Homotopies to solve Multilinear Systems

Yun Guan   Jan Verschelde

University of Illinois at Chicago
Department of Mathematics, Statistics, and Computer Science
http://www.math.uic.edu/˜guan   http://www.math.uic.edu/˜jan
guan@math.uic.edu   jan@math.uic.edu

ILAS 2010 – minisymposium on Applications of Linear and Multilinear Algebra in Life Sciences and Engineering
Palazzo dei Congressi, Pisa, Italy, 21-25 June 2010

Outline

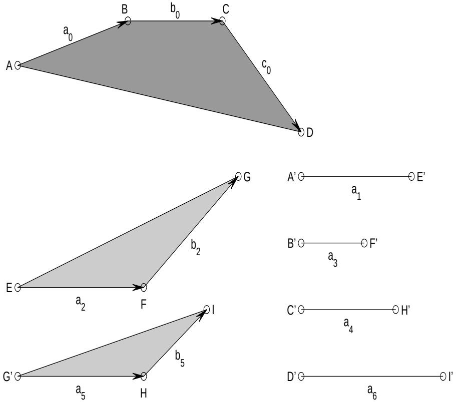
1. Assembling a Seven-Bar Mechanism
   - an application from mechanism design
   - solving a multilinear system
   - numerical representation of a space curve

2. Local Intrinsic Coordinates
   - conditioning of generic points
   - sampling in intrinsic coordinates
   - improving the numerical conditioning

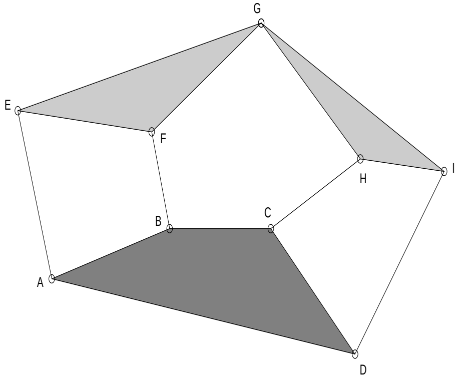
3. A Rescaling Algorithm and the Numerical Stability
   - sampling in local intrinsic coordinates
   - computational results on benchmark systems

Assembling a Seven-Bar Mechanism

an application



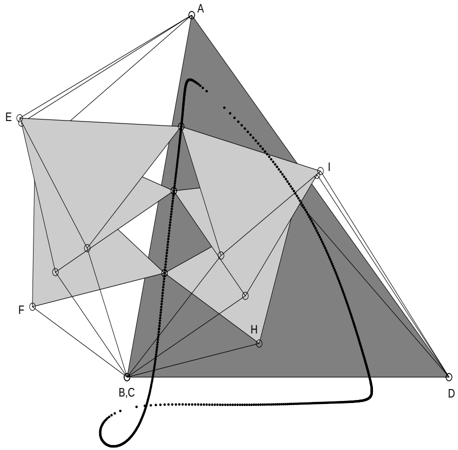
Problem: Find all possible assemblies of these pieces.



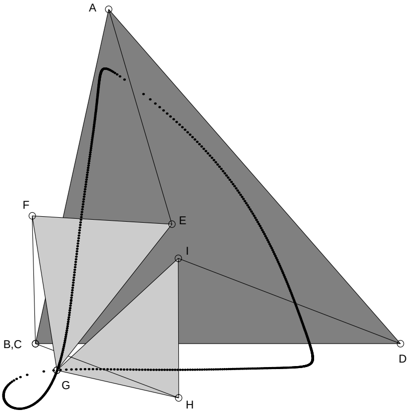
One possible assembly

- Generally, 18 solutions. (This example, 8 real, 10 complex.)
- Intersection of two four-bar coupler curves.

A Moving Seven-Bar Mechanism



Roberts cognate 7-bar moves on a degree-6 curve (coupler curve)
AND . . .



AND . . . has six isolated solutions

- two at each double point of coupler curve
- here, only 1 of 3 double points is real

A Multilinear System

```
 t1*T1 - 1; t2*T2 - 1; t3*T3 - 1;
 t4*T4 - 1; t5*T5 - 1; t6*T6 - 1;

0.71035834160605*t1 + 0.46*t2 - 0.41*t3
 + 0.24076130055512 + 1.07248215701824*i;

(-0.11+0.49*i)*t2 + 0.41*t3
- 0.50219518117959*t4 + 0.41*t5;

0.50219518117959*t4 + (-0.09804347826087
+ 0.43673913043478*i)*t5 - 0.77551855666366*t6 - 1.2;

0.71035834160605*T1 + 0.46*T2 - 0.41*T3
 + 0.24076130055512 - 1.07248215701824*i;

(-0.11-0.49*i)*T2 + 0.41*T3
 - 0.50219518117959*T4 + 0.41*T5;

0.50219518117959*T4 + (-0.09804347826087
- 0.43673913043478*i)*T5 - 0.77551855666366*T6 - 1.2;
```

What does solving mean?

