

Décomposition CP par diagonalisation conjointe : algorithmes et application à la spectroscopie de fluorescence

Xavier Luciani, Laurent Albera ⁽¹⁾

⁽¹⁾ LTSI, Université de Rennes 1 et INSERM U642, France

16 Janvier 2013



SOMMAIRE

- SALT : Un algorithme semi-algébrique pour la CP
- JD TM et JET : 2 algorithmes de diagonalisation conjointe par similarité.

① INTRODUCTION

② SALT, UN ALGORITHME POUR LA DÉCOMPOSITION CP

Principe

Conditions nécessaires

③ ALGORITHMES DE DIAGONALISATION CONJOINTE

Approche JD TM

Approche JET

④ OVER-FACTORIZING ET APPLICATION À LA SPECTROSCOPIE DE FLUORESCENCE

Over-factoring

Spectroscopie de Fluorescence

⑤ CONCLUSION

1 INTRODUCTION

2 SALT, UN ALGORITHME POUR LA DÉCOMPOSITION CP

3 ALGORITHMES DE DIAGONALISATION CONJOINTE

4 OVER-FACTORIZING ET APPLICATION À LA SPECTROSCOPIE DE FLUORESCENCE

5 CONCLUSION

DÉCOMPOSITION CP

Formulation du problème

- Tenseur \mathcal{T} , de rang R , d'ordre Q et de dimensions $I_1 \times I_2 \times \dots \times I_Q$
- $R \times Q$ facteurs $\mathbf{x}^{(q)}$

$$\mathcal{T} = \sum_{r=1}^R \mathbf{x}_r^{(1)} \circ \dots \circ \mathbf{x}_r^{(Q)}, \quad (1)$$

- ou encore Q matrices facteurs $\mathbf{X}^{(q)}$ de taille $I_q \times R$

$$\forall q = 1 \dots Q, \forall i_q = 1 \dots I_q \quad \mathcal{T}_{i_1 \dots i_Q} = \sum_{r=1}^R X_{i_1 r}^{(1)} X_{i_2 r}^{(2)} \dots X_{i_Q r}^{(Q)}. \quad (2)$$

- Retrouver les $\mathbf{X}^{(q)}$ à partir de \mathcal{T}

ALGORITHMES

Algorithmes itératifs

- Alternating Least Squares (ALS)
- Méthodes de descente : gradient, gradient conjugué, LM, BFGS...
- Enhanced Line Search (ELS)
- Avantages : Précision, possibilité d'ajouter des contraintes
- Défauts : Minima locaux, coût calcul, over-factoring, facteurs corrélés

Approche semi-algébrique

Idée : Réécrire la CP comme un problème de diagonalisation conjointe

- Par congruence (de Lathauwer, 2004)
- Par similarité : Closed Form Solution (CFS, Roemer 2008),
Semi-ALgebraic Tensor decomposition (SALT, Luciani 2011)

1 INTRODUCTION

2 SALT, UN ALGORITHME POUR LA DÉCOMPOSITION CP

Principe

Conditions nécessaires

3 ALGORITHMES DE DIAGONALISATION CONJOINTE

4 OVER-FACTORIZING ET APPLICATION À LA SPECTROSCOPIE DE FLUORESCENCE

5 CONCLUSION

INTÉRÊT DE L'APPROCHE SALT

par rapport aux algorithmes itératifs

- Convergence rapide et stable, faible coût calcul
- Peu sensible aux problèmes d'over-factoring et de facteurs corrélés

par rapport à CFS

- Coût calcul moindre
- Condition nécessaire de fonctionnement moins contraignante aux ordres supérieurs à 3

PRINCIPE

$$\mathcal{T}_{i_1 \dots i_Q} = \sum_{r=1}^R X_{i_1 r}^{(1)} X_{i_2 r}^{(2)} \dots X_{i_Q r}^{(Q)}. \quad (3)$$

- Tenseur \mathcal{T} d'ordre Q , de rang R et de dimensions $I_1 \times I_2 \times \dots \times I_Q$
- Q matrices facteurs $\mathbf{X}^{(q)}$ de taille $I_q \times R$

