X-rank and identifiability for a polynomial decomposition model

Pierre Comon, Yang Qi and Konstantin Usevich

GIPSA-lab, CNRS and Univ. Grenoble Alpes

SIGMA'2016 @ CIRM, Luminy, 03.11.2016

Supported by ERC grant #320594 "Decoda"

Overview

Overview

Matrix and tensor rank

A polynomial decomposition

Waring decomposition

X-rank and identifiablility

Results

Matrix rank

Rank of $M \in \mathbb{K}^{I \times J}$ ($\mathbb{K} = \mathbb{R}, \mathbb{C}$)
Two definitions:

1. $\operatorname{rank}(M) \stackrel{\text{def}}{=} \dim \operatorname{colspan} M = \dim \operatorname{rowspan} M$
2. $\operatorname{rank}(M) \stackrel{\text{def}}{=}$ minimal r such that

$$M = \mathbf{a}_1 \mathbf{b}_1^\top + \cdots + \mathbf{a}_r \mathbf{b}_r^\top, \quad \mathbf{a}_k \in \mathbb{K}^I, \mathbf{b}_k \in \mathbb{K}^J$$

sum of r rank-one matrices

Matrix rank

Rank of $M \in \mathbb{K}^{I \times J}$ ($\mathbb{K} = \mathbb{R}, \mathbb{C}$)
Two definitions:

1. $\operatorname{rank}(M) \stackrel{\text{def}}{=} \dim \operatorname{colspan} M = \dim \operatorname{rowspan} M$
2. $\operatorname{rank}(M) \stackrel{\text{def}}{=}$ minimal r such that

$$M = \mathbf{a}_1 \mathbf{b}_1^\top + \cdots + \mathbf{a}_r \mathbf{b}_r^\top, \quad \mathbf{a}_k \in \mathbb{K}^I, \mathbf{b}_k \in \mathbb{K}^J$$

sum of r rank-one matrices

Properties:

- Rank does not exceed dimensions ($r \leq \min(I, J)$), $r = \min(I, J)$ for general (random) M

Matrix rank

Rank of $M \in \mathbb{K}^{I \times J}$ ($\mathbb{K} = \mathbb{R}, \mathbb{C}$)
Two definitions:

1. $\operatorname{rank}(M) \stackrel{\text{def}}{=} \dim \operatorname{colspan} M = \dim \operatorname{rowspan} M$
2. $\operatorname{rank}(M) \stackrel{\text{def}}{=}$ minimal r such that

$$M = \mathbf{a}_1 \mathbf{b}_1^\top + \cdots + \mathbf{a}_r \mathbf{b}_r^\top, \quad \mathbf{a}_k \in \mathbb{K}^I, \mathbf{b}_k \in \mathbb{K}^J$$

sum of r rank-one matrices

Properties:

- Rank does not exceed dimensions ($r \leq \min(I, J)$), $r = \min(I, J)$ for general (random) M
- Not unique: $AB = ASS^{-1}B$ for any nonsingular S, we can take $\mathbf{a}_k$ and/or $\mathbf{b}_k$ orthogonal (SVD, QR)

Matrix rank

Rank of $M \in \mathbb{K}^{I \times J}$ ($\mathbb{K} = \mathbb{R}, \mathbb{C}$)
Two definitions:

1. $\operatorname{rank}(M) \stackrel{\mathrm{def}}{=} \dim \operatorname{colspan} M = \dim \operatorname{rowspan} M$
2. $\operatorname{rank}(M) \stackrel{\mathrm{def}}{=}$ minimal r such that

$$M = \mathbf{a}_1 \mathbf{b}_1^\top + \cdots + \mathbf{a}_r \mathbf{b}_r^\top, \quad \mathbf{a}_k \in \mathbb{K}^I, \mathbf{b}_k \in \mathbb{K}^J$$

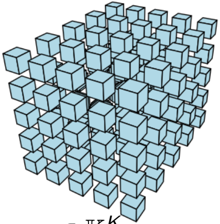
sum of r rank-one matrices

Properties:

- Rank does not exceed dimensions ($r \leq \min(I, J)$), $r = \min(I, J)$ for general (random) M
- Not unique: $AB = ASS^{-1}B$ for any nonsingular S, we can take $\mathbf{a}_k$ and/or $\mathbf{b}_k$ orthogonal (SVD, QR)
- $\operatorname{rank}_{\mathbb{C}}(M) = \operatorname{rank}_{\mathbb{R}}(M)$,

Tensor CP (canonical polyadic) rank

Tensor: d-dimensional array $\mathcal{T} = \left[\mathcal{T}_{i,j,\dots,k}\right]_{i,j,\dots,k=1}^{I,J,\dots,K}$



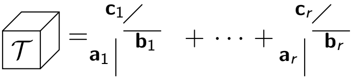
Rank-one tensor: $\mathcal{T}_{i,j,\cdots,k} = a_i b_j \cdots c_k$

