Homotopy techniques for determinantal systems

Éric Schost

University of Waterloo

joint work with

Jonathan Hauenstein, Mohab Safey El Din, **Xuan Thi Vu**

An example

Goal: minimize X_1 on $S : X_1^{100} + X_2^{100} + X_3^{100} = 1$.

The minima satisfy

$$X_1^{100} + X_2^{100} + X_3^{100} = 1$$

and

$$\operatorname{rank} \begin{bmatrix} 100X_1^{99} & 100X_2^{99} & 100X_3^{99} \\ 1 & 0 & 0 \end{bmatrix} < 2$$

Optimization, real algebraic geometry (polar varieties), ...

Our problem

$\mathbb{K}$ is a field of characteristic zero.

Given

- a matrix $\mathsf{F} \in \mathbb{K}[X_1, \ldots, X_n]^{p \times q}$
- polynomials $G = (g_1, \ldots, g_s)$ in $\mathbb{K}[X_1, \ldots, X_n]$,

such that

$$p \leqslant q \quad \text{and} \quad n = q - p + s + 1,$$

compute the set

$$\{\mathbf{x} \in \overline{\mathbb{K}}^n \mid G(\mathbf{x}) = 0, \ \operatorname{rank}(\mathsf{F}(\mathbf{x})) < p\}$$

Our problem

$\mathbb{K}$ is a field of characteristic zero.

Given

- a matrix $\mathsf{F} \in \mathbb{K}[X_1, \ldots, X_n]^{p \times q}$
- polynomials $G = (g_1, \ldots, g_s)$ in $\mathbb{K}[X_1, \ldots, X_n]$,

such that

$$p \leqslant q \quad \text{and} \quad n = q - p + s + 1,$$

compute the **isolated points of** the set

$$\{\mathbf{x} \in \overline{\mathbb{K}}^n \mid G(\mathbf{x}) = 0,\ \operatorname{rank}(\mathsf{F}(\mathbf{x})) < p\}$$

Why take $n = q - p + s + 1$?

This is because of known syzygies between minors.

Suppose $s = 0$, so $n = q - p + 1$.

Example (generic matrix)

$$\mathsf{F} = \begin{bmatrix} f_{1,1} & f_{1,2} & f_{1,3} \\ f_{2,1} & f_{2,2} & f_{2,3} \end{bmatrix}, \quad p = 2, \quad q = 3$$

One syzygy:

$$f_{1,1}(f_{1,2}f_{2,3} - f_{1,3}f_{2,2}) - f_{1,2}(f_{1,1}f_{2,3} - f_{1,3}f_{2,1}) + f_{1,3}(f_{1,1}f_{2,2} - f_{1,2}f_{2,1}) = 0$$

so if $f_{1,1} \neq 0$, only 2 equations (with no further relations)

In general: localize with top-left $(p-1)$-minor, $q-(p-1)$ equations

Why take $n = q - p + s + 1$?

This is because of known syzygies between minors.

Fact [Macaulay, Eagon-Northcott]

All irreducible components of the algebraic set

$$\{\mathbf{x} \in \overline{\mathbb{K}}^n \mid \operatorname{rank}(\mathsf{F}(\mathbf{x})) < p\}$$

have dimension at least $n - (q - p + 1)$

Why take $n = q - p + s + 1$?

This is because of known syzygies between minors.

Fact [Macaulay, Eagon-Northcott]

All irreducible components of the algebraic set

$$\{\mathbf{x} \in \overline{\mathbb{K}}^n \mid G(\mathbf{x}) = 0,\ \mathrm{rank}(\mathsf{F}(\mathbf{x})) < p\}$$

have dimension at least $n - (q - p + s + 1)$

Main result – column degrees

Suppose that:

- $n = q - p + s + 1$
- the degrees of $G = (g_1, \ldots, g_s)$ are at most $\gamma_1, \ldots, \gamma_s$,
- the column-degrees of F are at most $\alpha_1, \ldots, \alpha_q$
- all polynomials are given by a SLP of size L