On définit :

- les nombres : $\pi_a^b = I_a I_{a+1} \dots I_b$, ($b > a$)
- les matrices : $\mathbf{Y}_{(b)}^{(a)} = \mathbf{X}^{(b)} \odot \mathbf{X}^{(b-1)} \odot \dots \odot \mathbf{X}^{(a)}$, ($b > a$)

Et la matrice dépliée $\mathbf{T}(P)$ de taille $\pi_1^P \times \pi_{P+1}^Q$ de sorte que :

$$\mathbf{T}(P) = \mathbf{Y}_{(P)}^{(1)} \mathbf{Y}_{(Q)}^{(P+1)\top}. \quad (4)$$

PRINCIPE

Soit \mathbf{USV}^T la SVD de $\mathbf{T}(P)$ tronquée à l'ordre R , alors il existe une matrice inversible \mathbf{A} de taille $R \times R$ telle que

$$\mathbf{Y}_{(P)}^{(1)} = \mathbf{UA} \text{ et } \mathbf{Y}_{(Q)}^{(P+1)\top} = \mathbf{A}^{-1}\mathbf{SV}^T. \quad (5)$$

On définit alors I_Q matrices diagonales $\Phi^{(1)}, \dots, \Phi^{(I_Q)}$ à partir des I_Q lignes de $\mathbf{X}^{(Q)}$.

$$\mathbf{Y}_{(Q)}^{(P+1)\top} = \left[\Phi^{(1)}\mathbf{Y}_{(Q-1)}^{(P+1)\top}, \dots, \Phi^{(I_Q)}\mathbf{Y}_{(Q-1)}^{(P+1)\top} \right] \quad (6)$$

$$\mathbf{SV}^T = \left[\underbrace{\mathbf{A}\Phi^{(1)}\mathbf{Y}_{(Q-1)}^{(P+1)\top}}_{\Gamma^{(1)\top}}, \dots, \underbrace{\mathbf{A}\Phi^{(I_Q)}\mathbf{Y}_{(Q-1)}^{(P+1)\top}}_{\Gamma^{(I_Q)\top}} \right] \quad (7)$$

PRINCIPE

Réécriture comme un pb de diagonalisation conjointe

Soit, $\mathbf{M}^{(i_1, i_2)} = \Gamma^{(i_1)} \# \Gamma^{(i_2)}$ et $\Lambda^{(i_1, i_2)} = \Phi^{(i_1)} \# \Phi^{(i_2)}$ ($\#$: pseudo inverse), on a alors

$$\forall i_1 = 1 \cdots N-1, i_2 = 2 \cdots N, i_2 > i_1 \quad \mathbf{M}^{(i_1, i_2)} = \mathbf{A}^{-\top} \Lambda^{i_1, i_2} \mathbf{A}^{\top}. \quad (8)$$

Problème de diagonalisation conjointe par similarité

- Algorithme (itératif) de diagonalisation conjointe $\Rightarrow \mathbf{A}^{-\top}$
- $\mathbf{A}^{-\top} \Rightarrow \mathbf{Y}_{(Q)}^{(P+1)}$ et $\mathbf{Y}_{(P)}^{(1)}$ car $\mathbf{Y}_{(Q)}^{(P+1)\top} = \mathbf{A}^{-1} \mathbf{S} \mathbf{V}^{\top}$ et $\mathbf{Y}_{(P)}^{(1)} = \mathbf{U} \mathbf{A}$
- $\mathbf{Y}_{(Q)}^{(P+1)}$ et $\mathbf{Y}_{(P)}^{(1)} \Rightarrow \mathbf{X}^{(1)} \cdots \mathbf{X}^{(Q)}$ (HOSVD de rang 1)

CONDITIONS NÉCESSAIRES DES ALGORITHMES

- $(C_{SALT}) : \exists P \in [2; Q-1]_{\mathbb{N}} R \leq \min(\pi_1^P, \pi_{P+1}^{Q-1})$
- $(C_{CFS}) : \exists (q_1, q_2) \in [1; Q]_{\mathbb{N}}^2, q_1 \neq q_2$ tels que $I_{q_1} \geq R$ and $I_{q_2} \geq R$
- $(C_{ALS}) : \forall q \in [1; Q]_{\mathbb{N}}, \prod_{\substack{i=1 \\ i \neq q}}^Q I_i \geq R$

