On minimal ranks and the approximate block-term tensor decomposition

José Henrique de Morais Goulart[†,‡] and Pierre Comon[†]

[†] Laboratoire Gipsa-lab, UMR 5216, CNRS, UGA, France
[‡] I3S Laboratory, UMR 7271, CNRS, UCA, France
henrique.goulart@gipsa-lab.fr

ERC Project "DECODA" (Grant AdG-2013-320594)

Structured Matrix Days - May 14–15, 2018

Summary

Goal: study existence of soln's to the approximation problem

$$\min_{\mathbf{H}_i,\,\mathbf{w}_i} \left\| \mathcal{Y} - \sum_{i=1}^{k} \mathbf{H}_i \otimes \mathbf{w}_i \right\|, \quad \text{subj. to } \operatorname{rank} \mathbf{H}_i \leq r_i$$

- Block-term decomposition, a.k.a. BTD (De Lathauwer, 2008)
  - Useful for source separation (uniqueness results)
  - One block per source

Motivation: In practice, $\mathcal{Y} = (\text{BTD}) + \mathcal{E}$

- Computation: typically, optimization with fixed k and r_i
- But does a global minimum always exist?

Summary

Goal: study existence of soln's to the approximation problem

$$\min_{\mathbf{H}_i,\,\mathbf{w}_i} \left\| \mathcal{Y} - \sum_{i=1}^{k} \mathbf{H}_i \otimes \mathbf{w}_i \right\|, \quad \text{subj. to } \operatorname{rank} \mathbf{H}_i \leq r_i$$

- Block-term decomposition, a.k.a. BTD (De Lathauwer, 2008)
  - Useful for source separation (uniqueness results)
  - One block per source

Motivation: In practice, $\mathcal{Y} = (\mathrm{BTD}) + \mathcal{E}$

- Computation: typically, optimization with fixed k and r_i
- But does a global minimum always exist?
  - (Spoiler alert) No

Tensors

- Formally, elements of tensor products of abstract vector spaces

$$\mathcal{U} \otimes \mathcal{V} \otimes \mathcal{W}$$

- Can be viewed as multilinear maps

Tensors

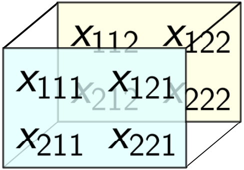
- Formally, elements of tensor products of abstract vector spaces

$$\mathcal{U} \otimes \mathcal{V} \otimes \mathcal{W}$$

- Can be viewed as multilinear maps
- Finite dimensions, fixed basis: tensors $\leftrightarrow$ multi-way arrays

$$\mathbb{R}^m \otimes \mathbb{R}^n \otimes \mathbb{R}^d \quad \simeq \quad \mathbb{R}^{m \times n \times d}$$

Example: $\mathfrak{X} \in \mathbb{R}^{2\times 2\times 2}$ (third order: three-way array)



$$\left(\begin{array}{cc|cc} x_{111} & x_{121} & x_{112} & x_{122} \\ x_{211} & x_{221} & x_{212} & x_{222} \end{array}\right)$$

Tensor product (Segre outer product)

Definition

- Binary operator

$$\otimes : \mathbb{R}^{m_1 \times \cdots \times m_p} \times \mathbb{R}^{n_1 \times \cdots \times n_q} \to \mathbb{R}^{m_1 \times \cdots \times m_p \times n_1 \times \cdots \times n_q}$$

$$[\mathcal{U} \otimes \mathcal{V}]_{i_1,\ldots,i_p,j_1,\ldots,j_q} = u_{i_1,\ldots,i_p} \, v_{j_1,\ldots,j_q}$$

- Associative, non-commutative and distributive w.r.t. $+$

$$(\mathbf{u} \otimes \mathbf{v}) \otimes \mathbf{w} = \mathbf{u} \otimes (\mathbf{v} \otimes \mathbf{w}) = \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w} \neq \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{w}$$

$$(\mathbf{u}_1 + \mathbf{u}_2) \otimes \mathbf{v} = \mathbf{u}_1 \otimes \mathbf{v} + \mathbf{u}_2 \otimes \mathbf{v} = \mathbf{u}_1 \mathbf{v}^\mathsf{T} + \mathbf{u}_2 \mathbf{v}^\mathsf{T}$$

Tensor product (Segre outer product)

Definition

- Binary operator

$$\otimes : \mathbb{R}^{m_1\times\cdots\times m_p} \times \mathbb{R}^{n_1\times\cdots\times n_q} \to \mathbb{R}^{m_1\times\cdots\times m_p\times n_1\times\cdots\times n_q}$$

$$[\mathcal{U}\otimes\mathcal{V}]_{i_1,\ldots,i_p,j_1,\ldots,j_q} = u_{i_1,\ldots,i_p}\, v_{j_1,\ldots,j_q}$$

- Associative, non-commutative and distributive w.r.t. $+$

$$(\mathbf{u}\otimes\mathbf{v})\otimes\mathbf{w} = \mathbf{u}\otimes(\mathbf{v}\otimes\mathbf{w}) = \mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w} \neq \mathbf{v}\otimes\mathbf{u}\otimes\mathbf{w}$$

$$(\mathbf{u}_1+\mathbf{u}_2)\otimes\mathbf{v} = \mathbf{u}_1\otimes\mathbf{v}+\mathbf{u}_2\otimes\mathbf{v} = \mathbf{u}_1\mathbf{v}^{\mathsf{T}}+\mathbf{u}_2\mathbf{v}^{\mathsf{T}}$$

