Symmetric Tensor descomposition

María López
Directors: M.E.Alonso,A.Díaz-Cano
Universidad Complutense de Madrid

Madrid, 25th novemeber 2011

UNIVERSIDAD COMPLUTENSE
MADRID

- Objective
- Binary case
- Problem Formulations
  - Polynomial descomposition
  - Veronese and secant varieties and The Big Waring Problem
  - Descomposition using duality
- Hankel operators and quotient algebra
- Truncated Hankel operators
- Symmetric tensor descomposition algorithm

- $\mathbb{K}$ Algebraically closed field of Characteristic 0
- $S := \mathbb{K}[x_0, ..., x_n]$, $S_d := \mathbb{K}[x_0, ..., x_n]_d$
- $R := \mathbb{K}[x_1, ..., x_n]$, $R_d := \mathbb{K}[x_1, ..., x_n]_d$
- R^* has a natural structure of $R-$module: $\forall\ \Lambda \in R^*$:

$$\begin{array}{rccc} p * \Lambda : & R & \longrightarrow & \mathbb{K} \\ & q & \longmapsto & \Lambda(pq) \end{array}$$

- Typical elements of R^* are the linear forms, s.t. for all $p = \sum p_\beta \overline{x}^\beta \in R$ and for all $\xi \in \mathbb{K}^n$:
  - $$\begin{array}{rccc} ev(\xi) : & R & \longrightarrow & \mathbb{K} \\ & p & \longmapsto & p(\xi) = \sum p_\beta \xi^\beta \end{array}$$
  - $$\begin{array}{rccc} \overline{\delta}_\xi^\alpha : & R & \longrightarrow & \mathbb{K} \\ & p & \longmapsto & \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}(p)(\xi) \end{array}$$

Definition

A tensor $x_{i_1} \otimes \cdots \otimes x_{i_d} \in \mathbb{C}^n \otimes \cdots \otimes \mathbb{C}^n$ is said to be symmetric if for any permutation σ of $\{1, ..., k\}$:

$$x_{i_1} \otimes \cdots \otimes x_{i_k} = x_{i_{\sigma(1)}} \otimes \cdots \otimes x_{i_{\sigma(k)}}$$

We will be interested in the decomposition of a symmetric tensor A into a minimal linear combination of symmetric outer products of vectors (i.e. of the form $v \otimes \cdots \otimes v$) such that:

$$A = \sum_{i=1}^{r} \lambda_i v \otimes \cdots \otimes v$$

$$S^d(\mathbb{C}^n) \cong \mathbb{C}[x_o, ..., x_n]_d$$

$$\begin{array}{ccc} \mathbb{C}^n \otimes \cdots \otimes \mathbb{C}^n & \longrightarrow & S_d = \mathbb{C}[x_0, ..., x_n]_d \\ [a_{i_1\cdots i_n}]_{i_1=0\cdots i_d=0}^{n\ \cdots\ n} & \longmapsto & \sum_{i_1=0\cdots i_d=0}^{n\ \cdots\ n} a_{i_1\cdots i_d} x_{i_1}\cdots x_{i_d} \end{array}$$

Reformulation of the problem:
The decomposition of a homogeneous polynomial f of degree d in $n+1$ variables as a sum of d-th powers of linear forms, i.e.:

$$f(\overline{x}) = \sum_{i=1}^{r} \lambda_i (k_{i0}x_0 + ... + k_{in}x_n)^d$$

Definition

The minimal r is called the symmetric rank of f.