Théorème

On montre que :

$$C_{CFS} \Rightarrow C_{SALT} \Rightarrow C_{ALS}$$

Exemple

- $\mathcal{T} : 3 \times 3 \times 5 \times 3, \text{rang } 5 \Rightarrow \text{ALS et SALT OK, CFS inapplicable.}$

EN PRATIQUE, CHOIX DU DÉPLIEMENT

Tous les modes n'ont pas le même rôle... mais on peut les permuter.
On doit donc bien choisir :

- ① La valeur de P (à l'ordre 3 on a forcément $P = 1$)
 - ② La permutation (ordre) des modes
- On choisira donc une valeur de P et une permutation qui maximise $\min(\pi_1^P, \pi_{P+1}^{Q-1})$.
 - Minimisation du coût calcul : plus petite dimension à la fin.
 - Exemples :
 - ① $\mathcal{T} : 5 \times 5 \times 20 \Rightarrow P = 1$ et $I_1 = 20, I_2 = 20, I_3 = 5$
 - ② $\mathcal{T} : 10 \times 100 \times 5 \times 10 \Rightarrow P = 2$ et $I_1 = 10, I_2 = 10, I_3 = 100, I_4 = 5$

1 INTRODUCTION

2 SALT, UN ALGORITHME POUR LA DÉCOMPOSITION CP

3 ALGORITHMES DE DIAGONALISATION CONJOINTE

Approche JD₁TM

Approche JET

4 OVER-FACTORIZING ET APPLICATION À LA SPECTROSCOPIE DE FLUORESCENCE

5 CONCLUSION

FORMULATION DU PROBLÈME

Trouver \mathbf{A} de taille $N \times N$ à partir des $\mathbf{M}^{(k)}$ telle que :

$$\forall k = 1 \dots K \quad \mathbf{M}^{(k)} = \mathbf{A} \mathbf{D}^{(k)} \mathbf{A}^{-1}. \quad (9)$$

$\mathbf{D}^{(1)} \dots \mathbf{D}^{(K)}$: matrices diagonales inconnues.

- On cherche \mathbf{A} comme un produit de matrices élémentaires
- Méthodes de balayage par paires (type Jacobi) pour minimiser un critère de « diagonalité »
- \mathbf{A} n'est pas forcément orthogonale \Rightarrow peu d'algorithmes dans ce cas.
- sh-rt (Fu, 2006) et JUST (Iferroudjene, 2009) basés sur la décomposition polaire de \mathbf{A}

Nous proposons 3 nouveaux algorithmes :

- JD TM, décomposition polaire, critère alternatif, précis
- JET-O ou JET-U (suivant le critère choisi), décomposition LU, rapide

APPROCHE JD TM (LUCIANI 2010)

Joint Diagonalization algorithm based on Targeting hyperbolic Matrices

Décomposition polaire

$$\mathbf{A} = \prod_{i=1}^{N-1} \prod_{j=i+1}^N \mathbf{G}(\theta_{ij}) \mathbf{H}(\phi_{ij}). \quad (10)$$

$$\mathbf{G}(\theta_{ij}) = \begin{pmatrix} 1 & \dots & i & \dots & j & \dots & N & & \\ 1 & \dots & 0 & \dots & 0 & \dots & 0 & 1 & \\ \vdots & \ddots & \vdots & & \vdots & & \vdots & & \\ \vdots & \ddots & \vdots & & \vdots & & \vdots & & \\ 1 & \dots & \cos(\theta_{ij}) & \dots & \sin(\theta_{ij}) & \dots & 0 & i & \\ \vdots & \ddots & \vdots & & \vdots & & \vdots & & \\ \vdots & \ddots & \vdots & & \vdots & & \vdots & & \\ 1 & \dots & -\sin(\theta_{ij}) & \dots & \cos(\theta_{ij}) & \dots & 0 & j & \\ \vdots & \ddots & \vdots & & \vdots & & \vdots & & \\ \vdots & \ddots & \vdots & & \vdots & & \vdots & & \\ 1 & \dots & 0 & \dots & 0 & \dots & 0 & N & \end{pmatrix}$$

