Solving Polynomial Systems

Jan Verschelde

Email: `jan@math.uic.edu`

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

Graduate Student Colloquium

11 March 2005

Outline of this Lecture

1. What does solving mean?
2. Numerical Polynomial Algebra
3. Numerical Algebraic Geometry
4. What is Symbolic-Numeric Computing?

What does solving mean?

Two vastly different views on polynomial systems.

Computer algebra views the polynomials in a system as generators of an ideal. A **Gröbner basis** allows to compute with the ideal and solve the system.

Numerical analysis views a polynomial system as a system of nonlinear equations. **Newton's method**, often with continuation, approximates the solutions.

Where and when do the algebra and analysis meet?

Ideals generated by Polynomial Systems

$\boldsymbol{P}(\mathbf{x}) = \mathbf{0}$ polynomial system:

$$P = (p_1, p_2, \ldots, p_N),\ p_k \in \mathbb{C}[\mathbf{x}],\ \mathbf{x} = (x_1, x_2, \ldots, x_n).$$

$\boldsymbol{I} = \langle \boldsymbol{P} \rangle$ ideal generated by P:

$$I = \{\ c_1(\mathbf{x})p_1 + c_2(\mathbf{x})p_2 + \cdots + c_N(\mathbf{x})p_N \mid c_k \in \mathbb{C}[\mathbf{x}]\ \}.$$

$\boldsymbol{Z}[\boldsymbol{I}]$ solution set:

$$Z[I] = \{\ \mathbf{z} \in \mathbb{C}^n \mid p(\mathbf{z}) = 0, \forall p \in I\ \}.$$

A Commutative Diagram

ideal

$$\begin{tikzcd}
 & \boldsymbol{I} \arrow[dl, shift right] \arrow[dr, shift left] & \\
\boldsymbol{R}[\boldsymbol{I}] \arrow[ur, shift right] \arrow[rr, shift left] & & \boldsymbol{D}[\boldsymbol{I}] \arrow[ul, shift left] \arrow[ll, shift left]
\end{tikzcd}$$

quotient ring

dual space

Ideals and Quotient Rings

The quotient ring $R[I]$ of a ideal I allows to calculate modulo I:

$R[I] = \{ \ [p]_I \mid p \in \mathbb{C}[\mathbf{x}] \ \}$ is the residue class ring,

with $[p]_I$ the **residue class of p mod I**:

$$\begin{aligned} [p]_I &= \{ \ r \in \mathbb{C}[\mathbf{x}] \mid p - r \in I \ \} \\ &= \{ \ r \in \mathbb{C}[\mathbf{x}] \mid p(\mathbf{z}) = r(\mathbf{z}), \forall \mathbf{z} \in Z[I] \ \}. \end{aligned}$$

Theorem: **If $\#Z[I] = m < \infty$, then $\dim(R[I]) = m$.**

$I \longrightarrow R[I]$: if $\dim(Z[I]) = 0$, the residue classes of Lagrange polynomials interpolating at $Z[I]$ give a basis for $R[I]$.

$R[I] \longrightarrow I$: for some basis $\mathbf{b}$ of $R[I]$: $A_k \mathbf{b} = x_k \mathbf{b}$, $k = 1, 2, \ldots, n$, means $x_k \mathbf{b} = A_k \mathbf{b}$ mod I, or $x_k \mathbf{b} - A_k \mathbf{b} = 0$ over $Z[I]$, and $x_k \mathbf{b} - A_k \mathbf{b} \in I$ leads to a border basis for I.

Example One

$$P(x,y) = \begin{cases} x^2 + 4xy + 4y^2 - 4 = 0 \\ 4x^2 - 4xy + y^2 - 4 = 0 \end{cases}$$

we expect 2×2
$= 4$ solutions

$$\xrightarrow[\text{on coefficients}]{\text{row reduction}} \begin{cases} 15x^2 - 20xy - 12 = 0 \\ 20y^2 + 20xy - 12 = 0 \end{cases}$$

A natural **basis for $R[I]$** is $\mathbf{b} = (1, x, y, xy)$.

$$x\mathbf{b} = x \begin{pmatrix} 1 \\ x \\ y \\ xy \end{pmatrix} = \begin{pmatrix} x \\ x^2 \\ xy \\ x^2y \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{12}{15} & 0 & 0 & \frac{20}{15} \\ 0 & 0 & 0 & 1 \\ 0 & \frac{48}{125} & \frac{36}{125} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ y \\ xy \end{pmatrix}$$

Observe: $\boldsymbol{A_x \mathrm{b} = x\mathrm{b}}$ is an eigenvalue problem.