Theorem (Sylvester, 1886)

A binary form $f(x_1, x_2) = \sum_{i=o}^{d} \binom{d}{i} c_i x_1^i x_2^{d-i}$ *can be written as a sum of dth powers of r distinct linear forms in* $\mathbb{C}$ *as:*

$$f(x_1, x_2) = \sum_{j=1}^{r} \lambda_j (\alpha_j x_1 + \beta_j x_2)^d$$

if and only if :

- *there exist a vector* $\overline{q} = (q_l)_{l=0}^{r}$ *s.t.*

$$\begin{bmatrix} c_0 & c_1 & \dots & c_r \\ . & & & . \\ c_{d-r} & \dots & & c_d \end{bmatrix} \left[\overline{q}\right] = \left[\overline{0}\right]$$

- *the polynomial* $q(x_1, x_2) = \sum_{l=0}^{r} q_l x_1^l x_2^{r-l}$ *admits r distint roots, i.e. can be written as* $q(x_1, x_2) = \prod_{j=1}^{r} (\beta_j x_1 - \alpha_j x_2)$.

Input:A binary polynomial $p(x_1, x_2)$ of degree d with coefficients $a_i = \binom{d}{i} c_i$, s.t. $0 \leq i \leq d$
Output: A descomposition of $p(x_1, x_2) = \sum_{j=1}^{r} \lambda_j k(x)_j^d$ with minimal r

- 1. Initialize $r = 0$
- 2. Increment $r := r + 1$
- 3. If the matrix $H[r]$ has $ker(H[r]) = \overline{0}$ go to step 2
- 4. Else compute a basis $k_1, .., k_l$ of the $ker(H[r])$
- 5. Specialization:
  - Take any vector in the kernel, eg $\overline{k}$
  - Compute the roots of the associated polynomial $k(x_1, x_2) = \sum_{l=0}^{r} k_l x_1^l x_2^{d-l}$
  - If the roots are not distint in $\mathbb{P}_2$, try another specialization. If cannot be obtained , go to step 2
  - Else if $k(x_1, x_2)$ admits r distinct roots then compute coefficients λ_j $1 \leq j \leq r$

$$\begin{bmatrix} \alpha_1^d & \dots & \alpha_r^d \\ \alpha_1^{d-1} & \dots & \alpha_r^{d-1}\beta_r \\ \dots & \dots & \dots \\ \beta_1^r & \dots & \beta_r^d \end{bmatrix} \overline{\lambda} = \begin{bmatrix} a_0 \\ . \\ . \\ a_d \end{bmatrix}$$

- 6. The descomposition is $p(x_1, x_2) = \sum_{j=1}^{r} \lambda_j (\alpha_j x_1 + \beta_j x_2)^d$

- Problem Formulations
    - Polynomial descomposition
    - Veronese and secant varieties and The Big Waring Problem
    - Descomposition using duality

Equating the coefficients of the same monomials:

$$f(\overline{x}) = \sum_{i=1}^{r} \lambda_i (k_{i0}x_0 + \ldots + k_{in}x_n)^d = \lambda_1 k_1(\overline{x})^d + \ldots + \lambda_r k_r(\overline{x})^d$$

- It introduces $r!$ redundant solutions, since every permutation of the linear form is a solution.
- We get an over-constrained polynomial system, where the polynomials involved are of high degree, that is, d.

Big Waring Problem

Which is the minimun integer s s.t. the generic degree d homogeneous polynomial $F \in S_d$ *is the sum of at most* s $d-th$ *powers of linear forms* $L_1, ..., L_s$*?*

$$F = L_1^d + ... + L_s^d$$

Answered by J.Alexander, A-Hirschowitz, 1995

Definition

The image of the following map is the d-th Veronese variety, $X_{n,d}$ *:*

$$\begin{aligned} \nu_d : \mathbb{P}(S_1) &\longrightarrow \mathbb{P}(S_d) \\ k(\overline{x}) &\longmapsto k(\overline{x})^d \end{aligned}$$

- The polynomials of rank one are exactly those lying on $X_{n,d}$
- Variety of polinomials of rank $r > 1$?