APPROCHE JD TM

$$\mathbf{D}^{(k)} = \left(\prod_{i=1}^{N-1} \prod_{j=i+1}^N \mathbf{H}(\phi_{ij})^{-1} \mathbf{G}(\theta_{ij})^\top \right) \mathbf{M}^{(k)} \left(\prod_{i=1}^{N-1} \prod_{j=i+1}^N \mathbf{G}(\theta_{ij}) \mathbf{H}(\phi_{ij}) \right), \quad (11)$$

Mise à jour

On répète pour chaque $(i, j), i < j$:

$$\forall k = 1 \dots K \quad \mathbf{N}^{(k)} \leftarrow \mathbf{G}(\theta_{ij})^\top \mathbf{M}^{(k)} \mathbf{G}(\theta_{ij}), \quad (12)$$

$$\forall k = 1 \dots K \quad \mathbf{M}^{(k)} \leftarrow \mathbf{H}(\phi_{ij})^{-1} \mathbf{N}^{(k)} \mathbf{H}(\phi_{ij}). \quad (13)$$

En pratique on répète l'opération jusqu'à convergence (sweep)

Originalité de l'approche

Simplification de la fonction de coût :

$$\zeta_G(\theta_{ij}) = \sum_k \sum_{\substack{p,q \\ p \neq q}}^{N,N} \left(N_{pq}^{(k)} \right)^2, \quad \zeta_H^{JD TM}(\phi_{ij}) = \sum_k \left(M_{ij}^{(k)} \right)^2 + \left(M_{ji}^{(k)} \right)^2. \quad (14)$$

APPROCHE JET (LUCIANI 2011)

JET : Joint Eigenvalue decomposition algorithm based on Triangular matrices

Décomposition LU

La matrice \mathbf{A} recherchée peut s'écrire $\mathbf{A} = \mathbf{L}\mathbf{V}$

- \mathbf{L} : Matrice triangulaire inférieure unité (1 sur la diagonal)
- \mathbf{V} : Matrice triangulaire supérieure unité (1 sur la diagonal)

On définit alors les matrices triangulaires supérieures $\mathbf{R}^{(k)}$:

$$\forall k = 1 \dots K \quad \mathbf{R}^{(k)} = \mathbf{V}\mathbf{D}^{(k)}\mathbf{V}^{-1} \quad (15)$$

d'où

$$\forall k = 1 \dots K \quad \mathbf{L}^{-1}\mathbf{M}^{(k)}\mathbf{L} = \mathbf{V}\mathbf{D}^{(k)}\mathbf{V}^{-1} \quad (16)$$

$$\mathbf{L}^{-1}\mathbf{M}^{(k)}\mathbf{L} = \mathbf{R}^{(k)} \quad (17)$$

APPROCHE JET

- \mathbf{L} « triangularise » conjointement les matrices $\mathbf{M}^{(k)}$
- \mathbf{V} se déduit algébriquement des $\mathbf{R}^{(k)}$

Définition

Une matrice triangulaire inférieure élémentaire $\mathbf{L}^{(ij)}(x)$ est une matrice triangulaire inférieure unité dont tous les éléments non diagonaux sont nuls sauf l'élément (i, j) qui est égal à x .

Proposition

Toute matrice triangulaire inférieure unité peut s'écrire comme un produit de matrices inférieures élémentaires

$\Rightarrow \exists \{x_{ij}\}, 1 \leq j \leq N-1, 2 \leq i \leq N, j < i$ tels que, $\forall k = 1 \dots K,$

$$\mathbf{R}^{(k)} = \prod_{i=1}^{N-1} \prod_{j=i+1}^N \left(\mathbf{L}^{(ij)}(x_{ij}) \right)^{-1} \mathbf{M}^{(k)} \prod_{i=1}^{N-1} \prod_{j=i+1}^N \mathbf{L}^{(ij)}(x_{ij}). \quad (18)$$