Elementary tensors

- Tensor products of vectors: $\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}$
- Every tensor can be written as a sum of elementary ones

$$\mathcal{X} = \textstyle\sum_i\sum_j\sum_l x_{ijl}\,\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_l \qquad \text{(trivial)}$$

Canonical polyadic (CP, tensor rank) decomposition

Definition

- Minimal decomposition as a sum of elementary tensors

$$\mathcal{X} = \sum_{i=1}^{r} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i$$

- r is minimal and defines rank $\mathcal{X}$

Canonical polyadic (CP, tensor rank) decomposition

Definition

- Minimal decomposition as a sum of elementary tensors

$$\mathcal{X} = \sum_{i=1}^{r} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i$$

- r is minimal and defines $\operatorname{rank} \mathcal{X}$

Example: diagonal tensor ($m = n = d$)

$\mathcal{X} = \sum_{i=1}^{n} x_{iii}\, \mathbf{e}_i \otimes \mathbf{e}_i \otimes \mathbf{e}_i \implies \operatorname{rank} \mathcal{X} = \#$ of nonzero x_{iii}

Canonical polyadic (CP, tensor rank) decomposition

Definition

- Minimal decomposition as a sum of elementary tensors

$$\mathcal{X} = \sum_{i=1}^{r} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i$$

- r is minimal and defines $\operatorname{rank} \mathcal{X}$

Example: diagonal tensor ($m = n = d$)

$\mathcal{X} = \sum_{i=1}^{n} x_{iii}\, \mathbf{e}_i \otimes \mathbf{e}_i \otimes \mathbf{e}_i \quad \Longrightarrow \quad \operatorname{rank} \mathcal{X} = \#$ of nonzero x_{iii}

Example: rank-one tensor

$$\mathcal{X} = \mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}_1 + \mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}_2 = \mathbf{a} \otimes \mathbf{b} \otimes (\mathbf{c}_1 + \mathbf{c}_2) \quad \Longrightarrow \quad \operatorname{rank} \mathcal{X} = 1$$

Multilinear transformation

Definition

Given

$$\mathbf{P} \in \mathbb{R}^{m' \times m}, \quad \mathbf{U} \in \mathbb{R}^{n' \times n}, \quad \mathbf{T} \in \mathbb{R}^{d' \times d}$$

$$\mathcal{X} = \sum_i \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \qquad \in \mathbb{R}^{m \times n \times d}$$

we define:

$$(\mathbf{P}, \mathbf{U}, \mathbf{T}) \cdot \mathcal{X} = \sum_i (\mathbf{P}\mathbf{a}_i) \otimes (\mathbf{U}\mathbf{b}_i) \otimes (\mathbf{T}\mathbf{c}_i) \in \mathbb{R}^{m' \times n' \times d'}$$

Multilinear transformation

Definition

Given

$$\mathbf{P} \in \mathbb{R}^{m' \times m}, \quad \mathbf{U} \in \mathbb{R}^{n' \times n}, \quad \mathbf{T} \in \mathbb{R}^{d' \times d}$$

$$\mathcal{X} = \sum_i \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \qquad \in \mathbb{R}^{m \times n \times d}$$

we define:

$$(\mathbf{P}, \mathbf{U}, \mathbf{T}) \cdot \mathcal{X} = \sum_i (\mathbf{P}\mathbf{a}_i) \otimes (\mathbf{U}\mathbf{b}_i) \otimes (\mathbf{T}\mathbf{c}_i) \in \mathbb{R}^{m' \times n' \times d'}$$

- If $(\mathbf{P}, \mathbf{U}, \mathbf{T}) \in \mathrm{GL}_m \times \mathrm{GL}_n \times \mathrm{GL}_d$: basis change (preserves rank)

Block-term decomposition

- Generally, a sum of tensors (blocks) having low multilinear ranks

We focus on the particular form:

$$\mathcal{X} = \sum_{i=1}^{k} \mathbf{H}_i \otimes \mathbf{w}_i \qquad \text{where} \qquad \operatorname{rank} \mathbf{H}_i \leq r_i$$

Block-term decomposition

- Generally, a sum of tensors (blocks) having low multilinear ranks

We focus on the particular form:

$$\mathcal{X} = \sum_{i=1}^{k} \mathbf{H}_i \otimes \mathbf{w}_i \qquad \text{where} \qquad \operatorname{rank} \mathbf{H}_i \leq r_i$$

- For $\mathbf{w} \in \mathbb{R}^d$:

$$\mathbf{H} \otimes \mathbf{w} = \mathbf{H} \otimes \left(\sum_i w_i \, \mathbf{e}_i \right) = \sum_i (w_i \, \mathbf{H}) \otimes \mathbf{e}_i$$

$$= \underbrace{\left(\begin{array}{ccc|c|ccc} w_1 h_{11} & \dots & w_1 h_{1n} & \dots & w_d h_{11} & \dots & w_d h_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ w_1 h_{m1} & \dots & w_1 h_{mn} & \dots & w_d h_{m1} & \dots & w_d h_{mn} \end{array} \right)}_{d \text{ "frontal slices"}}$$

