Tensor Decomposition, Low Rank Structured Matrix Approximation and Applications

B. Mourrain
GALAAD, INRIA Méditerranée, Sophia Antipolis
Bernard.Mourrain@inria.fr

16h-16h25, Thursday, August 1

Generalized Waring problem (1770)

Given a homogeneous polynomial T of degree d in the variables $\mathbf{x} = (x_0, x_1, \ldots, x_n)$:

$$T(\mathbf{x}) = \sum_{|\alpha|=d} T_\alpha \, \mathbf{x}^\alpha,$$

find a minimal decomposition of T of the form

$$T(\mathbf{x}) = \sum_{i=1}^{r} \gamma_i (\zeta_{i,0} x_0 + \zeta_{i,1} x_1 + \cdots + \zeta_{i,n} x_n)^d$$

for $\zeta_i = (\zeta_{i,0}, \zeta_{i,1}, \ldots, \zeta_{i,n}) \in \mathbb{C}^{n+1}$, $\gamma_i \in \mathbb{C}$.

The minimal r in such a decomposition is called the **rank** of T.

Geometric point of view

Definition (Veronese variety)

$$\begin{array}{rcl} \nu_d : \mathbb{P}(E) & \to & \mathbb{P}(S^d(E)) \\ [\mathbf{v}] & \mapsto & [\mathbf{v}(\mathbf{x})^{\mathbf{d}}] \end{array}$$

Its image is the **Veronese** variety, denoted $\chi_1(S^d(E))$.

Definition (Secant of the Veronese variety)

$\chi_r^{0,d} = \{[T] \in \mathbb{P}(S^d) \mid \exists\, \mathbf{w}_1, \ldots, \mathbf{w}_r \in \chi_1(S_d(E)) \text{ s.t. } T = \sum_{i=1}^r \mathbf{w}_i\}$
$\chi_r^d = \overline{\chi_r^{0,d}}$

- ☞ $r(T) =$ smallest r s.t. $T \in \chi_r^{0,d}$, called the **rank** of T.
- ☞ $r_\sigma(T) =$ smallest r s.t. $T \in \chi_r^d$, called the **border rank** of T.

Example: $T_\epsilon := \frac{1}{d\,\varepsilon}\left((x_0 + \varepsilon x_1)^d - x_0^d\right)$, $T_0 := x_0^{d-1} x_1$.

$$r(T_\epsilon) = r_\sigma(T_0) = 2 \text{ but } r(T_0) = d.$$

Algebraic point of view

Apolar product: For $f = \sum_{|\alpha|=d} f_\alpha \binom{d}{\alpha} \mathbf{x}^\alpha$, $g = \sum_{|\alpha|=d} g_\alpha \binom{d}{\alpha} \mathbf{x}^\alpha \in S^d$,

$$\langle f, g \rangle = \sum_{|\alpha|=d} f_\alpha \, g_\alpha \binom{d}{\alpha}.$$

For homogeneous polynomials $g(x_0, \ldots, x_n)$ of degree d and $\mathbf{k}(\mathbf{x}) = \mathbf{k}_0 \, x_0 + \cdots + \mathbf{k}_n x_n$,

$$\langle g(\mathbf{x}), \mathbf{k}(\mathbf{x})^d \rangle = g(\mathbf{k}_0, \ldots, \mathbf{k}_n) = g(\mathbf{k}).$$

☞ **If** $T = \mathbf{k}_1(\mathbf{x})^d + \cdots + \mathbf{k}_r(\mathbf{x})^d$ **and** $g(\mathbf{k}_i) = 0$ **for** $i = 1, \ldots, r$**, then** $\forall h$ **with** $\deg(h) = d - \deg(g)$

$$\langle g \, h, T \rangle = 0.$$

☞ **Find the polynomials apolar to T and compute their roots.**

Algebraic point of view

$S := \mathbb{K}[x_0, \ldots, x_n]$; $S^d := \{f \in S; \deg(f) = d\}$;

Definition (Apolar ideal)

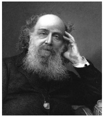
$(T^\perp) = \{g \in S \mid \forall h \in S^{d-\deg(g)},\ \langle g\, h, T\rangle = 0\} \supset S^{d+1}$.

