Numerical Local Rings and Local Solution of Nonlinear Systems

Barry H. Dayton

Northeastern Illinois University
Chicago, Illinois
bhdayton@neiu.edu

July 26, 2007
Symbolic-Numeric Computation '07
University of Western Ontario

Example

Consider the system (rounded here for display)

$$.5x^2+10xy-25xy^2-25x^4-166.6667xy^3-52.08333xy^4=0$$

$$5y-12.50x^2-20.833333y^3+26.0416667x^4+26.0416667y^5=0$$

Solution by BERTINI returns 21 finite solutions, one is (rounded)

$$\hat{\mathbf{x}}_B=(-.7250509\text{e-}11+.18546461\text{e-}11*I,\ .535604662\text{e-}12-.14049224\text{e-}12*I)$$

which BERTINI claims has multiplicity 2.

Problem: Testing multiplicity by DZ algorithm (Dayton-Zeng, ISSAC '05), tolerance set at 1e-06, the algorithm returns a multiplicity of 3. Which is correct?

Answer: Possibly both.

DZ algorithm

Given a non-linear system $F = [f_1, ..f_t]^\top$ in variables $x_1, \ldots, x_s$ write $\mathbf{x}^\mathbf{j} = x_1^{j_1} \ldots x_s^{j_s}$ and

$$\partial_{\mathbf{x}^\mathbf{j}} \equiv \partial_{x_1^{j_1} \cdots x_s^{j_s}} \equiv \frac{1}{j_1! \cdots j_s!} \frac{\partial^{j_1 + \cdots + j_s}}{\partial x_1^{j_1} \cdots \partial x_s^{j_s}}.$$

Let $\hat{\mathbf{x}}$ be an isolated zero of F. The *local array*, $L(k, \hat{\mathbf{x}})$ of, degree k at $\hat{\mathbf{x}}$ is the $t\binom{k+s-1}{s} \times \binom{k+s}{s}$ matrix with columns indexed by the monomials $\mathbf{x}^\mathbf{j}$ of degree $\leq k$ in degree lexicographical order. The rows are indexed by $\mathbf{x}^\mathbf{i} f_\alpha$ and grouped by degree of $\mathbf{x}^\mathbf{i}$. *The entry in the row indexed by* $\mathbf{x}^\mathbf{i} f_\alpha$ *and column indexed by* $\mathbf{x}^\mathbf{j}$ *is* $\partial_{\mathbf{x}^\mathbf{j}}((\mathbf{x}^\mathbf{i} - \hat{\mathbf{x}}^\mathbf{i}) f_\alpha)(\hat{\mathbf{x}})$.

$x_1 - x_2 + x_1^2, x_1 - x_2 + x_2^2$		1	x_1	x_2	x_1^2	$x_1 x_2$	x_2^2	x_1^3	$x_1^2 x_2$	$x_1 x_2^2$	x_2^3
	f_1	0	1	−1	1	0	0	0	0	0	0
$L(1, \hat{x})$	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
$L(2, \hat{x})$	$x_2 f_2$	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
$L(3, \hat{x})$	$x_2^2 f_2$	0	0	0	0	0	0	0	0	1	−1

Multiplicity and the Local Ring

In the exact polynomial case for an isolated zero $\hat{\mathbf{x}}$ the dimensions of the nullspaces of $L(k, \hat{\mathbf{x}})$, $\mathrm{null}(L(k, \hat{\mathbf{x}}))$, stabilize at some integer m as k increases. This is the *multiplicity* of $\hat{\mathbf{x}}$ as originally calculated by Macaulay in 1915. For numerical systems we used the dimension of the approxi-nullspace (eg. [Li-Zeng, 2005] or SVD).

