Back to the Roots: Solving Polynomial Systems with Numerical Linear Algebra Tools

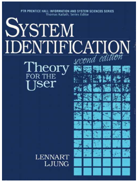
Bart De Moor

Katholieke Universiteit Leuven
Department of Electrical Engineering ESAT/SCD

Outline

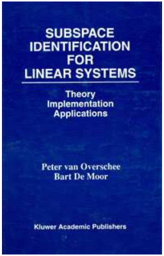
System Identification: PEM

- LTI models
- Non-convex optimization
- Considered 'solved' early nineties



Linear Algebra approach

⇒ **Subspace methods**



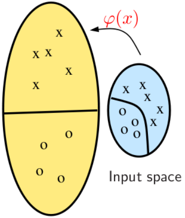
Nonlinear regression, modelling and clustering

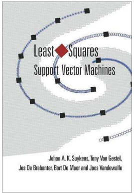
- Most regression, modelling and clustering problems are nonlinear when formulated in the input data space
- This requires nonlinear nonconvex optimization algorithms

Linear Algebra approach

⇒ **Least Squares Support Vector Machines**

- 'Kernel trick' = projection of input data to a high-dimensional feature space
- Regression, modelling, clustering problem becomes a large scale linear algebra problem (set of linear equations, eigenvalue problem)





Nonlinear Polynomial Optimization

- Polynomial object function + polynomial constraints
- Non-convex
- Computer Algebra, Homotopy methods, Numerical Optimization
- Considered 'solved' by mathematics community

Linear Algebra Approach

⇒ **Linear Polynomial Algebra**

Conceptual/Geometric Level

- Polynomial system solving is an eigenvalue problem!
- Row and Column Spaces: Ideal/Variety $\leftrightarrow$ Row space/Kernel of M, ranks and dimensions, nullspaces and orthogonality
- Geometrical: intersection of subspaces, angles between subspaces, Grassmann's theorem,...

Numerical Linear Algebra Level

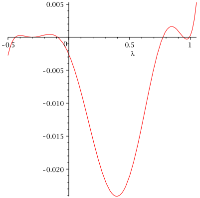
- Eigenvalue decompositions, SVDs,...
- Solving systems of equations (consistency, nb sols)
- QR decomposition and Gram-Schmidt algorithm

Numerical Algorithms Level

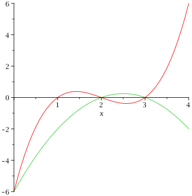
- Modified Gram-Schmidt (numerical stability), GS 'from back to front'
- Exploiting sparsity and Toeplitz structure (computational complexity $O(n^2)$ vs $O(n^3)$), FFT-like computations and convolutions,...
- Power method to find smallest eigenvalue (= minimizer of polynomial optimization problem)

Four instances of polynomial rooting problems

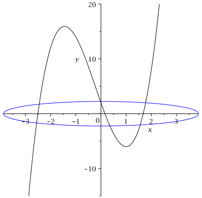
$$p(\lambda) = \det(A - \lambda I) = 0$$



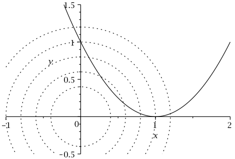
$$\begin{aligned} (x-1)(x-3)(x-2) &= 0 \\ -(x-2)(x-3) &= 0 \end{aligned}$$



$$\begin{aligned} x^2 + 3y^2 - 15 &= 0 \\ y - 3x^3 - 2x^2 + 13x - 2 &= 0 \end{aligned}$$



$$\begin{aligned} \min_{x,y} \quad & x^2 + y^2 \\ \text{s.t.} \quad & y - x^2 + 2x - 1 = 0 \end{aligned}$$

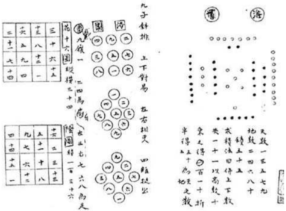


Outline

Solving Polynomial Systems: a long and rich history. . .

Diophantus (c200-c284) *Arithmetica*

Al-Khwarizmi (c780-c850)



Zhu Shijie (c1260-c1320) *Jade Mirror of the Four Unknowns*

Pierre de Fermat (c1601-1665)

René Descartes (1596-1650)

Isaac Newton (1643-1727)

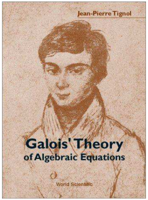
Gottfried Wilhelm Leibniz (1646-1716)

. . . leading to "Algebraic Geometry"

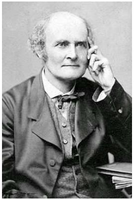
Etienne Bézout (1730-1783)

Jean-Victor Poncelet (1788-1867)

August Ferdinand Möbius (1790-1868)



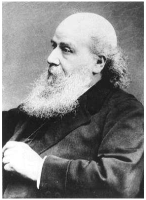
Evariste Galois (1811-1832)



