymbol{P}_k \boldsymbol{V}^T$$

CMTF Model

$$f(\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_Q, \boldsymbol{V}) = ||\boldsymbol{\mathcal{X}} - [\![\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_Q]\!]||_F^2 + ||\boldsymbol{Y} - \boldsymbol{P}_k \boldsymbol{V}^T||_F^2$$

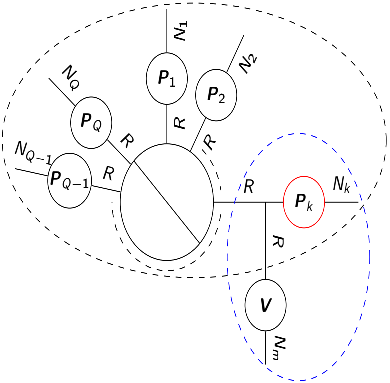


Figure: Factor graph modelization of CMTF

All-At-Once Optimization

Gradient-based approach : CMTF-OPT[1].

- All-at-once optimization.
- Gradient :

$$\nabla f = \begin{bmatrix} \text{vec}\left(\frac{\partial f}{\partial \boldsymbol{P}_1}\right) \\ \vdots \\ \text{vec}\left(\frac{\partial f}{\partial \boldsymbol{P}_Q}\right) \\ \text{vec}\left(\frac{\partial f}{\partial \boldsymbol{V}}\right) \end{bmatrix}$$

- Using Nonlinear Conjugate Gradient (NCG) algorithm to jointly-compute factor matrices $(\boldsymbol{P}_1, \ldots, \boldsymbol{P}_Q, \boldsymbol{V})$.

[1] E. Acar, T. G. Kolda, and D. M. Dunlavy. "All-at-once Optimization for Coupled Matrix and Tensor Factorizations". In: *in 9th Workshop on Mining and Learning with Graphs, San Diego, CA* (2011).

Curse of dimentionality

Limitation

- **Curse of dimensionality** : Case of coupled high-order tensor matrix factorization :

  $\rightarrow$ Number of elements increases exponentially in function of the dimensions.

  $\rightarrow$ Exponential increase of computation and memory requirements.

  $\Rightarrow$ **Limition in terms of order**

Dimensionality reduction : Tensor Train Model

D-order TT decomp. with TT-ranks $(R_1, \ldots, R_{D-1})$ [Oseledets, 2011]

$$\mathcal{X}(i_1, i_2, \cdots, i_D) = \sum_{r_1,\cdots,r_{D-1}=1}^{R_1,\cdots,R_{D-1}} \mathbf{G}_1(i_1, r_1)\mathcal{G}_2(r_1, i_2, r_2,)\mathcal{G}_3(r_2, i_3, r_3)\cdots$$
$$\cdots\mathcal{G}_{D-1}(r_{D-2}, i_{D-1}, r_{D-1})\mathbf{G}_D(r_{D-1}, i_D)$$

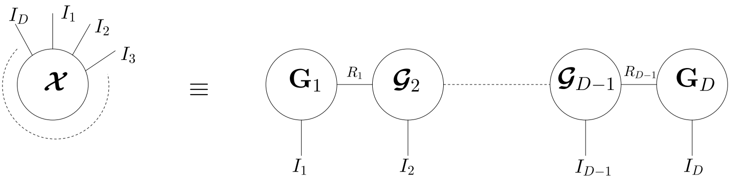


Fig. Forney-style factor graph [Forney, 2001]

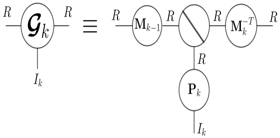
The most compact TN with a dominant storage cost : $O((D-2)IR^2)$. Linear in D.