Definition

The set that parameterizes polynomial homogeneus $f \in S_d$ of rank at most s is:

$$\sigma_s^0(X_{n,d}) := \bigcup_{[L_1^d],\ldots,[L_s^d] \in X_{n,d}} \langle [L_1^d], \ldots, [L_s^d] \rangle$$

but in general, $\sigma_s(X_{n,d})$ is not a variety

Definition

The $s-th$ secant variety of $X_{n,d} \subset \mathbb{P}(S_d)$, $\sigma_s(X_{n,d})$, is the Zariski closure of $\sigma_s^0(X_{n,d})$

- The integer s that solves the Big Waring Problem is the minimum integer s for which $\sigma_s(X_{n,d}) = \mathbb{P}(S_d)$

Definition

The minimum integer s for which $[F] \in \sigma_s(X_{n,d})$ is the symmetric border rank of F.

J.Alexander,A.Hirschowitz Theorem, 1995

$\sigma_s(X_{n,d})$ has always dimension $min(sn+s-1, \binom{n+d}{d}-1)$, except in the following cases:

- $d = 2,\ 2 \leq s \leq n$
- $n = 2, d = 4, s = 5$
- $n = 3, d = 4, s = 9$
- $n = 4, d = 4, s = 14$
- $n = 4, d = 3, s = 7$

The case of Veronese variety is the only one for which the defective case are completely classified.

Definition

Let $f, g \in S_d$ $f = \sum_{|\alpha|=d} f_\alpha x_0^{\alpha_0} \ldots x_n^{\alpha_n}$ *and* $g = \sum_{|\alpha|=d} g_\alpha x_0^{\alpha_0} \ldots x_n^{\alpha_n}$ *the apolar inner product on* S_d *is:*

$$\langle f, g \rangle = \sum_{|\alpha|=d} f_\alpha g_\alpha \binom{d}{\alpha_0, \ldots, \alpha_n}^{-1}$$

$$S_d \xrightarrow{\tau} S_d^* \xrightarrow{\pi} R_d^*$$
$$f \longmapsto f^* \longmapsto \Lambda_f$$
$$\text{such that } f^* : g \longmapsto \langle f, g \rangle$$

- $\left\langle f, k(\overline{x})^d \right\rangle = f(k)$
- $\tau(k(\overline{x})^d) = ev(k)$

Proposition

Let $f \in S_d$ and $k_1, ..., k_r \in \mathbb{C}^{n+1}$ such that $k_{i,0} = 1$ for all i. Then f can be written as:

$$f = \sum_{i=1}^{r} \lambda_i (x_0 k_{i,0} + ... + x_n k_{i,n})^d$$

if and only if $\Lambda_f \in R_d$ can be written as:

$$\Lambda_f = \sum_{i=1}^{r} \lambda_i ev(\underline{k}_i)$$

where $\underline{k}_i = (k_{i,1}, ..., k_{i,n})$.

The problem of decomposition can be restated: Let $\Lambda \in R_d^*$ find the minimal number of non-zero vectors $k_1, ..., k_r \in \mathbb{K}^n$ and non zero scalars $\lambda_1, ..., \lambda_r \in K$ such that $\Lambda = \sum_{i=1}^{r} \lambda_i ev(k_i)$.

Definition

For any $\Lambda \in R^*$ *we define the Hankel operator:*

$$H_\Lambda : R \longrightarrow R^*$$
$$p \longmapsto p * \Lambda$$
$$\mathbb{H}_\Lambda = (\Lambda(x^{\alpha+\beta}))_{\alpha,\beta} \alpha, \beta \in \mathbb{N}^n$$

Definition

Given $B = b_1, .., b_r$, $B' = b_1, .., b_r \subset R$ *we define:*

$$H_\Lambda^{B,B'} : \langle B \rangle \longrightarrow \langle B' \rangle^*$$

as the restriction of H_Λ *to the vector space* $\langle B \rangle$ *and the map goes from* R^* *to* $\langle B' \rangle^*$. *Let* $\mathbb{H}_\Lambda^{B,B'} = (\Lambda(b_i b_j'))$. *If* $B = B'$ *we also use* H_Λ^B *and* $\mathbb{H}_\Lambda^B$.