Notation: $\mathcal{T} = \mathbf{a} \otimes \mathbf{b} \otimes \cdots \otimes \mathbf{c}, \quad \mathbf{a} \in \mathbb{K}^I, \mathbf{b} \in \mathbb{K}^J, \cdots, \mathbf{c} \in \mathbb{K}^K$

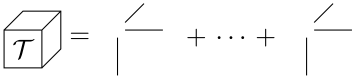
Definition (Hitchock,1927)

$\text{rank}(\mathcal{T}) \stackrel{\text{def}}{=}$ minimal r such that

$$\mathcal{T} = \sum_{\ell=1}^{r} \mathbf{a}_\ell \otimes \mathbf{b}_\ell \otimes \cdots \otimes \mathbf{c}_\ell \qquad \text{(CP decomposition)}$$

A picture: 

Usefulness of tensor CPD

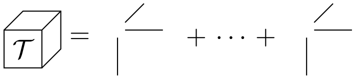


- Data mining (identification problems)

  Tensor = data/signal,   CP rank = # of components in a signal

  Examples: **fluorescence spectroscopy** (talk of Caroline Chaux), antenna array processing, hyperspectral imaging, etc.

Usefulness of tensor CPD

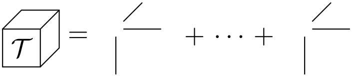


- Data mining (identification problems)

  Tensor = data/signal,   CP rank = # of components in a signal

  Examples: **fluorescence spectroscopy** (talk of Caroline Chaux), antenna array processing, hyperspectral imaging, etc.
- function approximation (talk of Anthony Nouy)

Usefulness of tensor CPD

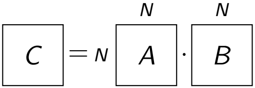


- Data mining (identification problems)

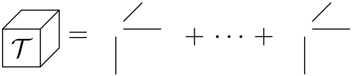
  Tensor = data/signal, CP rank = # of components in a signal

  Examples: **fluorescence spectroscopy** (talk of Caroline Chaux), antenna array processing, hyperspectral imaging, etc.
- function approximation (talk of Anthony Nouy)
- **Complexity of matrix multiplication**

# of operations: $O(N^3)$



Usefulness of tensor CPD



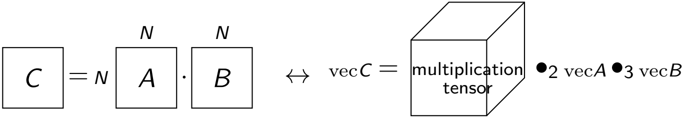
- Data mining (identification problems)

  Tensor = data/signal, CP rank = # of components in a signal

  Examples: **fluorescence spectroscopy** (talk of Caroline Chaux), antenna array processing, hyperspectral imaging, etc.
- function approximation (talk of Anthony Nouy)
- **Complexity of matrix multiplication**

  (Strassen, 1969):

  # of operations: $O(N^3) \quad \to O(N^{2.8})$



Symmetric tensor CPD

$$\mathcal{T} = \left[\mathcal{T}_{(i_1,\ldots,i_d)}\right]_{i_1,\ldots,i_d=1}^{m,\ldots,m} \in \mathbb{K}^{m\times\cdots\times m} \qquad \underbrace{\mathcal{T}_{(i_1,\ldots,i_d)} = \mathcal{T}_{\pi(i_1,\ldots,i_d)}, \forall\pi}_{\text{symmetric}}$$

Symmetric rank of the tensor: minimal r, such that

$$\mathcal{T} = \sum_{k=1}^{r} c_k \mathbf{a}_k \otimes \cdots \otimes \mathbf{a}_k \qquad (*)$$

Symmetric tensor CPD

$$\mathcal{T} = \left[\mathcal{T}_{(i_1,\ldots,i_d)}\right]_{i_1,\ldots,i_d=1}^{m,\ldots,m} \in \mathbb{K}^{m\times\cdots\times m} \qquad \underbrace{\mathcal{T}_{(i_1,\ldots,i_d)} = \mathcal{T}_{\pi(i_1,\ldots,i_d)}, \forall\pi}_{\text{symmetric}}$$

Symmetric rank of the tensor: minimal r, such that

$$\mathcal{T} = \sum_{k=1}^{r} c_k \mathbf{a}_k \otimes \cdots \otimes \mathbf{a}_k \qquad (*)$$

Useful for **blind source separation** (Comon, Jutten, 2010):

Mixing model: $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \underbrace{\boxed{A}}_{\text{unknown}} \underbrace{\begin{bmatrix} s_1 \\ \vdots \\ s_r \end{bmatrix}}_{\text{unknown}}.$

If s_k are independent (real) random variables, then the cumulant of $\mathbf{x}$ has the form $(*)$.