Dual Spaces of Quotient Rings

The **dual space** of the quotient ring $R[I]$:

$$(R[I])^* = \{ \ l : R[I] \to \mathbb{C} : [p]_I \mapsto l(p) := l(r), r \in R[I], p - r \in I \ \}.$$

The dual of I: $D[I] = \{ \ l : \mathbb{C}[\mathbf{x}] \to \mathbb{C} : l(p) = 0, \forall p \in I \ \}$. $\boldsymbol{I \longrightarrow D[I]}$

$$\boxed{\text{Theorem: } \boldsymbol{D[I] = (R[I])^*} \text{ and } \boldsymbol{R[I] = (D[I])^*}.}$$

The ideal of the dual: $\boldsymbol{D[I] \longrightarrow I}$

$$I[D[I]] = \ker(D[I]) = \{ \ p \in \mathbb{C}[\mathbf{x}] : l(p) = 0, \forall l \in D[I] \ \} = I.$$

Note: more satisfactory than $I[Z[I]] = \sqrt{I}$.

The dual defines the multiplicity structure of a multiple zero.

Multiplicity of an Isolated Zero

An isolated zero of multiplicity $\boldsymbol{m}$ occurs in numerical analysis as a cluster of $\boldsymbol{m}$ (ill-conditioned) regular zeros.

Problem: geometrical significance for overdetermined systems?
$\longrightarrow$ perturbed overdetermined system has no zeros!

Analogy with Univariate Case: z_0 is $\boldsymbol{m}$-fold zero of $f(x) = 0$:

$$\underbrace{f(z_0) = 0, \frac{\partial f}{\partial x}(z_0) = 0, \frac{\partial^2 f}{\partial x^2}(z_0) = 0, \ldots, \frac{\partial^{m-1} f}{\partial x^{m-1}}(z_0) = 0}\quad .$$

$\boldsymbol{m}$ = number of linearly independent polynomials annihilating z_0

The dual space $\boldsymbol{D_0}$ at $\mathbf{z}_0$ is spanned by $\boldsymbol{m}$ linear independent differentiation functionals annihilating $\mathbf{z}_0$.

$\boldsymbol{D_0}$ is **the multiplicity structure** of the $\boldsymbol{m}$-fold zero $\mathbf{z}_0$.

Example Two

Consider

$$f(x,y) = \begin{cases} x^2 = 0 \\ xy = 0 \\ y^2 = 0 \end{cases} \qquad \mathbf{z}_0 = (0,0).$$

The **multiplicity of $\mathbf{z}_0$ is 3** because

$$D_0 = \operatorname{span}\{\partial_{00}[\mathbf{z}_0], \partial_{10}[\mathbf{z}_0], \partial_{01}[\mathbf{z}_0]\}$$

with

$$\partial_{ij}[\mathbf{z}_0] = \frac{1}{i!j!}\,\frac{\partial^{i+j}}{\partial x^i \partial y^j} f(\mathbf{z}_0).$$

Solving means to compute $\mathbf{z}_0$ ***and*** D_0.

Newton and Continuation Methods

Consider a family of systems $h_k(x(t), y(t), t) = 0$, $k = 1, 2$.

By $\frac{\partial}{\partial t}$ on h_k: $\frac{\partial h_k}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial h_k}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial h_k}{\partial t}\frac{\partial t}{\partial t} = 0$, $k = 1, 2$.

Set $\Delta x := \frac{\partial x}{\partial t}$, $\Delta y := \frac{\partial y}{\partial t}$, and $\frac{\partial t}{\partial t} = 1$.

Increment $t := t + \Delta t$

Solve
$$\begin{bmatrix} \frac{\partial h_1}{\partial x} & \frac{\partial h_1}{\partial y} \\ \frac{\partial h_2}{\partial x} & \frac{\partial h_2}{\partial y} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = - \begin{bmatrix} \frac{\partial h_1}{\partial t} \\ \frac{\partial h_2}{\partial t} \end{bmatrix} \quad (Newton)$$

Update
$$\begin{cases} x := x + \Delta x \\ y := y + \Delta y \end{cases}$$

fails when the Jacobian matrix is singular!

An Application: cyclic 9-roots

The system

$$f(\mathbf{x}) = \begin{cases} f_i = \displaystyle\sum_{j=0}^{8} \prod_{k=1}^{i} x_{(k+j) \bmod 9} = 0, \quad i = 1, 2, \ldots, 8 \\ f_9 = x_0 x_1 x_2 x_3 x_4 x_5 x_6 x_7 x_8 - 1 = 0 \end{cases}$$

has 333×18 **isolated regular** zeros, 164 **isolated 4-fold** zeros, and 6 **cubic 2-dimensional** irreducible solution components.