Problem: find an ideal $I \subset S$ such that

- $I \subset (T^\perp)$;
- **I is saturated zero dimensional;**
- **I defines a minimal number r of simple points.**

☞ **necessary and sufficient conditions** that T is of rank r.

Sylvester approach (1886)



Theorem

The binary form $T(x_0,x_1)=\sum_{i=0}^{d} t_i\binom{d}{i}x_0^{d-i}x_1^{i}$ *can be decomposed as a sum of* r *distinct powers of linear forms*

$$T=\sum_{k=1}^{r}\lambda_k(\alpha_k x_0+\beta_k x_1)^d$$

iff there exists a polynomial q *such that*

$$\begin{bmatrix} t_0 & t_1 & \dots & t_r \\ t_1 & & & t_{r+1} \\ \vdots & & & \vdots \\ t_{d-r} & \dots & t_{d-1} & t_d \end{bmatrix}\begin{bmatrix} q_0 \\ q_1 \\ \vdots \\ q_r \end{bmatrix}=0$$

and of the form $\quad q(x_0,x_1):=\mu\prod_{k=1}^{r}(\beta_k x_0-\alpha_k x_1).$

Interpolation point of view

For $T = T(\underline{\mathbf{x}}) = \sum_{\substack{\alpha\in\mathbb{N}^n;\\ |\alpha|\le\delta}} T_\alpha\, \underline{\mathbf{x}}^\alpha \in \mathcal{T}$, let
$T^*(\mathbf{d}) = \sum_{\substack{\alpha\in\mathbb{N}^n;\\ |\alpha|\le\delta}} \binom{\delta}{\alpha}^{-1} T_\alpha\, \mathbf{d}^\alpha \in R^* = \mathbb{K}[[\mathbf{d}]]$ such that $\forall T' \in R_\delta$,

$$\langle T | T' \rangle = T^*(T').$$

Property: **Given $T \in R_\delta$, find $\gamma_i \neq 0, \zeta_i \in \mathbb{K}^n,\ i = 1, \dots, r$, such that**

$$T = \sum_{i=1}^{r} \gamma_i \, \langle \zeta_i, \mathbf{x} \rangle^d$$

iff

$$T^* \equiv \sum_{i=1}^{r} \gamma_i \, \mathbf{1}_{\zeta_i}.$$

on R_δ.

Flat extension point of view

Flattening map ϕ: For $\mathbf{x}^A \subset R$, $B, B' \subset A$ with $B \cdot B' \subset A$, and $\Lambda := T^* \in \langle \mathbf{x}^A \rangle^* : p = \sum_{\alpha \in A} p_\alpha \mathbf{x}^\alpha \mapsto \sum_{\alpha \in A} p_\alpha T_\alpha$, we define

$$
\begin{aligned}
H_{T^*}^{B',B} : \langle \mathbf{x}^{B'} \rangle &\to \langle \mathbf{x}^B \rangle^* \\
p = \sum_{\beta \in B'} p_\beta \mathbf{x}^\beta &\mapsto p \cdot T^*
\end{aligned}
$$

with $p \cdot \Lambda : q \in \langle B \rangle \mapsto T^*(p\, q) = \langle T, pq \rangle \in \mathbb{K}$.
Its matrix is $H_T^{B',B} := [T_{\beta'+\beta}]_{\beta' \in B', \beta \in B}$.
Problem: Given $T^* \in R^*_\delta$, find $\tilde{\Lambda} \in R^*$ such that

$$
\begin{aligned}
H_{\tilde{\Lambda}} : R &\to R^* \\
p &\mapsto p \cdot \tilde{\Lambda}
\end{aligned}
$$

- **extends $H_{T^*}^{B,B'}$**
- **has minimal rank and**
- **$I_{\tilde{\Lambda}} = \ker H_{\tilde{\Lambda}}$ is a radical ideal.**

Flat extension

$B^+ = B \cap x_1 B \cap \cdots \cap x_n B$. B connected to 1 if $b \in B$ is either 1 or $x_{i_0} b'$, $b' \in B$.