Approximate canonical polyadic decomposition

$$\arg\min_{\mathbf{a}_i,\mathbf{b}_i,\mathbf{c}_i} \left\| \mathcal{Y} - \sum_{i=1}^{k} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \right\|$$ may be empty for a given rank $k > 1$

Approximate canonical polyadic decomposition

$\arg\min\limits_{\mathbf{a}_i,\mathbf{b}_i,\mathbf{c}_i} \left\| \mathcal{Y} - \sum_{i=1}^{k} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \right\|$ may be empty for a given rank $k > 1$

Famous example: sequence of rank-2 tensors $\to$ rank-3 tensor

- If $\{\mathbf{a}_1, \mathbf{a}_2\}, \{\mathbf{b}_1, \mathbf{b}_2\}, \{\mathbf{c}_1, \mathbf{c}_2\}$ are l.i. :

$$\operatorname{rank}\left(\mathbf{a}_2 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1 + \mathbf{a}_1 \otimes \mathbf{b}_2 \otimes \mathbf{c}_1 + \mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_2\right) = 3$$

Approximate canonical polyadic decomposition

$$\underset{\mathbf{a}_i,\mathbf{b}_i,\mathbf{c}_i}{\arg\min} \left\| \mathcal{Y} - \sum_{i=1}^{k} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \right\|$$ may be empty for a given rank $k > 1$

Famous example: sequence of rank-2 tensors → rank-3 tensor

- If $\{\mathbf{a}_1, \mathbf{a}_2\}, \{\mathbf{b}_1, \mathbf{b}_2\}, \{\mathbf{c}_1, \mathbf{c}_2\}$ are l.i. :

$$\operatorname{rank}\left(\mathbf{a}_2 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1 + \mathbf{a}_1 \otimes \mathbf{b}_2 \otimes \mathbf{c}_1 + \mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_2\right) = 3$$

- But:

$$\mathcal{X}_p \triangleq p\left(\mathbf{a}_1 + \frac{1}{p}\mathbf{a}_2\right) \otimes \left(\mathbf{b}_1 + \frac{1}{p}\mathbf{b}_2\right) \otimes \left(\mathbf{c}_1 + \frac{1}{p}\mathbf{c}_2\right) - p\,\mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1$$

$$= \mathbf{a}_2 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1 + \mathbf{a}_1 \otimes \mathbf{b}_2 \otimes \mathbf{c}_1 + \mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_2 + \frac{1}{p}\mathcal{R}_1 + \frac{1}{p^2}\mathcal{R}_2$$

Approximate CPD: main ingredients

- Euclidean tensor norm:
  - $\langle \mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1, \mathbf{a}_2 \otimes \mathbf{b}_2 \otimes \mathbf{c}_2 \rangle \triangleq \langle \mathbf{a}_1, \mathbf{a}_2 \rangle \langle \mathbf{b}_1, \mathbf{b}_2 \rangle \langle \mathbf{c}_1, \mathbf{c}_2 \rangle$
  - $\|\mathcal{X}\|^2 \triangleq \langle \mathcal{X}, \mathcal{X} \rangle$

Approximate CPD: main ingredients

- Euclidean tensor norm:
  - $\langle \mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1, \mathbf{a}_2 \otimes \mathbf{b}_2 \otimes \mathbf{c}_2 \rangle \triangleq \langle \mathbf{a}_1, \mathbf{a}_2 \rangle \langle \mathbf{b}_1, \mathbf{b}_2 \rangle \langle \mathbf{c}_1, \mathbf{c}_2 \rangle$
  - $\|\mathcal{X}\|^2 \triangleq \langle \mathcal{X}, \mathcal{X} \rangle$
- Defining $\mathcal{S}_k = \{\mathcal{X} \mid \text{rank}(\mathcal{X}) \leq k\}$, we can write:

$$\arg\min_{\mathcal{X} \in \mathcal{S}_k} \|\mathcal{Y} - \mathcal{X}\| \qquad \text{(projection onto } \mathcal{S}_k\text{)}$$

Approximate CPD: main ingredients

- Euclidean tensor norm:
  - $\langle \mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1, \mathbf{a}_2 \otimes \mathbf{b}_2 \otimes \mathbf{c}_2 \rangle \triangleq \langle \mathbf{a}_1, \mathbf{a}_2 \rangle \langle \mathbf{b}_1, \mathbf{b}_2 \rangle \langle \mathbf{c}_1, \mathbf{c}_2 \rangle$
  - $\|\mathcal{X}\|^2 \triangleq \langle \mathcal{X}, \mathcal{X} \rangle$

- Defining $\mathcal{S}_k = \{\mathcal{X} \mid \text{rank}(\mathcal{X}) \leq k\}$, we can write:

$$\underset{\mathcal{X} \in \mathcal{S}_k}{\arg\min} \|\mathcal{Y} - \mathcal{X}\| \qquad \text{(projection onto } \mathcal{S}_k\text{)}$$

- Simple natural hierarchy: $\mathcal{S}_0 \subset \mathcal{S}_1 \subset \mathcal{S}_2 \subset \cdots \subset \mathcal{S}_{k_{\max}}$

