

Homotopy Methods for Solving Polynomial Systems

Jan Verschelde

Department of Math, Stat & CS
University of Illinois at Chicago
Chicago, IL 60607-7045, USA

Email: `jan@math.uic.edu`

URL: `http://www.math.uic.edu/~jan`

tutorial at ISSAC'05, Beijing, China, 24 July 2005

Structure of the Tutorial

Lecture I: Root Counting and Homotopy Methods *isolated*
Lecture II: Numerical Irreducible Decomposition *solution sets*
Lecture III: Software and Applications *hands-on session*

Three lectures of 40 minutes each. Two breaks of 15 minutes.

A.J. Sommese and C.W. Wampler: *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*. World Scientific Press, Singapore, 2005.

This material is based upon work supported by the National Science Foundation under Grant No. 0134611 and Grant No. 0410036.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Lecture I: Root Counting and Homotopy Methods

1. Deformation Methods avoid the Discriminant Variety
2. Product Structures generalize Bézout's Theorem
3. Newton Polytopes model Sparse Structures
4. Isolated Singularities computed by Deflation

A homotopy is *optimal* if every path leads to one isolated solution.

Bézout's theorem leads to optimal homotopies for **dense** systems,

Bernshtein's theorem gives optimal homotopies for **sparse** systems,

if the coefficients are random enough...

Newton's Method

Consider for example $f(\mathbf{x}) = \mathbf{0}$, $\mathbf{x} = (x_1, x_2)$:

$$f(\mathbf{x}) = \begin{cases} x_1^2 + x_2 - 3 = 0 \\ x_1 + 0.125x_2^2 - 1.5 = 0 \end{cases} \quad \frac{\partial f}{\partial \mathbf{x}} = \begin{bmatrix} 2x_1 & 1 \\ 1 & 0.125 \end{bmatrix}.$$

Start at \mathbf{x}^0 , solve $\frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}^0)\Delta\mathbf{x} = -f(\mathbf{x}^0)$, update $\mathbf{x}^1 = \mathbf{x}^0 + \Delta\mathbf{x}$.

Quadratic convergence if sufficiently close and regular Jacobian.

Residual $\|f(\mathbf{x}^0)\|$ measures **backward error**.

Condition number C of $\frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}^0)$ measures the **forward error**,

i.e.: if the coefficients are given with m digits precision,
then the error on \mathbf{x}^0 can be as large as $C10^{-m}$.

Homotopy with gamma trick

A natural choice of a homotopy to solve our example is $h(\mathbf{x}, t) =$

$$\gamma \underbrace{\begin{pmatrix} x_1^2 - 1 = 0 \\ x_2^2 - 1 = 0 \end{pmatrix}}_{\text{start system}} (1 - t) + \underbrace{\begin{pmatrix} x_1^2 + x_2 - 3 = 0 \\ x_1 + 0.125x_2^2 - 1.5 = 0 \end{pmatrix}}_{\text{target system}} t = \mathbf{0},$$

where t goes from 0 to 1, and $\gamma \in \mathbb{C}$.

For almost all choices of $\gamma \in \mathbb{C}$, every isolated solution of multiplicity m is reached by exactly m solution paths.

If we take $\gamma = 1$, then at $t \approx 0.92$ singular solutions occur.

Avoiding the Discriminant Variety

With Maple 9.5 we can compute the discriminant

```
[> f := [x^2 + y - 3, x + 1/8*y^2 - 3/2];
[> g := [x^2 - 1, y^2 - 1];
[> h := [g[1]*(1-t) + f[1]*t, g[2]*(1-t) + f[2]*t];
[> dh := Matrix(2,2, [[diff(h[1],x), diff(h[1],y)],
                      [diff(h[2],x), diff(h[2],y)]]);
[> d := LinearAlgebra[Determinant](dh);
[> gb := Groebner[gbasis]([h[1], h[2], d], plex(x,y,t));
[> fsolve(gb[1], t, complex);
```

Elimination of x and y gives a nonzero polynomial in t .

Elimination shows: only finitely many critical values for t .

Random choice of $\gamma \Rightarrow$ no critical values in $[0, 1]$.