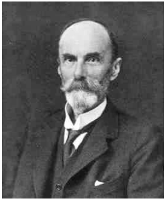
Arthur Cayley (1821-1895)

Leopold Kronecker (1823-1891)

Edmond Laguerre (1834-1886)



James Joseph Sylvester (1814-1897)

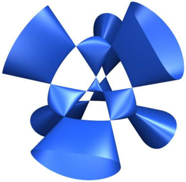


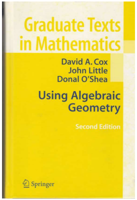
Francis Sowerby Macaulay (1862-1937)

David Hilbert (1862-1943)

Computational Algebraic Geometry

- Emphasis on symbolic manipulations
- Computer algebra
- Huge body of literature in Algebraic Geometry
- Computational tools: Gröbner Bases (next slide)





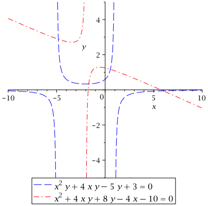
Wolfgang Gröbner
(1899-1980)

Bruno Buchberger

Example: Gröbner basis

Input system:

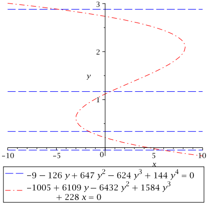
$$\begin{aligned} x^2y + 4xy - 5y + 3 &= 0 \\ x^2 + 4xy + 8y - 4x - 10 &= 0 \end{aligned}$$



- Generates simpler but equivalent system (same roots)
- Symbolic eliminations and reductions
- Monomial ordering (e.g., lexicographic)
- Exponential complexity
- Numerical issues! Coefficients become very large

Gröbner Basis:

$$\begin{aligned} -9 - 126y + 647y^2 - 624y^3 + 144y^4 &= 0 \\ -1005 + 6109y - 6432y^2 + 1584y^3 + 228x &= 0 \end{aligned}$$

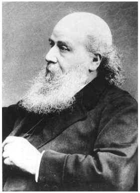


Outline

$$\underset{p\times q}{A}\ \underset{q\times(q-r)}{X} = \underset{p\times(q-r)}{0}$$

- $C(A^T) \perp C(X)$
- $\text{rank}(A) = r$
- $\dim N(A) = q - r = \text{rank}(X)$

$$\begin{aligned} A &= \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \\ &\Downarrow \\ X &= V_2 \end{aligned}$$



James Joseph Sylvester

$$\underset{p\times q}{A} \quad \underset{q\times(q-r)}{X} \quad = \quad \underset{p\times(q-r)}{0}$$

Reorder columns of A and partition

$$\overset{p\times q}{\overline{A}} = \left[\overset{p\times(q-r)}{\overline{A_1}} \quad \overset{p\times r}{\overline{A_2}}\right] \qquad \text{rank}(\overline{A_2}) = r \quad (\overline{A_2} \text{ full column rank})$$

Reorder rows of X and partition accordingly

$$\left[\overline{A_1} \quad \overline{A_2}\right] \overset{q-r}{\left[\begin{array}{c} \overline{X_1} \\ \hline \overline{X_2} \end{array}\right]} \begin{array}{l} q-r \\ r \end{array} = 0$$

$$\boxed{\begin{array}{ccc} \text{rank}(\overline{A_2}) & = & r \\ & \Updownarrow & \\ \text{rank}(\overline{X_1}) & = & q-r \end{array}}$$

$$\begin{bmatrix} \overline{A_1} & \overline{A_2} \end{bmatrix} \overset{q-r}{\begin{bmatrix} \overline{X_1} \\ \overline{X_2} \end{bmatrix}} \begin{matrix} q-r \\ r \end{matrix} = 0$$

- $\overline{X_1}$: independent variables
- $\overline{X_2}$: dependent variables

$$\begin{aligned} \overline{X_2} &= -\overline{A_2}^{\dagger}\, \overline{A_1}\, \overline{X_1} \\ \overline{A_1} &= -\overline{A_2}\, \overline{X_2}\, \overline{X_1}^{-1} \end{aligned}$$

- Number of different ways of choosing r linearly independent columns out of q columns (upper bound):

$$\binom{q}{q-r} = \frac{q!}{(q-r)!\; r!}$$

$$\underset{p\times q}{A}\ \underset{q\times(q-r_A)}{X} = \underset{p\times(q-r_A)}{0} \quad \text{and} \quad \underset{p\times t}{B}\ \underset{t\times(t-r_B)}{Y} = \underset{p\times(t-r_B)}{0}$$

What is the nullspace of $[\, A \quad B \,]$?

$$[\, A \quad B \,] \begin{bmatrix} \overset{q-r_A}{X} & \overset{t-r_B}{0} & \overset{?}{?} \\ 0 & Y & ? \end{bmatrix} = 0$$