On input is a system of 12 equations in 12 unknows.

We expect a curve of solutions . . .

. . . because there is an assembly that moves.

We also have rigid assemblies: isolated solutions.

Solutions to this system:

1. a curve of degree 6; and
2. 6 isolated solutions.

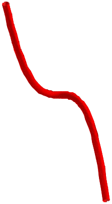
A.J. Sommese, J. Verschelde, and C.W. Wampler: **Numerical decomposition of the solution sets of polynomial systems into irreducible components.** *SIAM J. Numer. Anal.*, 38(6):2022–2046, 2001.

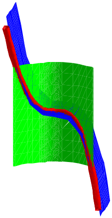
A.J. Sommese and C.W. Wampler. **The Numerical solution of systems of polynomials arising in engineering and science**. World Scientific, 2005.

Representing a Space Curve

Consider the twisted cubic:

$$\begin{cases} y - x^2 = 0 \\ z - x^3 = 0 \end{cases}$$





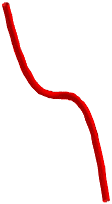
Important attributes are dimension and degree:

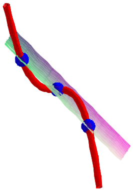
- dimension: cut with one random plane,
- degree: #points on the curve and in the plane.

Witness Set for a Space Curve

Consider the twisted cubic:

$$\begin{cases} y - x^2 = 0 \\ z - x^3 = 0 \end{cases} \qquad \begin{cases} y - x^2 = 0 \\ z - x^3 = 0 \\ c_0 + c_1 x + c_2 y + c_3 z = 0 \end{cases}$$





Intersect with a random plane $c_0 + c_1 x + c_2 y + c_3 z = 0$
$\rightarrow$ find three generic points on the curve.

Generic Points on Algebraic Sets

A polynomial system $f(\mathbf{x}) = \mathbf{0}$ defines an algebraic set $f^{-1}(\mathbf{0}) \subset \mathbb{C}^n$.

We assume

1. $f^{-1}(\mathbf{0})$ is pure dimensional, k is codimension; and moreover
2. $f(\mathbf{x}) = \mathbf{0}$ is a complete intersection, $k = $ #polynomials in f.

For example, consider all adjacent minors of a general 2-by-3 matrix:

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix} \quad f(\mathbf{x}) = \begin{cases} x_{11}x_{22} - x_{21}x_{12} = 0 \\ x_{12}x_{23} - x_{22}x_{13} = 0 \end{cases}$$

$n = 6$, $k = 2$: $\dim(f^{-1}(\mathbf{0})) = n - k = 4$.

To compute $\deg(f^{-1}(\mathbf{0}))$, add $n - k$ general linear equations $L(\mathbf{x}) = \mathbf{0}$ to $f(\mathbf{x}) = \mathbf{0}$ and solve $\{f(\mathbf{x}) = \mathbf{0}, L(\mathbf{x}) = \mathbf{0}\}$.

$\rightarrow$ 4 generic points for all adjacent minors of a general 2-by-3 matrix.

Intrinsic Coordinates save Work

Generic points for all adjacent minors of a general 2-by-3 matrix satisfy (for random coefficients $c_{ij} \in \mathbb{C}$):

$$\left.\begin{array}{r} x_{11}x_{22} - x_{21}x_{12} = 0 \\ x_{12}x_{23} - x_{22}x_{13} = 0 \\ c_{10} + c_{11}x_{11} + c_{12}x_{12} + c_{13}x_{13} + c_{14}x_{21} + c_{15}x_{22} + c_{16}x_{23} = 0 \\ c_{20} + c_{21}x_{11} + c_{22}x_{12} + c_{23}x_{13} + c_{24}x_{21} + c_{25}x_{22} + c_{26}x_{23} = 0 \\ c_{30} + c_{31}x_{11} + c_{32}x_{12} + c_{33}x_{13} + c_{34}x_{21} + c_{35}x_{22} + c_{36}x_{23} = 0 \\ c_{40} + c_{41}x_{11} + c_{42}x_{12} + c_{43}x_{13} + c_{44}x_{21} + c_{45}x_{22} + c_{46}x_{23} = 0 \end{array}\right\}$$

