Exact solutions in Structured Low-Rank Approximation

Giorgio Ottaviani, Pierre-Jean Spaenlehauer, Bernd Sturmfels

Inria Nancy – Grand Est/LORIA/CNRS, project CARAMEL

SLRA 2015, Grenoble, 06/01, 2015

Problem Statement

$p, q, r \in \mathbb{N}$
E a **linear/affine subspace** of $p \times q$ matrices with real entries
M a $p \times q$ matrix, $\Lambda = (\lambda_{i,j})$ a $p \times q$ positive matrix

$$\|M\|_\Lambda = \sqrt{\sum_{i,j} \lambda_{i,j} M_{i,j}^2}$$

Problem Statement

$p, q, r \in \mathbb{N}$
E a **linear/affine subspace** of $p \times q$ matrices with real entries
M a $p \times q$ matrix, $\Lambda = (\lambda_{i,j})$ a $p \times q$ positive matrix

$$\|M\|_\Lambda = \sqrt{\sum_{i,j} \lambda_{i,j} M_{i,j}^2}$$

Structured (and weighted) Low-Rank Approximation

Given $U \in E$, compute **a matrix** $M \in E$ such that

- $\mathrm{Rank}(M) \leq r$;
- $\|U - M\|_\Lambda$ is **minimum**.

Some applications in symbolic-numeric computations

- E = **Sylvester matrices** $\rightsquigarrow$ univariate approximate GCD

$$\begin{bmatrix} a_3 & 0 & b_2 & 0 & 0 \\ a_2 & a_3 & b_1 & b_2 & 0 \\ a_1 & a_2 & b_0 & b_1 & b_2 \\ a_0 & a_1 & 0 & b_0 & b_1 \\ 0 & a_0 & 0 & 0 & b_0 \end{bmatrix}$$

Some applications in symbolic-numeric computations

- $E =$ **Sylvester matrices** $\rightsquigarrow$ univariate approximate GCD
- $E =$ **Hankel matrices** $\rightsquigarrow$ denoising, signal processing, tensors

$$\begin{bmatrix} a & b & c & d & e \\ b & c & d & e & f \\ c & d & e & f & g \\ d & e & f & g & h \\ e & f & g & h & i \end{bmatrix}$$

Some applications in symbolic-numeric computations

- $E =$**Sylvester matrices** $\rightsquigarrow$ univariate approximate GCD
- $E =$**Hankel matrices** $\rightsquigarrow$ denoising, signal processing, tensors
- $E =$affine **coordinate spaces** $\rightsquigarrow$ matrix completion

$$\begin{bmatrix} 3 & ? & ? & 5 & 5 \\ 1 & 2 & 3 & 2 & ? \\ 10 & 4 & ? & 9 & -4 \\ 6 & ? & 3 & 9 & 10 \\ ? & 5 & -2 & ? & 9 \end{bmatrix}$$

Some applications in symbolic-numeric computations

- E =**Sylvester matrices** $\rightsquigarrow$ univariate approximate GCD
- E =**Hankel matrices** $\rightsquigarrow$ denoising, signal processing, tensors
- E =affine **coordinate spaces** $\rightsquigarrow$ matrix completion
- E =**Ruppert matrices** $\rightsquigarrow$ multivariate factorization

$$\begin{bmatrix} 0 & -2 & -a & 0 & -2b & -d \\ -1 & 0 & c & -b & 0 & e \\ a & 2c & 0 & d & 2e & 0 \\ 0 & 0 & 0 & 1 & a & c \\ 0 & 0 & 0 & -b & -d & -e \end{bmatrix}$$

$XY^2 + aXY + bY^2 + cX + dY + e \in \mathbb{C}[X, Y]$ factors $\Leftrightarrow \mathrm{rank} \leq 4$

Motivations

Several **approaches** to **SLRA**:

Structured Total Least Norm (*Park, Kaltofen, Zhi*), Alternating projections (*Cadzow, Condat, Hirabayashi*), Riemannian optimization (*Absil, Amodei, Meyer, Vandereycken*), Matrix Factorization(*Ishteva, Usevich, Markovsky*), Newton iteration (*Schost, S.*)...

Motivations

Several **approaches** to **SLRA**:

Structured Total Least Norm (*Park, Kaltofen, Zhi*), Alternating projections (*Cadzow, Condat, Hirabayashi*), Riemannian optimization (*Absil, Amodei, Meyer, Vandereycken*), Matrix Factorization(*Ishteva, Usevich, Markovsky*), Newton iteration (*Schost, S.*). . .

