

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



Philippe Dreesen



Kim Batselier



Bart De Moor

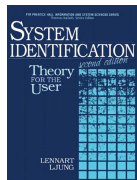
**KU Leuven**  
Department of Electrical Engineering  
ESAT-STADIUS

# Outline

- 1 Introduction
- 2 History
- 3 Linear Algebra
- 4 Multivariate Polynomials
- 5 Algebraic Optimization
- 6 Applications
- 7 Conclusions

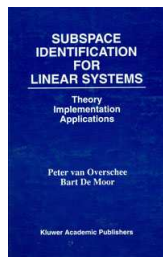
## System Identification: PEM

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



## Linear Algebra approach

⇒ **Subspace methods**



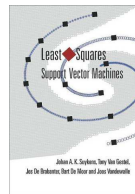
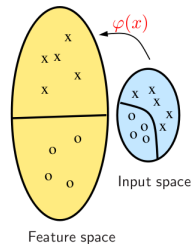
## Nonlinear regression, modelling and clustering

- 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

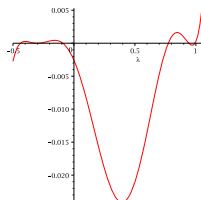
- Eigenvalue decompositions, SVDs,...
- Solving systems of equations (consistency, nb sols)
- QR decomposition and Gram-Schmidt algorithm

## 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)

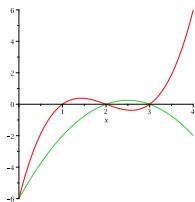
## Four instances of polynomial rooting problems

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



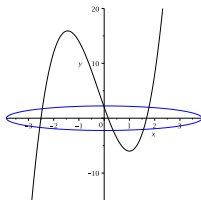
$$(x - 1)(x - 3)(x - 2) = 0$$

$$-(x - 2)(x - 3) = 0$$



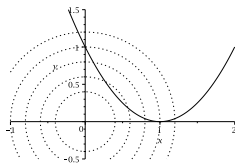
$$x^2 + 3y^2 - 15 = 0$$

$$y - 3x^3 - 2x^2 + 13x - 2 = 0$$



$$\min_{x,y} \quad x^2 + y^2$$

$$\text{s. t.} \quad y - x^2 + 2x - 1 = 0$$



# Outline

- 1 Introduction
- 2 History**
- 3 Linear Algebra
- 4 Multivariate Polynomials
- 5 Algebraic Optimization
- 6 Applications
- 7 Conclusions



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

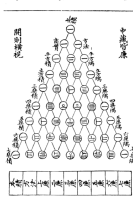


**Diophantus**  
(c200-c284)  
*Arithmetica*



**Al-Khwarizmi**  
(c780-c850)

## 圖方算七法古



**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)



**Carl Friedrich Gauss**  
(1777-1755)



**Jean-Victor Poncelet**  
(1788-1867)



**Evariste Galois**  
(1811-1832)



**Arthur Cayley**  
(1821-1895)



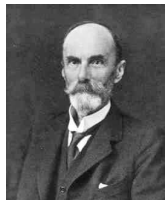
**Leopold Kronecker**  
(1823-1891)



**Edmond Laguerre**  
(1834-1886)



**James J. Sylvester**  
(1814-1897)



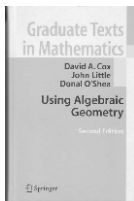
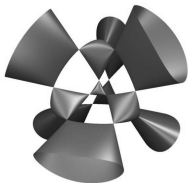
**Francis S. Macaulay**  
(1862-1937)



**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

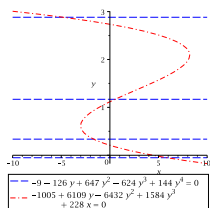
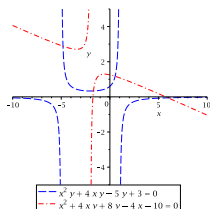
### Input system:

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

- 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

- 1 Introduction
- 2 History
- 3 Linear Algebra**
- 4 Multivariate Polynomials
- 5 Algebraic Optimization
- 6 Applications
- 7 Conclusions

$$\begin{array}{ccc}
 A & X & = & 0 \\
 p \times q & q \times (q-r) & & p \times (q-r)
 \end{array}$$