$L^{-1}(\mathbf{0})$ is a 2-plane in $\mathbb{C}^6$, spanned by

$$\begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{21} \\ x_{22} \\ x_{23} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \end{bmatrix} + \xi_1 \begin{bmatrix} v_{11} \\ v_{12} \\ v_{13} \\ v_{14} \\ v_{15} \\ v_{16} \end{bmatrix} + \xi_2 \begin{bmatrix} v_{21} \\ v_{22} \\ v_{23} \\ v_{24} \\ v_{25} \\ v_{26} \end{bmatrix}$$

b is offset point
$\mathbf{v}_1, \mathbf{v}_2$ orthonormal basis

(ξ_1, ξ_2) intrinsic coordinates

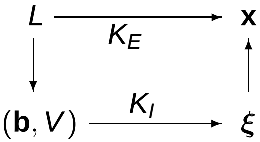
A Commutative Diagram

- $f(\mathbf{x}) = 0$ a system of k polynomials in n variables $\mathbf{x}$,
- $L(\mathbf{x}) = 0$ a system of $n - k$ general linear equations in $\mathbf{x}$,
- $\mathbf{b} \in \mathbb{C}^n$ is offset point, $V = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_k]$, $V^*V = I_k$.

Intrinsic coordinates $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_k)$ for $\mathbf{x}$:

$$\mathbf{x} = \mathbf{b} + \xi_1\mathbf{v}_1 + \xi_2\mathbf{v}_2 + \cdots + \xi_k\mathbf{v}_k = \mathbf{b} + V\boldsymbol{\xi}.$$

Use $f(\mathbf{x} = \mathbf{b} + V\boldsymbol{\xi}) = \mathbf{0}$ to compute generic points:



$$\frac{||\Delta\mathbf{x}||}{||\mathbf{x}||} \leq K_E \frac{||\Delta L||}{||L||}$$

$$\frac{||\Delta\boldsymbol{\xi}||}{||\boldsymbol{\xi}||} \leq K_I \frac{||\Delta(\mathbf{b},V)||}{||(\mathbf{b},V)||}$$

We observe worsening of the numerical conditioning: $K_I \gg K_E$.

Sampling in Intrinsic Coordinates

Represent L via $(\mathbf{b}, V)$ and use intrinsic coordinates $\boldsymbol{\xi} \in \mathbb{C}^k$.

Moving from $(\mathbf{b}, V)$ to $(\mathbf{c}, W)$, as t goes from 0 to 1, homotopy:

$$f\left(\mathbf{x} = \underbrace{(1-t)\mathbf{b} + t\mathbf{c}}_{\text{moving offset point}} + \underbrace{((1-t)V + tW)}_{\text{moving basis vectors}} \boldsymbol{\xi} \right) = \mathbf{0}.$$

Track paths $\boldsymbol{\xi}(t)$ via predictor-corrector methods.

Binomial expansion destroys sparse monomial structure of f.
For example, evaluate $x_1^{a_1} x_2^{a_2}$ at $x_1 = b_1 + \xi_1 v_1$ and $x_2 = b_2 + \xi_2 v_2$:

$$\left(\sum_{i=0}^{a_1} \binom{a_1}{i} b_1^i (\xi_1 v_1)^{a_1 - i} \right) \left(\sum_{j=0}^{a_2} \binom{a_2}{j} b_2^j (\xi_2 v_2)^{a_2 - j} \right).$$

In general: $f(\mathbf{b} + V(\boldsymbol{\xi} + \Delta\boldsymbol{\xi})) = f(\mathbf{b} + V\boldsymbol{\xi}) + \Delta f$, with very large $||\Delta f||$.

Local Intrinsic Coordinates

What if we could keep $||\boldsymbol{\xi}||$ small?

$$
\begin{aligned}
&(b_1 + \xi_1 v_1)^{a_1} (b_2 + \xi_2 v_2)^{a_2} \\
&= \left(b_1^{a_1} + a_1 b_1^{a_1 - 1} \xi_1 v_1 + O(\xi_1^2) \right) \left(b_2^{a_2} + a_2 b_2^{a_2 - 1} \xi_2 v_2 + O(\xi_2^2) \right) \\
&= b_1^{a_1} b_2^{a_2} + a_1 b_1^{a_1 - 1} b_2^{a_2} \xi_1 v_1 + a_2 b_1^{a_1} b_2^{a_2 - 1} \xi_2 v_2 + O(\xi_1^2, \xi_1 \xi_2, \xi_2^2)
\end{aligned}
$$

Now we have: $f(\mathbf{b} + V\boldsymbol{\xi}) = f(\mathbf{b}) + \Delta f$,
where $||\Delta f||$ is $O(||V\boldsymbol{\xi}||) = O(||\boldsymbol{\xi}||)$ as V is orthonormal basis.