Global polynomial optimization (*Abril Bucero, Greuet, Lasserre, Mourrain, Nie, Parrilo, Safey, Schost, Sturmfels,. . .*)

Motivations

Several **approaches** to **SLRA**:

Structured Total Least Norm (*Park, Kaltofen, Zhi*), Alternating projections (*Cadzow, Condat, Hirabayashi*), Riemannian optimization (*Absil, Amodei, Meyer, Vandereycken*), Matrix Factorization(*Ishteva, Usevich, Markovsky*), Newton iteration (*Schost, S.*)...

Global polynomial optimization (*Abril Bucero, Greuet, Lasserre, Mourrain, Nie, Parrilo, Safey, Schost, Sturmfels,...*)

The EDdegree, algebraic degree of optimization of Euclidean distances on algebraic varieties:
Draisma/Horobet/Ottaviani/Sturmfels/Thomas'13

Motivations

Several **approaches** to **SLRA**:

Structured Total Least Norm (*Park, Kaltofen, Zhi*), Alternating projections (*Cadzow, Condat, Hirabayashi*), Riemannian optimization (*Absil, Amodei, Meyer, Vandereycken*), Matrix Factorization(*Ishteva, Usevich, Markovsky*), Newton iteration (*Schost, S.*)...

Global polynomial optimization (*Abril Bucero, Greuet, Lasserre, Mourrain, Nie, Parrilo, Safey, Schost, Sturmfels,...*)

The EDdegree, algebraic degree of optimization of Euclidean distances on algebraic varieties:
Draisma/Horobet/Ottaviani/Sturmfels/Thomas'13

Goals:

- **Certified and global SLRA** using symbolic (Gröbner bases) and symbolic-numeric algorithms (homotopy continuation methods) methods
- a priori estimates of the "algebraic difficulty" of the problem
  $\rightsquigarrow$ explicit formulas for **EDdegree** of SLRA
- **Applications**: **low-rank tensor approximation** from diffusion magnetic resonance imaging (*Schultz*), Hankel matrices, approximate GCD

A symbolic approach to SLRA

$p, q, r \in \mathbb{N}$, E a **linear/affine subspace** of $p \times q$ matrices with real entries

M a $p \times q$ matrix, Λ a $p \times q$ positive matrix, $\|M\|_\Lambda = \sqrt{\sum_{i,j} \lambda_{i,j} M_{i,j}^2}$

$\mathcal{D}_r$: variety of $p \times q$ matrices of rank at most r

The minimizers of SLRA are algebraic

Minimizing a polynomial function $M \mapsto \sum_{i,j} \lambda_{i,j}(U_{i,j} - M_{i,j})^2$ on an algebraic variety $\mathcal{D}_r \cap E$

$\rightsquigarrow$ SLRA can be modeled by polynomial system solving

Many possible **approaches**: Gröbner bases, border bases, homotopy methods, resultants, triangular sets, geometric resolution, ...

A symbolic approach to SLRA

$p, q, r \in \mathbb{N}$, E a **linear/affine subspace** of $p \times q$ matrices with real entries
M a $p \times q$ matrix, Λ a $p \times q$ positive matrix, $\|M\|_\Lambda = \sqrt{\sum_{i,j} \lambda_{i,j} M_{i,j}^2}$
$\mathcal{D}_r$: variety of $p \times q$ matrices of rank at most r

The minimizers of SLRA are algebraic

Minimizing a polynomial function $M \mapsto \sum_{i,j} \lambda_{i,j}(U_{i,j} - M_{i,j})^2$ on an algebraic variety $\mathcal{D}_r \cap E$
$\rightsquigarrow$ SLRA can be modeled by polynomial system solving

Many possible **approaches**: Gröbner bases, border bases, homotopy methods, resultants, triangular sets, geometric resolution, ...

First step: model the problem as a polynomial system
$\rightsquigarrow$ Ideal vanishing on the **regular critical points**

A symbolic approach to SLRA

$p, q, r \in \mathbb{N}$, E a **linear/affine subspace** of $p \times q$ matrices with real entries
M a $p \times q$ matrix, Λ a $p \times q$ positive matrix, $\|M\|_\Lambda = \sqrt{\sum_{i,j} \lambda_{i,j} M_{i,j}^2}$
$\mathcal{D}_r$: variety of $p \times q$ matrices of rank at most r

The minimizers of SLRA are algebraic

Minimizing a polynomial function $M \mapsto \sum_{i,j} \lambda_{i,j}(U_{i,j} - M_{i,j})^2$ on an algebraic variety $\mathcal{D}_r \cap E$
$\rightsquigarrow$ SLRA can be modeled by polynomial system solving

Many possible **approaches**: Gröbner bases, border bases, homotopy methods, resultants, triangular sets, geometric resolution, ...