Tensors ($d \geq 3$): bad news

- Set $\{\mathcal{T} \mid \text{rank}(\mathcal{T}) \leq r\}$ is not closed.

$$\mathcal{T} = \mathbf{a} \otimes \mathbf{b} \otimes \mathbf{b} + \mathbf{b} \otimes \mathbf{a} \otimes \mathbf{b} + \mathbf{b} \otimes \mathbf{b} \otimes \mathbf{a}$$

rank$(\mathcal{T}) = 3$, but

$$\mathcal{T} = \frac{1}{\varepsilon}\left((\mathbf{a} + \varepsilon \mathbf{b})^{\otimes 3} - \mathbf{a}^{\otimes 3}\right) + O(\varepsilon)$$

- No polynomial time algorithm to determine rank (Hastad, 1990).
- $\text{rank}_{\mathbb{C}}(\mathcal{T}) \leq \text{rank}_{\mathbb{R}}(\mathcal{T})$, may be strict
- For **symmetric** tensors: symmetric rank $\stackrel{?}{=}$ rank (Comon conjecture)

Tensors ($d \geq 3$): good (or interesting) news

- Rank can exceed dimensions
  Example: $\mathcal{T} = \left[\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right]$, $\text{rank}(\mathcal{T}) = 3$
- CP decomposition is often unique, doesn't need to be orthogonal
- Unusual rank properties (take $2 \times 2 \times 2$ tensor):
  - maximal rank is 3
  - random (Gaussian i.i.d.) real tensor:
    $P(\text{rank}(\mathcal{T}) = 2) = \pi/4$, $P(\text{rank}(\mathcal{T}) = 3) = 1 - \pi/4$.
  - random complex tensor: rank 2

Overview

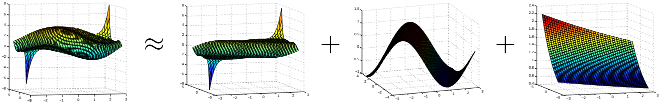
A polynomial decomposition

Given a multivariate polynomial $f(u_1, \ldots, u_m) = \sum_{i_1,\ldots,i_m=0}^{|i_1+\ldots+i_m|\leq d} f_{i_1,\ldots,i_m} u_1^{i_1} \cdots u_m^{i_m}$,

find its shortest representation

$$f(\mathbf{u}) = g_1(\mathbf{v}_1^\top \mathbf{u}) + \cdots + g_1(\mathbf{v}_r^\top \mathbf{u}),$$

where $\mathbf{v}_k \in \mathbb{K}^m$, $g_k(t) = c_{1,k}t + c_{2,k}t^2 + \cdots + c_{d,k}t^d$, $\deg g_k \leq d$



Appears in

- approximation theory (ridge approximation, (Lin, Pinkus, 1993))
- machine learning (polynomial neural networks, (Shin, Ghosh, 1995))
- blind source separation

Decomposition of polynomial maps

Given a polynomial map $\mathbf{f} : \mathbb{K}^m \to \mathbb{K}^n$, degree d, $\mathbf{f}(\mathbf{u}) = \left[f_1(\mathbf{u}) \cdots f_n(\mathbf{u})\right]^\top$ find its shortest representation

$$\mathbf{f}(\mathbf{u}) = \mathbf{w}_1 g_1(\mathbf{v}_1^\top \mathbf{u}) + \cdots + \mathbf{w}_r g_r(\mathbf{v}_r^\top \mathbf{u}),$$

where $\mathbf{v}_k \in \mathbb{K}^m$, $\mathbf{w}_k \in \mathbb{K}^n$, $g_k(t) = c_{1,k} t + c_{2,k} t^2 + \cdots + c_{d,k} t^d$

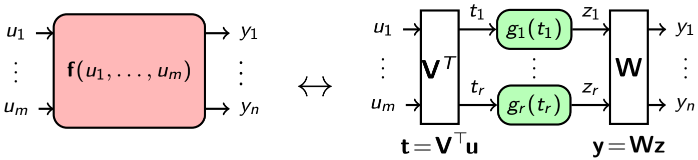
Decomposition of polynomial maps

Given a polynomial map $\mathbf{f} : \mathbb{K}^m \to \mathbb{K}^n$, degree d, $\mathbf{f}(\mathbf{u}) = \left[f_1(\mathbf{u}) \cdots f_n(\mathbf{u})\right]^\top$ find its shortest representation

$$\mathbf{f}(\mathbf{u}) = \mathbf{w}_1 g_1(\mathbf{v}_1^\top \mathbf{u}) + \cdots + \mathbf{w}_r g_r(\mathbf{v}_r^\top \mathbf{u}),$$

where $\mathbf{v}_k \in \mathbb{K}^m$, $\mathbf{w}_k \in \mathbb{K}^n$, $g_k(t) = c_{1,k}t + c_{2,k}t^2 + \cdots + c_{d,k}t^d$

Block-structured nonlinear system identification (Schoukens et al., 2014):



where $\mathbf{V} = \begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_r\end{bmatrix}^\top$ and $\mathbf{W} = \begin{bmatrix}\mathbf{w}_1 & \cdots & \mathbf{w}_r\end{bmatrix}^\top$.