Theorem

- there are at most

$$c = \gamma_1 \cdots \gamma_s E_{n-s}(\alpha_1, \ldots, \alpha_q)$$

  isolated solutions, counted with multiplicities
  (E_k = k-th elementary symmetric function)
- can be computed in time $(cL)^{O(1)}$ (randomized algorithm)

Main result – column degrees

Suppose that:

- $n = q - p + s + 1$
- the degrees of $G = (g_1, \ldots, g_s)$ are at most $\gamma_1, \ldots, \gamma_s$,
- the column-degrees of F are at most $\alpha_1, \ldots, \alpha_q$
- all polynomials are given by a SLP of size L

Example

$$\mathsf{F} = \begin{bmatrix} [2] & [1] & [5] & [7] \\ [2] & [1] & [5] & [7] \\ [2] & [1] & [5] & [7] \end{bmatrix} \quad p = 3, q = 4, s = 0 \implies n = 2$$

$$c = E_2(2, 1, 5, 7) = 2 \cdot 1 + 2 \cdot 5 + 2 \cdot 7 + 1 \cdot 5 + 1 \cdot 7 + 5 \cdot 7 = 73$$

Main result – row degrees

Suppose that:

- $n = q - p + s + 1$
- the degrees of $G = (g_1, \ldots, g_s)$ are at most $\gamma_1, \ldots, \gamma_s$,
- the row-degrees of F are at most $\beta_1, \ldots, \beta_p$
- all polynomials are given by a SLP of size L

Theorem

- there are at most

$$c' = \gamma_1 \cdots \gamma_s S_{n-s}(\beta_1, \ldots, \beta_p)$$

  isolated solutions, counted with multiplicities
  (S_k = k-th complete symmetric function)
- can be computed in time $(c'L)^{O(1)}$ (randomized algorithm)

Main result – row degrees

Suppose that:

- $n = q - p + s + 1$
- the degrees of $G = (g_1, \ldots, g_s)$ are at most $\gamma_1, \ldots, \gamma_s$,
- the row-degrees of F are at most $\beta_1, \ldots, \beta_p$
- all polynomials are given by a SLP of size L

Example

$$\mathsf{F} = \begin{bmatrix} [2] & [2] & [2] & [2] \\ [1] & [1] & [1] & [1] \\ [5] & [5] & [5] & [5] \end{bmatrix} \quad p = 3, q = 4, s = 0 \implies n = 2$$

$$c' = S_2(2, 1, 5) = 2^2 + 2 \cdot 1 + 2 \cdot 5 + 1^2 + 1 \cdot 5 + 5^2 = 47$$

Previous work

[Giambelli] (see also [Miller-Sturmfels])

- Hilbert function of determinantal ideals of generic matrices
- more refined type of degrees

[Nie-Ranestad]

- degree bounds $E_{n-s}(\cdots)$ and $S_{n-s}(\cdots)$ sharp for generic polynomials (we reuse their construction for column degrees)

[Faugère-Safey El Din-Spaenlehauer], [Spaenlehauer]

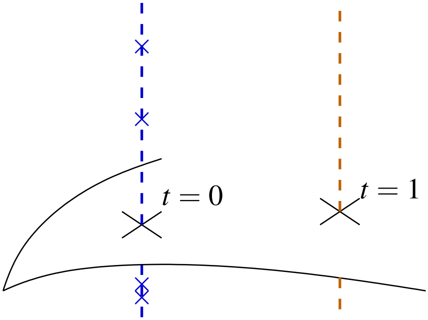
- complexity of Gröbner basis for generic polynomials

[Huber-Sottile-Sturmfels], [Verschelde], ...