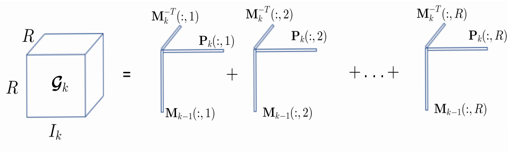
Model equivalence - Computational relation (TT-SVD)

Recall that a D-order CPD of rank-R is defined as $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{D,R} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2 \times_3 \ldots \times_D \mathbf{P}_D$

Key result [Zniyed, Boyer *et al*, submitted to LAA]

- $\boldsymbol{\mathcal{G}}_k = \boldsymbol{\mathcal{I}}_{3,R} \times_1 \mathbf{M}_{k-1} \times_2 \mathbf{P}_k \times_3 \mathbf{M}_k^{-T}$
- $\mathbf{G}_1 = \mathbf{P}_1 \mathbf{M}_1^{-1}$
- $\mathbf{G}_D = \mathbf{M}_{D-1} \mathbf{P}_D^T$





Impact on CMTF : Equivalence CMTF / Coupled-TT

- Reminding the Tensor-Train model

$$\boldsymbol{\mathcal{X}} = \boldsymbol{G}_1 \times_2^1 \boldsymbol{\mathcal{G}}_2 \times_3^1 \ldots \times_{Q-1}^1 \boldsymbol{\mathcal{G}}_{Q-1} \times_Q^1 \boldsymbol{G}_Q$$

- The matrix $\boldsymbol{Y}$

$$\boldsymbol{Y} = \boldsymbol{P}_k \boldsymbol{V}^T$$

- The TT-core $\boldsymbol{\mathcal{G}}_k$ shares the common factor $\boldsymbol{P}_k$ with $\boldsymbol{Y}$

$$\boldsymbol{\mathcal{G}}_k = [\![\boldsymbol{M}_{k-1}, \boldsymbol{P}_k, \boldsymbol{M}_k^{-T}]\!]$$

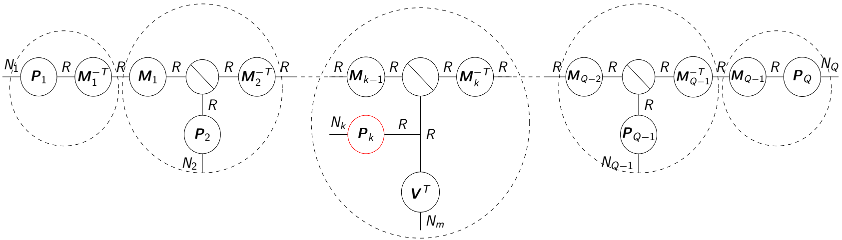


Figure: Factor graph modelization of coupled tensor-train

k-th 3-order CMTF problem:

$$f_2(\boldsymbol{M}_{k-1}, \boldsymbol{P}_k, \boldsymbol{M}_k, \boldsymbol{V}) = \lambda_1 ||\mathcal{G}_k - [\![\boldsymbol{M}_{k-1}, \boldsymbol{P}_k, \boldsymbol{M}_k^{-T}]\!]||_F^2 + \lambda_2 ||\boldsymbol{Y} - \boldsymbol{P}_k \boldsymbol{V}^T||_F^2$$

Breaking the curse of dimensionality

The high-order CMTF problem has been reduced into a 3-order CMTF and a collection of 3-order CPDs

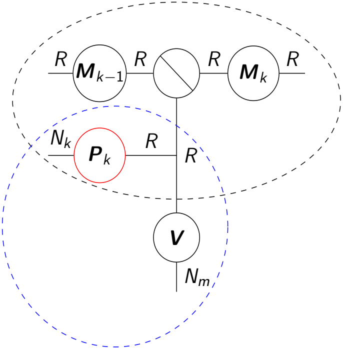


Figure: Factor graph modelization of 3-order CMTF

Estimation of the common factor $\boldsymbol{P}_k$ and matrix $\boldsymbol{V}$

- Least-squares solution for $\boldsymbol{P}_k$

$$\boldsymbol{P}_k = \left[\lambda_1 \boldsymbol{G}_k^{(2)} \quad \lambda_2 \boldsymbol{Y}\right] \left[\lambda_1 (\boldsymbol{M}_k^{-T} \odot \boldsymbol{M}_{k-1})^T \quad \lambda_2 \boldsymbol{V}^T\right]^{\dagger}$$

- $\boldsymbol{V}$ is given by

$$\boldsymbol{V} = (\boldsymbol{P}_k^T)^{\dagger} \boldsymbol{Y}$$

C-JIRAFE

Input: Tensor $\boldsymbol{\mathcal{X}}$, matrix $\boldsymbol{Y}$, order Q, rank R and common mode k
Output: Estimated factor matrices $\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_Q, \boldsymbol{V}$