Use extrinsic coordinates of generic point as offset point for k-plane: for $d = \deg(f^{-1}(\mathbf{0}))$ and d generic points $\{\mathbf{z}_1, \mathbf{z}_1, \ldots, \mathbf{z}_d\}$:

$$
\mathbf{x} = \mathbf{z}_\ell + V\boldsymbol{\xi}, \quad \ell = 1, 2, \ldots, d.
$$

The local intrinsic coordinates are defined by $(\{\mathbf{z}_1, \mathbf{z}_1, \ldots, \mathbf{z}_d\}, V)$.

Improved Numerical Conditioning

Condition number K_E of zero $\mathbf{z}$ of $F(\mathbf{x}) = \left\{ \begin{array}{l} f(\mathbf{x}) = \mathbf{0} \\ L(\mathbf{x}) = \mathbf{0} \end{array} \right.$:

$$\underbrace{F'(\mathbf{z})}_{=A} \Delta \mathbf{z} = -f(\mathbf{z}), \quad K_E = \kappa(A),$$

where $\kappa(A)$ is the condition number of the Jacobian matrix A of F at $\mathbf{z}$.

In local intrinsic coordinates, $\mathbf{x} = \mathbf{z} + V\xi$:

$$\underbrace{f'(\xi)}_{=B} \Delta \xi = -f(\xi), \quad K_{LI} = \kappa(B),$$

where $\kappa(A)$ is the condition number of the Jacobian matrix B of f at ξ and K_{LI} is the condition number for the local intrinsic coordinates.

$$\xi = \mathbf{0} \leftrightarrow \mathbf{x} = \mathbf{z} \text{ and } f' \subset F' \quad \Rightarrow \quad K_{LI} \leq K_E$$

Sampling in Local Intrinsic Coordinates

Generic points $\{\mathbf{z}_1, \mathbf{z}_1, \ldots, \mathbf{z}_d\}$ are offset points for k-plane L with directions in the orthonormal matrix V.

Moving from $(\mathbf{z}_\ell, V)$ to $(\mathbf{b}, W)$, as t goes from 0 to 1, homotopy:

$$f(\mathbf{x} = (1-t)\mathbf{z}_\ell + t\mathbf{b} + W\xi) = \mathbf{0}$$

$\rightarrow$ *only the offset point moves!*

Instead of moving to $\mathbf{b}$, let $\mathbf{c}$ be the orthogonal projection of $\mathbf{z}_\ell$ onto the k-plane L.

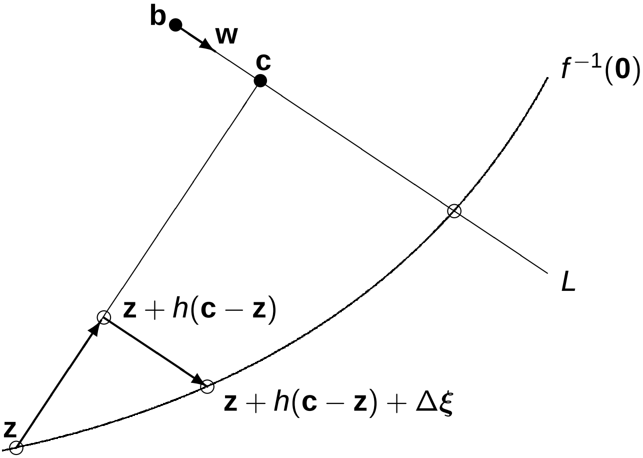
For some step size h, consider:

$$f(\mathbf{x} = \mathbf{z}_\ell + h(\mathbf{c} - \mathbf{z}_\ell) + W\xi) = \mathbf{0}$$

and apply Newton's method to find the correction $\Delta\xi$.