- Schubert calculus
- minors of generic matrices

Symbolic homotopies

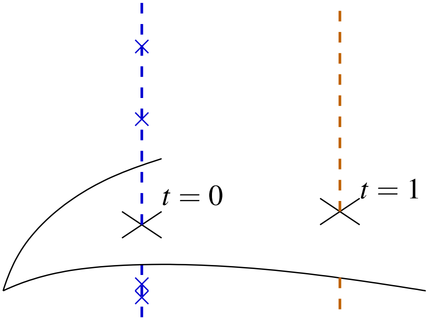
- want to solve $(f_i(\mathbf{x}))_{i \leqslant m}$, know the solutions of $(b_i(\mathbf{x}))_{i \leqslant m}$



$r(T) = 0, X_1 = v_1(T), \ldots, X_n = v_n(T)$ in $\mathbb{K}[T]$

Symbolic homotopies

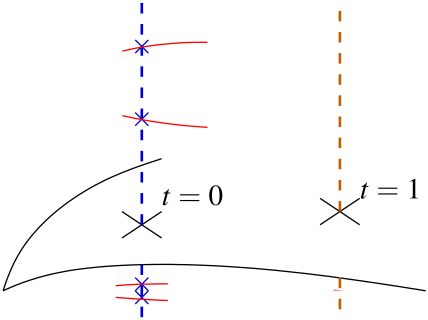
- want to solve $(f_i(\mathbf{x}))_{i \leqslant m}$, know the solutions of $(b_i(\mathbf{x}))_{i \leqslant m}$
- find $\mathbf{h}(t, \mathbf{X})$ such that $\mathbf{h}(1, \mathbf{X}) = \mathbf{f}$ and $\mathbf{h}(0, \mathbf{X}) = \mathbf{b}$



$r(T) = 0, X_1 = v_1(T), \ldots, X_n = v_n(T)$ in $\mathbb{K}[T]$

Symbolic homotopies

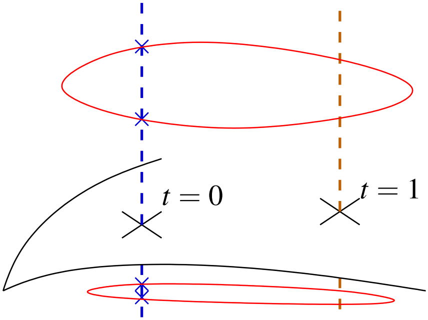
- want to solve $(f_i(\mathbf{x}))_{i \leqslant m}$, know the solutions of $(b_i(\mathbf{x}))_{i \leqslant m}$
- find $\mathbf{h}(t, \mathbf{X})$ such that $\mathbf{h}(1, \mathbf{X}) = \mathbf{f}$ and $\mathbf{h}(0, \mathbf{X}) = \mathbf{b}$
- compute a description of the solution curve



$R(t, T) = 0, X_1 = V_1(t, T), \ldots, X_n = V_n(t, T)$ in $\mathbb{K}[[t]][T]$

Symbolic homotopies

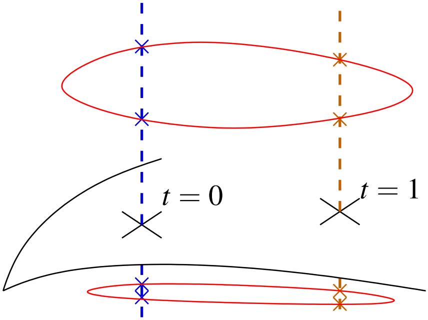
- want to solve $(f_i(\mathbf{x}))_{i \leqslant m}$, know the solutions of $(b_i(\mathbf{x}))_{i \leqslant m}$
- find $\mathbf{h}(t, \mathbf{X})$ such that $\mathbf{h}(1, \mathbf{X}) = \mathbf{f}$ and $\mathbf{h}(0, \mathbf{X}) = \mathbf{b}$
- compute a description of the solution curve



$\rho(t, T) = 0, X_1 = \phi_1(t, T), \ldots, X_n = \phi_n(t, T)$ in $\mathbb{K}(t)[T]$