Newton's method with 64 decimal places, tolerance is 10^{-60}:

regular : 4 iterations (**quadratic convergence**)

4-fold : 79 iterations (>1 **step for one correct decimal place**)

about 20 times slower to reach same magnitude of residual ...

Newton's method for Overdetermined Systems

Singular Value Decomposition of N-by-n Jacobian matrix J_f:

$J_f = U\Sigma V^T$, $\quad U$ and V are orthogonal: $U^TU = I_N, V^TV = I_n$,

and singular values $\boldsymbol{\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n}$ as the only nonzero elements on the diagonal of the N-by-n matrix Σ $(N > n)$.

Moore-Penrose inverse: $J_f^+ = V\Sigma^+U^T$, with $\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \ldots, \frac{1}{\sigma_n}$ on the diagonal of the n-by-N matrix Σ^+.

Then $\Delta\mathbf{z} = -J_f(\mathbf{z})^+f(\mathbf{z})$ is the least squares solution.

The **condition number** $\mathrm{cond}(J_f(\mathbf{z})) = \frac{\sigma_1}{\sigma_n}$.

$\mathrm{Rank}(J_f(\mathbf{z})) = R \Longleftrightarrow \Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_R, 0, \ldots, 0)$.

At a **multiple root** $\mathbf{z}_0$: $\mathrm{Rank}(J_f(\mathbf{z}_0)) = R < n$.

Close to $\mathbf{z}_0$, $\mathbf{z} \approx \mathbf{z}_0 : \sigma_{R+1} \approx 0$.

Newton with Deflation – Example Two revisited

$$f(x,y) = \begin{cases} x^2 = 0 \\ xy = 0 \\ y^2 = 0 \end{cases} \qquad J_f(x,y) = \begin{bmatrix} 2x & 0 \\ 1 & 1 \\ 0 & 2y \end{bmatrix} \qquad \begin{array}{l} \mathbf{z}_0 = (0,0), m = 3 \\ \text{Rank}(J_f(\mathbf{z}_0)) = 1 \end{array}$$

A nontrivial linear combination of the columns of $J_f(\mathbf{z}_0)$ is zero.

$$G(x,y,\lambda_1,\lambda_2) = \begin{cases} \begin{array}{rcl} f(x,y) & = & 0 \\ \begin{bmatrix} 2x & 0 \\ 1 & 1 \\ 0 & 2y \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} & = & \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ c_1\lambda_1 + c_2\lambda_2 & = & 1, \qquad \text{random } c_1, c_2 \in \mathbb{C} \end{array} \end{cases}$$

The system $G(x,y,\lambda_1,\lambda_2) = 0$ has $(0,0,\lambda_1^*,\lambda_2^*)$ as **regular** zero!

The Deflation Operator Dfl

Suppose $\text{Rank}(J_f(\mathbf{z}_0)) = R$ 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.