Properties of the Hankel operators

- $I_\Lambda := ker H_\Lambda$ is an ideal
- If $rank(\mathbb{H}_\Lambda) = r < \infty$
  - $A_\Lambda = R/I_\Lambda$ is an algebra of dimension r over $\mathbb{K}$.
  - A_Λ is a Gorenstein algebra (i.e. A_Λ^* is a free module of rank 1), such that $A_\Lambda^* = A_\Lambda * \Lambda$.
  - Let $I_\Lambda = Q_1 \cap \cdots \cap Q_d$ and let $A_i = ann(Q_i)$, then $A_\Lambda = A_1 \oplus \cdots \oplus A_d$ and $A_i^* = A_i * (e_i * \Lambda)$ where $1 = e_1 + \cdots + e_d$.

Definition

For any $\Lambda \in R^*$ s.t. $dim_{\mathbb{K}} A_\Lambda < \infty$ and $a \in A_\Lambda$ we define the operators of multiplication in A_Λ, M_a: $A_\Lambda \longrightarrow A_\Lambda$ such that $b \in A_\Lambda$ $M_a(b) = ab$. And its transposed M_a^*:$A_\Lambda^* \longrightarrow A_\Lambda^*$ such that $\forall \gamma \in A_\Lambda^*$ $M_a^*(\gamma) = a * \gamma$

Theorem

Let $Z(I_\Lambda)$=$\{\xi_1, ..., \xi_d\}$ the variety defined by the ideal I_Λ:

- *The eigenvalues of the operators M_a and M_a^* are given by $\{a(\xi_1), .., a(\xi_d)\}$*
- *The common eigenvectors of the operators $(M_{x_i}^*)\forall i$ are (up to scalar) $ev(\xi_i)$*

Inverse systems

- $R^* \approx \mathbb{K}[[\overline{\delta}_\xi]]$
- The map:
  $\perp: \quad \{I \subset R \text{ an ideal s.t. } Z(I) < \infty\} \quad \to \quad \{L\ R - module\ s.t.\ \dim_{\mathbb{K}} L < \infty\}$
  is bijective
- Let $I = Q_{\xi_1} \cap \cdots \cap Q_{\xi_d}$
  - $(I^\perp \cap \mathbb{K}[\overline{\delta}_\xi])^\perp = Q_\xi$.
  - $I^\perp = Q_{\xi_1}^\perp \oplus \cdots \oplus Q_{\xi_d}^\perp$ and for all $\Lambda \in I^\perp$:

$$\Lambda = \sum_{i=1}^{d} ev(\xi_i) \circ p_i(\overline{\partial})$$

Theorem

If $\mathrm{rank}\, H_\Lambda = r < \infty$ *,then:*

- A_Λ *is of dimension* r *over* $\mathbb{K}$ *and the set of roots* $Z(I_\Lambda) = \{\xi_1, .., \xi_d\}$ *is finite with* $d \leq r$*.*
- *There exists* $p_i \in \mathbb{K}[\partial_1, .., \partial_n]$ *such that* $\Lambda = \sum_{i=1}^{d} ev(\xi_i) \circ p_i(\overline{\alpha})$*.*
  Moreover, the multiplicity of ξ_i *is the dimension of the vector space generated by* $ev(\xi_i) \circ p_i(\overline{\alpha})$*.*

Theorem

Let $\Lambda \in R^*$ $\Lambda = \sum_{i=1}^{r} \lambda_i ev(\xi_i)$ *with* $\lambda_i \neq 0$ *and* ξ_i *distint points of* $\mathbb{K}^n$*, iff* $\mathrm{rank} H_\Lambda = r$ *and* I_Λ *is a radical ideal.*

Reformulation of the problem: Given $f^* \in R_d^*$ find the smallest r such that there exist $\Lambda \in R^*$ which extends f^* with rank H_Λ of rank r and I_Λ a radical ideal

Definition

Let $B \subset R_d$ be a set of monomials of degree at most d and $f^* \in R_d^*$. The Hankel matrix are:

$$\mathbf{H}_{\mathbf{\Lambda}}^{\mathbf{B}}(\overline{h}) = (h_{\alpha+\beta})_{\alpha,\beta} \alpha, \beta \in B$$

where $h_\alpha = f^*(\overline{x}^\alpha)$ if card(α) $\leq d$; otherwise h_α is a variable.