Coefficient-Parameter Polynomial Continuation

Consider this “nontrivial example”:

$$f(\mathbf{x}) = \begin{cases} 0.5(x_2^2 + 4x_2x_3 + x_3^2) + a(x_2^2x_3^2 - 1) = 0 \\ 0.5(x_3^2 + 4x_3x_1 + x_1^2) + a(x_3^2x_1^2 - 1) = 0 \\ 0.5(x_1^2 + 4x_1x_2 + x_2^2) + a(x_1^2x_2^2 - 1) = 0 \end{cases}$$

from molecular chemistry, with parameter a .

**For generic choices of the parameter a ,
there are 16 regular isolated solutions.**

Cheater's homotopy: substituting the parameters by random complex values gives a start system for an optimal homotopy to compute all isolated solutions.

Multihomogenous Bézout numbers

Partition $\mathbf{x} = (x_1, x_2, x_3)$ into $((x_1), (x_2), (x_3))$ and form degree matrix A , $A = (a_{ij})$, $a_{ij} = \deg(f_i, x_j)$:

$$f(\mathbf{x}) = \begin{cases} 0.5(x_1^2 + 4x_1x_2 + x_2^2) + x_1^2x_2^2 - 1 = 0 \\ 0.5(x_2^2 + 4x_2x_3 + x_3^2) + x_2^2x_3^2 - 1 = 0 \\ 0.5(x_3^2 + 4x_3x_1 + x_1^2) + x_3^2x_1^2 - 1 = 0 \end{cases} \quad A = \begin{bmatrix} 2 & 2 & 0 \\ 0 & 2 & 2 \\ 2 & 0 & 2 \end{bmatrix}$$

Permanent of A via row expansion of A as in $\det(A)$ only with $+$:

$$\begin{aligned} \text{per}(A, ((x_1), (x_2), (x_3))) &= 2_{(1,1)} \times 2_{(2,2)} \times 2_{(3,3)} \\ &+ 2_{(1,2)} \times 2_{(2,3)} \times 2_{(3,1)} = 16. \end{aligned}$$

The system $f(\mathbf{x}) = 0$ can have at most 16 isolated solutions.

Notice: $16 < 64 = 4^3$, the 1-homogeneous Bézout number.

Linear-Product Start Systems

Computing a permanent corresponds to solving a start system:

$$g(\mathbf{x}) = \begin{cases} (c_{31} + c_{32}x_1)(c_{33} + c_{34}x_1) \cdot (c_{35} + c_{36}x_2)(c_{37} + c_{38}x_2) \cdot 1 = 0 \\ 1 \cdot (c_{11} + c_{12}x_2)(c_{13} + c_{14}x_2) \cdot (c_{15} + c_{16}x_3)(c_{17} + c_{18}x_3) = 0 \\ (c_{21} + c_{22}x_1)(c_{23} + c_{24}x_1) \cdot 1 \cdot (c_{25} + c_{26}x_3)(c_{27} + c_{28}x_3) = 0 \end{cases}$$

where the coefficients c_{ij} are random complex numbers.

$$A = \begin{bmatrix} 2 & 2 & 0 \\ 0 & 2 & 2 \\ 2 & 0 & 2 \end{bmatrix} \quad \begin{aligned} & \text{per}(A, ((x_1), (x_2), (x_3))) \\ &= 2_{(1,1)} \times 2_{(2,2)} \times 2_{(3,3)} \\ &+ 2_{(1,2)} \times 2_{(2,3)} \times 2_{(3,1)} = 16 \end{aligned}$$

**For almost all choices of the coefficients $c_{ij} \in \mathbb{C}$,
 $g(\mathbf{x}) = 0$ has exactly 16 regular isolated solutions.**

Recommended Background Literature

E.L. Allgower and K. Georg: Numerical Continuation Methods, an Introduction. Springer 1990.

Volume 45 of Classics in Applied Mathematics, SIAM 2003.

L. Blum, F. Cucker, M. Shub, and S. Smale:

Complexity and Real Computation. Springer 1998.

A. Morgan: Solving polynomial systems using continuation for engineering and scientific problems. Prentice-Hall, 1987.

T.Y. Li: Solving polynomial systems.

The Mathematical Intelligencer 9(3):33–39, 1987.

T.Y. Li: Numerical solution of multivariate polynomial systems by homotopy continuation methods. *Acta Numerica* 6:399–436, 1997.