First step: model the problem as a polynomial system
$\rightsquigarrow$ Ideal vanishing on the **regular critical points**

Technical assumptions for this talk:

- Finitely-many complex critical points on the smooth locus of $\mathcal{D}_r \cap E$.
- Minimum is reached on the smooth locus of $\mathcal{D}_r \cap E$.

Weighted low-rank approximation of the 4×4 determinant

$D(\mathbf{x}) \in \mathbb{Q}[x_{11}, \dots, x_{44}]$: determinant of the matrix (x_{ij})
U: 4×4 matrix picked at random
Λ: positive 4×4 matrix

$$D(\mathbf{x}) = 0$$

$$\mathrm{Rank} \begin{bmatrix} \partial D / \partial x_{11} & \dots & \partial D / \partial x_{44} \\ \lambda_{11}(x_{11} - u_{11}) & \dots & \lambda_{44}(x_{44} - u_{44}) \end{bmatrix} \leq 1$$

$$\mathrm{Rank}(x_{ij}) = 3$$

Weighted low-rank approximation of the 4×4 determinant

$D(\mathbf{x}) \in \mathbb{Q}[x_{11}, \ldots, x_{44}]$: determinant of the matrix (x_{ij})
U: 4×4 matrix picked at random
Λ: positive 4×4 matrix

$$D(\mathbf{x}) = 0$$

$$\begin{bmatrix} y & 1 \end{bmatrix} \begin{bmatrix} \partial D/\partial x_{11} & \ldots & \partial D/\partial x_{44} \\ \lambda_{11}(x_{11} - u_{11}) & \ldots & \lambda_{44}(x_{44} - u_{44}) \end{bmatrix} = \begin{bmatrix} 0 & \ldots & 0 \end{bmatrix}$$

variables: $x_{11}, \ldots, x_{44}, y$. 17 equations.

Weighted low-rank approximation of the 4×4 determinant

$D(\mathbf{x}) \in \mathbb{Q}[x_{11}, \ldots, x_{44}]$: determinant of the matrix (x_{ij})
U: 4×4 matrix picked at random
Λ: positive 4×4 matrix

$$D(\mathbf{x}) = 0$$

$$\begin{bmatrix} y & 1 \end{bmatrix} \begin{bmatrix} \partial D/\partial x_{11} & \ldots & \partial D/\partial x_{44} \\ \lambda_{11}(x_{11} - u_{11}) & \ldots & \lambda_{44}(x_{44} - u_{44}) \end{bmatrix} = \begin{bmatrix} 0 & \ldots & 0 \end{bmatrix}$$

variables: $x_{11}, \ldots, x_{44}, y$. 17 equations.

Timings with FGb (*Faugère*):

- Λ generic, over $\mathbb{Q}$: > 1 day
- $\Lambda = \mathbf{1}$, over $\mathbb{Q}$: 0.3s

Weighted low-rank approximation of the 4×4 determinant

$D(\mathbf{x}) \in \mathbb{Q}[x_{11}, \ldots, x_{44}]$: determinant of the matrix (x_{ij})
U: 4×4 matrix picked at random
Λ: positive 4×4 matrix

$$D(\mathbf{x}) = 0$$

$$\begin{bmatrix} y & 1 \end{bmatrix} \begin{bmatrix} \partial D/\partial x_{11} & \ldots & \partial D/\partial x_{44} \\ \lambda_{11}(x_{11} - u_{11}) & \ldots & \lambda_{44}(x_{44} - u_{44}) \end{bmatrix} = \begin{bmatrix} 0 & \ldots & 0 \end{bmatrix}$$

variables: $x_{11}, \ldots, x_{44}, y$. 17 equations.

Timings with FGb (*Faugère*):

- Λ generic, over $\mathbb{Q}$: > 1 day
- $\Lambda = \mathbf{1}$, over $\mathbb{Q}$: 0.3s

Can we **explain** these timings and/or find a **better modeling**?