Remarks:

- Degree 1 — matrix factorization
- Can be also interpreted as polynomial neural network.

A tensor-based algorithm*

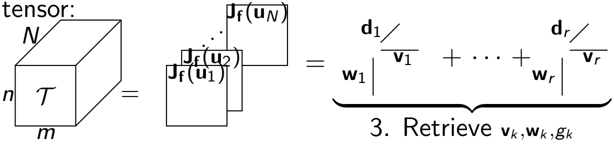
Given a polynomial map $\mathbf{f}$ (of degree d), find r, $\mathbf{w}_k$, $\mathbf{v}_k$, g_k such that

$$\mathbf{f}(\mathbf{u}) = \mathbf{w}_1 g_1(\mathbf{v}_1^\top \mathbf{u}) + \cdots + \mathbf{w}_r g_r(\mathbf{v}_r^\top \mathbf{u}),$$

(Dreesen et al, 2015), (Van Mulders et al, 2014): transform to a tensor CPD

Algorithm.

1. Evaluate $\mathbf{J_f}(\mathbf{u})$ at N points $\mathbf{u}_1, \ldots, \mathbf{u}_N \in \mathbb{K}^m$
2. Stack it into a tensor:



Our questions:

- When is the decomposition unique? (identifiability of the model)
- What is the maximal/typical number of terms?

Existing results

For f of degree d:

$$f(u_1, \ldots, u_m) = g_1(\mathbf{v}_1^\top \mathbf{u}) + \cdots + g_1(\mathbf{v}_r^\top \mathbf{u}),$$

- (Schinzel, 2002), $m = 2$:
  a general ("random") polynomial in $\mathbb{C}$ has $r = \lceil \frac{2d+5-\sqrt{8d+17}}{2} \rceil$ terms
- (Białynicki-Birula, Schinzel, 2008):
  any f can be represented with $\boxed{r \leq \binom{m+d-2}{d-1}}$ terms
- nothing about identifiability

Overview

Special case: homogeneous polynomials

Given a homogeneous polynomial
$$f(u_1, \ldots, u_m) = \sum_{|i_1+\ldots+i_m|=d} f_{i_1,\ldots,i_m} u_1^{i_1} \cdots u_m^{i_m}$$
find its shortest representation of the form

$$f(\mathbf{u}) = g_1(\mathbf{v}_1^\top \mathbf{u}) + \cdots + g_1(\mathbf{v}_r^\top \mathbf{u}),$$

Special case: homogeneous polynomials

Given a homogeneous polynomial
$f(u_1, \ldots, u_m) = \sum_{|i_1+\ldots+i_m|=d} f_{i_1,\ldots,i_m} u_1^{i_1} \cdots u_m^{i_m}$
find its shortest representation of the form

$$f(\mathbf{u}) = c_1 \cdot (\mathbf{v}_1^\top \mathbf{u})^d + \cdots + c_r \cdot (\mathbf{v}_r^\top \mathbf{u})^d,$$

Waring decomposition

Special case: homogeneous polynomials

Given a homogeneous polynomial

$$f(u_1, \ldots, u_m) = \sum_{|i_1+\ldots+i_m|=d} f_{i_1,\ldots,i_m} u_1^{i_1} \cdots u_m^{i_m}$$

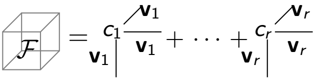
find its shortest representation of the form

$$f(\mathbf{u}) = c_1 \cdot (\mathbf{v}_1^\top \mathbf{u})^d + \cdots + c_r \cdot (\mathbf{v}_r^\top \mathbf{u})^d,$$

Waring decomposition

$$\Updownarrow \quad f(\mathbf{u}) = \mathcal{F} \circ_1 \mathbf{u} \cdots \circ_d \mathbf{u}$$

symmetric tensor CP decomposition



Special case: homogeneous polynomials

Given a homogeneous polynomial

$$f(u_1,\ldots,u_m) = \sum_{|i_1+\ldots+i_m|=d} f_{i_1,\ldots,i_m} u_1^{i_1}\cdots u_m^{i_m}$$

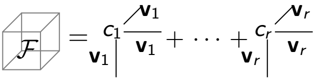
find its shortest representation of the form

$$f(\mathbf{u}) = c_1\cdot(\mathbf{v}_1^\top\mathbf{u})^d + \cdots + c_r\cdot(\mathbf{v}_r^\top\mathbf{u})^d,$$

Waring decomposition

$$\Updownarrow \quad f(\mathbf{u}) = \mathcal{F}\circ_1\mathbf{u}\cdots\circ_d\mathbf{u}$$

symmetric tensor CP decomposition



Example. $(d=2)$:

$$\mathbf{u}^\top T\mathbf{u} = c_1\cdot(..)^2 + \cdots + c_r\cdot(..)^2$$

diagonalize a quadratic form

$$\Updownarrow$$

diagonalize a symmetric matrix

$$T = A\begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_r \end{pmatrix} A^\top$$