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



James Joseph Sylvester

$$\begin{array}{l}
 A = [U_1 \quad U_2] \begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \\
 \Downarrow \\
 X = V_2
 \end{array}$$

$$\begin{array}{ccc} A & X & = & 0 \\ p \times q & q \times (q-r) & & p \times (q-r) \end{array}$$

Reorder columns of  $A$  and partition

$$\begin{array}{ccc} p \times q & p \times (q-r) & p \times r \\ \bar{A} & = & [\bar{A}_1 \quad \bar{A}_2] \quad \text{rank}(\bar{A}_2) = r \quad (\bar{A}_2 \text{ full column rank}) \end{array}$$

Reorder rows of  $X$  and partition accordingly

$$[\bar{A}_1 \quad \bar{A}_2] \begin{array}{c} \begin{array}{c} q-r \\ \bar{X}_1 \\ r \\ \bar{X}_2 \end{array} \end{array} = 0$$

$$\begin{array}{l} \text{rank}(\bar{A}_2) = r \\ \updownarrow \\ \text{rank}(\bar{X}_1) = q - r \end{array}$$

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

- $\overline{X_1}$ : 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):

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



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

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

$$[A \ B] \begin{array}{c} q - r_A \quad t - r_B \quad ? \\ \left[ \begin{array}{ccc} X & 0 & ? \\ 0 & Y & ? \end{array} \right] = 0$$

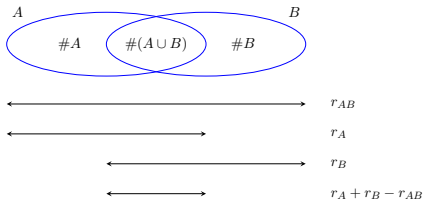
Let  $\text{rank}([A \ B]) = r_{AB}$

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

$$[A \quad B] \begin{bmatrix} X & 0 & Z_1 \\ 0 & Y & Z_2 \end{bmatrix} = 0$$

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

$$x^3 + a_1x^2 + a_2x + a_3 = 0,$$

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
- corank = 3
- Observability matrix-like
- Realization theory!

Consider

$$x^3 + a_1x^2 + a_2x + a_3 = 0$$

$$x^2 + b_1x + b_2 = 0$$

Build the Sylvester Matrix:

$$\begin{bmatrix} 1 & a_1 & a_2 & a_3 & 0 \\ 0 & 1 & a_1 & a_2 & a_3 \\ 1 & b_1 & b_2 & 0 & 0 \\ 0 & 1 & b_1 & b_2 & 0 \\ 0 & 0 & 1 & b_1 & b_2 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \\ x^3 \\ x^4 \end{bmatrix} = 0$$

Row Space	Null Space
<b>Ideal</b> = union of ideals = multiply rows with powers of $x$	<b>Variety</b> = 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  $S(f, g) = 0$

$$S(f, g) = \det \begin{bmatrix} -6 & 11 & -6 & 1 & 0 \\ 0 & -6 & 11 & -6 & 1 \\ \hline -6 & 5 & -1 & 0 & 0 \\ 0 & -6 & 5 & -1 & 0 \\ 0 & 0 & -6 & 5 & -1 \end{bmatrix}$$



James Joseph Sylvester

The corank of the Sylvester matrix is 2!

Sylvester's construction can be understood from

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

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

$$\begin{aligned} S_1KD &= S_2K \\ S_1ZVD &= S_2ZV \end{aligned}$$

leading to the generalized eigenvalue problem

$$(S_2Z)V = (S_1Z)VD$$



# Outline

- 1 Introduction
- 2 History
- 3 Linear Algebra
- 4 Multivariate Polynomials**
- 5 Algebraic Optimization
- 6 Applications
- 7 Conclusions

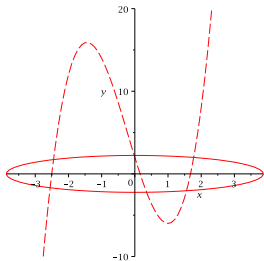
- Consider

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

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

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

$$\begin{array}{l} p(x, y) \\ q(x, y) \\ x \cdot p(x, y) \\ y \cdot p(x, y) \end{array} \begin{bmatrix} 1 & x & y & x^2 & xy & y^2 & x^3 & x^2y & xy^2 & y^3 \\ -15 & & & 1 & & 3 & & & & \\ -2 & 13 & 1 & -2 & & & -3 & & & \\ -15 & & & & & & 1 & & 3 & \\ & & -15 & & & & & 1 & & 3 \end{bmatrix}$$