The Euclidean Distance degree

The Euclidean distance degree
Draisma/Horobet/Ottaviani/Sturmfels/Thomas 13

$V \in \mathbb{C}^n$ an algebraic variety, $\mathbf{u} \in \mathbb{C}^n$ a generic point. The **EDdegree$_\Lambda$** of V is the number of **complex critical points** of the function

$$\lambda_1(x_1 - u_1)^2 + \cdots + \lambda_n(x_n - u_n)^2$$

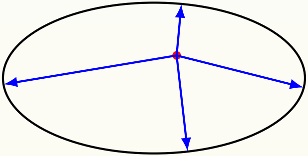
on the smooth locus of V.

The Euclidean distance degree
Draisma/Horobet/Ottaviani/Sturmfels/Thomas 13

$V \in \mathbb{C}^n$ an algebraic variety, $\mathbf{u} \in \mathbb{C}^n$ a generic point. The **EDdegree**$_\Lambda$ of V is the number of **complex critical points** of the function

$$\lambda_1(x_1 - u_1)^2 + \cdots + \lambda_n(x_n - u_n)^2$$

on the smooth locus of V.



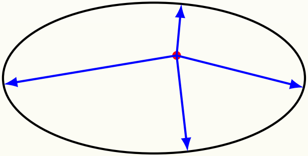
$\mathrm{EDdegree}(\text{ellipse}) = 4.$

The Euclidean distance degree
Draisma/Horobet/Ottaviani/Sturmfels/Thomas 13

$V \in \mathbb{C}^n$ an algebraic variety, $\mathbf{u} \in \mathbb{C}^n$ a generic point. The **EDdegree$_\Lambda$** of V is the number of **complex critical points** of the function

$$\lambda_1(x_1 - u_1)^2 + \cdots + \lambda_n(x_n - u_n)^2$$

on the smooth locus of V.

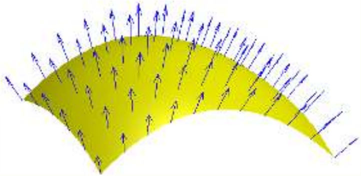


$\mathrm{EDdegree}(\text{ellipse}) = 4.$

Solution of SLRA:

critical point of the distance function on a **linear section of a determinantal variety** $\mathscr{D}_r \cap E$.

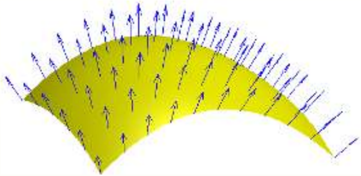
The conormal variety



Let $X \subset \mathbb{C}^n$ be an affine cone (the vanishing locus of homogeneous polynomials). The *conormal variety* $\mathcal{N}_X \subset \mathbb{C}^n \times \mathbb{C}^n$ is defined as

$$\mathcal{N}_X = \overline{\{(\mathbf{x}, \mathbf{v}) : \mathbf{x} \in X_{\mathrm{smooth}}, \mathbf{v} \in N_{\mathbf{x}} X\}}.$$

The conormal variety



Let $X \subset \mathbb{C}^n$ be an affine cone (the vanishing locus of homogeneous polynomials). The *conormal variety* $\mathcal{N}_X \subset \mathbb{C}^n \times \mathbb{C}^n$ is defined as

$$\mathcal{N}_X = \overline{\{(\mathbf{x}, \mathbf{v}) : \mathbf{x} \in X_{\mathrm{smooth}}, \mathbf{v} \in N_{\mathbf{x}} X\}}.$$

$$\mathbf{x} \in X_{\mathrm{smooth}} \text{ critical point of } \sum \lambda_i (x_i - u_i)^2$$

$$\Updownarrow$$

$$\nabla \sum \lambda_i (x_i - u_i)^2 \in N_x X$$

$$\Updownarrow$$

$$\begin{bmatrix} 2\lambda_1 x_1 - v_1 \\ \vdots \\ 2\lambda_n x_n - v_n \end{bmatrix} = \begin{bmatrix} 2\lambda_1 u_1 \\ \vdots \\ 2\lambda_n u_n \end{bmatrix} \text{ for } (\mathbf{x}, \mathbf{v}) \in \mathcal{N}_X$$

Duality

Proposition (Draisma/Horobet/Ottaviani/Sturmfels/Thomas)

The **EDdegree** of a projective variety is bounded by the **sum of the degrees of its polar classes**. Equality holds when the diagonal of the **conormal variety** is empty.

Duality:

$$\mathcal{N}_X = \overline{\{(\mathbf{x}, \mathbf{v}) : \mathbf{x} \in X_{\text{smooth}}, \mathbf{v} \in N_{\mathbf{x}} X\}}.$$

Duality

Proposition (Draisma/Horobet/Ottaviani/Sturmfels/Thomas)

The **EDdegree** of a projective variety is bounded by the **sum of the degrees of its polar classes**. Equality holds when the diagonal of the **conormal variety** is empty.