Waring problem for $n = 2$ (binary forms)

Theorem (J.J. Sylvester)

A polynomial $f(x,y) = \sum_{j=0}^{d} \binom{d}{j} f_j x^{d-j} y^j$ *has the Waring decomposition*

$$f(x,y) = \sum_{k=1}^{r} c_k \cdot (x + \lambda_k y)^d,$$

if and only if there exist q_j, $j = 0, \ldots, r$ *such that*

$$\begin{bmatrix} q_0 & \cdots & q_r \end{bmatrix} \begin{bmatrix} f_0 & f_1 & \cdots & f_{d-r} \\ f_1 & f_2 & \iddots & \vdots \\ \vdots & \iddots & \iddots & \vdots \\ f_r & \cdots & \cdots & f_d \end{bmatrix} = 0,$$

and $q(t) = \sum_{k=0}^{r} q_k t^k$ *has* r *distinct roots* $\lambda_k \in \mathbb{K} \cup \{\infty\}$.

Proof. follows from $(x + \lambda y)^d = \sum_{j=0}^{d} \binom{d}{j} \lambda^j x^{d-j} y^j$

Overview

X-rank: definitions

Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$, A — N-dim vector space over $\mathbb{K}$.
$\mathrm{rank}_X(\mathbf{v}) \stackrel{\mathrm{def}}{=}$ the minimal r such that

$$\mathbf{v} = \mathbf{p}_1 + \cdots + \mathbf{p}_r \text{ for some } \mathbf{p}_i \in \widehat{X},$$

where $\widehat{X}$ is the set of rank-one elements ($\mathrm{rank}_X(\mathbf{0}) = 0$)
Similar to atomic/sparse decomposition.

Convenient conditions:

- $\widehat{X}$ is scale-invariant ($\alpha\widehat{X} = \widehat{X}$ for $\alpha \in \mathbb{K}$)
- $\widehat{X}$ nondegenerate, i.e., $\widehat{X}$ does not lie in a linear subspace of A;
- $\widehat{X}$ (irreducible) algebraic variety.

Inroduced by (Zak, 2004) ("rank w.r.t. a projective variety")

Some definitions

- $Z \subseteq A$ is an affine algebraic variety, if

$$Z = \{\mathbf{v} \in A | h_1(\mathbf{v}) = \cdots = h_M(\mathbf{v}) = 0\},$$

  for some polynomials h_k.
- Zariski closure of $Y \subseteq A$:
  $\stackrel{\text{def}}{=}$ the smallest algebraic variety Z containing Y.
- A nonempty Z is irreducible, if it cannot be decomposed as a union $Z = Z_1 \cup Z_2$ of distinct (i.e. $Z_1 \not\subseteq Z_2$) varieties.
- Dimension of Z — dimension of the tangent space at smooth points,
  Tangent space $\stackrel{\text{def}}{=}$ kernel of $J_{\mathbf{h}}(\mathbf{v}) = [\frac{\partial h_i}{\partial v_j}(\mathbf{v})]_{i,j=1}^{M,N}$.

X-rank: examples

$$\mathrm{rank}_X(v) \stackrel{def}{=} \min r : \quad \mathbf{v} = \mathbf{p}_1 + \cdots + \mathbf{p}_r, \quad \mathbf{p}_k \in \widehat{X}.$$

object $(v \in A)$	$\dim(A)$	variety $\widehat{X}$	$\dim(\widehat{X})$
$\mathcal{T} \in \mathbb{K}^{I_1 \times \ldots \times I_d}$ tensor	$I_1 \cdots I_d$	$\{\mathbf{a}_1 \otimes \cdots \otimes \mathbf{a}_d\}$ Segre	$\sum_{k=1}^{d} I_k - d + 1$
$f \in \mathbb{K}_{=d}[x,y]$ binary form	$d+1$	$\{c \cdot (ax+by)^d\}$ rational normal curve	2
$f \in \mathbb{K}_{=d}[\mathbf{u}]$ m-ary form	$\binom{m+d-1}{d}$	$\{c \cdot (\mathbf{a}^\top \mathbf{u})^d\}$ Veronese	m
$f \in \mathbb{K}_{\leq d}[x,y]$ polynomial	$\binom{2+d}{d} - 1$	$\{g(ax+by)\}$ rational normal scroll	$d+1$
$f \in \mathbb{K}_{\leq d}[\mathbf{u}]$ polynomial	$\binom{m+d}{d} - 1$	$\{g(\mathbf{a}^\top \mathbf{u})\}$ Veronese scroll	$m+d-1$

Maximum and typical ranks: definitions

- Maximal rank

$$r_{max} \stackrel{\text{def}}{=} \max_{v \in \mathbb{K}^N} \{\text{rank}_X(\mathbf{v}) = k\}.$$