## Null space based Root-finding

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

Continue to enlarge  $M$ :

it #	form	1	x	y	x <sup>2</sup>	xy	y <sup>2</sup>	x <sup>3</sup>	x <sup>2</sup> y	xy <sup>2</sup>	y <sup>3</sup>	x <sup>4</sup>	x <sup>3</sup> y	yx <sup>2</sup>	y <sup>2</sup> x	xy <sup>3</sup>	y <sup>4</sup>	x <sup>5</sup>	x <sup>4</sup> y	yx <sup>3</sup>	y <sup>2</sup> x <sup>2</sup>	y <sup>3</sup> x	y <sup>4</sup>	y <sup>5</sup>	
d = 3	p xp yp q	-15	-15	-15	1	3		1	3																
d = 4	x <sup>2</sup> p xyp y <sup>2</sup> p xq yq		-2	-2	-15	-15		-2	-2			1	3												
d = 5	x <sup>3</sup> p x <sup>2</sup> yp xy <sup>2</sup> p y <sup>3</sup> p x <sup>2</sup> q xyq y <sup>2</sup> q				-15	-15		13	1			-2	-2						1	3					

- # rows grows faster than # cols  $\Rightarrow$  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  $(x_i, y_i)$ :

1	1	...	1
$x_1$	$x_2$	...	$x_s$
$y_1$	$y_2$	...	$y_s$
$x_1^2$	$x_2^2$	...	$x_s^2$
$x_1 y_1$	$x_2 y_2$	...	$x_s y_s$
$y_1^2$	$y_2^2$	...	$y_s^2$
$x_1^3$	$x_2^3$	...	$x_s^3$
$x_1^2 y_1$	$x_2^2 y_2$	...	$x_s^2 y_s$
$x_1 y_1^2$	$x_2 y_2^2$	...	$x_s y_s^2$
$y_1^3$	$y_2^3$	...	$y_s^3$
$x_1^4$	$x_2^4$	...	$x_s^4$
$x_1^3 y_1$	$x_2^3 y_2$	...	$x_s^3 y_s$
$x_1^2 y_1^2$	$x_2^2 y_2^2$	...	$x_s^2 y_s^2$
$x_1 y_1^3$	$x_2 y_2^3$	...	$x_s y_s^3$
$y_1^4$	$y_2^4$	...	$y_s^4$
⋮	⋮	⋮	⋮

- Choose  $s$  linear independent rows in  $K$

$$S_1 K$$

- This corresponds to finding linear dependent columns in  $M$

1	1	...	1
$x_1$	$x_2$	...	$x_s$
$y_1$	$y_2$	...	$y_s$
$x_1^2$	$x_2^2$	...	$x_s^2$
$x_1 y_1$	$x_2 y_2$	...	$x_s y_s$
$y_1^2$	$y_2^2$	...	$y_s^2$
$x_1^3$	$x_2^3$	...	$x_s^3$
$x_1^2 y_1$	$x_2^2 y_2$	...	$x_s^2 y_s$
$x_1 y_1^2$	$x_2 y_2^2$	...	$x_s y_s^2$
$y_1^3$	$y_2^3$	...	$y_s^3$
$x_1^4$	$x_2^4$	...	$x_s^4$
$x_1^3 y_1$	$x_2^3 y_2$	...	$x_s^3 y_s$
$x_1^2 y_1^2$	$x_2^2 y_2^2$	...	$x_s^2 y_s^2$
$x_1 y_1^3$	$x_2 y_2^3$	...	$x_s y_s^3$
$y_1^4$	$y_2^4$	...	$y_s^4$
⋮	⋮	⋮	⋮