Theorem (LM'09, BCMT'10, BBCM'11)

Let B, B' be connected to 1 of size r, E, E' connected to 1 with $B^+ \subset E$, $B'^+ \subset E'$ and $\Lambda \in \langle E \cdot E' \rangle^*$. The following conditions are equivalent:

1. *there exists a unique element $\tilde{\Lambda} \in R^*$ which extends Λ and such that B and B' are basis of $\mathcal{A}_\Lambda = R/I_\Lambda$.*
2. $\mathrm{rank} H_\Lambda^{E,E'} = \mathrm{rank} H_\Lambda^{B,B'} = r$.
3. *$H_\Lambda^{B,B'}$ is invertible and the matrices $M_i := H_\Lambda^{B,x_i B'} (H_\Lambda^{B,B'})^{-1}$ satisfy*

$$M_i \circ M_j = M_j \circ M_i \quad (1 \le i, j \le n).$$

In this case,

- $I_\Lambda = (\ker H_\Lambda^{E,E'})$,
- *$\tilde{\Lambda}$ is supported on the points $\mathcal{V}_\Lambda^{\mathbb{C}} = \mathcal{V}^{\mathbb{C}}(I_\Lambda) = \{\zeta_1, \ldots, \zeta_{r'}\}$ with $r' < r$.*

Recovering the decomposition

If the flat extension problem has a solution $\tilde{\Lambda}$, then:

- $r = \mathrm{rank} H_{\tilde{\Lambda}} = \dim R/(\ker H_{\tilde{\Lambda}})$ where $r = |B| = |B'|$;
- Let B, B' be maximal sets of monomials $\subset R$ s.t. $H_{\tilde{\Lambda}}^{B',B}$ invertible, then
$$M_i^{B',B} := H_{\tilde{\Lambda}}^{B',x_i B} (H_{\tilde{\Lambda}}^{B',B})^{-1}.$$
is the matrix of multiplication by x_i in $\mathcal{A}_{\tilde{\Lambda}} = R/(\ker H_{\tilde{\Lambda}})$.
- If the decomposition is of size r, the eigenvectors of the operators $(M_i^t)_i$ are simple and equal (up to scalar) to $\{\mathbf{1}_{\zeta_1}, \ldots, \mathbf{1}_{\zeta_r}\}$.
- $\ker H_{\tilde{\Lambda}} = (\ker H_{\tilde{\Lambda}}^{B'^+,B^+})$ where $B^+ = B \cup x_1\, B \cup \cdots x_n B$;
- For each $x^\alpha \in \partial B = B^+ \setminus B$, there exists a unique
$$f_\alpha = x^\alpha - \Sigma_\beta z_{\alpha,\beta}\, x^\beta \in \ker H_\Lambda^{B'^+,B^+}.$$
The $(f_\alpha)_{\alpha \in \partial B}$ form a **border basis** of $I_{\tilde{\Lambda}} = \ker H_{\tilde{\Lambda}}$ with respect to B.

Example in $\chi_6(S^4(\mathbb{K}^3))$

The tensor:

$$T = 79\, x_0 x_1^3 + 56\, x_0^2 x_2^2 + 49\, x_1^2 x_2^2 + 4\, x_0 x_1 x_2^2 + 57\, x_0^3 x_1.$$

The 15×15 Hankel matrix:

	1	x_1	x_2	x_1^2	x_1x_2	x_2^2	x_1^3	$x_1^2x_2$	$x_1x_2^2$	x_2^3	x_1^4	$x_1^3x_2$	$x_1^2x_2^2$	$x_1x_2^3$	x_2^4
1	0	$\frac{57}{4}$	0	0	0	$\frac{28}{3}$	$\frac{79}{4}$	0	$\frac{1}{3}$	0	0	0	$\frac{49}{6}$	0	0
x_1	$\frac{57}{4}$	0	0	$\frac{79}{4}$	0	$\frac{1}{3}$	0	0	$\frac{49}{6}$	0	h_{500}	h_{410}	h_{320}	h_{230}	h_{140}
x_2	0	0	$\frac{28}{3}$	0	$\frac{1}{3}$	0	0	$\frac{49}{6}$	0	0	h_{410}	h_{320}	h_{230}	h_{140}	h_{050}
x_1^2	0	$\frac{79}{4}$	0	0	0	$\frac{49}{6}$	h_{500}	h_{410}	h_{320}	h_{230}	h_{600}	h_{510}	h_{420}	h_{330}	h_{240}
x_1x_2	0	0	$\frac{1}{3}$	0	$\frac{49}{6}$	0	h_{410}	h_{320}	h_{230}	h_{140}	h_{510}	h_{420}	h_{330}	h_{240}	h_{150}
x_2^2	$\frac{28}{3}$	$\frac{1}{3}$	0	$\frac{49}{6}$	0	0	h_{320}	h_{230}	h_{140}	h_{050}	h_{420}	h_{330}	h_{240}	h_{150}	h_{060}
x_1^3	$\frac{79}{4}$	0	0	h_{500}	h_{410}	h_{320}	h_{600}	h_{510}	h_{420}	h_{330}	h_{700}	h_{610}	h_{520}	h_{430}	h_{340}
$x_1^2x_2$	0	0	$\frac{49}{6}$	h_{410}	h_{320}	h_{230}	h_{510}	h_{420}	h_{330}	h_{240}	h_{610}	h_{520}	h_{430}	h_{340}	h_{250}
$x_1x_2^2$	$\frac{1}{3}$	$\frac{49}{6}$	0	h_{320}	h_{230}	h_{140}	h_{420}	h_{330}	h_{240}	h_{150}	h_{520}	h_{430}	h_{340}	h_{250}	h_{160}
x_2^3	0	0	0	h_{230}	h_{140}	h_{050}	h_{330}	h_{240}	h_{150}	h_{060}	h_{430}	h_{340}	h_{250}	h_{160}	h_{070}
x_1^4	0	h_{500}	h_{410}	h_{600}	h_{510}	h_{420}	h_{700}	h_{610}	h_{520}	h_{430}	h_{800}	h_{710}	h_{620}	h_{530}	h_{440}
$x_1^3x_2$	0	h_{410}	h_{320}	h_{510}	h_{420}	h_{330}	h_{610}	h_{520}	h_{430}	h_{340}	h_{710}	h_{620}	h_{530}	h_{440}	h_{350}
$x_1^2x_2^2$	$\frac{49}{6}$	h_{320}	h_{230}	h_{420}	h_{330}	h_{240}	h_{520}	h_{430}	h_{340}	h_{250}	h_{620}	h_{530}	h_{440}	h_{350}	h_{260}
$x_1x_2^3$	0	h_{230}	h_{140}	h_{330}	h_{240}	h_{150}	h_{430}	h_{340}	h_{250}	h_{160}	h_{530}	h_{440}	h_{350}	h_{260}	h_{170}
x_2^4	0	h_{140}	h_{050}	h_{240}	h_{150}	h_{060}	h_{340}	h_{250}	h_{160}	h_{070}	h_{440}	h_{350}	h_{260}	h_{170}	h_{080}

Extract a (6×6) principal minor of full rank:

$$\mathtt{H}_{\Lambda}^{B} = \begin{bmatrix} 0 & \frac{57}{4} & 0 & 0 & 0 & \frac{28}{3} \\ \frac{57}{4} & 0 & 0 & \frac{79}{4} & 0 & \frac{1}{3} \\ 0 & 0 & \frac{28}{3} & 0 & \frac{1}{3} & 0 \\ 0 & \frac{79}{4} & 0 & 0 & 0 & \frac{49}{6} \\ 0 & 0 & \frac{1}{3} & 0 & \frac{49}{6} & 0 \\ \frac{28}{3} & \frac{1}{3} & 0 & \frac{49}{6} & 0 & 0 \end{bmatrix}$$

The columns (and the rows) of the matrix correspond to the monomials $\{1, x_1, x_2, x_1^2, x_1x_2, x_2^2\}$.