- r is called a typical rank if the set $\{\mathbf{v} \in A\colon \text{rank}_{\mathcal{X}}(\mathbf{v}) = r\}$ contains an open ball

Properties:

- If $\mathbb{K} = \mathbb{C}$ there exists only one typical rank, called generic rank r_{gen}
- (Bernardi et al, 2015): If $\mathbb{K} = \mathbb{R}$, typical ranks form a contiguous set $\{r_{typ}^{min}, \ldots, r_{typ}^{max}\}$
- (Blekherman, Teitler, 2014):
  If $\widehat{X}_{\mathbb{C}} \subset \mathbb{C}^N$ is a complexification of $\widehat{X}_{\mathbb{R}} \subset \mathbb{R}^N$ and $\widehat{X}_{\mathbb{C}}$ has a smooth real point, then $r_{gen}(\widehat{X}_{\mathbb{C}}) = r_{typ}^{min}(\widehat{X}_{\mathbb{R}})$

Example: binary forms

Example: (binary forms of degree d, $\widehat{X}$ — rational normal curve)

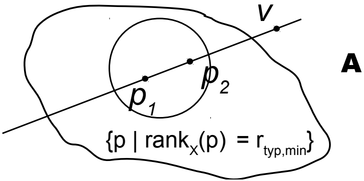
- $r_{max} = d$
- $r_{gen} = \lceil \frac{d+1}{2} \rceil$ (Sylvester)
- (Blekherman, 2014): all ranks $\{r_{gen}, \ldots, r_{max}\}$ are real typical

Relations between maximal, typical and generic rank

Theorem ((Blekherman, Teitler, 2014))

- $r_{max}(\widehat{X}_{\mathbb{R}}) \leq 2r_{typ}^{min}(\widehat{X}_{\mathbb{R}})$
- $r_{max}(\widehat{X}_{\mathbb{C}}) \leq 2r_{gen}(\widehat{X}_{\mathbb{C}})$

Proof.



□

- Sharp for binary forms $d \leq 2\lceil \frac{d+1}{2} \rceil$
- Improved upped bounds on r_{max} for:
  - Waring decomposition for *m*-ary forms
  - rank of $2 \times \cdots \times 2$ non-symmetric tensors

Secant varieties and generic rank ($\mathbb{K} = \mathbb{C}$)

r-th secant variety = closure of set of elements of rank $\leq r$

$$\sigma_r(\widehat{X}) \stackrel{\text{def}}{=} \overline{\{\mathbf{p}_1 + \dots + \mathbf{p}_r | \mathbf{p}_i \in \widehat{X}\}}$$

Properties:

- $\widehat{X} = \sigma_1(\widehat{X}) \subset \sigma_2(\widehat{X}) \subset \dots \subset \sigma_{r_{gen}}(\widehat{X}) = \sigma_{r_{gen}+1}(\widehat{X}) = \dots = A.$
  $\Rightarrow r_{gen}$ — the smallest r such that $\dim \sigma_r(\widehat{X}) = \dim(A)$

Expected dimension: $\exp\dim \sigma_r(\widehat{X}) \stackrel{\text{def}}{=} \min(r \dim \widehat{X}, \dim A)$

- In general, $\exp\dim \sigma_r(\widehat{X}) \geq \dim \sigma_r(\widehat{X})$.
- If "$>$", $\widehat{X}$ is called r-defective.
- If $\widehat{X}$ is not r-defective for all r , then $r_{gen} = \lceil \frac{\dim(A)}{\dim \widehat{X}} \rceil$

Generic rank: examples

$\dim(A)$	variety $\widehat{X}$	$\dim(\widehat{X})$	$r_{gen} \overset{?}{=} \lceil \frac{\dim(A)}{\dim \widehat{X}} \rceil$
$l_1 \cdots l_d$	$\{\mathbf{a}_1 \otimes \cdots \otimes \mathbf{a}_d\}$ Segre	$\sum_{k=1}^{d} l_k - d + 1$	conjecture* (Chiantini et al., 2014)
$d+1$	$\{c \cdot (ax+by)^d\}$ rat. norm. curve	2	yes (Sylvester)
$\binom{m+d-1}{d}$	$\{c \cdot (\mathbf{a}^\top \mathbf{u})^d\}$ Veronese	m	yes* (Alexander, Hirschowitz, 1995)
$\binom{2+d}{d} - 1$	$\{g(ax+by)\}$ rat. norm scroll	$d+1$	no, but $\exists$ formula (Schinzel, 2002)
$\binom{m+d}{d} - 1$	$\{g(\mathbf{a}^\top \mathbf{u})\}$ Veronese scroll	$m+d-1$	no

$*$ = except finite exceptions and degenerate cases

Identifiability (definition)

$$\mathbf{v} = \mathbf{p}_1 + \cdots + \mathbf{p}_r, \quad \mathbf{p}_k \in \widehat{X} \qquad (*)$$

- Uniqueness of $(*)$ — up to permutation
- A general point in $\sigma_r(\widehat{X})$ is of the form $(*)$

Definition (Global identifiability)

$\widehat{X}$ is r-identifiable if a general $\mathbf{v} \in \sigma_r(\widehat{X})$ has unique decomposition $(*)$.