Let $\mathrm{rank}([\, A \quad B \,]) = r_{AB}$

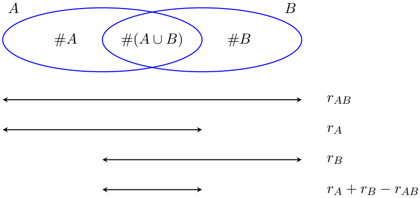
$$(q - r_A) + (t - r_B) + ? = (p + t) - r_{AB} \quad \Rightarrow \quad ? = r_A + r_B - r_{AB}$$

KATHOLIEKE UNIVERSITEIT LEUVEN ESAT SCD

$$[\,A \quad B\,]\begin{bmatrix} \overset{q-r_A}{X} & \overset{t-r_B}{0} & \overset{r_A+r_B-r_{AB}}{Z_1} \\ 0 & Y & Z_2 \end{bmatrix} = 0$$

Intersection between column space of A and B:

$$AZ_1 = -BZ_2$$



Hermann Grassmann

$$\#(A\cup B)=\#A+\#B-\#(A\cap B)$$

- **Characteristic Polynomial**
  The eigenvalues of A are the roots of

$$p(\lambda) = \det(A - \lambda I) = 0$$

- **Companion Matrix**
  Solving

$$q(x) = 7x^3 - 2x^2 - 5x + 1 = 0$$

leads to

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1/7 & 5/7 & 2/7 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} = x \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix}$$

Consider the univariate equation

$$x^3 + a_1 x^2 + a_2 x + a_3 = 0,$$

having three distinct roots x_1, x_2 and x_3

$$\begin{bmatrix} a_3 & a_2 & a_1 & 1 & 0 & 0 \\ 0 & a_3 & a_2 & a_1 & 1 & 0 \\ 0 & 0 & a_3 & a_2 & a_1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \\ x_1^3 & x_2^3 & x_3^3 \\ x_1^4 & x_2^4 & x_3^4 \\ x_1^5 & x_2^5 & x_3^5 \end{bmatrix} = 0$$

- Homogeneous linear system
- Rectangular Vandermonde
- $\text{corank} = 3$
- Observability matrix-like
- Realization theory!

Consider

$$\begin{aligned} x^3 + a_1 x^2 + a_2 x + a_3 &= 0 \\ x^2 + b_1 x + b_2 &= 0 \end{aligned}$$

Build the Sylvester Matrix:

$$\left[\begin{array}{ccccc} 1 & a_1 & a_2 & a_3 & 0 \\ 0 & 1 & a_1 & a_2 & a_3 \\ \hline 1 & b_1 & b_2 & 0 & 0 \\ 0 & 1 & b_1 & b_2 & 0 \\ 0 & 0 & 1 & b_1 & b_2 \end{array}\right] \left[\begin{array}{c} 1 \\ x \\ x^2 \\ x^3 \\ x^4 \end{array}\right] = 0$$

Row Space	Null Space
Ideal =union of ideals =multiply rows with powers of x	**Variety** =intersection of null spaces

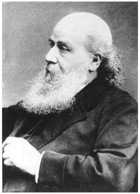
- Corank of Sylvester matrix = number of common zeros
- null space = intersection of null spaces of two Sylvester matrices
- common roots follow from realization theory in null space
- notice 'double' Toeplitz-structure of Sylvester matrix

- **Sylvester Resultant**
  Consider two polynomials $f(x)$ and $g(x)$:

$$f(x) = x^3 - 6x^2 + 11x - 6 = (x-1)(x-2)(x-3)$$
$$g(x) = -x^2 + 5x - 6 = -(x-2)(x-3)$$

Common roots iff $S(f,g) = 0$

$$S(f,g) = \det \left[\begin{array}{rrrrr} -6 & 11 & -6 & 1 & 0 \\ 0 & -6 & 11 & -6 & 1 \\ \hline -6 & 5 & -1 & 0 & 0 \\ 0 & -6 & 5 & -1 & 0 \\ 0 & 0 & -6 & 5 & -1 \end{array} \right]$$



James Joseph Sylvester

The corank of the Sylvester matrix is 2!

Sylvester's construction can be understood from

$$
\begin{array}{l} f(x)=0 \\ x\cdot f(x)=0 \\ g(x)=0 \\ x\cdot g(x)=0 \\ x^2\cdot g(x)=0 \end{array}
\quad
\begin{array}{c} \begin{array}{ccccc} 1 & x & x^2 & x^3 & x^4 \end{array} \\
\begin{bmatrix} -6 & 11 & -6 & 1 & 0 \\ & -6 & 11 & -6 & 1 \\ -6 & 5 & -1 & & \\ & -6 & 5 & -1 & \\ & & -6 & 5 & -1 \end{bmatrix} \end{array}
\begin{bmatrix} 1 & 1 \\ x_1 & x_2 \\ x_1^2 & x_2^2 \\ x_1^3 & x_2^3 \\ x_1^4 & x_2^4 \end{bmatrix} = 0
$$

where $x_1 = 2$ and $x_2 = 3$ are the common roots of f and g

The vectors in the canonical kernel K obey a 'shift structure':

$$\begin{bmatrix} 1 \\ x \\ x^2 \\ x^3 \end{bmatrix} x = \begin{bmatrix} x \\ x^2 \\ x^3 \\ x^4 \end{bmatrix}$$