Solving Binomial Systems

$$f(\mathbf{x}) = \begin{cases} x_1^3 x_2^2 = 1 \\ x_1^5 x_2^1 = 1 \end{cases} \Leftrightarrow \mathbf{x}^A = \mathbf{c} : [x_1 \ x_2] \begin{bmatrix} 3 & 5 \\ 2 & 1 \end{bmatrix} = [1 \ 1].$$

Compute $\gcd(3, 2) = 1 = (+1) \cdot 3 + (-1) \cdot 2$ to define M :

$$M = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}, \det(M) = 1 \quad MA = \begin{bmatrix} 1 & 4 \\ 0 & -7 \end{bmatrix} = U.$$

M defines a **coordinate change** $\mathbf{y}^M = \mathbf{x} \Rightarrow \mathbf{x}^A = \mathbf{y}^{MA} = \mathbf{y}^U$.

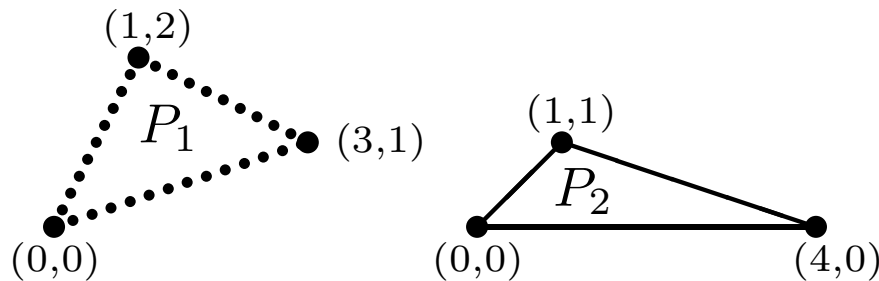
$$\begin{cases} x_1^3 x_2^2 = (y_1 y_2^{-2})^3 (y_1^{-1} y_2^3)^2 = y_1^{1 \cdot 3 + (-2) \cdot 3} y_2^{(-2) \cdot 3 + 3 \cdot 2} = y_1 \\ x_1^5 x_2^1 = (y_1 y_2^{-2})^5 (y_1^{-1} y_2^3)^1 = y_1^{1 \cdot 5 + (-1) \cdot 1} y_2^{-2 \cdot 5 + 3 \cdot 1} = y_1^4 y_2^{-7} \end{cases}$$

$f(\mathbf{x}) = \mathbf{0}$ has exactly $7(= |\det(A)|)$ isolated regular solutions

Supports and Newton Polytopes

$$\begin{aligned}
 f &= (f_1, f_2) & \mathcal{A} &= (A_1, A_2), \text{ supports of } f \\
 &= \begin{cases} x_1^3 x_2 + x_1 x_2^2 + 1 = 0 \\ x_1^4 + x_1 x_2 + 1 = 0 \end{cases} & A_1 &= \{(3, 1), (1, 2), (0, 0)\} \\
 & & A_2 &= \{(4, 0), (1, 1), (0, 0)\}
 \end{aligned}$$

The *Newton polytopes*:

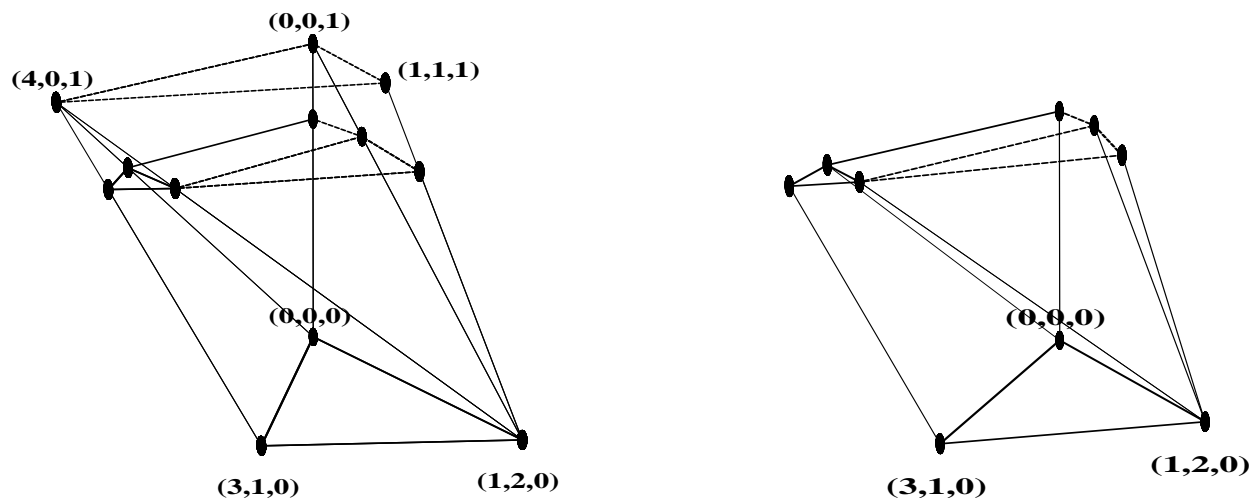


The Cayley Trick

$$f = (f_1, f_2) \quad \mathcal{A} = (A_1, A_2), \text{ supports of } f$$

$$= \begin{cases} x_1^3 x_2 + x_1 x_2^2 + 1 = 0 \\ x_1^4 + x_1 x_2 + 1 = 0 \end{cases} \quad \begin{aligned} A_1 &= \{(3, 1), (1, 2), (0, 0)\} \\ A_2 &= \{(4, 0), (1, 1), (0, 0)\} \end{aligned}$$

A triangulation of the Cayley polytope has three cells:



The cross section contains a **mixed subdivision** of \mathcal{A} .

Minkowski's Theorem and Mixed volumes

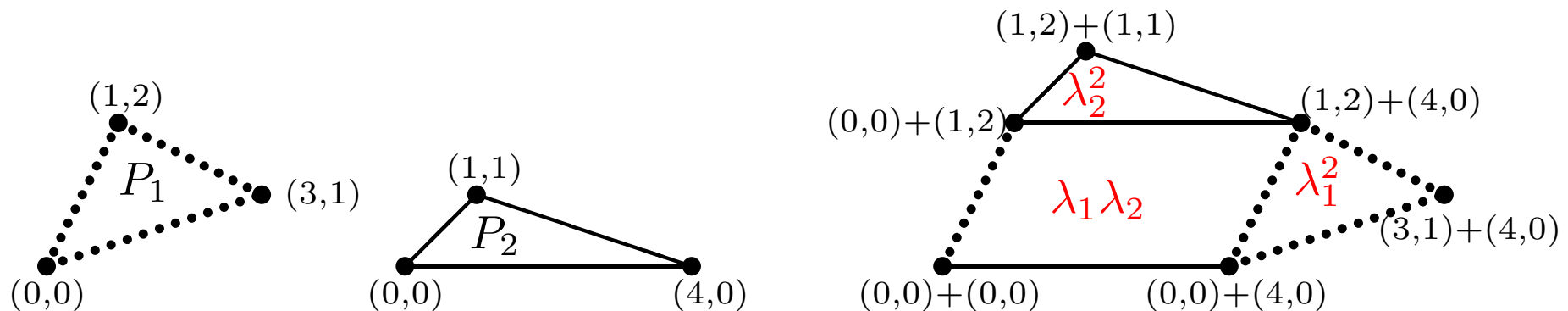
Supports $A_1 = \{(3, 1), (1, 2), (0, 0)\}$ and $A_2 = \{(4, 0), (1, 1), (0, 0)\}$.

$P_1 = \text{conv}(A_1)$ and $P_2 = \text{conv}(A_2)$ are Newton polygons.

Minkowski's theorem applied: $V(P_1, P_2)$ is *mixed volume*

$$\begin{aligned} \text{area}(\lambda_1 P_1 + \lambda_2 P_2) &= V(P_1, P_1)\lambda_1^2 + V(P_1, P_2)\lambda_1\lambda_2 + V(P_2, P_2)\lambda_2^2 \\ &= 5\lambda_1^2 + 8\lambda_1\lambda_2 + 4\lambda_2^2. \end{aligned}$$

Note: the area of the unit triangle equals one.