Approximate CPD: main ingredients

- Euclidean tensor norm:
  - $\langle \mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1, \mathbf{a}_2 \otimes \mathbf{b}_2 \otimes \mathbf{c}_2 \rangle \triangleq \langle \mathbf{a}_1, \mathbf{a}_2 \rangle \langle \mathbf{b}_1, \mathbf{b}_2 \rangle \langle \mathbf{c}_1, \mathbf{c}_2 \rangle$
  - $\|\mathcal{X}\|^2 \triangleq \langle \mathcal{X}, \mathcal{X} \rangle$

- Defining $\mathcal{S}_k = \{\mathcal{X} \mid \text{rank}(\mathcal{X}) \leq k\}$, we can write:

$$\arg\min_{\mathcal{X} \in \mathcal{S}_k} \|\mathcal{Y} - \mathcal{X}\| \qquad (\text{projection onto } \mathcal{S}_k)$$

- Simple natural hierarchy: $\mathcal{S}_0 \subset \mathcal{S}_1 \subset \mathcal{S}_2 \subset \cdots \subset \mathcal{S}_{k_{\max}}$

- $\mathcal{S}_k$ is generally not closed

Approximate CPD: main ingredients

- Euclidean tensor norm:
  - $\langle \mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1, \mathbf{a}_2 \otimes \mathbf{b}_2 \otimes \mathbf{c}_2 \rangle \triangleq \langle \mathbf{a}_1, \mathbf{a}_2 \rangle \langle \mathbf{b}_1, \mathbf{b}_2 \rangle \langle \mathbf{c}_1, \mathbf{c}_2 \rangle$
  - $\|\mathcal{X}\|^2 \triangleq \langle \mathcal{X}, \mathcal{X} \rangle$
- Defining $\mathcal{S}_k = \{\mathcal{X} \mid \operatorname{rank}(\mathcal{X}) \leq k\}$, we can write:

$$\arg\min_{\mathcal{X} \in \mathcal{S}_k} \|\mathcal{Y} - \mathcal{X}\| \qquad \text{(projection onto } \mathcal{S}_k\text{)}$$

- Simple natural hierarchy: $\mathcal{S}_0 \subset \mathcal{S}_1 \subset \mathcal{S}_2 \subset \cdots \subset \mathcal{S}_{k_{\max}}$
- $\mathcal{S}_k$ is generally not closed
- Nonexistence example: $\mathcal{Y} \notin \mathcal{S}_k$ and $\arg\min_{\mathcal{X} \in \mathcal{S}_k} \|\mathcal{Y} - \mathcal{X}\| = \emptyset$

Approximate CPD: main ingredients

- Euclidean tensor norm:
  - $\langle \mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1, \mathbf{a}_2 \otimes \mathbf{b}_2 \otimes \mathbf{c}_2 \rangle \triangleq \langle \mathbf{a}_1, \mathbf{a}_2 \rangle \langle \mathbf{b}_1, \mathbf{b}_2 \rangle \langle \mathbf{c}_1, \mathbf{c}_2 \rangle$
  - $\|\mathcal{X}\|^2 \triangleq \langle \mathcal{X}, \mathcal{X} \rangle$
- Defining $\mathcal{S}_k = \{\mathcal{X} \mid \text{rank}(\mathcal{X}) \leq k\}$, we can write:

$$\arg\min_{\mathcal{X} \in \mathcal{S}_k} \|\mathcal{Y} - \mathcal{X}\| \qquad (\text{projection onto } \mathcal{S}_k)$$

- Simple natural hierarchy: $\mathcal{S}_0 \subset \mathcal{S}_1 \subset \mathcal{S}_2 \subset \cdots \subset \mathcal{S}_{k_{\text{max}}}$
- $\mathcal{S}_k$ is generally not closed
- Nonexistence example: $\mathcal{Y} \notin \mathcal{S}_k$ and $\arg\min_{\mathcal{X} \in \mathcal{S}_k} \|\mathcal{Y} - \mathcal{X}\| = \emptyset$
- We need similar concepts in the BTD case

Sets of BTDs

- For $r_1 \geq r_2 \geq \cdots \geq r_d \geq 0$, we define:

$$\mathcal{B}_{r_1,\ldots,r_d} \triangleq \left\{ \sum_{i=1}^{d} \mathbf{H}_i \otimes \mathbf{w}_i \;\middle|\; \text{rank}(\mathbf{H}_i) \leq r_i \text{ and } \mathbf{w}_1, \ldots, \mathbf{w}_d \text{ are l.i.} \right\}$$

Sets of BTDs

- For $r_1 \geq r_2 \geq \cdots \geq r_d \geq 0$, we define:

$$\mathcal{B}_{r_1,\ldots,r_d} \triangleq \left\{ \sum_{i=1}^{d} \mathbf{H}_i \otimes \mathbf{w}_i \;\middle|\; \text{rank}(\mathbf{H}_i) \leq r_i \text{ and } \mathbf{w}_1, \ldots, \mathbf{w}_d \text{ are l.i.} \right\}$$

Example:

$$\left(\sum_{i=1}^{3} \mathbf{u}_i \otimes \mathbf{v}_i\right) \otimes \mathbf{d}_1 + \mathbf{x} \otimes \mathbf{y} \otimes \mathbf{d}_2 \quad \in \mathcal{B}_{3,1} \subset \mathcal{B}_{3,2} \subset \mathcal{B}_{4,4}$$