The canonical kernel K is not available directly, instead we compute Z, for which $ZV = K$. We now have

$$\begin{aligned} S_1 K D &= S_2 K \\ S_1 Z V D &= S_2 Z V \end{aligned}$$

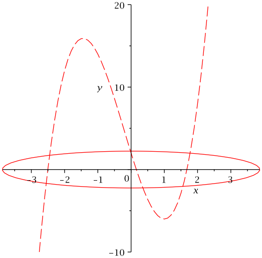
leading to the generalized eigenvalue problem

$$(S_2 Z) V = (S_1 Z) V D$$

Outline

- Consider

$$\begin{cases} p(x,y) & = & x^2 + 3y^2 - 15 = 0 \\ q(x,y) & = & y - 3x^3 - 2x^2 + 13x - 2 = 0 \end{cases}$$



- Fix a monomial order, e.g., $1 < x < y < x^2 < xy < y^2 < x^3 < x^2y < \ldots$

- Construct M: write the system in matrix-vector notation:

	1	x	y	x^2	xy	y^2	x^3	x^2y	xy^2	y^3
$p(x,y)$	-15			1		3				
$q(x,y)$	-2	13	1	-2			-3			
$x \cdot p(x,y)$		-15					1		3	
$y \cdot p(x,y)$			-15					1		3

$$\begin{cases} p(x,y) &= x^2 + 3y^2 - 15 = 0 \\ q(x,y) &= y - 3x^3 - 2x^2 + 13x - 2 = 0 \end{cases}$$

Continue to enlarge M:

it #	form	1	x	y	x^2	xy	y^2	x^3	x^2y	xy^2	y^3	x^4	x^3y	x^2y^2	xy^3	y^4	x^5	x^4y	x^3y^2	x^2y^3	xy^4	y^5	→
$d=3$	p	-15			1		3																
	xp		-15					1		3													
	yp			-15					1		3												
	q	-2	13	1	-2			-3															
$d=4$	x^2p				-15							1		3									
	xyp					-15							1		3								
	y^2p						-15							1		3							
	xq		-2		13	1		-2				-3											
	yq			-2		13	1		-2				-3										
$d=5$	x^3p							-15									1		3				
	x^2yp								-15									1		3			
	xy^2p									-15									1		3		
	y^3p										-15									1		3	
	x^2q				-2			13	1			-2					-3						
	xyq					-2			13	1			-2					-3					
	y^2q						-2			13	1			-2					-3				
	↓							⋱	⋱	⋱	⋱	⋱	⋱	⋱	⋱	⋱	⋱	⋱	⋱	⋱	⋱	⋱	⋱

- # rows grows faster than # cols ⇒ overdetermined system
- rank deficient by construction!

- Coefficient matrix M:

$$M = \begin{bmatrix} \times & \times & \times & \times & 0 & 0 & 0 \\ 0 & \times & \times & \times & \times & 0 & 0 \\ 0 & 0 & \times & \times & \times & \times & 0 \\ 0 & 0 & 0 & \times & \times & \times & \times \end{bmatrix}$$

- Solutions generate vectors in kernel of M:

$$Mk = 0$$

- Number of solutions s follows from corank

Canonical nullspace K built from s solutions (x_i, y_i):

$$\left[\begin{array}{c|c|c|c} 1 & 1 & \ldots & 1 \\ \hline x_1 & x_2 & \ldots & x_s \\ y_1 & y_2 & \ldots & y_s \\ \hline x_1^2 & x_2^2 & \ldots & x_s^2 \\ x_1 y_1 & x_2 y_2 & \ldots & x_s y_s \\ y_1^2 & y_2^2 & \ldots & y_s^2 \\ \hline x_1^3 & x_2^3 & \ldots & x_s^3 \\ x_1^2 y_1 & x_2^2 y_2 & \ldots & x_s^2 y_s \\ x_1 y_1^2 & x_2 y_2^2 & \ldots & x_s y_s^2 \\ y_1^3 & y_2^3 & \ldots & y_s^3 \\ \hline x_1^4 & x_2^4 & \ldots & x_4^4 \\ x_1^3 y_1 & x_2^3 y_2 & \ldots & x_s^3 y_s \\ x_1^2 y_1^2 & x_2^2 y_2^2 & \ldots & x_s^2 y_s^2 \\ x_1 y_1^3 & x_2 y_2^3 & \ldots & x_s y_s^3 \\ y_1^4 & y_2^4 & \ldots & y_s^4 \\ \hline \vdots & \vdots & \vdots & \vdots \end{array}\right]$$