The Lift and Prune Algorithm

Assign an integer **lifting** to the points in \mathcal{A} to obtain $\widehat{\mathcal{A}} = (\widehat{A}_1, \widehat{A}_2)$:

$$\widehat{A}_1 = \{(3, 1, \mathbf{1}), (1, 2, \mathbf{1}), (0, 0, \mathbf{0})\}, \widehat{A}_2 = \{(4, 0, \mathbf{0}), (1, 1, \mathbf{1}), (0, 0, \mathbf{1})\}.$$

A mixed cell is spanned by two edges. Since it corresponds to a facet of the lower hull of the sum of the polytopes, it is defined by one **inner normal** $\mathbf{v} \neq \mathbf{0}$. One mixed cell is $\widehat{\mathcal{C}}^{(1)} = (\widehat{C}_1, \widehat{C}_2)$:

$$\begin{array}{l} \widehat{C}_1 = \{(1, 2, 1), (0, 0, 0)\}, \\ \widehat{C}_2 = \{(4, 0, 0), (1, 1, 1)\}, \\ \mathbf{v} = (1, -4, 7), V(\mathcal{C}^{(1)}) = 7 \end{array} \quad \left\{ \begin{array}{l} \langle (1, 2, 1), \mathbf{v} \rangle = \langle (0, 0, 0), \mathbf{v} \rangle < \langle (3, 1, 1), \mathbf{v} \rangle \\ \langle (4, 0, 0), \mathbf{v} \rangle = \langle (1, 1, 1), \mathbf{v} \rangle < \langle (0, 0, 1), \mathbf{v} \rangle \end{array} \right.$$

v satisfies linear inequality system

Linear programming **prunes** the tree of all edge-edge combinations.

Polyhedral Homotopies induced by Lifting

$\widehat{A}_1 = \{(3, 1, \mathbf{1}), (1, 2, \mathbf{1}), (0, 0, \mathbf{0})\}$ $\widehat{A}_2 = \{(4, 0, \mathbf{0}), (1, 1, \mathbf{1}), (0, 0, \mathbf{1})\}$
are lifted supports, defining exponents of t :

$$\widehat{g}(\mathbf{x}, t) = \begin{cases} c_{31}^{(1)} x_1^3 x_2 t^{\mathbf{1}} + c_{12}^{(1)} x_1 x_2^2 t^{\mathbf{1}} + c_{00}^{(1)} t^{\mathbf{0}} = 0 \\ c_{40}^{(2)} x_1^4 t^{\mathbf{0}} + c_{11}^{(2)} x_1 x_2 t^{\mathbf{1}} + c_{00}^{(2)} t^{\mathbf{1}} = 0 \end{cases} \quad c_{ij}^{(k)} \in \mathbb{C}.$$

Inner normal $\mathbf{v} = (1, -4, 7)$ defines $x_1 = y_1 s^1$, $x_2 = y_2 s^{-4}$, $t = s^7$:

$$g(\mathbf{y}, s) = \begin{cases} c_{31}^{(1)} y_1^3 y_2 s^6 + c_{12}^{(1)} y_1 y_2^2 + c_{00}^{(1)} = 0 \\ \left(c_{40}^{(2)} y_1^4 + c_{11}^{(2)} y_1 y_2 + c_{00}^{(2)} s^3 \right) s^4 = 0 \end{cases}$$

After clearing the s^4 from the second equation, we observe:

**exponent vectors of those terms which occur with s^0
belong to the support of the mixed cell with inner normal \mathbf{v} .**

Bernshtein's first theorem

Let $g(\mathbf{x}) = \mathbf{0}$ have the same Newton polytopes \mathcal{P} as $f(\mathbf{x}) = \mathbf{0}$, but with randomly chosen complex coefficients.

I. Compute $V_n(\mathcal{P})$:

I.1 lift polytopes

I.2 mixed cells

I.3 volume of mixed cell

II. Solve $g(\mathbf{x}) = \mathbf{0}$:

II.1 introduce parameter t

II.2 start systems

II.3 path following

III. Coefficient-parameter continuation to solve $f(\mathbf{x}) = \mathbf{0}$:

$$h(\mathbf{x}, t) = (1 - t)g(\mathbf{x}) + tf(\mathbf{x}) = \mathbf{0}, \quad \text{for } t \text{ from } 0 \text{ to } 1.$$

#isolated solutions in $(\mathbb{C}^*)^n$, $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$, of $f(\mathbf{x}) = \mathbf{0}$ is bounded by the mixed volume of the Newton polytopes of f .