APPROCHE JET

Mise à jour (à répéter à chaque sweep)

On répète pour chaque $(i, j), j < i$:

$$\forall k = 1 \cdots K \quad \mathbf{M}^{(k)} \leftarrow \left(\mathbf{L}^{(ij)}(x_{ij}) \right)^{-1} \mathbf{M}^{(k)} \mathbf{L}^{(ij)}(x_{ij}) \quad (19)$$

Fonction de coût

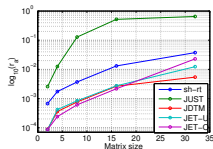
$$\zeta_O(x_{ij}) = \sum_{k=1}^K \sum_{q=1}^{N-1} \sum_{p=q+1}^N \left| M_{p,q}^{(k)} \right|^2 \quad (20)$$

$$\zeta_U(x_{ij}) = \sum_{k=1}^K \left| M_{i,j}^{(k)} \right|^2 \quad (21)$$

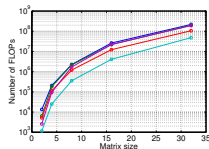
Matrice \mathbf{V}

On montre que les éléments de \mathbf{V} peuvent se déduire arithmétiquement un par un des éléments des matrices $\mathbf{R}^{(k)}$.

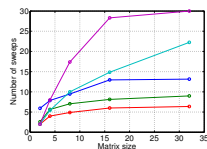
QUELQUES RÉSULTATS



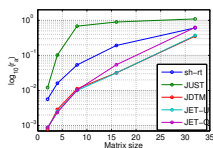
(a) NMSE sur la matrice A



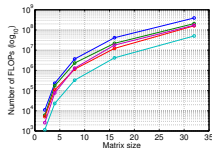
(b) Nombre de FLOP



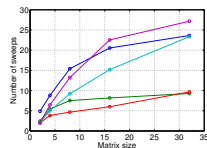
(c) Nombre de sweeps



(d) NMSE sur la matrice A



(e) Nombre de FLOP



(f) Nombre de sweeps

FIGURE: Évolution des critères de comparaison en fonction de la taille des matrices pour un jeu de 64 matrices en haut 60 dB, en bas 40 dB.

- 1 INTRODUCTION
- 2 SALT, UN ALGORITHME POUR LA DÉCOMPOSITION CP
- 3 ALGORITHMES DE DIAGONALISATION CONJOINTE
- 4 OVER-FACTORIZING ET APPLICATION À LA SPECTROSCOPIE DE FLUORESCENCE**
Over-factorizing
Spectroscopie de Fluorescence
- 5 CONCLUSION

OVER-FACTORIZING

- En pratique le rang du tenseur est très grand devant le rang utile R de la décomposition (nombre de sources), à cause du bruit.
- R est souvent inconnu et on a tendance à le surestimer (modélisation du bruit)
- Deux possibilités si on surestime R (over-factoring) :
 - ① Les facteurs supplémentaires n'ont pas d'influence sur les facteurs réels et ont des contributions nulles \Rightarrow Pas de problème dans ce cas.
 - ② Les facteurs réels estimés sont déformés et les facteurs supplémentaires ne sont pas négligeables \Rightarrow problème pour choisir le rang du modèle.

OVER-FACTORIZING

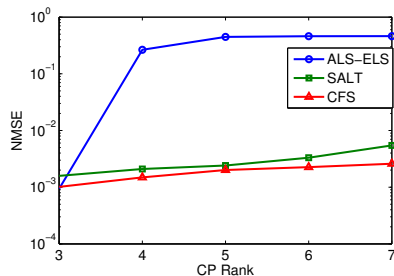
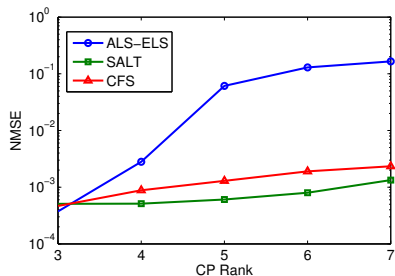
(a) tenseurs de taille $7 \times 7 \times 7$ et de rang utile 3.(b) tenseurs de taille $7 \times 7 \times 7 \times 7$ et de rang utile 3.