Symbolic homotopies

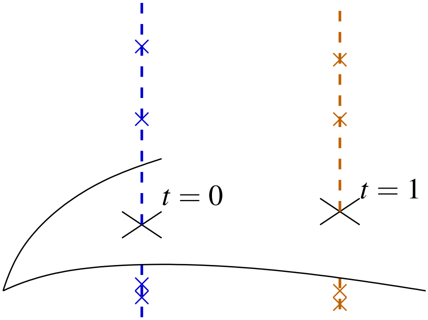
- want to solve $(f_i(\mathbf{x}))_{i\leqslant m}$, know the solutions of $(b_i(\mathbf{x}))_{i\leqslant m}$
- find $\mathbf{h}(t, \mathbf{X})$ such that $\mathbf{h}(1, \mathbf{X}) = \mathbf{f}$ and $\mathbf{h}(0, \mathbf{X}) = \mathbf{b}$
- compute a description of the solution curve and let $t = 1$



$s(T) = 0, X_1 = w_1(T), \ldots, X_n = w_n(T)$ in $\mathbb{K}[T]$

Symbolic homotopies

- want to solve $(f_i(\mathbf{x}))_{i \leqslant m}$, know the solutions of $(b_i(\mathbf{x}))_{i \leqslant m}$
- find $\mathbf{h}(t, \mathbf{X})$ such that $\mathbf{h}(1, \mathbf{X}) = \mathbf{f}$ and $\mathbf{h}(0, \mathbf{X}) = \mathbf{b}$
- compute a description of the solution curve and let $t = 1$



$s(T) = 0, X_1 = w_1(T), \ldots, X_n = w_n(T)$ in $\mathbb{K}[T]$

Previous work

[Giusti-Lecerf-Salvy]

- symbolic Newton iteration

[Jeronimo *et al.*]

- symbolic polyhedral homotopies

[Safey El Din-Schost]

- bit complexity of multi-homogeneous homotopies

Homotopies for non-square systems

Assumptions

- all components of $V(\mathbf{h}) \subset \overline{\mathbb{K}}^{n+1}$ have dimension at least 1
- if a localization $\mathbf{h} \cdot \mathbb{K}[t, \mathbf{X}]_{\mathfrak{m}}$ has height n, it is unmixed
- $\mathbf{X}$-degree($\mathbf{b}$)=$\mathbf{X}$-degree($\mathbf{h}$), with no solution at infinity
- the ideal generated by $\mathbf{b}$ is radical

Theorem

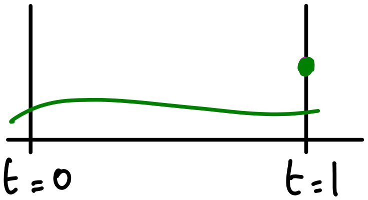
Let κ be the number of solutions of the start system.

- target system has at most κ solutions (with multiplicities)
- they can be computed by a symbolic homotopy

Homotopies for non-square systems

Assumptions

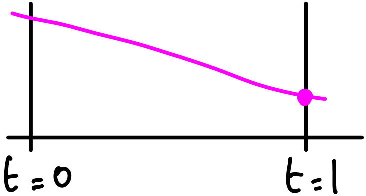
- all components of $V(\mathfrak{h}) \subset \overline{\mathbb{K}}^{n+1}$ have dimension at least 1
- if a localization $\mathfrak{h} \cdot \mathbb{K}[t, \mathsf{X}]_{\mathfrak{m}}$ has height n, it is unmixed
- X-degree($\mathfrak{b}$)=X-degree($\mathfrak{h}$), with no solution at infinity
- the ideal generated by $\mathfrak{b}$ is radical



Homotopies for non-square systems

Assumptions

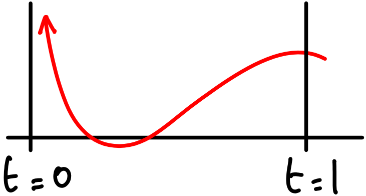
- all components of $V(\mathbf{h}) \subset \overline{\mathbb{K}}^{n+1}$ have dimension at least 1
- if a localization $\mathbf{h} \cdot \mathbb{K}[t, \mathbf{X}]_{\mathfrak{m}}$ has height n, it is unmixed
- $\mathbf{X}$-degree($\mathbf{b}$)=$\mathbf{X}$-degree($\mathbf{h}$), with no solution at infinity
- the ideal generated by $\mathbf{b}$ is radical