Sets of BTDs

- For $r_1 \geq r_2 \geq \cdots \geq r_d \geq 0$, we define:

$$\mathcal{B}_{r_1,\ldots,r_d} \triangleq \left\{ \sum_{i=1}^{d} \mathbf{H}_i \otimes \mathbf{w}_i \;\middle|\; \mathrm{rank}(\mathbf{H}_i) \leq r_i \text{ and } \mathbf{w}_1, \ldots, \mathbf{w}_d \text{ are l.i.} \right\}$$

Example:

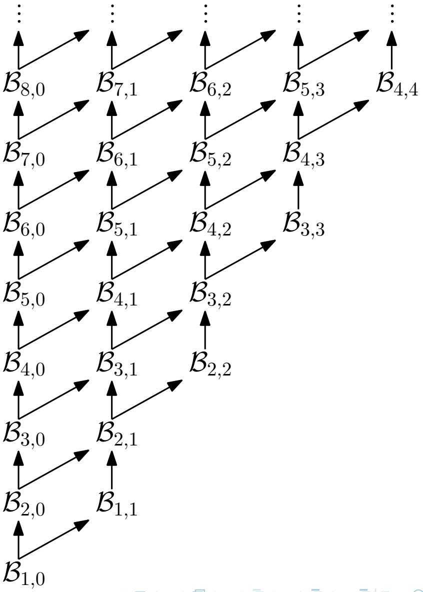
$$\left(\sum_{i=1}^{3} \mathbf{u}_i \otimes \mathbf{v}_i\right) \otimes \mathbf{d}_1 + \mathbf{x} \otimes \mathbf{y} \otimes \mathbf{d}_2 \quad \in \mathcal{B}_{3,1} \subset \mathcal{B}_{3,2} \subset \mathcal{B}_{4,4}$$

Richer than the hierarchy induced by tensor rank

Hierarchy of BTD sets

Notation:

$\mathcal{B} \rightarrow \mathcal{C} \equiv \mathcal{B} \subset \mathcal{C}$



Minimal ranks

- Need parallel to: "$\mathcal{X}$ has rank r" $\quad \Leftrightarrow \quad \mathcal{X} \in \mathcal{S}_r \setminus \mathcal{S}_{r-1}$

Definition: Minimal ranks of $m \times n \times d$ tensor $\mathcal{X}$

$$\rho(\mathcal{X}) = (r_1, \ldots, r_d) \quad \text{when} \quad \mathcal{X} \in \mathcal{B}_{s_1,\ldots,s_d} \Leftrightarrow \forall\, i, s_i \geq r_i$$

Minimal ranks

- Need parallel to: "$\mathcal{X}$ has rank r" $\Leftrightarrow$ $\mathcal{X} \in \mathcal{S}_r \setminus \mathcal{S}_{r-1}$

Definition: Minimal ranks of $m \times n \times d$ tensor $\mathcal{X}$

$$\rho(\mathcal{X}) = (r_1, \ldots, r_d) \quad \text{when} \quad \mathcal{X} \in \mathcal{B}_{s_1,\ldots,s_d} \Leftrightarrow \forall\, i, s_i \geq r_i$$

$\rho(\mathcal{X})$ is well-defined (with $r_i \geq r_{i+1}$)

- By definition, can't be true for $(r_1', \ldots, r_d') \neq (r_1, \ldots, r_d)$
- Must be true for some $(r_1, \ldots, r_d)$

Minimal ranks

- Need parallel to: "$\mathcal{X}$ has rank r" $\quad\Leftrightarrow\quad \mathcal{X} \in \mathcal{S}_r \setminus \mathcal{S}_{r-1}$

Definition: Minimal ranks of $m \times n \times d$ tensor $\mathcal{X}$

$$\rho(\mathcal{X}) = (r_1, \ldots, r_d) \quad \text{when} \quad \mathcal{X} \in \mathcal{B}_{s_1,\ldots,s_d} \Leftrightarrow \forall\, i, s_i \geq r_i$$

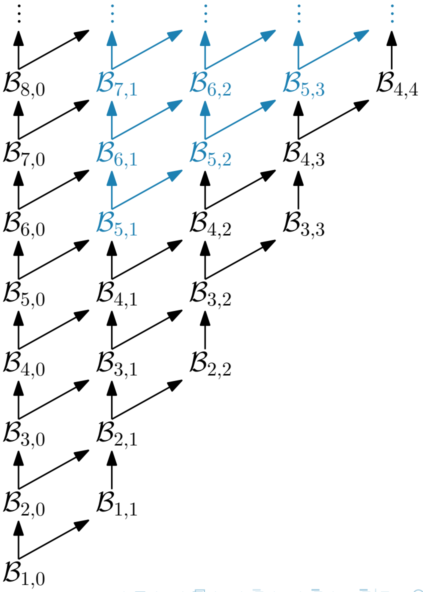
$\rho(\mathcal{X})$ is well-defined (with $r_i \geq r_{i+1}$)

- By definition, can't be true for $(r'_1, \ldots, r'_d) \neq (r_1, \ldots, r_d)$
- Must be true for some $(r_1, \ldots, r_d)$