1. TT cores estimation

$$[\boldsymbol{G}_1, \boldsymbol{\mathcal{G}}_2, \ldots, \boldsymbol{\mathcal{G}}_{Q-1}, \boldsymbol{G}_Q)] = \textbf{TT-SVD}(\boldsymbol{\mathcal{X}}, R)$$

2. Joint estimation of $\boldsymbol{P}_k$ and $\boldsymbol{V}$:

$$[\hat{\boldsymbol{M}}_{k-1}, \hat{\boldsymbol{P}}_k, \hat{\boldsymbol{M}}_k^{-T}, \hat{\boldsymbol{V}}] = \textbf{Exp-ALS}(\boldsymbol{\mathcal{G}}_k, \boldsymbol{Y}, R)$$

3. The rest of factors
   for $q = k-1 \ldots 2$

$$[\hat{\boldsymbol{M}}_{q-1}, \hat{\boldsymbol{P}}_q] = \textbf{Bi-ALS}(\boldsymbol{\mathcal{G}}_q, \hat{\boldsymbol{M}}_q^{-T}, R)$$

   end for
   for $q = k+1 \ldots Q-1$

$$[\hat{\boldsymbol{P}}_q, \hat{\boldsymbol{M}}_q^{-T}] = \textbf{Bi-ALS}(\boldsymbol{\mathcal{G}}_q, \hat{\boldsymbol{M}}_{q-1}, R)$$

   end for
   $\hat{\boldsymbol{P}}_1 = \boldsymbol{G}_1 \hat{\boldsymbol{M}}_1$ and $\hat{\boldsymbol{P}}_Q = \boldsymbol{G}_Q^T \hat{\boldsymbol{M}}_{Q-1}^{-T}$

C-JIRAFE

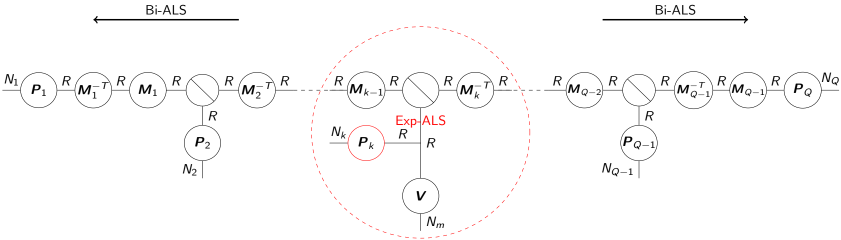


Figure: Coupled JIRAFE

Bi-ALS loops can be done in parallel independently

Simulation

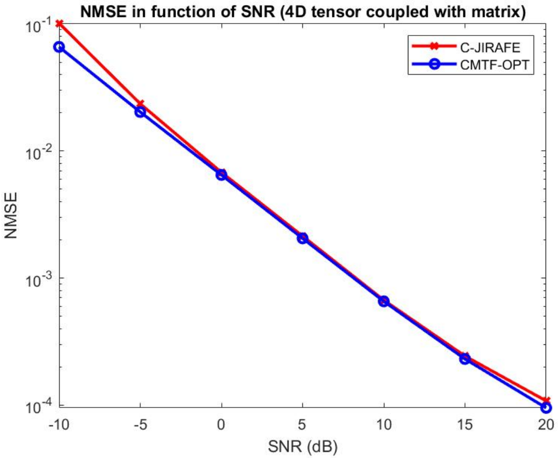


Figure: NMSE (log-scale) in function of signal to noise ratio (dB) for $Q = 4$, $R = 2$, $N = 10$

Simulation

Tensor order	C-JIRAFE	CMTF-OPT	Gain
4	0.0201 (s)	0.1124 (s)	5.5920
5	0.0310 (s)	1.3369 (s)	43.1258
6	0.1057 (s)	20.5572 (s)	194.4863

Table: Execution time in function of order for a fixed signal to noise ratio $SNR = 10dB$

Conclusion

- Breaking the curse of dimensionality for HO-CMTF.
- Extracting the common mode by solving a 3-order CMTF only.
- Using double coupling (physical model coupling and the TT coupling).
- Reducing execution time.
- Possibility of parallel processing.