Duality:

$$\mathcal{N}_X = \overline{\{(\mathbf{x}, \mathbf{v}) : \mathbf{x} \in X_{\mathrm{smooth}}, \mathbf{v} \in N_{\mathbf{x}} X\}}.$$

$$\begin{array}{rccc} \pi_2 : & \mathcal{N}_X & \to & \mathbb{C}^n \\ & (\mathbf{x}, \mathbf{v}) & \mapsto & \mathbf{v} \end{array} \qquad\qquad X^* = \mathrm{Im}(\pi_2)$$

Rank r matrices are dual to corank r matrices.

Duality

Proposition (Draisma/Horobet/Ottaviani/Sturmfels/Thomas)

The **EDdegree** of a projective variety is bounded by the **sum of the degrees of its polar classes**. Equality holds when the diagonal of the **conormal variety** is empty.

Duality:

$$\mathcal{N}_X = \overline{\{(\mathbf{x}, \mathbf{v}) : \mathbf{x} \in X_{\mathrm{smooth}}, \mathbf{v} \in N_{\mathbf{x}}X\}}.$$

$$\begin{array}{rccc} \pi_2 : & \mathcal{N}_X & \to & \mathbb{C}^n \\ & (\mathbf{x}, \mathbf{v}) & \mapsto & \mathbf{v} \end{array} \qquad X^* = \mathrm{Im}(\pi_2)$$

Rank-deficient matrices are dual to rank 1 matrices
$\rightsquigarrow$ **Segre varieties.**

Back to the 4×4 determinant: duality

$D(\mathbf{x}) \in \mathbb{Q}[x_{11}, \ldots, x_{44}]$: determinant of the matrix (x_{ij})
U: 4×4 matrix picked at random
Λ: positive 4×4 matrix
Projective dual to $\{D(\mathbf{x}) = 0\}$: rank 1 matrices

$$\begin{array}{rccc}
\varphi : & \mathbb{C}^3 \times \mathbb{C}^4 & \to & \mathbb{C}^{16} \\
& \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ 1 \end{bmatrix}, \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \end{bmatrix} & \mapsto & \begin{bmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 & a_1 b_4 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 & a_2 b_4 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 & a_3 b_4 \\ b_1 & b_2 & b_3 & b_4 \end{bmatrix}
\end{array}$$

Back to the 4×4 determinant: duality

$D(\mathbf{x}) \in \mathbb{Q}[x_{11}, \ldots, x_{44}]$: determinant of the matrix (x_{ij})
U: 4×4 matrix picked at random
Λ: positive 4×4 matrix
Projective dual to $\{D(\mathbf{x}) = 0\}$: rank 1 matrices

$$\begin{array}{cccc} \varphi : & \mathbb{C}^3 \times \mathbb{C}^4 & \to & \mathbb{C}^{16} \\ & \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ 1 \end{bmatrix}, \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \end{bmatrix} & \mapsto & \begin{bmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 & a_1 b_4 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 & a_2 b_4 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 & a_3 b_4 \\ b_1 & b_2 & b_3 & b_4 \end{bmatrix} \end{array}$$

Dual optimization problem:

$$\nabla \|\varphi(a_1, a_2, a_3, b_1, b_2, b_3, b_4) - U'\|^2_{\Lambda'} = 0$$
$$\lambda'_{ij} = 1/\lambda_{ij} \qquad u'_{ij} = \lambda_{ij} u_{ij}$$

Primal vs. dual

Timings with FGb (primal/dual):

- Λ generic, over $\mathbb{Q}$: >1day/**891s**
- $\Lambda = \mathbf{1}$, over $\mathbb{Q}$: 0.3s/0.2s

Explanation of the gap between timings:

$$EDdegree_{\mathbf{1}} = 4 \qquad EDdegree_{gen} = 284.$$

+ general polynomial modeling for SLRA.

Primal vs. dual

Timings with FGb (primal/dual):

- Λ generic, over $\mathbb{Q}$: $>$1day/**891s**
- $\Lambda = \mathbf{1}$, over $\mathbb{Q}$: 0.3s/0.2s

Explanation of the gap between timings:

$$EDdegree_{\mathbf{1}} = 4 \qquad EDdegree_{gen} = 284.$$

+ general polynomial modeling for SLRA.

Strong correlation between timings and EDdegree of the problem.
A priori estimates of the EDdegree?