The shifted matrix $\mathtt{H}^{B,B}_{x_1\cdot\Lambda}$ is

$$\mathtt{H}^{B,B}_{x_1\cdot\Lambda} = \mathtt{H}^{B,x_1 B}_{\Lambda} = \begin{bmatrix} \frac{57}{4} & 0 & 0 & \frac{79}{4} & 0 & \frac{1}{3} \\ 0 & \frac{79}{4} & 0 & 0 & 0 & \frac{49}{6} \\ 0 & 0 & \frac{1}{3} & 0 & \frac{49}{6} & 0 \\ \frac{79}{4} & 0 & 0 & h_{500} & h_{410} & h_{320} \\ 0 & 0 & \frac{49}{6} & h_{410} & h_{320} & h_{230} \\ \frac{1}{3} & \frac{49}{6} & 0 & h_{320} & h_{230} & h_{140} \end{bmatrix}$$

The columns of the matrix correspond to the monomials
$\{x_1, x_1^2, x_1x_2, x_1^3, x_1^2x_2, x_1x_2^2\} = \{1, x_1, x_2, x_1^2, x_1x_2, x_2^2\} \times x_1.$
Similarly,

$$\mathtt{H}^{B,B}_{x_2\cdot\Lambda} = \mathtt{H}^{B,x_2 B}_{\Lambda} = \begin{bmatrix} 0 & 0 & \frac{28}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{49}{6} & 0 \\ \frac{28}{3} & \frac{1}{3} & 0 & \frac{49}{6} & 0 & 0 \\ 0 & 0 & \frac{49}{6} & h_{410} & h_{320} & h_{230} \\ \frac{1}{3} & \frac{49}{6} & 0 & h_{320} & h_{230} & h_{140} \\ 0 & 0 & 0 & h_{230} & h_{140} & h_{050} \end{bmatrix}$$

We form (all) the possible matrix equations:

$$\mathtt{M}_{x_i}\mathtt{M}_{x_j} - \mathtt{M}_{x_j}\mathtt{M}_{x_i} = \mathtt{H}_{x_1\cdot\Lambda}\,\mathtt{H}_{\Lambda}^{-1}\,\mathtt{H}_{x_2\cdot\Lambda}\,\mathtt{H}_{\Lambda}^{-1} - \mathtt{H}_{x_2\cdot\Lambda}\,\mathtt{H}_{\Lambda}^{-1}\,\mathtt{H}_{x_1\cdot\Lambda}\,\mathtt{H}_{\Lambda}^{-1} = 0.$$

Many of the resulting equations are trivial. We have 6 unknonws: $h_{500}, h_{410}, h_{320}, h_{230}, h_{140}, h_{050}$ and 15 non-trivial equations.

A solution of the system is:

$$h_{500} = 1, h_{410} = 2, h_{320} = 3, h_{230} = 1.5060, h_{140} = 4.960, h_{050} = 0.056.$$

We subsitute these values to $\mathtt{H}_{x_1\cdot\Lambda}$ and solve the generalized eigenvalue problem $(\mathtt{H}_{x_1\cdot\Lambda} - \zeta\,\mathtt{H}_{\Lambda})\,\mathbf{v} = 0$.The normalized eigenvectors are

$$\begin{bmatrix} 1 \\ -0.830 + 1.593\,i \\ -0.326 - 0.0501\,i \\ -1.849 - 2.645\,i \\ 0.350 - 0.478\,i \\ 0.103 + 0.0327\,i \end{bmatrix}, \begin{bmatrix} 1 \\ -0.830 - 1.593\,i \\ -0.326 + 0.050\,i \\ -1.849 + 2.645\,i \\ 0.350 + 0.478\,i \\ 0.103 - 0.032\,i \end{bmatrix}, \begin{bmatrix} 1 \\ 1.142 \\ 0.836 \\ 1.305 \\ 0.955 \\ 0.699 \end{bmatrix},$$