Homotopies for non-square systems

Assumptions

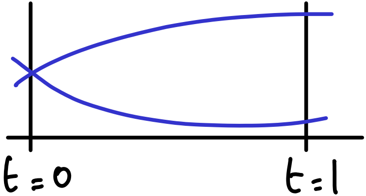
- all components of $V(\mathfrak{h}) \subset \overline{\mathbb{K}}^{n+1}$ have dimension at least 1
- if a localization $\mathfrak{h} \cdot \mathbb{K}[t, \mathsf{X}]_{\mathfrak{m}}$ has height n, it is unmixed
- X-degree($\mathfrak{b}$)=X-degree($\mathfrak{h}$), with no solution at infinity
- the ideal generated by $\mathfrak{b}$ is radical



Homotopies for non-square systems

Assumptions

- all components of $V(\mathfrak{h}) \subset \overline{\mathbb{K}}^{n+1}$ have dimension at least 1
- if a localization $\mathfrak{h} \cdot \mathbb{K}[t, \mathfrak{X}]_{\mathfrak{m}}$ has height n, it is unmixed
- $\mathfrak{X}$-degree($\mathfrak{b}$)=$\mathfrak{X}$-degree($\mathfrak{h}$), with no solution at infinity
- the ideal generated by $\mathfrak{b}$ is radical



Subroutines

Mainly classical (Newton iteration, rational reconstruction, ...)

Another ingredient: local dimension test

- we are given $\mathbf{x}$ such that $\mathbf{h}(\mathbf{x}) = 0$
- either $\mathbf{x}$ belongs to a positive-dimensional component of $V(\mathbf{h})$, or $\mathbf{x}$ is isolated with multiplicity at most κ
- $\mathbf{h}$ is given by a straight-line program of length L.

Proposition

We can decide whether $\mathbf{x}$ is an isolated root of $V(\mathbf{f})$ using $(\kappa L m)^{O(1)}$ operations in $\mathbb{K}$.

Previous work by [Marinari *et al.*], [Mourrain] and [Bates *et al.*] adapted to our SLP model.

Column degrees (aka the easy case)

When there are no polynomials G (to simplify), let

$$\mathsf{H} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^{q-1} \\ \vdots & \vdots & & \vdots \\ 1 & p & \cdots & p^{q-1} \end{bmatrix} \begin{bmatrix} \ell_{1,1} \cdots \ell_{1,\alpha_1} & & & \\ & \ell_{2,1} \cdots \ell_{2,\alpha_2} & & \\ & & \ddots & \\ & & & \ell_{q,1} \cdots \ell_{q,\alpha_q} \end{bmatrix}$$

with $\ell_{i,j}$ random linear forms, and h be the p-minors of $(1-t)\mathsf{H} + t\mathsf{F}$.

Fact: $\operatorname{rank}(\mathsf{H}) < p \iff n$ products $\ell_{i,1} \cdots \ell_{i,\alpha_i}$ vanish

This leads us to solve $E_n(\alpha_1, \ldots, \alpha_q)$ linear systems of size n

Remark:

- H already used in [Nie and Ranestad] for degree bounds
- Lagrange multipliers + bihomogeneous homotopy give similar results

Row degrees (aka the useful case)

Let

$$\mathsf{H} = \begin{bmatrix} L_{1,1} & & & & L_{1,p+1} & \cdots & L_{1,q} \\ & L_{2,2} & & & L_{2,p+1} & \cdots & L_{2,q} \\ & & \ddots & & \vdots & & \vdots \\ & & & L_{p,p} & L_{p,p+1} & \cdots & L_{p,q} \end{bmatrix}$$

with $L_{i,j}$ a product of β_i random linear forms, and h as before.