$\rho(\mathcal{X})$ is $\mathrm{GL}_m \times \mathrm{GL}_n \times \mathrm{GL}_d$-invariant

$$\mathcal{X} \in \mathcal{B}_{r_1,\ldots,r_d} \quad \Leftrightarrow \quad (\mathbf{P}, \mathbf{U}, \mathbf{T}) \cdot \mathcal{X} \in \mathcal{B}_{r_1,\ldots,r_d}$$

Example of minimal ranks



$\rho(\mathfrak{X}) = (5, 1)$

Notation:

$\mathcal{B} \rightarrow \mathcal{C} \equiv \mathcal{B} \subset \mathcal{C}$

Third-order tensors as matrix polynomials

- Consider the isomorphism $\Phi : \mathbb{R}^{m \times n \times d} \to \mathcal{P}_{d-1}[\mathbb{R}^{m \times n}]$

$$\mathcal{X} = \sum_{i=1}^{d} \mathbf{X}_i \otimes \mathbf{e}_i \quad \mapsto \quad X(\lambda) = \sum_{i=1}^{d} \lambda^{i-1} \mathbf{X}_i$$

- Tensor slices are mapped to matrix coefficients
- In particular: $\mathbf{A} \otimes \mathbf{e}_1 + \mathbf{B} \otimes \mathbf{e}_2 \quad \mapsto \quad \mathbf{A} + \lambda \mathbf{B}$

Third-order tensors as matrix polynomials

- Consider the isomorphism $\Phi : \mathbb{R}^{m \times n \times d} \to \mathcal{P}_{d-1}[\mathbb{R}^{m \times n}]$

$$\mathcal{X} = \sum_{i=1}^{d} \mathbf{X}_i \otimes \mathbf{e}_i \quad \mapsto \quad X(\lambda) = \sum_{i=1}^{d} \lambda^{i-1} \mathbf{X}_i$$

- Tensor slices are mapped to matrix coefficients
- In particular: $\mathbf{A} \otimes \mathbf{e}_1 + \mathbf{B} \otimes \mathbf{e}_2 \quad \mapsto \quad \mathbf{A} + \lambda \mathbf{B}$

Minimal ranks of matrix polynomial

$$\rho(X) \triangleq \rho\left(\Phi^{-1}(X)\right)$$

Minimal ranks of matrix pencil ($\simeq$ 2-block BTD)

- Basic idea: given $\mathbf{A} + \lambda \mathbf{B}$, find $\mathbf{T} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} \in \mathrm{GL}_2$ that yields

$$\mathbf{A}' + \lambda \mathbf{B}' = (t_{11}\,\mathbf{A} + t_{12}\,\mathbf{B}) + \lambda\,(t_{21}\,\mathbf{A} + t_{22}\mathbf{B})$$

such that the ranks of $\mathbf{A}'$ and $\mathbf{B}'$ are minimal:

$$(t_{21}, t_{22}) = \arg\min_{(y,z)} \operatorname{rank}(y\mathbf{A} + z\mathbf{B})$$
$$(t_{11}, t_{12}) = \arg\min_{(y,z) \neq \alpha(t_{21}, t_{22})} \operatorname{rank}(y\mathbf{A} + z\mathbf{B})$$

Kronecker canonical form of matrix pencil

- $P(\lambda) = \mathbf{A} + \lambda\mathbf{B} \sim P'(\lambda) = \mathbf{A}' + \lambda\mathbf{B}' \implies \rho(P) = \rho(P')$
- Easier to determine $\rho(P')$

$$P'(\lambda) = \underbrace{\color{red}{S(\lambda)}}_{\text{singular part}} \oplus \underbrace{\color{blue}{R(\lambda)}}_{\text{regular part}}$$

$$R(\lambda) = \underbrace{\color{blue}{N(\lambda)}}_{\substack{\text{infinite elementary}\\ \text{divisors}}} \oplus \underbrace{\color{blue}{J(\lambda)}}_{\substack{\text{finite elementary}\\ \text{divisors}}} \oplus \underbrace{\color{red}{Q(\lambda)}}_{\substack{\text{finite elementary divisors } \neq\\ \text{powers of linear forms}}}$$

- Legend: Can reduce rank / Cannot reduce rank

Minimal ranks of pencils

Theorem (G. & C., 2017)

$$\rho(P) = \rho(\mathbf{A}, \mathbf{B}) = (r_0 + r', \; r_0 + s')$$

where:

- r_0 = sum of minimal indices = rank of singular part
- $s' = \text{size}(R) -$ largest number of elementary divisors of $N \oplus J$ having a common factor
- $r' = \text{size}(R) -$ second largest number of elementary divisors of $N \oplus J$ having a common factor

… and some examples

Example 1

$$\mathbf{H} = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}, \quad \mathbf{H}' = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix} \quad \text{and} \quad \mathbf{H}'' = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

$$\rho(\mathbf{H}, \mathbf{E}) = (3, 2), \quad \rho(\mathbf{H}', \mathbf{E}) = (3, 1), \quad \rho(\mathbf{H}'', \mathbf{E}) = (2, 2)$$