Generic E, Generic weights, corank 1

critical points of $\lambda_{1,1}(x_{1,1} - u_{1,1})^2 + \cdots + \lambda_{p,q}(x_{p,q} - u_{p,q})^2$ on $(\mathcal{D}_r \cap E)_{smooth}$.

Generic E, Generic weights, corank 1

critical points of $\lambda_{1,1}(x_{1,1}-u_{1,1})^2+\cdots+\lambda_{p,q}(x_{p,q}-u_{p,q})^2$ on $(\mathcal{D}_r\cap E)_{smooth}$.

Proposition

Let E be a **generic codimension** s linear space of $p\times q$ matrices, and $\mathcal{D}_r$ be the variety of **rank-deficient matrices**. The **generic EDdegree** of $\mathcal{D}_r\cap E$ equals

$$\delta_0+\cdots+\delta_{pq-2-s}.$$

where

$$\begin{aligned}\delta_\ell &= \sum_{k=\ell}^{p+q-2}(-1)^{p+q-k}\binom{k+1}{\ell+1}v_k\\ v_k &= \left[s^{p-1}t^{q-1}\right](1+s)^p(1+t)^q(t+s)^k.\end{aligned}$$

Generic E, Generic weights, corank 1

critical points of $\lambda_{1,1}(x_{1,1}-u_{1,1})^2+\cdots+\lambda_{p,q}(x_{p,q}-u_{p,q})^2$ on $(\mathcal{D}_r\cap E)_{smooth}$.

Proposition

Let E be a **generic codimension** s linear space of $p\times q$ matrices, and $\mathcal{D}_r$ be the variety of **rank-deficient matrices**. The **generic EDdegree** of $\mathcal{D}_r\cap E$ equals

$$\delta_0+\cdots+\delta_{pq-2-s}.$$

where

$$\begin{aligned}\delta_\ell &= \sum_{k=\ell}^{p+q-2}(-1)^{p+q-k}\binom{k+1}{\ell+1}v_k\\ v_k &= \left[s^{p-1}t^{q-1}\right](1+s)^p(1+t)^q(t+s)^k.\end{aligned}$$

+ similar statement for rank 1 matrices.

Generic E, Generic weights, corank 1

critical points of $\lambda_{1,1}(x_{1,1} - u_{1,1})^2 + \cdots + \lambda_{p,q}(x_{p,q} - u_{p,q})^2$ on $(\mathcal{D}_r \cap E)_{smooth}$.

Proposition

Let E be a **generic codimension** s linear space of $p \times q$ matrices, and $\mathcal{D}_r$ be the variety of **rank-deficient matrices**. The **generic EDdegree** of $\mathcal{D}_r \cap E$ equals

$$\delta_0 + \cdots + \delta_{pq-2-s}.$$

where

$$\begin{aligned} \delta_\ell &= \sum_{k=\ell}^{p+q-2} (-1)^{p+q-k} \binom{k+1}{\ell+1} v_k \\ v_k &= \left[s^{p-1} t^{q-1}\right] (1+s)^p (1+t)^q (t+s)^k. \end{aligned}$$

+ similar statement for rank 1 matrices.

Intermediate coranks $\rightsquigarrow$ Schubert calculus

E generic, $\Lambda = \mathbf{1}$?

EDdegree dramatically decreases.
Role of the **isotropic quadric**, the self-dual hypersurface $\sum x_{ij}^2 = 0$:

$$EDdegree_{\mathbf{1}}(V) = EDdegree_{gen}(V)$$

$$\Updownarrow$$

V intersects tranversely the isotropic quadric.

E generic, $\Lambda = \mathbf{1}$?

EDdegree dramatically decreases.
Role of the **isotropic quadric**, the self-dual hypersurface $\sum x_{ij}^2 = 0$:

$$EDdegree_{\mathbf{1}}(V) = EDdegree_{gen}(V)$$

$$\Updownarrow$$

V intersects tranversely the isotropic quadric.

Conjecture

Let $r = \min(p,q) - 1$ and Z be the locus of non-tranverse intersection between $\mathcal{D}_r \cap E$ and the isotropic quadric.

$$EDdegree_{\mathbf{1}}(\mathcal{D}_r \cap E) = EDdegree_{\mathbf{gen}}(\mathcal{D}_r \cap E) - EDdegree_{\mathbf{gen}}(Z).$$

E generic, $\Lambda = \mathbf{1}$?