- Choose s linear independent rows in K

$$S_1 K$$

- This corresponds to finding linear dependent columns in M

1	1	$\ldots$	1
x_1	x_2	$\ldots$	x_s
y_1	y_2	$\ldots$	y_s
x_1^2	x_2^2	$\ldots$	x_s^2
$x_1 y_1$	$x_2 y_2$	$\ldots$	$x_s y_s$
y_1^2	y_2^2	$\ldots$	y_s^2
x_1^3	x_2^3	$\ldots$	x_s^3
$x_1^2 y_1$	$x_2^2 y_2$	$\ldots$	$x_s^2 y_s$
$x_1 y_1^2$	$x_2 y_2^2$	$\ldots$	$x_s y_s^2$
y_1^3	y_2^3	$\ldots$	y_s^3
x_1^4	x_2^4	$\ldots$	x_4^4
$x_1^3 y_1$	$x_2^3 y_2$	$\ldots$	$x_s^3 y_s$
$x_1^2 y_1^2$	$x_2^2 y_2^2$	$\ldots$	$x_s^2 y_s^2$
$x_1 y_1^3$	$x_2 y_2^3$	$\ldots$	$x_s y_s^3$
y_1^4	y_2^4	$\ldots$	y_s^4
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Shift property in monomial basis

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix} x = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix} y = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}$$

- Finding the x-roots: let $D = \mathrm{diag}(x_1, x_2, \ldots, x_s)$, then

$$S_1 K D = S_2 K,$$

  where S_1 and S_2 select rows from K wrt. shift property
- Reminiscent of **Realization Theory**

Nullspace of M

Find a basis for the nullspace of M using an SVD:

$$M = \begin{bmatrix} \times & \times & \times & \times & 0 & 0 & 0 \\ 0 & \times & \times & \times & \times & 0 & 0 \\ 0 & 0 & \times & \times & \times & \times & 0 \\ 0 & 0 & 0 & \times & \times & \times & \times \end{bmatrix} = \begin{bmatrix} X & Y \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} W^T \\ Z^T \end{bmatrix}$$

Hence,

$$MZ = 0$$

We have

$$S_1 K D = S_2 K$$

However, K is not known, instead a basis Z is computed as

$$ZV = K$$

Which leads to

$$(S_2 Z)V = (S_1 Z)VD$$

Algorithm

1 Fix a monomial ordering scheme

2 Construct coefficient matrix M

3 Compute basis for nullspace of M, Z

4 Find s linear independent rows in Z

5 Choose shift function, e.g., x

6 Write down shift relation in monomial basis k for the chosen shift function using row selection matrices S_1 and S_2

7 The construction of above gives rise to a generalized eigenvalue problem

$$(S_2 Z)V = (S_1 Z)VD$$

of which the eigenvalues correspond to the, e.g., x-solutions of the system of polynomial equations.

8 Reconstruct canonical kernel $K = ZV$

Data-Driven root-finding

- Dual version of Kernel-based root-finding
- All operations are done on coefficient matrix M
  - Find linear dependent columns of M instead of linear independent rows of K (corank)
  - Write down eigenvalue problem in terms of partitiong of M
  - Allows sparse representation of M
  - Rank-revealing QR instead of SVD

$$\begin{cases} p(x,y) &= x^2 + 3y^2 - 15 = 0 \\ q(x,y) &= y - 3x^3 - 2x^2 + 13x - 2 = 0 \end{cases}$$

Finding linear dependent columns of M

	1	x	y	x^2	xy	y^2	x^3	x^2y	xy^2	y^3	...
p	-15			1		3					
xp		-15					1		3		
yp			-15					1		3	
q	-2	13	1	-2			-3				
x^2p				-15							
xyp					-15						
y^2p						-15					
xq		-2		13	1		-2				
yq			-2		13	1		-2			
x^3p							-15				
x^2yp								-15			
xy^2p									-15		
y^3p										-15	
x^2q				-2			13	1			
xyq					-2			13	1		
y^2q						-2			13	1	
$\vdots$							$\ddots$	$\ddots$	$\ddots$	$\ddots$	

$$\begin{cases} p(x,y) &= x^2+3y^2-15=0 \\ q(x,y) &= y-3x^3-2x^2+13x-2=0 \end{cases}$$