Bernshtein's second theorem

- Face $\partial_{\mathbf{v}} f = (\partial_{\mathbf{v}} f_1, \partial_{\mathbf{v}} f_2, \dots, \partial_{\mathbf{v}} f_n)$ of system $f = (f_1, f_2, \dots, f_n)$ with Newton polytopes $\mathcal{P} = (P_1, P_2, \dots, P_n)$ and mixed volume $V(\mathcal{P})$.

$$\partial_{\mathbf{v}} f_i(\mathbf{x}) = \sum_{\mathbf{a} \in \partial_{\mathbf{v}} A_i} c_{i\mathbf{a}} \mathbf{x}^{\mathbf{a}} \quad \begin{array}{l} \partial_{\mathbf{v}} P_i = \text{conv}(\partial_{\mathbf{v}} A_i) \\ \text{face of Newton polytope} \end{array}$$

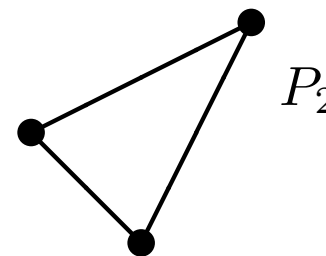
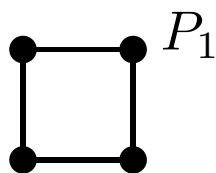
Theorem: **If $\forall \mathbf{v} \neq 0$, $\partial_{\mathbf{v}} f(\mathbf{x}) = 0$ has no solutions in $(\mathbb{C}^*)^n$, then $V(\mathcal{P})$ is exact and all solutions are isolated. Otherwise, for $V(\mathcal{P}) \neq 0$: $V(\mathcal{P}) > \#$ isolated solutions.**

- Newton polytopes *in general position*:
 $V(\mathcal{P})$ is **exact** for every nonzero choice of the coefficients.

Newton polytopes in general position

$$\text{Consider } f(\mathbf{x}) = \begin{cases} c_{111}x_1x_2 + c_{110}x_1 + c_{101}x_2 + c_{100} = 0 \\ c_{222}x_1^2x_2^2 + c_{210}x_1 + c_{201}x_2 = 0 \end{cases}$$

The Newton polytopes:



$$\forall \mathbf{v} \neq \mathbf{0} : \partial_{\mathbf{v}} A_1 + \partial_{\mathbf{v}} A_2 \leq 3 \quad \Rightarrow \quad V(P_1, P_2) = 4 \text{ \textbf{always exact}} \\ \text{for all nonzero coefficients}$$

Polyhedral Endgames: the direction \mathbf{v} of a diverging solution path is normal to the faces of the Newton polytopes defining $\partial_{\mathbf{v}} f(\mathbf{x}) = 0$ and solution at ∞ .

Recommended Background Literature

- I.M. Gel'fand, M.M. Kapranov, and A.V. Zelevinsky: **Discriminants, Resultants and Multidimensional Determinants**. Birkhäuser, 1994.
- J.E. Goodman and J. O'Rourke (editors): **Handbook of Discrete and Computational Geometry**. CRC Press, 1997.
- R. Schneider: **Convex Bodies: The Brunn-Minkowski Theory**. Cambridge University Press, 1993.
- B. Sturmfels: **Gröbner Bases and Convex Polytopes**. AMS, 1996.
- B. Sturmfels: **Polynomial equations and convex polytopes**. *Amer. Math. Monthly*, 105(10):907–922, 1998.
- B. Sturmfels: **Solving Systems of Polynomial Equations**. AMS, 2002.
- G.M. Ziegler: **Lectures on Polytopes**. Springer, 1995.

Deflation Operator **Dfl** reduces to Corank One

Consider $f(\mathbf{x}) = \mathbf{0}$, N equations in n unknowns, $N \geq n$.

Suppose $\text{Rank}(A(\mathbf{z}_0)) = R < n$ for \mathbf{z}_0 an isolated zero of $f(\mathbf{x}) = 0$.

Choose $\mathbf{h} \in \mathbb{C}^{R+1}$ and $B \in \mathbb{C}^{n \times (R+1)}$ at random.

Introduce $R + 1$ new multiplier variables $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_{R+1})$.

$$\mathbf{Dfl}(f)(\mathbf{x}, \boldsymbol{\lambda}) := \begin{cases} f(\mathbf{x}) = \mathbf{0} & \text{Rank}(A(\mathbf{x})) = R \\ A(\mathbf{x})B\boldsymbol{\lambda} = \mathbf{0} & \Downarrow \\ \mathbf{h}\boldsymbol{\lambda} = 1 & \text{corank}(A(\mathbf{x})B) = 1 \end{cases}$$

Compared to the deflation of Ojika, Watanabe, and Mitsui:

- (1) we do not compute a maximal minor of the Jacobian matrix;
- (2) we only add new equations, we never replace equations.