$$\begin{bmatrix} 1 \\ 0.956 \\ -0.713 \\ 0.914 \\ -0.682 \\ 0.509 \end{bmatrix}, \begin{bmatrix} 1 \\ -0.838 + 0.130\,i \\ 0.060 + 0.736\,i \\ 0.686 - 0.219\,i \\ -0.147 - 0.610\,i \\ -0.539 + 0.089\,i \end{bmatrix}, \begin{bmatrix} 1 \\ -0.838 - 0.130\,i \\ 0.060 - 0.736\,i \\ 0.686 + 0.219\,i \\ -0.147 + 0.610\,i \\ -0.539 - 0.089\,i \end{bmatrix}.$$

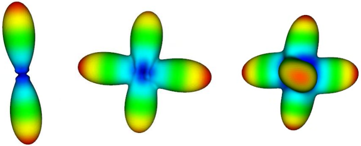
As the coordinates of the eigenvectors correspond to the evaluations of $\{1, x_1, x_2, x_1^2, x_1x_2, x_2^2\}$, we can recover the values of (x_1, x_2).

After solving the over-constrained linear system obtained by expansion and looking coefficient-wise, we deduce the decomposition:

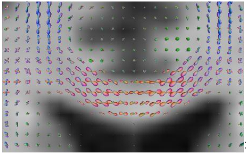
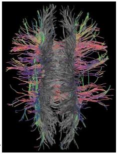
$$
\begin{aligned}
&(0.517+0.044\,i)\,(x_0-(0.830-1.593\,i)x_1-(0.326+0.050\,i)x_2)^4\\
&+(0.517-0.044\,i)\,(x_0-(0.830+1.593\,i)x_1-(0.326-0.050\,i)x_2)^4\\
&+2.958\,(x_0+(1.142)x_1+0.836x_2)^4\\
&+4.583\,(x_0+(0.956)x_1-0.713x_2)^4\\
&-(4.288+1.119\,i)\,(x_0-(0.838-0.130\,i)x_1+(0.060+0.736\,i)x_2)^4\\
&-(4.288-1.119\,i)\,(x_0-(0.838+0.130\,i)x_1+(0.060-0.736\,i)x_2)^4
\end{aligned}
$$

Recovering branching structures

- At points, perform directional measurements (of water diffusion):



- Approximate the set of measurements on the sphere by a (symmetric) tensor $T = (T_{i,j,k,l})$.
- Decompose/approximate it as a minimal sum of tensors of rank 1:
$$T = \sum_{i=1}^{r} \lambda_i \, v_i^4.$$
to identify the main directions of diffusion.
- From the decomposition of tensors, deduce the geometric structure:

 $\Rightarrow$ 

cf. [T. Schultz, H.P. Seidel'08], [A. Ghosh, R. Deriche, ...'09]

Concluding remarks/questions

- What about **nearest r-decomposition(s)** and SVD-like properties ?
- Applies to **multi-symmetric** tensors. Can we extend it to **anti-symmetric** tensors ?
- Can we extend it to minimal decomposition **over the real** ? to minimal **positive** decomposition ?

Collaborators: A. Bernardi, J. Brachat, P. Comon, M. Dien, E. Tsigaridas.

References:

J. Brachat, P. Comon, B. Mourrain, and E. Tsigaridas.
Symmetric tensor decomposition.
Linear Algebra and Applications, 433:1851–1872, 2010.

A. Bernardi, J. Brachat, P. Comon, and B. Mourrain.
General tensor decomposition, moment matrices and applications.
Journal of Symbolic Computation, (52):51–71, 2013.

A. Bernardi, J. Brachat, and B. Mourrain.
A comparison of different notions of ranks of symmetric tensors.
Preprint, 2012.

Thanks for your attention