EDdegree dramatically decreases.
Role of the **isotropic quadric**, the self-dual hypersurface $\sum x_{ij}^2 = 0$:

$$EDdegree_{\mathbf{1}}(V) = EDdegree_{gen}(V)$$

$$\Updownarrow$$

V intersects tranversely the isotropic quadric.

Conjecture

Let $r = \min(p, q) - 1$ and Z be the locus of non-tranverse intersection between $\mathcal{D}_r \cap E$ and the isotropic quadric.

$$EDdegree_{\mathbf{1}}(\mathcal{D}_r \cap E) = EDdegree_{\mathbf{gen}}(\mathcal{D}_r \cap E) - EDdegree_{\mathbf{gen}}(Z).$$

$+$ explicit formula for $EDdegree_{\mathbf{gen}}(Z)$. Tested on many examples.

Special linear space: k-th Sylvester matrices

$$\begin{aligned} f(X) &= f_m X^m + \cdots + f_1 X + f_0 \\ g(X) &= g_n X^n + \cdots + g_1 X + f_0 \\ \|(f,g)\|^2 &= \alpha_m f_m^2 + \cdots + \alpha_0 f_0^2 + \beta_n g_n^2 + \cdots + \beta_0 g_0^2 \end{aligned}$$

Approximate GCD problem:
find nearest pair (f^*, g^*) such that $\deg(\mathrm{GCD}(f^*, g^*)) \geq k$.

Special linear space: k-th Sylvester matrices

$$
\begin{aligned}
f(X) &= f_m X^m + \cdots + f_1 X + f_0 \\
g(X) &= g_n X^n + \cdots + g_1 X + f_0 \\
\|(f,g)\|^2 &= \alpha_m f_m^2 + \cdots + \alpha_0 f_0^2 + \beta_n g_n^2 + \cdots + \beta_0 g_0^2
\end{aligned}
$$

Approximate GCD problem:

find nearest pair (f^*, g^*) such that $\deg(\mathrm{GCD}(f^*, g^*)) \geq k$.

$$
\left.\underbrace{\left[\begin{array}{ccccc} f_m & 0 & \dots & 0 & 0 \\ f_{m-1} & f_m & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \dots & \vdots \\ 0 & 0 & \dots & f_0 & f_1 \\ 0 & 0 & \dots & 0 & f_0 \end{array}\right.}_{n-k+1}
\underbrace{\left.\begin{array}{ccccc} g_n & 0 & \dots & 0 & 0 \\ g_{n-1} & g_n & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \dots & \vdots \\ 0 & 0 & \dots & g_0 & g_1 \\ 0 & 0 & \dots & 0 & g_0 \end{array}\right]}_{m-k+1}\right\} m+n-2k+2
$$

Special linear space: k-th Sylvester matrices

$$\begin{aligned} f(X) &= f_m X^m + \cdots + f_1 X + f_0 \\ g(X) &= g_n X^n + \cdots + g_1 X + f_0 \\ \|(f,g)\|^2 &= \alpha_m f_m^2 + \cdots + \alpha_0 f_0^2 + \beta_n g_n^2 + \cdots + \beta_0 g_0^2 \end{aligned}$$

Approximate GCD problem:
find nearest pair (f^*, g^*) such that $\deg(\mathrm{GCD}(f^*, g^*)) \geq k$.

$$\left.\underbrace{\begin{bmatrix} f_m & 0 & \dots & 0 & 0 \\ f_{m-1} & f_m & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \dots & \vdots \\ 0 & 0 & \dots & f_0 & f_1 \\ 0 & 0 & \dots & 0 & f_0 \end{bmatrix}}_{n-k+1}\underbrace{\begin{bmatrix} g_n & 0 & \dots & 0 & 0 \\ g_{n-1} & g_n & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \dots & \vdots \\ 0 & 0 & \dots & g_0 & g_1 \\ 0 & 0 & \dots & 0 & g_0 \end{bmatrix}}_{m-k+1}\right\} m+n-2k+2$$

Rank deficient if and only if $\deg(\mathrm{GCD}(f,g)) \geq k$
$\rightsquigarrow$ SLRA problem

Special linear subspaces: resultant and approximate GCD

$$\begin{aligned} f(X) &= f_m X^m + \cdots + f_1 X + f_0 \\ g(X) &= g_n X^n + \cdots + g_1 X + f_0 \\ \|(f,g)\|^2 &= \alpha_m f_m^2 + \cdots + \alpha_0 f_0^2 + \beta_n g_n^2 + \cdots + \beta_0 g_0^2 \end{aligned}$$

$V_k \subset \mathbb{P}^{m+n+1}$:

pairs of pols sharing a GCD of degree at least k.

$$\begin{aligned} f(X) &= f_m X^m + \cdots + f_1 X + f_0 \\ g(X) &= g_n X^n + \cdots + g_1 X + f_0 \\ \|(f,g)\|^2 &= \alpha_m f_m^2 + \cdots + \alpha_0 f_0^2 + \beta_n g_n^2 + \cdots + \beta_0 g_0^2 \end{aligned}$$

$V_k \subset \mathbb{P}^{m+n+1}$:

pairs of pols sharing a GCD of degree at least k.