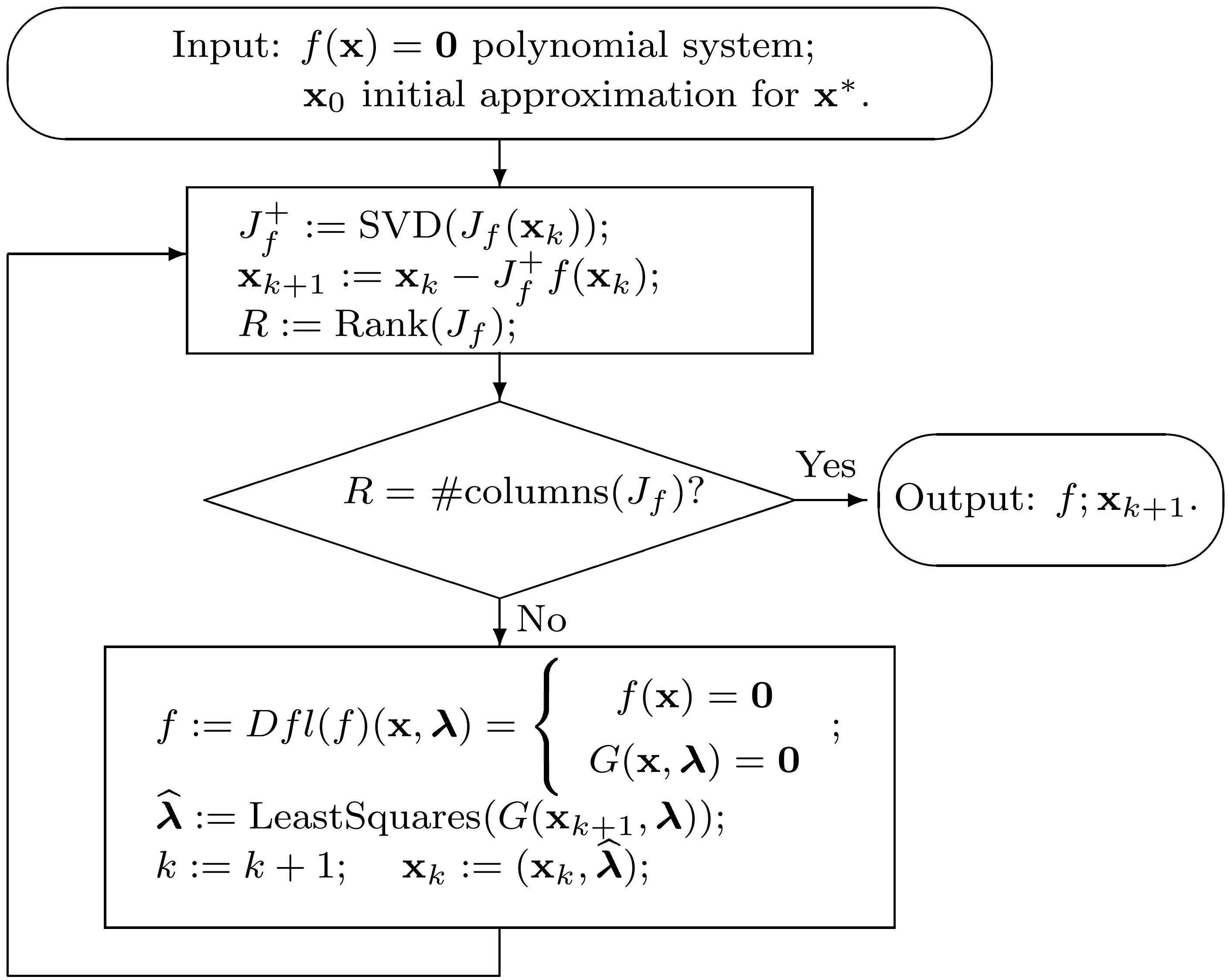
Introduce $R+1$ new multiplier variables $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_{R+1})$.

$$Dfl(f)(\mathbf{x}, \boldsymbol{\lambda}) := \left\{ \begin{array}{rcl} f(\mathbf{x}) & = & \mathbf{0} \\ J_f(\mathbf{x})B\boldsymbol{\lambda} & = & \mathbf{0} \\ \mathbf{h}\boldsymbol{\lambda} & = & 1 \end{array} \right.$$

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.

Newton's method with Deflation



cyclic 9-roots revisited

Recall:

$$f(\mathbf{x}) = \begin{cases} f_i = \displaystyle\sum_{j=0}^{8} \prod_{k=1}^{i} x_{(k+j) \bmod 9} = 0, \quad i = 1, 2, \ldots, 8 \\ f_9 = x_0 x_1 x_2 x_3 x_4 x_5 x_6 x_7 x_8 - 1 = 0 \end{cases}$$

has 164 solutions of multiplicity 4.

One deflation suffices to restore quadratic convergence.

The **condition number** drops from 1.8E+9 to 5.6E+2.

$\rightarrow$ **deflation <u>reconditions</u> the system**

Computing the Multiplicity Structure

following B.H. Dayton and Z. Zeng

Looking for differentiation functionals $d[\mathbf{z}_0] = \sum_{\mathbf{a}} c_{\mathbf{a}} \partial_{\mathbf{a}}[\mathbf{z}_0]$,

$$\text{with} \quad \partial_{\mathbf{a}}[\mathbf{z}_0](p) = \frac{1}{a_1! a_2! \cdots a_n!} \left(\frac{\partial^{a_1+a_2+\cdots+a_n}}{\partial x_1^{a_1} \partial x_2^{a_2} \cdots \partial x_n^{a_n}} p \right) (\mathbf{z}_0).$$

Membership criterium for $d[\mathbf{z}_0]$:

$$\boldsymbol{d[\mathrm{z}_0] \in D_0 \Leftrightarrow d[\mathrm{z}_0](p f_i) = 0, \forall p \in \mathbb{C}[\mathrm{x}], i = 1, 2, \ldots, N.}$$

To turning this criterium into an **algorithm**, observe:

1. since $d[\mathbf{z}_0]$ is linear, restrict p to $\mathbf{x}^{\mathbf{k}} = x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$; and
2. limit degrees $k_1 + k_2 + \cdots + k_n \leq a_1 + a_2 + \cdots + a_n$, as $\mathbf{z}_0 = \mathbf{0}$ vanishes trivially if not annihilated by $\partial_{\mathbf{a}}$.