Newton with Deflation – A Simple Example

$$f(x, y) = \begin{cases} x^2 = 0 \\ xy = 0 \\ y^2 = 0 \end{cases} \quad A(x, y) = \begin{bmatrix} 2x & 0 \\ y & x \\ 0 & 2y \end{bmatrix} \quad \begin{aligned} \mathbf{z}_0 &= (0, 0), m = 3 \\ \text{Rank}(A(\mathbf{z}_0)) &= 0 \end{aligned}$$

We use a 2-by-1 random matrix B and one multiplier λ_1 :

$$\mathbf{Df}(f)(x, y, \lambda_1) = \begin{cases} f(x, y) = 0 \\ \begin{bmatrix} 2x & 0 \\ y & x \\ 0 & 2y \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} \lambda_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ h_1 \lambda_1 = 1, \quad \text{random } h_1 \in \mathbb{C} \end{cases}$$

$\mathbf{Df}(f)(x, y, \lambda_1) = \mathbf{0}$ has $(0, 0, \lambda_1^*)$ as **regular** zero!

Newton's Method with Deflation

Input: $f(\mathbf{x}) = \mathbf{0}$ polynomial system;
 \mathbf{x}_0 initial approximation for \mathbf{x}^* ;
 ϵ tolerance for numerical rank.

Newton's Method with Deflation

Input: $f(\mathbf{x}) = \mathbf{0}$ polynomial system;
 \mathbf{x}_0 initial approximation for \mathbf{x}^* ;
 ϵ tolerance for numerical rank.

$$\begin{aligned} [A^+, R] &:= \text{SVD}(A(\mathbf{x}_k), \epsilon); \\ \mathbf{x}_{k+1} &:= \mathbf{x}_k - A^+ f(\mathbf{x}_k); \end{aligned}$$

Gauss-Newton

Newton's Method with Deflation

Input: $f(\mathbf{x}) = \mathbf{0}$ polynomial system;
 \mathbf{x}_0 initial approximation for \mathbf{x}^* ;
 ϵ tolerance for numerical rank.

$[A^+, R] := \text{SVD}(A(\mathbf{x}_k), \epsilon);$
 $\mathbf{x}_{k+1} := \mathbf{x}_k - A^+ f(\mathbf{x}_k);$

Gauss-Newton

$R = \#\text{columns}(A)?$

Yes

Output: $f; \mathbf{x}_{k+1}$.

Newton's Method with Deflation

Input: $f(\mathbf{x}) = \mathbf{0}$ polynomial system;
 \mathbf{x}_0 initial approximation for \mathbf{x}^* ;
 ϵ tolerance for numerical rank.

$[A^+, R] := \text{SVD}(A(\mathbf{x}_k), \epsilon);$
 $\mathbf{x}_{k+1} := \mathbf{x}_k - A^+ f(\mathbf{x}_k);$

Gauss-Newton

$R = \# \text{columns}(A)?$

Yes

Output: $f; \mathbf{x}_{k+1}.$

No

$f := \text{Dfl}(f)(\mathbf{x}, \boldsymbol{\lambda}) = \begin{cases} f(\mathbf{x}) = \mathbf{0} \\ G(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0} \end{cases};$
 $\hat{\boldsymbol{\lambda}} := \text{LeastSquares}(G(\mathbf{x}_{k+1}, \boldsymbol{\lambda}));$
 $k := k + 1; \quad \mathbf{x}_k := (\mathbf{x}_k, \hat{\boldsymbol{\lambda}});$

Deflation Step

A Bound on the Number of Deflations

Theorem (Anton Leykin, JV, Ailing Zhao):

The number of deflations needed to restore the quadratic convergence of Newton's method converging to an isolated solution is strictly less than the multiplicity.

Duality Analysis of Barry H. Dayton and Zhonggang Zeng:

- (1) tighter bound on number of deflations; and
- (2) special case algorithms, for corank = 1.

(Proceedings of ISSAC 2005, pages 116-123.)