Theorem

The **generic EDdegree** of V_k equals that of the Segre variety of $(k+1) \times (n+m-2k+2)$ matrices of rank 1.

Special linear subspaces: resultant and approximate GCD

$$\begin{array}{rcl} f(X) & = & f_m X^m + \cdots + f_1 X + f_0 \\ g(X) & = & g_n X^n + \cdots + g_1 X + f_0 \\ \|(f,g)\|^2 & = & \alpha_m f_m^2 + \cdots + \alpha_0 f_0^2 + \beta_n g_n^2 + \cdots + \beta_0 g_0^2 \end{array}$$

$V_k \subset \mathbb{P}^{m+n+1}$:

pairs of pols sharing a GCD of degree at least k.

Theorem

The **generic EDdegree** of V_k equals that of the Segre variety of $(k+1) \times (n+m-2k+2)$ matrices of rank 1.

Sketch of proof:

$$\begin{array}{rcl} \mathbb{P}^k \times \mathbb{P}^{n+m-2k+1} & \to & V_k \\ (A(x), (B(x), C(x)) & \mapsto & (A(x)B(x), A(x)C(x)) \end{array}$$

is a desingularization and it factors through the Segre embedding of $\mathbb{P}^k \times \mathbb{P}^{n+m-2k+1} \rightsquigarrow$ EDdegree of rank 1 matrices.

Open questions

What about the **number** of real critical points/local minima?

Open questions

What about the **number** of real critical points/local minima?

Open even for **unstructured** weighted low-rank approximation!

Question (*Rey'13*): is the number of local minima of rank 1 (resp. corank 1) approximation bounded by $\min(p, q)$?

Open questions

What about the **number** of real critical points/local minima?

Open even for **unstructured** weighted low-rank approximation!

Question (*Rey'13*): is the number of local minima of rank 1 (resp. corank 1) approximation bounded by $\min(p, q)$?

Ottaviani/S./Sturmfels'13: **negative answer**

$$U = \begin{bmatrix} -59 & 11 & 59 \\ 11 & 59 & -59 \\ 59 & -59 & 11 \end{bmatrix} \quad \Lambda = \begin{bmatrix} 9 & 6 & 1 \\ 6 & 1 & 9 \\ 1 & 9 & 6 \end{bmatrix}$$

Rank 1 approximation of U has 7 local minima. $\mathrm{EDdegree} = 39$, number of real critical points: 19.
Can we find **more real critical points/local minima**?

Conclusion

Algebraic geometry techniques: analysis of singularities, characteristic class computations

$\Updownarrow$

Computational aspects, complexity of SLRA.

Conclusion

Algebraic geometry techniques: analysis of singularities, characteristic class computations

$$\Updownarrow$$

Computational aspects, complexity of SLRA.

Also in the paper: Hankel matrices, LRA of some tensors.

Conclusion

Algebraic geometry techniques: analysis of singularities, characteristic class computations

$$\Updownarrow$$

Computational aspects, complexity of SLRA.

Also in the paper: Hankel matrices, LRA of some tensors.

SLRA: hard problem with (relatively) low algebraic degree.

Conclusion

Algebraic geometry techniques: analysis of singularities, characteristic class computations

$\Updownarrow$

Computational aspects, complexity of SLRA.

Also in the paper: Hankel matrices, LRA of some tensors.

SLRA: hard problem with (relatively) low algebraic degree.

Questions:

- conjecture for the formula of the EDdegree of SLRA for non-generic weights?
- algos: exploiting duality for SLRA?
- real critical points?

Conclusion

Algebraic geometry techniques: analysis of singularities, characteristic class computations

$\Updownarrow$

Computational aspects, complexity of SLRA.

Also in the paper: Hankel matrices, LRA of some tensors.

SLRA: hard problem with (relatively) low algebraic degree.

Questions:

- conjecture for the formula of the EDdegree of SLRA for non-generic weights?
- algos: exploiting duality for SLRA?
- real critical points?

Thank you!