Definition

Suppose that $\mathbf{H}_{\mathbf{\Lambda}}^{\mathbf{B}}(\overline{h})$ is invertible in $\mathbb{K}(\overline{h})$.We define the multiplication operators:

$$\mathbf{M}_{\mathbf{i}}^{\mathbf{B}}(\overline{h}) := (\mathbf{H}_{\mathbf{\Lambda}}^{\mathbf{B}}(\overline{h}))^{-1} \mathbf{H}_{\mathbf{x_i * \Lambda}}^{\mathbf{B}}(\overline{h})$$

Theorem

Let $B \subset R$ be a set of monomials of degree at most d connected to $1 (m \in B \neq 1$ implies $m = x_i m'$ with $m' \in B)$ and let Λ be a linear form in $\langle B.B^+\rangle^*_d$. Let $\Lambda(\overline{h})$ be the linear form of $\langle B.B^+\rangle^*$ defined by $\Lambda(\overline{h})(\overline{x}_\alpha) = \Lambda(\overline{x}_\alpha)$ if α is at most d and $h_\alpha \in \mathbb{K}$ otherwise. Then $\Lambda(\overline{h})$ admits an extension $\Lambda_e \in R_*$ such that H_{Λ_e} is of rank r with B a basis of A_{Λ_e} iff:

- $\mathbf{M_i^B} \circ \mathbf{M_j^B}(\overline{\mathbf{h}}) - \mathbf{M_j^B} \circ \mathbf{M_i^B}(\overline{\mathbf{h}}) = \mathbf{0}$
- $det(\mathbf{H_\Lambda^B})(\overline{\mathbf{h}}) \neq 0.$

Moreover, such Λ_e is unique.

Theorem

Let $B = \{\overline{x}^{\beta_1}, ..., \overline{x}^{\beta_r}\}$ be a set of monomials of degree at most d, connected to 1, and let $\Lambda \in \langle B^+B^+\rangle^*_{\leq d}$ and $\Lambda(\overline{h}) \in \langle B^+B^+\rangle^*$ defined as follows:

$$\Lambda(\overline{h})(\overline{x}^{\gamma}) = \begin{cases} \Lambda(\overline{x}^{\gamma}) & \textit{if } |\gamma| \leq d; \\ h_{\gamma} & \textit{in other case.} \end{cases}$$

Then, Λ admits an extension $\widetilde{\Lambda} \in R^*$ such that $H_{\widetilde{\Lambda}}$ is of rank r, with B a basis of $A_{\widetilde{\Lambda}}$ if and only if there exists a solution $\overline{h}$ to the problem:

- *i) All $(r+1) \times (r+1)$ minors of $\mathbf{H_{\Lambda}}^{\mathbf{B}^+}(\overline{\mathbf{h}})$ vanish.*
- *ii)det$(\mathbf{H}^{\mathbf{B}}_{\mathbf{\Lambda}})(\overline{\mathbf{h}}) \neq 0$.*

Moreover, for every solution $\overline{h}_0 \in \mathbb{K}^N$ an extension such $\widetilde{\Lambda} = \Lambda(\overline{h_0})$ over $\langle B^+B^+\rangle$ is unique.