Theorem (Strassen,1983)

- *If* $\dim \sigma_r(\widehat{X}) < \operatorname{expdim} \sigma_r(\widehat{X})$, *a general point in* $\sigma_r(\widehat{X})$ *has infinite number of decompositions.*
- *If* $\dim \sigma_r(\widehat{X}) = \operatorname{expdim} \sigma_r(\widehat{X})$, *a general point in* $\sigma_r(\widehat{X})$ *has a finite number of decompositions.*

Non-defectivity can be checked numerically:

- For r random points $\mathbf{p}_1, \ldots, \mathbf{p}_r \in \widehat{X}$, check if $\dim \operatorname{Span}\langle T_{\mathbf{p}_1}\widehat{X}, \ldots, T_{\mathbf{p}_r}\widehat{X}\rangle = \operatorname{expdim} \sigma_r(\widehat{X})$;

Hierarchy of properties

X not r-weakly-defective $\Rightarrow$ X not r-tangentially weakly defective $\Rightarrow$ X is r-identifiable $\Rightarrow$ X is non-defective

Definition (Chiantini-Ottaviani)

$\widehat{X}$ is called not r-tangentially weakly defective if there exists a set of r points $\mathbf{p}_1, \ldots, \mathbf{p}_r \in \widehat{X}$, such that the span $\mathrm{Span}\langle T_{\mathbf{p}_1}\widehat{X}, \ldots, T_{\mathbf{p}_r}\widehat{X}\rangle$ does not contain $T_{\mathbf{p}}\widehat{X}$ for $\mathbf{p} \notin \{\mathbf{p}_1, \ldots, \mathbf{p}_r\}$.

We can construct a certificate to check global identifiability (Chiantini et. al, 2014)

Overview

Matrix and tensor rank

A polynomial decomposition

Waring decomposition

X-rank and identifiablility

Results

(s, r)-partial identifiability

$=$ generic uniqueness of rank-r decomposition (except $\{c_{j,k}\}_{j<s}$)

$$\mathbf{f}(\mathbf{u}) = \mathbf{w}_1 g_1(\mathbf{v}_1^\top \mathbf{u}) + \cdots + \mathbf{w}_r g_r(\mathbf{v}_r^\top \mathbf{u}),$$

where $\mathbf{v}_k \in \mathbb{K}^m$, $\mathbf{w}_k \in \mathbb{K}^n$, $g_k(t) = c_{1,k} t + c_{2,k} t^2 + \cdots + c_{d,k} t^d$

Proposition ((Comon, Q., U., 2016), simplified)

Let d, m, n be such that $d > 3$, $m \geq 2$, and s be a number $1 \leq s < d$

If $r \leq \min\left(\binom{m+s-1}{s}, \left\lceil \frac{\binom{m+d-1}{d}}{m+n-1} - 1 \right\rceil\right) n$

then the decomposition $(\star)$ is (s, r)-partially identifiable.

(s, r)-partial identifiability

$=$ generic uniqueness of rank-r decomposition (except $\{c_{j,k}\}_{j<s}$)

$$\mathbf{f}(\mathbf{u}) = \mathbf{w}_1 g_1(\mathbf{v}_1^\top \mathbf{u}) + \cdots + \mathbf{w}_r g_r(\mathbf{v}_r^\top \mathbf{u}),$$

where $\mathbf{v}_k \in \mathbb{K}^m$, $\mathbf{w}_k \in \mathbb{K}^n$, $g_k(t) = c_{1,k}t + c_{2,k}t^2 + \cdots + c_{d,k}t^d$

Proposition ((Comon, Q., U., 2016), simplified)

Let d, m, n be such that $d>3$, $m \geq 2$, and s be a number $1 \leq s < d$
If $r \leq \min\left(\binom{m+s-1}{s}, \left\lceil \frac{\binom{m+d-1}{d}}{m+n-1} - 1 \right\rceil\right) n$
then the decomposition $(\star)$ is (s, r)-partially identifiable.

Sketch of the proof:

- Prove identifiability of d-th homogeneous part of general $\mathbf{f}(\mathbf{u})$.
- Determine $c_{j,k}$, $j \geq s$ from the linear system:

$$\mathbf{f}^{(j)}(\mathbf{u}) = c_{j,1} \cdot \mathbf{w}_1(\mathbf{v}_1^\top \mathbf{u})^j + \cdots + c_{j,r} \cdot \mathbf{w}_r(\mathbf{v}_r^\top \mathbf{u})^j$$

Identifiability: examples

Corollary

1. *The decomposition* **cannot be** r**-identifiable for** $r > mn$.
2. *Gven* m, n*: the decomposition is mn-identifiable for large* d

Table: Our bound for maximal identifiable ranks, for $d = 3$.