… and some examples

Example 1

$$\mathbf{H} = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}, \quad \mathbf{H}' = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix} \quad \text{and} \quad \mathbf{H}'' = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

$$\rho(\mathbf{H}, \mathbf{E}) = (3, 2), \quad \rho(\mathbf{H}', \mathbf{E}) = (3, 1), \quad \rho(\mathbf{H}'', \mathbf{E}) = (2, 2)$$

Example 2

$$\lambda\mathbf{E} - \mathbf{Q} = \lambda\mathbf{E} - \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \quad \text{has one elem. divisor: } (\lambda - a)^2 + b^2$$

$$\rho(-\mathbf{Q}, \mathbf{E}) = (2, 2) \text{ in } \mathbb{R} \quad \left(\text{but} \quad \rho(-\mathbf{Q}, \mathbf{E}) = (1, 1) \text{ in } \mathbb{C}! \right)$$

... and some examples

Example 1

$$\mathbf{H} = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}, \quad \mathbf{H}' = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix} \quad \text{and} \quad \mathbf{H}'' = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

$$\rho(\mathbf{H}, \mathbf{E}) = (3, 2), \quad \rho(\mathbf{H}', \mathbf{E}) = (3, 1), \quad \rho(\mathbf{H}'', \mathbf{E}) = (2, 2)$$

Example 2

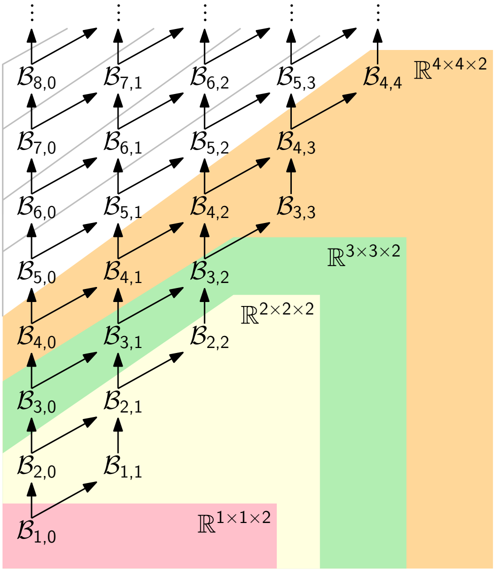
$$\lambda\mathbf{E} - \mathbf{Q} = \lambda\mathbf{E} - \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \quad \text{has one elem. divisor: } (\lambda - a)^2 + b^2$$

$$\rho(-\mathbf{Q}, \mathbf{E}) = (2, 2) \text{ in } \mathbb{R} \quad \left(\text{but} \quad \rho(-\mathbf{Q}, \mathbf{E}) = (1, 1) \text{ in } \mathbb{C}! \right)$$

Example 3

$$\rho\left(\begin{bmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda + a \end{bmatrix}\right) = \rho\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = (2, 2)$$

Spaces of real $n \times n$ pencils



Corollary

$X(\lambda)$ has full ranks

$\Leftrightarrow$

$X(\lambda) = Q(\lambda)$

$\Leftrightarrow$

$X(\lambda)$ is regular with all eigenvalues in $\mathbb{C}$

Generalization of celebrated CPD example

- Let $\mathbf{A}, \mathbf{B} : m \times s$ and $\mathbf{C}, \mathbf{D} : n \times s$ and consider

$$X(\lambda) = \left(\mathbf{AC}^\mathsf{T} + \mathbf{BD}^\mathsf{T}\right) + \lambda\, \mathbf{BC}^\mathsf{T}$$

Generalization of celebrated CPD example

- Let $\mathbf{A}, \mathbf{B} : m \times s$ and $\mathbf{C}, \mathbf{D} : n \times s$ and consider

$$X(\lambda) = \left(\mathbf{A}\mathbf{C}^\mathsf{T} + \mathbf{B}\mathbf{D}^\mathsf{T}\right) + \lambda\, \mathbf{B}\mathbf{C}^\mathsf{T}$$

- Define the sequence of pencils $X_p(\lambda) \in \mathcal{B}_{s,s}$ converging to $X(\lambda)$:

$$X_p(\lambda) \triangleq p\left[\left(\mathbf{B} + \frac{1}{p}\mathbf{A}\right)\left(\mathbf{C} + \frac{1}{p}\mathbf{D}\right)^\mathsf{T}\right]\left(1 + \frac{1}{p}\lambda\right) - p\left(\mathbf{B}\mathbf{C}^\mathsf{T}\right)$$
$$= X(\lambda) + \frac{1}{p}R_1(\lambda) + \frac{1}{p^2}R_2(\lambda)$$

Generalization of celebrated CPD example

- Let $\mathbf{A}, \mathbf{B} : m \times s$ and $\mathbf{C}, \mathbf{D} : n \times s$ and consider

$$X(\lambda) = \left(\mathbf{A}\mathbf{C}^\mathsf{T} + \mathbf{B}\mathbf{D}^\mathsf{T}\right) + \lambda\, \mathbf{B}\mathbf{C}^\mathsf{T}$$

- Define the sequence of pencils $X_p(\lambda) \in \mathcal{B}_{s,s}$ converging to $X(\lambda)$:

$$X_p(\lambda) \triangleq p \left[\left(\mathbf{B} + \frac{1}{p}\mathbf{A}\right) \left(\mathbf{C} + \frac{1}{p}\mathbf{D}\right)^\mathsf{T} \right] \left(1 + \frac{1}{p}\lambda\right) - p\left(\mathbf{B}\mathbf{C}^\mathsf{T}\right)$$

$$= X(\lambda) + \frac{1}{p} R_1(\lambda) + \frac{1}{p^2} R_2(\lambda)$$

- If $\min\left\{\operatorname{rank}\begin{bmatrix}\mathbf{A} & \mathbf{B}\end{bmatrix}, \operatorname{rank}\begin{bmatrix}\mathbf{C} & \mathbf{D}\end{bmatrix}\right\} > 3s/2$, then $X \notin \mathcal{B}_{s,s}$.