Input: A homogeneous polynomial $f(x_0, .., x_n)$ of degree d
Output: A decomposition of f as $f = \sum_{i=1}^{r} \lambda_i k_i(\overline{x})^d$ with r minimal

1. Compute the coefficients of $\underline{f^*}$: $c_\alpha = a_\alpha \binom{d}{\alpha}^{-1}$.
2. Initialize $r := 0$
3. Increment $r := r + 1$
4. Specialization:
   - Take any basis B connected to 1 with $|B| = r$
   - Build the matrix $H_{\underline{f^*}(\overline{h})}^{B^+}$ with the coefficients c_α.
   - If there exists any minor of order $r+1$ in $H_{\underline{f^*}(\overline{h})}^{B^+}$, without coefficients depending on $\overline{h}$, different to zero, try another specialization. If cannot be obtained go to step 3.
   - Else if all minors of order $r+1$ in $H_{\underline{f^*}(\overline{h})}^{B^+}$, without coefficients depending on $\overline{h}$, vanish, compute $\overline{h}$ s.t:
     - $\det(H_{\underline{f^*}(\overline{h})}^{B}) \neq 0$
     - the operators $M_i^B(\overline{h}) := (H_{\underline{f^*}(\overline{h})}^{B})^{-1}(H_{x_i * \underline{f^*}(\overline{h})})$ commute
     - the eigenvalues of $M_i^B(\overline{h})$ are simple

     If there not exist such $\overline{h}$ try another specialization. If cannot be obtained go to step 3.
   - Else if there exists such $\overline{h}$ compute the eigenvalues $\xi_{i,j}$ and the eigenvectors v_j s.t $M_i^B v_j = \xi_{i,j} v_j$ for $i = 1, \dots, n$ and $j = 1, \dots, r$.
5. Solve the linear system in (λ_j) s.t $f(\overline{x}) = \sum_{j=1}^{r} \lambda_j k_j(\overline{x})^d$ where $k_i(\overline{x}) = (x_0 + v_{i,1}x_1 + \dots + v_{i,n}x_n)$.

Let $f(x,y,z) =$
$3x^4+4x^3y-4x^3z+6x^2y^2-12x^2yz+18x^2z^2+4xy^3-12xy^2z+12xyz^2-4xz^3+y^4-4y^3z+6y^2z^2-4yz^3+3z^4$

- To $r=3$ and $B=\{1,y,z\}$.

$$H^{B^+}_{\underline{f*}} = \begin{pmatrix} 3 & 1 & -1 & 1 & 3 & -1 \\ 1 & 1 & -1 & 1 & 1 & -1 \\ -1 & -1 & 3 & -1 & -1 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1 \\ 3 & 1 & -1 & 1 & 3 & -1 \\ -1 & -1 & 1 & -1 & -1 & 1 \end{pmatrix}$$

$$M^B_y = (H^B_{\underline{f*}})^{-1} H_{y*\underline{f*}} = \begin{pmatrix} \frac{1}{2} & \frac{-1}{2} & 0 \\ \frac{-1}{2} & 2 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$M^B_z = (H^B_{\underline{f*}})^{-1} H_{z*\underline{f*}} = \begin{pmatrix} \frac{1}{2} & \frac{-1}{2} & 0 \\ \frac{-1}{2} & 2 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} -1 & -1 & 3 \\ -1 & -1 & 1 \\ 3 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

- The eigenvalues of $(M_z^B)^t$ are $x_1 = -1$, $x_2 = -1$ and $x_3 = 1$,
- The eigenvalues of $(M_y^B)^t$ are $x_1 = 0$, $x_2 = 0$ and $x_3 = 1$.
- The eigenvalues of M_p^B are $x_1 = 2$, $x_2 = -2$ and $x_3 = 0$, with $p = y + z$.

The eigenvectors of M_z^t which are:

$$\xi_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \xi_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \xi_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

- $f = \lambda_1(x+z)^4 + \lambda_2(x+y-z)^4 + \lambda_3(x-z)^4$

$\mathbf{f(x,y,z) = (x+z)^4 + (x+y-z)^4 + (x-z)^4}$

thanks!