Schematic of the new Sampling Algorithm

one predictor-corrector step



pseudocode for one predictor-corrector step

Input: $\mathbf{b} \in \mathbb{C}^n$, $W = [\mathbf{w}_1 \ \mathbf{w}_2 \ \cdots \ \mathbf{w}_k] \in \mathbb{C}^{n \times k}$, $W^* W = I_k$
$\mathbf{z} \in \mathbb{C}^n$, $f(\mathbf{z}) = \mathbf{0}$, $K(\mathbf{z}) = \mathbf{0}$, $h > 0$, $\epsilon > 0$, some L.

Output: $\widehat{\mathbf{z}}$, $f(\widehat{\mathbf{z}}) = \mathbf{0}$: $\widehat{\mathbf{z}}$ closer to L.

$$\mathbf{v} := \mathbf{z} - \mathbf{b}; \qquad \mathbf{v} := \mathbf{v} - \sum_{i=1}^{k} (\overline{\mathbf{w}_i}^T \mathbf{v}) \mathbf{w}_i; \qquad \mathbf{v} := \mathbf{v} / ||\mathbf{v}||;$$
$$\widetilde{\mathbf{z}} := \mathbf{z} + h\,\mathbf{v}; \qquad \widehat{\mathbf{z}} := \widetilde{\mathbf{z}}; \qquad \xi := \mathbf{0};$$

while $||f(\widehat{\mathbf{z}} + W\xi)|| > \epsilon$ do
$\quad \Delta\xi := f(\widehat{\mathbf{z}} + W\xi) / f'(\widehat{\mathbf{z}} + W\xi)$;
$\quad \xi := \xi + \Delta\xi$.

Numerical Stability

For some step size h, we evaluate

$$f(\mathbf{x} = \mathbf{z}_\ell + h(\mathbf{c} - \mathbf{z}_\ell)) = f(\mathbf{z}_\ell) + O(h) = O(h).$$

If step size h is too large, then Newton is unlikely to converge.

If step size h is too large, then $f(\mathbf{x} = \mathbf{z}_\ell + h(\mathbf{c} - \mathbf{z}_\ell)) \gg h$.

If $f(\mathbf{x} = \mathbf{z}_\ell + h(\mathbf{c} - \mathbf{z}_\ell)) \gg h$, then reduce h immediately.

Do not wait for (costly) Newton corrector to fail.

We can control size of residual $||f(\xi)||$ *to be always* $O(h)$.

Implementation and Benchmark Systems

Available since version 2.3.53 of PHCpack

Algorithm 795: PHCpack: A general-purpose solver for polynomial systems by homotopy continuation. *ACM Trans. Math. Softw.*, 25(2):251–276, 1999.
`http://www.math.uic.edu/~jan/download.html`

Three classes, families of systems:

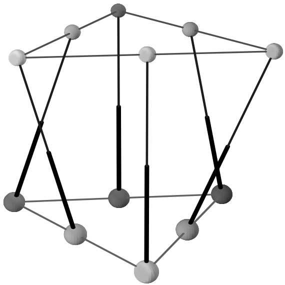
1. all adjacent minors of a general 2-by-n matrix, $n = 3, 4, \ldots, 13$
2. cyclic n-roots, $n = 4, 8, 9$ (an academic benchmark)
3. Griffis-Duffy platforms and other systems from mechanical design

Computational experimental setup:

- given one set of generic points, generate another random k-plane
- move the given set of generic points to the new random k-plane
- check results for accuracy, #predictor-corrector steps, timings

Architecturally Singular Platforms Move

M. Griffis and J. Duffy: **Method and apparatus for controlling geometrically simple parallel mechanisms with distinctive connections.** US Patent 5,179,525, 1993.



end plate, the platform

is connected by legs to

a stationary base

- Base and endplate are equilateral triangles.
- Legs connect vertices to midpoints.

Computational Results

Characteristics of three families of polynomial systems:

	polynomial system	n	$n-k$	d
1	Griffis-Duffy platform	8	1	40
2	cyclic 8-roots system	8	1	144
3	all adjacent minors of 2-by-11 matrix	22	12	1,024

n: number of variables, k: codimension, d: degree

Sampling in global intrinsic/local intrinsic coordinates:

system	#iterations	timings
1	207/164	550/535 μsec
2	319/174	5.3/3.2 sec
3	285/219	44.6/40.3 sec

Done on a Mac OS X 3.2 Ghz Intel Xeon, using 1 core.

Conclusions

Advantages of using local intrinsic coordinates:

- only offset point moves during sampling
- keep sparse structure of the polynomials
- control step size by evaluation

Applications to numerical algebraic geometry:

- implicitization via interpolation
- monodromy breakup algorithm
- diagonal homotopies to intersect solution sets