- Writing down the eigenvalue problem in terms of a **re-ordered** partitioning of M
- all linear dependent columns of M corresponding with monomials of the lowest possible degree are grouped in M_1

$$M = \begin{bmatrix} \times & \times & \times & \times & 0 & 0 & 0 \\ 0 & \times & \times & \times & \times & 0 & 0 \\ 0 & 0 & \times & \times & \times & \times & 0 \\ 0 & 0 & 0 & \times & \times & \times & \times \end{bmatrix} = [M_1 \; M_2]$$

$$[M_1 \quad M_2] \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} = 0$$

$$K_2 = -M_2^{\dagger} \, M_1 \, K_1$$

($\dagger$: Moore-Penrose pseudoinverse)

$$\begin{cases} p(x,y) &= x^2 + 3y^2 - 15 = 0 \\ q(x,y) &= y - 3x^3 - 2x^2 + 13x - 2 = 0 \end{cases}$$

Writing down the eigenvalue problem in terms of a partitioning of M

$$K_1 \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_s \end{bmatrix} = S_x \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}$$

$$K_1 \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_s \end{bmatrix} = S_x \begin{bmatrix} I_{tcr} \\ -M_2^{\dagger} M_1 \end{bmatrix} K_1$$

There are 3 kinds of roots:

1. Roots in zero
2. Finite nonzero roots
3. Roots at infinity

Applying Grassmann's Dimension theorem on the Kernel allows to write the following partitioning

$$[M_1 \; M_2] \begin{bmatrix} X_1 & 0 & X_2 \\ 0 & Y_1 & Y_2 \end{bmatrix} = 0$$

- X_1 corresponds with the roots in zero (multiplicities included!)
- Y_1 corresponds with the roots at infinity (multiplicities included!)
- $[X_2; Y_2]$ corresponds with the finite nonzero roots (multiplicities included!)

Roots at infinity: univariate case

$$0\,x^2 + x - 2 = 0$$

transform $x \rightarrow \frac{1}{X}$

$$\Rightarrow X(1 - 2X) = 0$$

- 1 affine root $x = 2$ $(X = \frac{1}{2})$
- 1 root at infinity $x = \infty$ $(X = 0)$

Roots at infinity: multivariate case

$$\begin{cases} (x-2)y &= 0 \\ y - 3 &= 0 \end{cases}$$

transform $x \rightarrow \frac{X}{T}$, $y \rightarrow \frac{Y}{T}$

$$\Rightarrow \begin{cases} XY - 2YT &= 0 \\ Y - 3T &= 0 \end{cases}$$

- 1 affine root (2,3,1) $(T = 1)$
- 1 root at infinity (1,0,0) $(T = 0)$

General Canonical null space K

- Multiplicities of roots $\rightarrow$ multiplicity structure of kernel K
- Partial derivatives

$$\partial_{j_1 j_2 \ldots j_s} \equiv \partial_{x_1^{j_1} x_2^{j_2} \ldots x_s^{j_s}} \equiv \frac{1}{j_1! j_2! \ldots j_s!} \frac{\partial^{j_1+j_2+\ldots+j_s}}{\partial x_1^{j_1} \partial x_2^{j_2} \ldots \partial x_s^{j_s}}$$

  needed to describe extra columns of K
- Currently investigating technicalities
- Possibility of trading in multiplicities for extra equations (Radical Ideal)

Univariate case

$$f(x) = (x-1)^3 = 0$$

triple root in $x = 1$: $f'(1) = 0$ and $f''(1) = 0$

$$f \begin{bmatrix} 1 & 3 & -3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ x & 1 & 0 \\ x^2 & 2x & 1 \\ x^3 & 3x^2 & 3x \end{bmatrix} = 0$$

or

$$\begin{matrix} \partial_0 f \\ \partial_1 f \\ \partial_2 f \end{matrix} \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \\ x^3 \end{bmatrix} = 0$$

Multivariate case

- Polynomial system in 2 unknowns (x, y) with
  - 1 affine root $z_1 = (x_1, y_1)$ with multiplicity 3: $[\partial_{00}|_{z_1} \ \partial_{10}|_{z_1} \ \partial_{01}|_{z_1}]$
  - 1 root $z_2 = (x_2, y_2)$ at infinity: $\partial_{00}|_{z_2}$
  - M matrix of degree 4

Canonical Kernel K

$$K = \begin{array}{c} \\ \end{array}\begin{array}{cccc} \partial_{00}|_{z_1} & \partial_{10}|_{z_1} & \partial_{01}|_{z_1} & \partial_{00}|_{z_2} \end{array}$$

$$K = \begin{pmatrix} 1 & 0 & 0 & 0 \\ x_1 & 1 & 0 & 0 \\ y_1 & 0 & 1 & 0 \\ x_1^2 & 2x_1 & 0 & 0 \\ x_1y_1 & y_1 & x_1 & 0 \\ y_1^2 & 0 & 2y_1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ x_1^4 & 4x_1 & 0 & 0 \\ x_1^3y_1 & 3x_1^2y_1 & x_1^3 & 1 \\ x_1^2y_1^2 & 2x_1y_1^2 & 2x_1^2y_1 & 0 \\ x_1y_1^3 & y_1^3 & 3x_1y_1^2 & 0 \\ y_1^4 & 0 & 4y_1^3 & 0 \end{pmatrix}$$

Polynomial Optimization Problems

If
$$A_1 b = xb$$
and
$$A_2 b = yb$$
then
$$(A_1^2 + A_2^2)b = (x^2 + y^2)b.$$
(choose any polynomial objective function as an eigenvalue!)