FIGURE: Sensibilité à l'over-factoring des algorithmes ALS-ELS, CFS et SALT, NMSE sur les vrais facteurs (moyenne sur tous les facteurs)

SPECTROSCOPIE DE FLUORESCENCE

- Mesure de l'intensité de fluorescence de différents mélanges des mêmes composés fluorescents (fluorophores)
- Mesure fonction des :
 - ① concentrations dans les mélanges (mode 1)
 - ② longueurs d'onde d'excitation (mode 2)
 - ③ longueurs d'onde d'émission de fluorescence (mode 3)
- Faibles concentrations : les mesures peuvent être modélisé par une CP
- Facteurs à estimer, pour chaque fluorophore (source) :
 - ① évolution de la concentration du fluorophore dans les mélanges (mode 1)
 - ② spectre d'excitation du fluorophore (mode 2)
 - ③ spectre d'émission du fluorophore (mode 3)

SPECTROSCOPIE DE FLUORESCENCE

- 2 fluorophores \Rightarrow Tenseur de rang réel 2
- 3 Mélanges, 47 longueurs d'onde d'excitation, 61 longueurs d'onde d'émission \Rightarrow Tenseur de taille $3 \times 46 \times 61$
- Que se passe t'il si l'on utilise une CP de rang 3 comme modèle ?

SPECTROSCOPIE DE FLUORESCENCE

- 2 fluorophores \Rightarrow Tenseur de rang réel 2
- 3 Mélanges, 47 longueurs d'onde d'excitation, 61 longueurs d'onde d'émission \Rightarrow Tenseur de taille $3 \times 46 \times 61$
- Que se passe t'il si l'on utilise une CP de rang 3 comme modèle ?

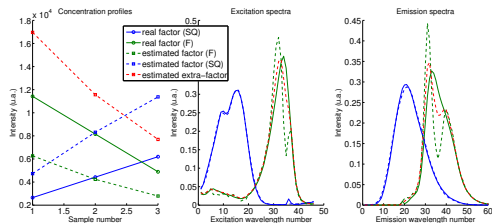


FIGURE: Facteurs de la décomposition estimés par ALS.

SPECTROSCOPIE DE FLUORESCENCE

- 2 fluorophores \Rightarrow Tenseur de rang réel 2
- 3 Mélanges, 47 longueurs d'onde d'excitation, 61 longueurs d'onde d'émission \Rightarrow Tenseur de taille $3 \times 46 \times 61$
- Que se passe t'il si l'on utilise une CP de rang 3 comme modèle ?

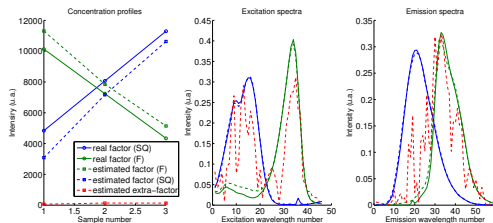


FIGURE: Facteurs de la décomposition estimés par SALT.

- 1 INTRODUCTION
- 2 SALT, UN ALGORITHME POUR LA DÉCOMPOSITION CP
- 3 ALGORITHMES DE DIAGONALISATION CONJOINTE
- 4 OVER-FACTORIZING ET APPLICATION À LA SPECTROSCOPIE DE FLUORESCENCE
- 5 CONCLUSION

CONCLUSION

Décomposition CP

- Algorithme SALT, semi-algébrique, cas réel et complexe
- Peu coûteux et conditions d'application moins contraignante que CFS
- Facteurs très corrélés, ordre élevé, over-factoring
- Possibilité de combiner avec ALS (compromis précision/coût calcul)

Diagonalisation conjointe

- 3 algorithmes : JD_{TM}, JET-U, JET-O, cas réel et complexe (sauf JET-O)
- JD_{TM} : précis, efficace dans les cas les plus difficiles, très peu de sweeps
- JET-U : original, très peu coûteux, très efficace dans les cas simples
- JET-O : souvent le plus coûteux mais aussi le plus précis