Let d be such that $\mathrm{null}(L(k, \hat{\mathbf{x}}))$ has stabilized for $k \geq d$. For $k > d$ view the rows of $L(k, \hat{\mathbf{x}})$ as polynomials in $\mathbb{C}[x_1, \ldots, x_s]$ by multiplying each entry by its column index. Then the rowspace of $L(k, \hat{\mathbf{x}})$ generates a polynomial ideal $\mathcal{I}$. In the exact case we have:

Lemma

Let $\hat{\mathbf{x}}$ *be an isolated zero of a polynomial system. The ring* $\mathcal{O}_{\hat{\mathbf{x}}} = \mathbb{C}[x_1, \ldots, x_s]/\mathcal{I}$ *is the* local ring *of* F *at* $\hat{\mathbf{x}}$. *In particular* $\dim_{\mathbb{C}} \mathcal{O}_{\hat{\mathbf{x}}} = m$ *where* m *is the multiplicity of* $\hat{\mathbf{x}}$.

Example: $x_1 - x_2 + x_1^2, x_1 - x_2 + x_2^2$

The presentation from $L(3, \hat{\mathbf{x}})$ is not very instructive.
This can be improved by applying the RREF form to $L(3, \hat{\mathbf{x}})$, starting from the bottom right and moving towards the upper left.
I call this the Reverse Reduced Row Echelon Form (RRREF).

1	x_1	x_2	x_1^2	x_1x_2	x_2^2	x_1^3	$x_1^2x_2$	$x_1x_2^2$	x_2^3
0	1	−1	1	0	0	0	0	0	0
0	1	−1	0	1	0	0	0	0	0
0	1	−1	0	0	1	0	0	0	0
0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	1

Thus the presentation is

$$x_1 - x_2 + x_1^2,\ x_1 - x_2 + x_1x_2,\ x_1 - x_2 + x_2^2,\ x_1^3,\ x_1^2x_2,\ x_1x_2^2,\ x_2^3$$

This is a global Gröbner basis! Note also that the 3 non-pivot columns tell us that $\mathcal{O}_{\hat{\mathbf{x}}}$ has $\mathbb{C}$-dimension 3.

ARRREF

If the system is given numerically RRREF will generally give poor results, in particular, an incorrect nullspace. For numeric systems we use ARRREF, Approximate Reduced Row Echelon Form of L for tolerance τ as follows:

- Find row space A of L by SVD, setting singular values less than τ to zero.
- Decide if the i^{th} column of A from the right is approxi-independent from the columns to the right of it, with tolerance τ, if so, that column is a *pivot* column, $i = 2, 3, \ldots, n$, (n is number of columns).
- Form the matrix B of the pivot columns, if B is invertible then ARRREF $= B^{-1}A$

A similar construction for an Approximate Reduced Row Echelon Form has been proposed independently by Robin Scott, a graduate student here at the University of Western Ontario.

Matrix Representation

For a numeric system the resulting "Gröbner basis" has redundant elements which are generally not exactly consistent. It is better to look at the matrix representation.

Classically, any finite dimensional complex associative algebra $\mathcal{A}$ is isomorphic to a matrix algebra as follows: Let $b_1, \ldots, b_m$ be an $\mathbb{C}$-basis for $\mathcal{A}$ and $a \in \mathcal{A}$. Let M_a be the $m \times m$ complex *multiplication* matrix given by

$$M_a \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} ab_1 \\ ab_2 \\ \vdots \\ ab_m \end{bmatrix}$$

Then $a \mapsto M_a^\top$ is an algebra isomorphism. In particular if $ab_i = c_{i1}b_1 + \cdots + c_{im}b_m$ then the i^{th} row of M_a is $[c_{i1}, \ldots, c_{im}]$. This formulation makes it easy to read off the multiplication matrices M_{x_i} from the ARRREF matrix.

Example Continued

Consider the system $x_1 - x_2 + x_1^2, x_1 - x_2 + x_2^2$ at $\hat{\mathbf{x}} = (0, 0)$ above. The basis for $\mathcal{O}_{\hat{\mathbf{x}}}$ is $1, x_1, x_2$. The first part of RRREF, and M_{x_1}, M_{x_2} augmented with column index are

$$\begin{bmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 & x_1^3 \\ 0 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & x_1 & x_2 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & x_1 & x_2 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{bmatrix}$$

$\mathcal{O}_{\hat{\mathbf{x}}}$ is generated as a matrix algebra by I_3, M_{x_1} and M_{x_2}, I_3 the identity matrix. Check that $M_{x_1}^2 = M_{x_1}M_{x_2} = M_{x_2}^2 = -M_{x_1} + M_{x_2}$ and that $M_{x_1}^3 = M_{x_1}^2 M_{x_2} = M_{x_1}M_{x_2}^2 = M_{x_2}^3 = \mathbf{0}$.