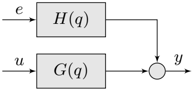
Polynomial optimization problems with a polynomial objective function and polynomial constraints can always be written as eigenvalue problems where we search for the minimal eigenvalue!

→ 'Convexification' of polynomial optimization problems

Outline

- PEM System identification
- Measured data $\{u_k, y_k\}_{k=1}^{N}$
- Model structure

$$y_k = G(q)u_k + H(q)e_k$$



- Output prediction

$$\hat{y}_k = H^{-1}(q)G(q)u_k + (1 - H^{-1})y_k$$

- Model classes: ARX, ARMAX, OE, BJ

$$A(q)y_k = B(q)/F(q)u_k + C(q)/D(q)e_k$$

Class	Polynomials
ARX	$A(q), B(q)$
ARMAX	$A(q)$, $B(q)$, $C(q)$
OE	$B(q)$, $F(q)$
BJ	$B(q)$, $C(q)$, $D(q)$, $F(q)$

- Minimize the prediction errors $y - \hat{y}$, where

$$\hat{y}_k = H^{-1}(q)G(q)u_k + (1 - H^{-1})y_k,$$

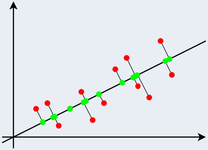
subject to the model equations

- Example

ARMAX identification: $G(q) = B(q)/A(q)$ and $H(q) = C(q)/A(q)$, where $A(q) = 1 + aq^{-1}$, $B(q) = bq^{-1}$, $C(q) = 1 + cq^{-1}$, $N = 5$

$$\begin{array}{ll} \min\limits_{\hat{y},a,b,c} & (y_1 - \hat{y}_1)^2 + \ldots + (y_5 - \hat{y}_5)^2 \\ \text{s. t.} & \hat{y}_5 - c\hat{y}_4 - bu_4 - (c - a)y_4 = 0, \\ & \hat{y}_4 - c\hat{y}_3 - bu_3 - (c - a)y_3 = 0, \\ & \hat{y}_3 - c\hat{y}_2 - bu_2 - (c - a)y_2 = 0, \\ & \hat{y}_2 - c\hat{y}_1 - bu_1 - (c - a)y_1 = 0, \end{array}$$

Static Linear Modeling



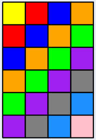
- Rank deficiency
- minimization problem:

$$\begin{array}{ll} \min & \|[\,\Delta A \quad \Delta b\,]\|_F^2, \\ \text{s. t.} & (A+\Delta A)v = b+\Delta b, \\ & v^T v = 1 \end{array}$$

- Singular Value Decomposition:
  find (u, σ, v) which minimizes σ^2
  Let $M = \begin{bmatrix} A & b \end{bmatrix}$

$$\left\{ \begin{array}{rcl} Mv &=& u\sigma \\ M^T u &=& v\sigma \\ v^T v &=& 1 \\ u^T u &=& 1 \end{array} \right.$$

Dynamical Linear Modeling



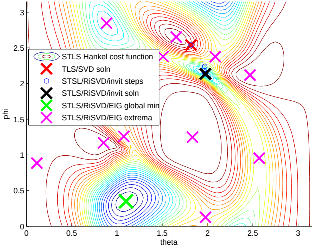
- Rank deficiency
- minimization problem:

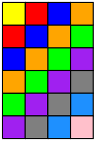
$$\begin{array}{ll} \min & \|[\Delta A \quad \Delta b]\|_F^2, \\ \text{s. t.} & (A+\Delta A)v = b+\Delta b, \\ & v^T v = 1 \\ & [\Delta A \quad \Delta b] \text{ structured} \end{array}$$

- Riemannian SVD:
  find (u, τ, v) which minimizes τ^2

$$\left\{ \begin{array}{rcl} Mv &=& D_v u\tau \\ M^T u &=& D_u v\tau \\ v^T v &=& 1 \\ u^T D_v u &=& 1 \;(= v^T D_u v) \end{array} \right.$$

$$\min_{v} \quad \tau^2 = v^T M^T D_v^{-1} M v$$
$$\text{s. t.} \quad v^T v = 1.$$





method	TLS/SVD	STLS inv. it.	STLS eig
v_1	.8003	.4922	**.8372**
v_2	-.5479	-.7757	**.3053**
v_3	.2434	.3948	**.4535**
τ^2	4.8438	3.0518	**2.3822**
global solution?	no	no	**yes**

CpG Islands

- genomic regions that contain a high frequency of sites where a cytosine (C) base is followed by a guanine (G)
- rare because of methylation of the C base
- hence CpG islands indicate functionality