Fact: $\operatorname{rank}(\mathsf{H}) < p \iff$ some (say τ) of the $L_{i,i}$ vanish and

Row degrees (aka the useful case)

Let

$$\mathbf{H} = \begin{bmatrix} L_{1,1} & & & & L_{1,p+1} & \cdots & L_{1,q} \\ & L_{2,2} & & & L_{2,p+1} & \cdots & L_{2,q} \\ & & \ddots & & \vdots & & \vdots \\ & & & L_{p,p} & L_{p,p+1} & \cdots & L_{p,q} \end{bmatrix}$$

with $L_{i,j}$ a product of β_i random linear forms, and $\mathbf{h}$ as before.

Fact: $\mathrm{rank}(\mathbf{H}) < p \iff$ some (say τ) of the $L_{i,i}$ vanish and

Row degrees (aka the useful case)

Let

$$\mathsf{H} = \begin{bmatrix} L_{1,1} & & & & L_{1,p+1} & \cdots & L_{1,q} \\ & L_{2,2} & & & L_{2,p+1} & \cdots & L_{2,q} \\ & & \ddots & & \vdots & & \vdots \\ & & & L_{p,p} & L_{p,p+1} & \cdots & L_{p,q} \end{bmatrix}$$

with $L_{i,j}$ a product of β_i random linear forms, and h as before.

Fact: $\operatorname{rank}(\mathsf{H}) < p \iff$ some (say τ) of the $L_{i,i}$ vanish and the corresponding right-block is rank-deficient

Row degrees (aka the useful case)

Let

$$\mathsf{H} = \begin{bmatrix} L_{1,1} & & & & L_{1,p+1} & \cdots & L_{1,q} \\ & L_{2,2} & & & L_{2,p+1} & \cdots & L_{2,q} \\ & & \ddots & & \vdots & & \vdots \\ & & & L_{p,p} & L_{p,p+1} & \cdots & L_{p,q} \end{bmatrix}$$

with $L_{i,j}$ a product of β_i random linear forms, and h as before.

Fact: $\mathrm{rank}(\mathsf{H}) < p \iff$ some (say τ) of the $L_{i,i}$ vanish and the corresponding right-block is rank-deficient $\implies \tau \times (n-1)$ matrix in $n - \tau$ variables

Row degrees (aka the useful case)

Let

$$\mathsf{H} = \begin{bmatrix} L_{1,1} & & & & L_{1,p+1} & \cdots & L_{1,q} \\ & L_{2,2} & & & L_{2,p+1} & \cdots & L_{2,q} \\ & & \ddots & & \vdots & & \vdots \\ & & & L_{p,p} & L_{p,p+1} & \cdots & L_{p,q} \end{bmatrix}$$

with $L_{i,j}$ a product of β_i random linear forms, and h as before.

Fact: $\mathrm{rank}(\mathsf{H}) < p \iff$ some (say τ) of the $L_{i,i}$ vanish and the corresponding right-block is rank-deficient $\implies \tau \times (n-1)$ matrix in $n - \tau$ variables

Theorem: for generic $L_{i,j}$, the homotopy works

Row degrees (aka the useful case)

Let

$$\mathsf{H} = \begin{bmatrix} L_{1,1} & & & & L_{1,p+1} & \cdots & L_{1,q} \\ & L_{2,2} & & & L_{2,p+1} & \cdots & L_{2,q} \\ & & \ddots & & \vdots & & \vdots \\ & & & L_{p,p} & L_{p,p+1} & \cdots & L_{p,q} \end{bmatrix}$$

with $L_{i,j}$ a product of β_i random linear forms, and h as before.

Fact: $\operatorname{rank}(\mathsf{H}) < p \iff$ some (say τ) of the $L_{i,i}$ vanish and
the corresponding right-block is rank-deficient
$\implies \tau \times (n-1)$ matrix in $n - \tau$ variables

Theorem: for generic $L_{i,j}$, the homotopy works

Caveat: We recursively use a homotopy to solve the start system

Conclusions

Done

- symbolic algorithm, complexity
- prototype

To do

- error probability
- bit complexity
- numerical version
- higher rank defect