Shifting the selected rows gives (shown for 3 columns)

$$\begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ \hline x_1^2 & x_2^2 & x_3^2 \\ x_1 y_1 & x_2 y_2 & x_3 y_3 \\ y_1^2 & y_2^2 & y_3^2 \\ \hline x_1^3 & x_2^3 & x_3^3 \\ x_1^2 y_1 & x_2^2 y_2 & x_3^2 y_3 \\ x_1 y_1^2 & x_2 y_2^2 & x_3 y_3^2 \\ y_1^3 & y_2^3 & y_3^3 \\ \hline x_1^4 & x_2^4 & x_3^4 \\ x_1^3 y_1 & x_2^3 y_2 & x_3^3 y_3 \\ x_1^2 y_1^2 & x_2^2 y_2^2 & x_3^2 y_3^2 \\ x_1 y_1^3 & x_2 y_2^3 & x_3 y_3^3 \\ y_1^4 & y_2^4 & y_3^4 \\ \hline \vdots & \vdots & \vdots \\ \hline \end{array} \rightarrow \text{"shift with } x" \rightarrow \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ \hline x_1^2 & x_2^2 & x_3^2 \\ x_1 y_1 & x_2 y_2 & x_3 y_3 \\ y_1^2 & y_2^2 & y_3^2 \\ \hline x_1^3 & x_2^3 & x_3^3 \\ x_1^2 y_1 & x_2^2 y_2 & x_3^2 y_3 \\ x_1 y_1^2 & x_2 y_2^2 & x_3 y_3^2 \\ y_1^3 & y_2^3 & y_3^3 \\ \hline x_1^4 & x_2^4 & x_3^4 \\ x_1^3 y_1 & x_2^3 y_2 & x_3^3 y_3 \\ x_1^2 y_1^2 & x_2^2 y_2^2 & x_3^2 y_3^2 \\ x_1 y_1^3 & x_2 y_2^3 & x_3 y_3^3 \\ y_1^4 & y_2^4 & y_3^4 \\ \hline \vdots & \vdots & \vdots \\ \hline \end{array}$$

simplified:

$$\begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ \hline x_1 y_1 & x_2 y_2 & x_3 y_3 \\ x_1^3 & x_3^3 & x_3^3 \\ x_1^2 & x_2^2 & x_3^2 \\ x_1 y_1 & x_2 y_2 & x_3 y_3 \\ \hline \end{array} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{array}{|c|c|c|} \hline x_1^2 & x_2^2 & x_3^2 \\ x_1 & x_2 & x_3 \\ x_1 y_1 & x_2 y_2 & x_3 y_3 \\ x_1^2 y_1 & x_2^2 y_2 & x_3^2 y_3 \\ x_1^4 & x_3^4 & x_3^4 \\ x_1 & x_2 & x_3 \\ x_1 y_1 & x_2 y_2 & x_3 y_3 \\ \hline \end{array}$$

- finding the  $x$ -roots: let  $D_x = \text{diag}(x_1, x_2, \dots, x_s)$ , then

$$S_1 K D_x = S_x K,$$

where  $S_1$  and  $S_x$  select rows from  $K$  wrt. shift property

- reminiscent of **Realization Theory**

We have

$$S_1 KD_x = S_x K$$

However,  $K$  is not known, instead a basis  $Z$  is computed that satisfies

$$ZV = K$$

Which leads to

$$(S_x Z)V = (S_1 Z)VD_x$$



It is possible to shift with  $y$  as well. . .

We find

$$S_1 K D_y = S_y K$$

with  $D_y$  diagonal matrix of  $y$ -components of roots, leading to

$$(S_y Z) V = (S_1 Z) V D_y$$

Some interesting results:

- same eigenvectors  $V$ !
- $(S_3 Z)^{-1}(S_1 Z)$  and  $(S_2 Z)^{-1}(S_1 Z)$  commute

## Nullspace of $M$

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

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

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

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

# Realization Theory and Polynomial System Solving

- Attasi model

$$v(k_1, \dots, k_{i-1}, \mathbf{k}_i + \mathbf{1}, k_{i+1}, \dots, k_n) = \mathbf{A}_i v(k_1, \dots, k_n)$$

- Null space of Macaulay matrix:  $n$ D state sequence

$$\left( \begin{array}{c|c|c|c|c|c|c|c|c|c|c} v_{00} & v_{10} & v_{01} & v_{20} & v_{11} & v_{02} & v_{30} & v_{21} & v_{12} & v_{03} & \\ \hline & & & & & & & & & & \\ \hline & & & & & & & & & & \end{array} \right) =$$

$$\left( \begin{array}{c|c|c|c|c|c|c|c|c|c|c} v_{00} & A_1 v_{00} & A_2 v_{00} & \cdots & A_1^3 v_{00} & A_1^2 A_2 v_{00} & A_1 A_2^2 v_{00} & A_2^3 v_{00} & & & \\ \hline & & & & & & & & & & \\ \hline & & & & & & & & & & \end{array} \right)$$