Note that because of the nilpotence of M_{x_1}, M_{x_2} all the eigenvalues of these matrices are 0. This is true for the matrix representations of all exact zero dimensional local rings.

Example 3

A more complicated example is given as Example 3 in the paper. Here the system is

$$F = \begin{bmatrix} x^3 + 5z \\ x^2y + y^4 \\ z + 7xy^4 - 6y^5 \end{bmatrix}$$

Here we use $L(7, \mathbf{0})$ to find the local ring at the origin. Although the system is exact, because of the 252×120 matrix it is advisable to use ARRREF rather than RRREF. A tolerance of 1e-06 is sufficient to deduce the exact 108×120 RRREF matrix as the fractions involved have denomenators 2 or 6.

The $\mathbb{C}$-basis obtained is

$$1, x, y, z, x^2, xy, yz, y^2, x^2y, xy^2, y^3, xy^3$$

with non-redundant relations from rows of ARRREF

$$x^3 + 5z,\ y^4 + x^2y,\ z^2,\ xz,\ y^2z,\ 6x^2y^2 + 35yz + z$$

To represent $\mathcal{O}_\mathbf{0}$ as a matrix algebra it is enough to give M_x, M_y as $M_z = -\frac{1}{5}M_x^3$. (Note: this last equation was incorrectly given in the paper). It is hard to see from the relations, but easy from ARRREF that $y^6 \neq 0$ in this local ring but all monomials of degree 7 or higher do vanish, so the nil-index here is 6.

Representation Approach to Möller - Stetter Method

An simplistic approach is the following: If $I_m, M_{x_1}, \ldots, M_{x_s}$ generate a matrix representation by $m \times m$ matrices for the ring $R = \mathbb{C}[x_1, \ldots, x_s]/\langle f_1, \ldots, f_t\rangle$ of $\mathbb{C}$-dimension m, then because these matrices commute they have a joint set of eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$. For any given one of these eigenvectors $\mathbf{v}$ let λ_j be the eigenvalue of M_{x_j} associated with $\mathbf{v}$, i.e. $M_{x_j}\mathbf{v} = \lambda_j\mathbf{v}$. In particular for any monomial in the M_{x_j},

$$M_{x_1}^{e_1} \cdots M_{x_s}^{e_s}\mathbf{v} = \lambda_1^{e_1} \cdots \lambda_s^{e_s}\mathbf{v}$$

For each $\alpha = 1, \ldots, t$ $f_\alpha = 0$ in R so adding over all monomials gives

$$f_\alpha(\lambda_1, \ldots, \lambda_s)\mathbf{v} = f_\alpha(M_{x_1}, \ldots, M_{x_s})\mathbf{v} = 0$$

But $\mathbf{v} \neq \mathbf{0}$ so $(\lambda_1, \ldots, \lambda_s)$ is a zero of the system.

A more sophisticated approach gives the converse, zeros of the system are all of this form.

Solving a numeric system by Möller-Stetter

Using an inexact Gröbner basis such as the one obtained by ARRREF the multiplication matrices may not exactly commute, so one may not exactly have joint eigenvectors. The following works much of the time:

Common Eigenvector Calculation, CEC

- For random numbers $r_1, \ldots, r_s$ let $M = r_1 M_{x_1} + \ldots r_s M_{x_s}$.
- Find eigenvectors of M, normalize to make first coordinate 1.
- For each normalized eigenvector $\mathbf{v}$ solve quadratic least squares problem

$$\begin{bmatrix} M_{x_1}\mathbf{u} - \mu_1\mathbf{u} \\ M_{x_2}\mathbf{u} - \mu_2\mathbf{u} \\ \vdots \\ M_{x_s}\mathbf{u} - \mu_s\mathbf{u} \\ u_1 - 1 \end{bmatrix} = \mathbf{0}$$

  in variables $u_1, \ldots, u_m, \mu_1, \ldots, \mu_s$ where $\mathbf{u} = [u_1, u_2, \ldots, u_m]^\top$. This seems best solved by one iteration of Gauss-Newton using inital values $\mathbf{u} = \mathbf{v}$, μ_i the first coordinate of $M_{x_i}\mathbf{v}$, $i = 1..s$. Return zero $\hat{\mathbf{x}} = [\mu_1, \ldots, \mu_s]^\top$