Computing the Multiplicity Structure – An Example

$f_1 = x_1 - x_2 + x_1^2,\ f_2 = x_1 - x_2 + x_2^2$ following B.H. Dayton and Z. Zeng

		$\lvert a\rvert=0$	$\lvert a\rvert=1$		$\lvert a\rvert=2$			$\lvert a\rvert=3$			
		∂_{00}	∂_{10}	∂_{01}	∂_{20}	∂_{11}	∂_{02}	∂_{30}	∂_{21}	∂_{12}	∂_{03}
	f_1	0	1	–1	1	0	0	0	0	0	0
S_1	f_2	0	1	–1	0	0	1	0	0	0	0
	$x_1 f_1$	0	0	0	1	–1	0	1	0	0	0
	$x_1 f_2$	0	0	0	1	–1	0	0	0	1	0
	$x_2 f_1$	0	0	0	0	1	–1	0	1	0	0
S_2	$x_2 f_2$	0	0	0	0	1	–1	0	0	0	1
	$x_1^2 f_1$	0	0	0	0	0	0	1	–1	0	0
	$x_1^2 f_2$	0	0	0	0	0	0	1	–1	0	0
	$x_1 x_2 f_1$	0	0	0	0	0	0	0	1	–1	0
	$x_1 x_2 f_2$	0	0	0	0	0	0	0	1	–1	0
	$x_2^2 f_1$	0	0	0	0	0	0	0	0	1	–1
S_3	$x_2^2 f_2$	0	0	0	0	0	0	0	0	1	–1

Nullity(S_2) = Nullity(S_3) $\Rightarrow$ stop algorithm

$D_0 = \text{span}\{\ \partial_{00}, \partial_{10} + \partial_{01}, -\partial_{10} + \partial_{20} + \partial_{11} + \partial_{02}\ \} \Rightarrow$ multiplicity $= 3$

cyclic 9-roots once more

Recall:

$$f(\mathbf{x}) = \begin{cases} f_i = \displaystyle\sum_{j=0}^{8} \prod_{k=1}^{i} x_{(k+j) \bmod 9} = 0, \quad i = 1, 2, \ldots, 8 \\ f_9 = x_0 x_1 x_2 x_3 x_4 x_5 x_6 x_7 x_8 - 1 = 0 \end{cases}$$

has 164 solutions of multiplicity 4.

Running the algorithm of Dayton and Zeng:

$$H[1] = 1, H[2] = 2, H[3] = 1, H[4] = 0,$$
$$\text{with } H[i] = \text{Nullity}(S_i) - \text{Nullity}(S_{i-1}), i > 0,$$

so we compute the multiplicity as 4.

What is Symbolic-Numeric Computing?

A puristic point of view:

Computer algebra rewrites the problem, producing "easier" equations of the ideal, but **"easier" $\neq$ numerically better**.

Numerical analysis produces approximate numbers for a fixed system of equations, but **many problems are "ill-posed"**.

The synergistic approach

Symbolic-Numeric Computing rewrites an ill-conditioned numerical problem into a well-conditioned formulation.

works very well in Newton's method with deflation

Chronological List of References

T. Ojika, S. Watanabe, and T. Mitsui: **Deflation algorithm for the multiple roots of a system of nonlinear equations.** *J. Math. Anal. Appl.*, 96:463–479, 1983.

B. Mourrain: **Isolated points, duality and residues.** Journal of Pure and Applied Algebra 117/118:469–493, 1997.

H.J. Stetter and G.H. Thallinger: **Singular Systems of Polynomials.** In *ISSAC'98*, pages 9–16, ACM 1998.

H.J. Stetter: **Numerical Polynomial Algebra.** SIAM, 2004.

A. Leykin, J. Verschelde, and A. Zhao: **Newton's Method with Deflation for Isolated Singularities of Polynomial Systems.** Manuscript, 2004.

B.H. Dayton and Z. Zeng: **Computing the Multiplicity Structure in Solving Polynomial Systems.** Manuscript, 2005.