Given observed sequence of DNA:

CTCACGTGATGAGAGCATTCTCAGA
CCGTGACGCGTGTAGCAGCGGCTCA

Problem

Decide whether the observed sequence came from a CpG island

The model

- 4-dimensional state space $[m] = \{\mathrm{A}, \mathrm{C}, \mathrm{G}, \mathrm{T}\}$
- Mixture model of 3 distributions on $[m]$
  1. : CG rich DNA
  2. : CG poor DNA
  3. : CG neutral DNA
- Each distribution is characterised by probabilities of observing base A,C,G or T

Table: Probabilities for each of the distributions (Durbin; Pachter & Sturmfels)

DNA Type	A	C	G	T
CG rich	0.15	0.33	0.36	0.16
CG poor	0.27	0.24	0.23	0.26
CG neutral	0.25	0.25	0.25	0.25

- The probabilities of observing each of the bases A to T are given by

$$\begin{aligned} p(A) &= -0.10\,\theta_1 + 0.02\,\theta_2 + 0.25 \\ p(C) &= +0.08\,\theta_1 - 0.01\,\theta_2 + 0.25 \\ p(G) &= +0.11\,\theta_1 - 0.02\,\theta_2 + 0.25 \\ p(T) &= -0.09\,\theta_1 + 0.01\,\theta_2 + 0.25 \end{aligned}$$

- θ_i is probability to sample from distribution i $(\theta_1 + \theta_2 + \theta_3 = 1)$
- Maximum Likelihood Estimate:

$$(\hat{\theta_1}, \hat{\theta_2}, \hat{\theta_3}) = \arg\max_{\theta} \; l(\theta)$$

where the log-likelihood $l(\theta)$ is given by

$$l(\theta) = 11 \log p(A) + 14 \log p(C) + 15 \log p(G) + 10 \log p(T)$$

- Need to solve the following polynomial system

$$\begin{cases} \frac{\partial l(\theta)}{\partial \theta_1} = \sum_{i=1}^{4} \frac{u_i}{p(i)} \frac{\partial p(i)}{\partial \theta_1} = 0 \\ \frac{\partial l(\theta)}{\partial \theta_2} = \sum_{i=1}^{4} \frac{u_i}{p(i)} \frac{\partial p(i)}{\partial \theta_2} = 0 \end{cases}$$

Solving the Polynomial System

- corank$(M) = 9$
- Reconstructed Kernel

$$K = \begin{bmatrix} 1 & 1 & 1 & 1 & \ldots \\ 0.52 & 3.12 & -5.00 & 10.72 & \ldots \\ 0.22 & 3.12 & -15.01 & 71.51 & \ldots \\ 0.27 & 9.76 & 25.02 & 115.03 & \ldots \\ 0.11 & 9.76 & 75.08 & 766.98 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{matrix} 1 \\ \theta_1 \\ \theta_2 \\ \theta_1^2 \\ \theta_1\theta_2 \\ \vdots \end{matrix} .$$

- θ_i's are probabilities: $0 \leq \theta_i \leq 1$
- Could have introduced slack variables to impose this constraint!
- Only solution that satisfies this constraint is $\hat{\theta} = (0.52, 0.22, 0.26)$

Applications are found in

- Polynomial Optimization Problems
- Structured Total Least Squares
- Model order reduction
- Analyzing identifiability nonlinear model structures
- Robotics: kinematic problems
- Computational Biology: conformation of molecules
- Algebraic Statistics
- Signal Processing
- ...

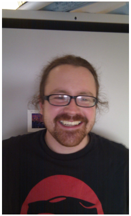
Outline

- Finding roots of multivariate polynomials is linear algebra and realization theory!
- Finding minimizing zero of a polynomial optimization problem is extremal eigenvalue problem
- (Numerical) linear algebra/systems theory version of results in algebraic geometry/symbolic algebra (Gröbner bases, resultants, rings, ideals, varieties,. . .)
- These relations in principle 'convexify'/linearize many problems
  - Algebraic geometry
  - System identification (PEM)
  - Numerical linear algebra (STLS, affine EVP $Ax = x\lambda + a$, etc.)
  - Multilinear algebra (tensor least squares approximation problems)
  - Algebraic statistics (HMM, Bayesian networks, discrete probabilities)
  - Differential algebra (Glad/Ljung)
- Convexification occurs by projecting up to higher dimensional vector space (difficult in low number of dimensions; 'easy' in high number of dimensions: an eigenvalue problem)
- Many challenges remain:
  - Efficient construction of the eigenvalue problem - exploiting sparseness and structure
  - Algorithms to find the minimizing solution directly (inverse power method)
  - . . .

Questions?

Bart De Moor
(1960-...)

Kim Batselier
(1981-...)



Philippe Dreesen
(1982-...)

"At the end of the day,
the only thing we really understand,
is linear algebra".