Opening Example again

$$F = \begin{bmatrix} .5x^2 + 10xy - 25xy^2 - 25x^4 - 166.6667xy^3 - 52.08333xy^4 \\ 5y - 12.50x^2 - 20.8333y^3 + 26.04167x^4 + 26.0416667y^5 \end{bmatrix}$$

$$\hat{\mathbf{x}}_{\mathrm{B}} = (-.7250509\text{e-}11 + .18546461\text{e-}11i \;\; .535604662\text{e-}12 - .14049224\text{e-}12i)$$

Calculating $L(5, \hat{\mathbf{x}})$ gives a 30×21 matrix. Using ARRREF with tolerance $\tau = 1\text{e-}06$ gives 18×21 matrix with non-pivot columns indexed by $1, x, y$. The multiplication matrices are calculated as approximately

$$M_x = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & .40032 \\ 0 & 0 & -.01968 \end{bmatrix}, \quad M_y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & -.01968 \\ 0 & 0 & 0 \end{bmatrix}$$

We see that M_x, M_y each have one non-zero eigenvalue, so this representation is not exactly a local ring. Calculating zeros by the CEC gives a double zero at approximately zero and a zero at approximately $(-0.0196788, 0.0009673)$, this is close to another BERTINI zero $\hat{\mathbf{y}} = (-0.0196765, 0.0009671)$.

Conclusion

- There is a double zero at $\hat{\mathbf{x}} \approx \mathbf{0}$ and a simple zero at $\hat{\mathbf{y}}$ OR
- There is a zero of multiplicity 3 at $\hat{\mathbf{x}}$ on the basis of clustered solutions (see [Kahan, 1972], [Janovitz-Freireich, Ronyai, Szanto, ISSAC '06]).

Numerical Local Ring, NLR

The steps taken in the previous slides produce an NLR:

- Start with an approximate zero $\hat{\mathbf{x}}$ of a numerical system that may indicate multiplicity. It is a good idea to refine using LVZ deflation.
- Calculate local matrices $L(k, \hat{\mathbf{x}})$, estimating numerical co-rank, look for stabilization, e.g. last $\binom{k+s-1}{k}$ columns are independent.
- Apply ARRREF with appropriate tolerance, construct multiplication matrices.
- Find zeros of resulting matrix algebra via CEC.

An NLR is not a local ring in the algebraic sense, in fact not even a ring in a technical sense. Rather it is local in an analytic sense in that all zeros are close to the center $\hat{\mathbf{x}}$.

Motivated by the previous example one can deliberately use a loose tolerance to force a larger dimension of the NLR in an attempt to reveal solutions of the original system near $\hat{\mathbf{x}}$.

An Analytic Example

Consider the system given by Griewank and Osborne (1983) in their study of real Newton's Method. They noticed a multiple zero at the origin as well as a real zero at approximately $(0.1159, 0.0328, -0.0345)$.

$$\begin{aligned} 0.5x^2\cos(10x) + x\sin(10y) + 10x\sinh(z) &= 0 \\ \cos(5x) + \sin(5y) - 1 &= 0 \\ \cosh(5x) + 5\sinh(z) - 1 &= 0 \end{aligned}$$

To force NLR to find zeros in a large neighborhood of the origin, use $L(7, \mathbf{0})$, a 252×120 matrix, apply ARRREF with tolerance 0.08 to get a 113×120 matrix so the NLR is of dimension 7 with basis $1, x, y, z, x^2, xy, xz$. Then CEC verifies that the origin has multiplicity 2 and finds the real zero above. But it also finds close approximations to the following 4 complex zeros.