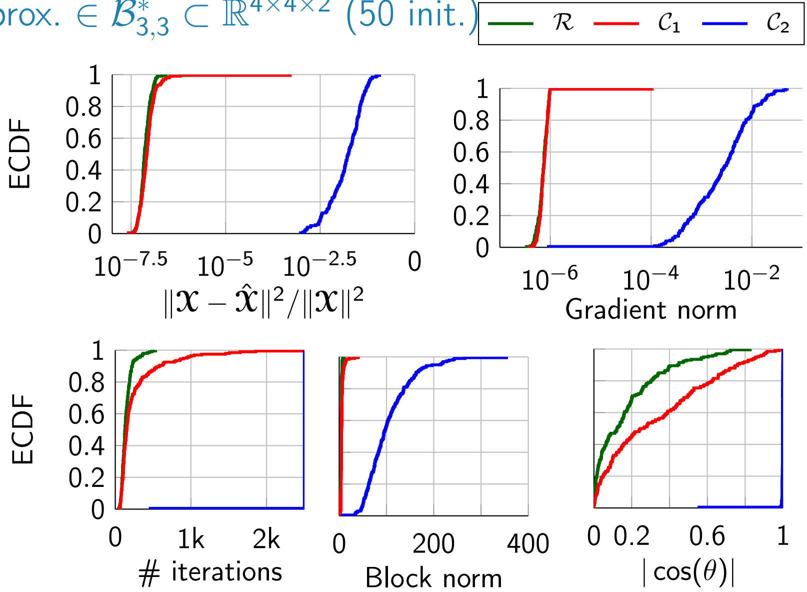
Therefore: $\underset{Z(\lambda) \in \mathcal{B}_{s,s}}{\arg\min} \; \|X(\lambda) - Z(\lambda)\| = \emptyset$

Much trickier: does it happen for positive-volume sets?

Seems true: one can generate random examples where

- Optimization algorithm never seems to converge
- Keeps improving error but ultimately gives ill-conditioned estimate (nearly collinear blocks)
- Norms of these blocks "blow up," but overall error stays bounded

Approx. $\in \mathcal{B}^*_{3,3} \subset \mathbb{R}^{4\times4\times2}$ (50 init.)



Main result

Theorem (G. & C., 2017)

If $\mathcal{X}$ is a real $2k \times 2k \times 2$ tensor satisfying $\rho(\mathcal{X}) = (2k, 2k)$, then $\mathcal{X}$ has no best approximation in $\mathcal{B}_{2k-1,2k-1}$.

Main result

Theorem (G. & C., 2017)

If $\mathcal{X}$ is a real $2k \times 2k \times 2$ tensor satisfying $\rho(\mathcal{X}) = (2k, 2k)$, then $\mathcal{X}$ has no best approximation in $\mathcal{B}_{2k-1,2k-1}$.

- Full minimal ranks $\Leftrightarrow$ all eigenvalues are complex-valued
  - Open nonempty set $\implies$ has positive volume
- Still true over $\{\mathbf{H}_1 \otimes \mathbf{w}_1 + \mathbf{H}_2 \otimes \mathbf{w}_2 \ : \ \operatorname{rank} \mathbf{H}_i \leq 2k-1\}$

Proof reasoning

- $\mathcal{C}_k = \mathcal{B}_{2k,2k}/\mathcal{B}_{2k,2k-1}$ (set of full-minimal rank tensors) is open
- $\implies \mathcal{B}_{2k,2k-1} = \mathbb{R}^{m\times n\times d} \setminus \mathcal{B}_{2k,2k}$ is closed
- Best approximation of $\mathcal{X} \in \mathcal{C}_k$ in $\mathcal{B}_{2k,2k-1}$ exists and has minimal ranks $(2k, 2k-1)$
- But tensors having minimal ranks $(2k, 2k-1)$ can be arbitrarily well approximated by tensors in $\mathcal{B}_{2k-1,2k-1}$
- $\implies \mathcal{X}$ has no best approximation in $\mathcal{B}_{2k-1,2k-1}$

Summary

Tensors and tensor decompositions

Minimal ranks of BTDs and of matrix polynomials

Non-existence of best approximate BTD

Final remarks

Contributions and possible extensions

Minimal ranks

- Study of BTD properties
- Coordinate-free: invariant under action of GL
- Connection with matrix pencils/matrix polynomials
- Computation of ρ for matrix pencils based on KCF
- What about matrix polynomials?

Ill-posedness of best approximate BTD

- Extension of CPD degeneracy example
- Over $\mathbb{R}^{2k\times 2k\times 2}$, it happens for a positive-volume set

For more info, see arXiv:1712.05742

henrique.goulart @ gipsa-lab.fr