- shift-invariance property, e.g., for  $x_2$ :

$$\begin{pmatrix} -v_{00} \\ -v_{10} \\ -v_{01} \\ -v_{20} \\ -v_{11} \\ -v_{02} \end{pmatrix} A_2^T = \begin{pmatrix} -v_{01} \\ -v_{11} \\ -v_{02} \\ -v_{21} \\ -v_{12} \\ -v_{03} \end{pmatrix},$$

- corresponding  $nD$  system realization

$$\begin{aligned} v(k+1, l) &= A_1 v(k, l) \\ v(k, l+1) &= A_2 v(k, l) \\ v(0, 0) &= v_{00} \end{aligned}$$

- choice of basis null space leads to different system realizations
- eigenvalues of  $A_1$  and  $A_2$  invariant:  $x_1$  and  $x_2$  components

There are 3 kinds of roots:

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

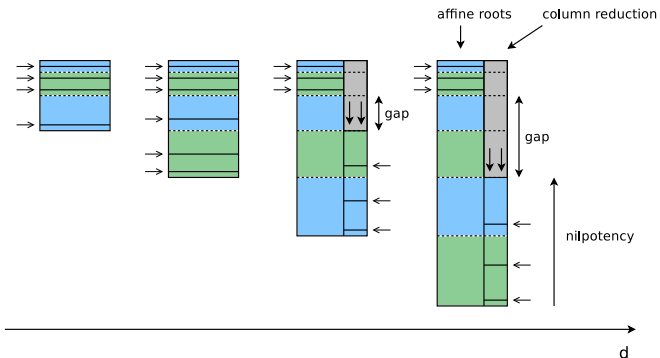
Applying Grassmann's Dimension theorem on the kernel allows to write the following partitioning

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

- $X_1$  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!)

## Mind the Gap!

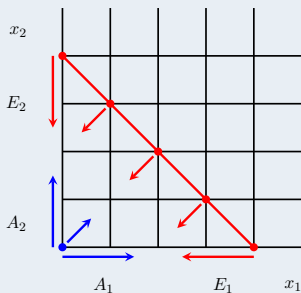
- dynamics in the null space of  $M(d)$  for increasing degree  $d$
- nilpotency gives rise to a ‘gap’
- mechanism to count and separate affine from infinity



- Kronecker Canonical Form decoupling affine and infinity roots

$$\begin{pmatrix} v(k+1) \\ w(k-1) \end{pmatrix} = \left( \begin{array}{c|c} A & 0 \\ \hline 0 & E \end{array} \right) \begin{pmatrix} v(k) \\ w(k) \end{pmatrix},$$

- Action of  $A_i$  and  $E_i$  represented in grid of monomials



## Roots at Infinity: $nD$ Descriptor Systems

Weierstrass Canonical Form decouples affine/infinity

$$\begin{bmatrix} v(k+1) \\ w(k-1) \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & E \end{bmatrix} \begin{bmatrix} v(k) \\ w(k) \end{bmatrix}$$

Singular  $nD$  Attasi model (for  $n = 2$ )

$$v(k+1, l) = A_x v(k, l)$$

$$v(k, l+1) = A_y v(k, l)$$

$$w(k-1, l) = E_x w(k, l)$$

$$w(k, l-1) = E_y w(k, l)$$

with  $E_x$  and  $E_y$  nilpotent matrices.



# Summary

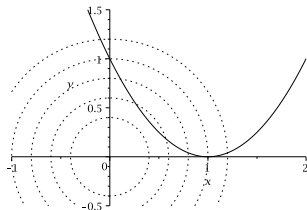
- solving multivariate polynomials
  - question in linear algebra
  - realization theory in null space of Macaulay matrix
  - $nD$  autonomous (descriptor) Attasi model
- decisions made based upon (numerical) rank
  - # roots (nullity)
  - # affine roots (column reduction)
- mind the gap phenomenon: affine vs. infinity roots
- not discussed
  - multiplicity of roots
  - column-space based method
  - over-constrained systems

# Outline

- 1 Introduction
- 2 History
- 3 Linear Algebra
- 4 Multivariate Polynomials
- 5 Algebraic Optimization**
- 6 Applications
- 7 Conclusions