$$(0.0653 \mp 0.2606i, -0.1534 \mp 0.1401i, 0.1440 \pm 0.0635)$$
$$(-0.1297 \pm 0.0818i, 0.0263 \mp 0.0507i, -0.0235 \pm 0.0553)$$

Griewank-Osborne continued

By observation there are further zeros at $\hat{\mathbf{x}} = (0, \pm\frac{\pi}{5}, 0)$, note $\frac{\pi}{5} \approx 0.6283$. Similar to the opening example starting with $L(5, \hat{\mathbf{x}})$ and tolerance 1e-06 NLR returns two zeros at $\mathbf{0}$ and one at approximately $\hat{\mathbf{y}} = (.00996, -.000247, -.000247)$. Recalling that NLR moves the origin to $\hat{\mathbf{x}}$, the new zero is at $\hat{\mathbf{x}} + \hat{\mathbf{y}} = (.0099595, .62807, -.00024758)$. As above we could regard the multiplicity at $(0, \frac{\pi}{5}, 0)$ as either 2 or 3.

One might question whether it is appropriate to apply this method in calculating the multiplicity of an analytic system. One can prove, using Rouche's Theorem, that the polynomial system consisting of high enough order Taylor approximations of the system gives the correct multiplicity calculation. The advantage of the DZ or NLR method is that the correct order is found without having to verify Rouche Theorem hypotheses.

A cautionary example, Example 1

Consider the system

$$F = \begin{bmatrix} x - \frac{8}{7}y - \frac{1}{7}x^3 + 8x^4 \\ x - \frac{8}{7}y + y^3 \end{bmatrix}$$

There is a zero of multiplicity 3 at the origin. However setting a loose tolerance $L(8, \hat{\mathbf{x}})$ has co-rank 6 but is stable. Applying ARRREF we get a $\mathbb{C}$-basis of $1, x, y, x^2, y^2, x^3$. The multiplication matrices are far from commuting, but we can still apply the CEC giving 3 zeros close to the origin plus

Point	Calculated Point	Nearby Solution
(a)	$\begin{bmatrix} 0.10345 \\ 0.09118 \end{bmatrix}$	$\begin{bmatrix} 0.10345 \\ 0.09118 \end{bmatrix}$
(b)	$\begin{bmatrix} 0.47028 \\ 0.64427 \end{bmatrix}$	$\begin{bmatrix} 0.45802 \\ 0.69682 \end{bmatrix}$
(c)	$\begin{bmatrix} -0.44685 \\ -0.83917 \end{bmatrix}$	no nearby solution

So not every point calculated by the NLR method corresponds to an actual zero of the system. However the point (c) is Euclidean distance 0.95 from $\hat{\mathbf{x}}$ so it is not unreasonable that this local method does not work.

Selected References

- B.H. Dayton, *Numerical Local Rings and Local Solution of Nonlinear Systems*, Proceedings of SNC '07, ACM Press, pp. 79–86, 2007.
- B. H. Dayton and Z. Zeng,*Computing the Multiplicity Structure in Solving Polynomial systems.*, Proc. of ISSAC '05, ACM Press, pp. 116–123.
- A. Griewank and M.R. Osborne, *Analysis of Newton's Method at irregular singularities*. SIAM J. Numer. Anal. 20(4), pp. 747-773, 1983.
- I. Janovitz-Freireich, L. Ronyai, A. Szanto, *Approximate Radical of Ideals with Clusters of Roots*. Proc of ISSAC '06, ACM press, pp. 146–153.
- W. Kahan, *Conserving confluence curbs ill-condition*. Technical report 6, Computer Science, UC Berkeley, 1972.
- T. Y. Li and Z. Zeng, *A rank-revealing method with updating, downdating and applications*, SIAM J. Matrix Anal. Appl., 2005, 26(4), pp. 918-946.
- R. Scott, *Approximate Gröbner bases – a backwards approach*. Master of Science Thesis, Greg Reid supervisor, Graduate Program in Applied Mathematics, University of Western Ontario, London, Ontario, Canada, 2006.