$m \setminus n$	1	2	3	4	5	6	7	8	9	10	11	12
2	1	2	0	0	0	0	0	0	0	0	0	0
3	**3**	4	3	4	5	6	7	0	0	0	0	0
4	**4**	6	9	8	10	12	7	8	9	10	11	12
5	**5**	**10**	12	16	15	18	21	16	18	20	22	24
6	**6**	**12**	**18**	**24**	25	30	28	32	27	30	33	36
7	**7**	**14**	**21**	**28**	**35**	36	42	40	45	50	44	48
8	**8**	**16**	**24**	**32**	**40**	**48**	**56**	56	63	70	66	72
9	**9**	**18**	**27**	**36**	**45**	**54**	**63**	**72**	**81**	**90**	88	96
10	**10**	**20**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**100**	**110**	**120**
11	**11**	**22**	**33**	**44**	**55**	**66**	**77**	**88**	**99**	**110**	**121**	**132**
12	**12**	**24**	**36**	**48**	**60**	**72**	**84**	**96**	**108**	**120**	**132**	**144**

Identifiability: examples

Corollary

1. *The decomposition* **cannot be r-identifiable for** $r > mn$.
2. *Gven m, n: the decomposition is mn-identifiable for large d*

Table: Our bound for maximal identifiable ranks, for $d = 4$.

$m \setminus n$	1	2	3	4	5	6	7	8	9	10	11	12
2	**2**	2	3	0	0	0	0	0	0	0	0	0
3	**3**	**6**	6	8	10	6	7	8	9	10	11	12
4	**4**	**8**	**12**	**16**	**20**	18	21	24	18	20	22	24
5	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	40	44	48
6	**6**	**12**	**18**	**24**	**30**	**36**	**42**	**48**	**54**	**60**	**66**	**72**
7	**7**	**14**	**21**	**28**	**35**	**42**	**49**	**56**	**63**	**70**	**77**	**84**
8	**8**	**16**	**24**	**32**	**40**	**48**	**56**	**64**	**72**	**80**	**88**	**96**
9	**9**	**18**	**27**	**36**	**45**	**54**	**63**	**72**	**81**	**90**	**99**	**108**
10	**10**	**20**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**100**	**110**	**120**
11	**11**	**22**	**33**	**44**	**55**	**66**	**77**	**88**	**99**	**110**	**121**	**132**
12	**12**	**24**	**36**	**48**	**60**	**72**	**84**	**96**	**108**	**120**	**132**	**144**

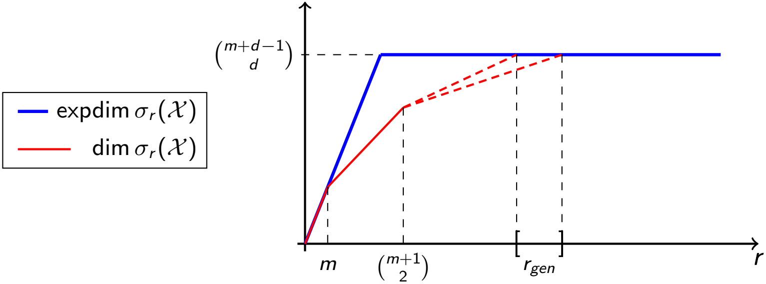
Generic and maximal ranks (univariate poly)

Proposition ((Comon, Qi, U., 2016), simplified)

Let $d \geq 4$, $n = 1$, and $m > (d-2)(d-1)$. Then

$$\left\lceil \frac{\binom{m+d-2}{d-1} + \binom{m+d-1}{d}}{m+1} \right\rceil \leq r_{gen} \leq \left\lceil \frac{\binom{m+d-2}{d-1} + (m-1)\left\lceil \frac{\binom{m+d-1}{d}}{m} \right\rceil}{m} \right\rceil$$

Proof (sketch). from the result on partial identifiability, we have that:



Maximal ranks

Recall that $r_{max} \leq 2r_{gen}$ (Blekherman, Teitler, 2014)

Corollary

$$r_{max} \leq 2 \left\lceil \frac{\binom{m+d-2}{d-1} + (m-1)\left\lceil \frac{\binom{m+d-1}{d}}{m} \right\rceil}{m} \right\rceil$$

Best known previous bound (Białynicki-Birula, Schinzel, 2008):

$$r_{max} \leq \binom{m+d-2}{d-1}$$

Our bound is better, the ratio $\sim \frac{2}{d}$ (as $m \to \infty$)

Summary

arXiv:1603.01566

- X-rank is very useful (or at least interesting).
- Can be extended to semialgebraic X, e.g. nonnegative tensors (Qi et al, 2016).
- Hopefully, useful for other sparse decompositions.

Summary

arXiv:1603.01566

- X-rank is very useful (or at least interesting).
- Can be extended to semialgebraic X, e.g. nonnegative tensors (Qi et al, 2016).
- Hopefully, useful for other sparse decompositions.

Thank you for your attention!