# Polynomial Optimization Problems

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



**Lagrange multipliers** give conditions for optimality:

$$L(x, y, z) = x^2 + y^2 + z(y - x^2 + 2x - 1)$$

we find

$$\partial L / \partial x = 0 \rightarrow 2x - 2xz + 2z = 0$$

$$\partial L / \partial y = 0 \rightarrow 2y + z = 0$$

$$\partial L / \partial z = 0 \rightarrow y - x^2 + 2x - 1 = 0$$

Observations:

- everything remains polynomial
- system of polynomial equations
- shift with objective function to find minimum/maximum

Let

$$A_x V = xV$$

and

$$A_y V = yV$$

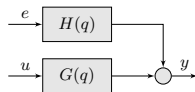
then find min/max eigenvalue of

$$(A_x^2 + A_y^2)V = (x^2 + y^2)V$$

# Outline

- 1 Introduction
- 2 History
- 3 Linear Algebra
- 4 Multivariate Polynomials
- 5 Algebraic Optimization
- 6 Applications**
- 7 Conclusions

- 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

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

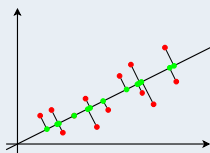
subject to the model equations

- Example

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

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

## Static Linear Modeling



- Rank deficiency
- minimization problem:

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

- Singular Value Decomposition:  
find  $(u, \sigma, v)$  which minimizes  $\sigma^2$   
Let  $M = [A \quad b]$

$$\begin{cases} Mv = u\sigma \\ M^T u = v\sigma \\ v^T v = 1 \\ u^T u = 1 \end{cases}$$

## Dynamical Linear Modeling



- Rank deficiency
- minimization problem:

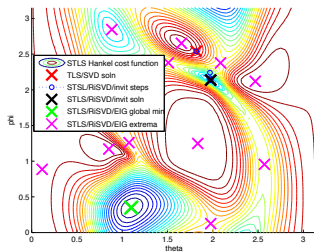
$$\begin{aligned} \min \quad & \| [\Delta A \quad \Delta b] \|_F^2, \\ \text{s. t.} \quad & (A + \Delta A)v = b + \Delta b, \\ & v^T v = 1 \\ & [\Delta A \quad \Delta b] \text{ structured} \end{aligned}$$

- Riemannian SVD:  
find  $(u, \tau, v)$  which minimizes  $\tau^2$

$$\begin{cases} Mv = D_v u \tau \\ M^T u = D_u v \tau \\ v^T v = 1 \\ u^T D_v u = 1 (= v^T D_u v) \end{cases}$$



$$\begin{aligned} \min_v \quad & \tau^2 = v^T M^T D_v^{-1} M v \\ \text{s. t.} \quad & v^T v = 1. \end{aligned}$$



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

## 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  $[m] = \{A, C, G, T\}$
- 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

$$p(A) = -0.10 \theta_1 + 0.02 \theta_2 + 0.25$$

$$p(C) = +0.08 \theta_1 - 0.01 \theta_2 + 0.25$$

$$p(G) = +0.11 \theta_1 - 0.02 \theta_2 + 0.25$$

$$p(T) = -0.09 \theta_1 + 0.01 \theta_2 + 0.25$$

- $\theta_i$  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

- $\text{corank}(M) = 9$
- Reconstructed Kernel

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

- $\theta_i$ '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

- 1 Introduction
- 2 History
- 3 Linear Algebra
- 4 Multivariate Polynomials
- 5 Algebraic Optimization
- 6 Applications
- 7 Conclusions**

# Conclusions

- Finding roots: linear algebra and realization theory!
- Polynomial optimization: extremal eigenvalue problems
- (Numerical) linear algebra/systems theory translation of algebraic geometry/symbolic algebra
- These relations ‘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)
  - Algebraic statistics (HMM, Bayesian networks, discrete probabilities)
  - Differential algebra (Glad/Ljung)
- Convexification: projecting up to higher dimensional space (difficult in low number of dimensions; ‘easy’ in high number of dimensions)



# Open Problems

Many challenges remain!

- Efficient construction of the eigenvalue problem - exploiting sparseness and structure
- Algorithms to find the minimizing solution directly (inverse power method?)
- Unraveling structure at infinity (realization theory)
- Positive dimensional solution set: parametrization eigenvalue problem
- $nD$  version of Cayley-Hamilton theorem
- ...

# Questions?

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