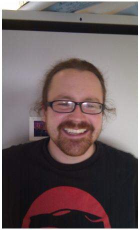
Back to the Roots

at the occasion of Anders Lindquist 75 !



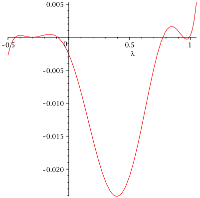
Philippe Dreesen **Kim Batselier** **Bart De Moor**

KU Leuven
Department of Electrical Engineering
ESAT-STADIUS

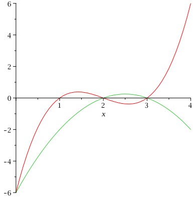
KU LEUVEN

Outline

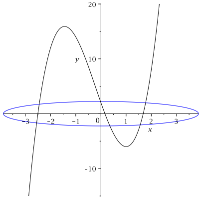
$$p(\lambda) = \det(A - \lambda I) = 0$$



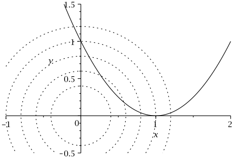
$$\begin{aligned}(x-1)(x-3)(x-2) &= 0\\ -(x-2)(x-3) &= 0\end{aligned}$$



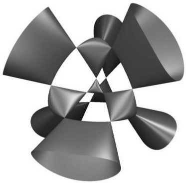
$$\begin{aligned}x^2 + 3y^2 - 15 &= 0\\ y - 3x^3 - 2x^2 + 13x - 2 &= 0\end{aligned}$$

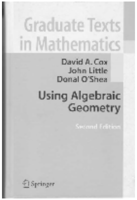


$$\begin{aligned}\min_{x,y} \quad & x^2 + y^2\\ \text{s.t.} \quad & y - x^2 + 2x - 1 = 0\end{aligned}$$



- Algebraic Geometry: ‘Queen of mathematics’ (literature = huge !)
- Computer algebra: symbolic manipulations
- Computational tools: Gröbner Bases, Buchberger algorithm





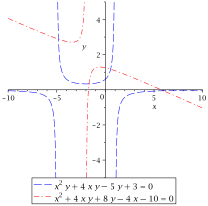
Wolfgang Gröbner
(1899-1980)

Bruno Buchberger

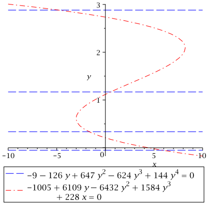
Example: Gröbner basis

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
- Exponential complexity
- Numerical issues
  - NO floating point but integer arithmetic
  - Coefficients become very large



Gröbner Basis:

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

KU LEUVEN

Outline

- **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

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

- Banded Toeplitz; linear homogeneous equations
- Null space: (Confluent) Vandermonde structure
- Corank (nullity) = number of solutions
- Realization theory in null space: eigenvalue problem

KU LEUVEN

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

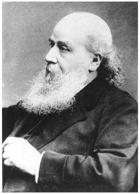
- **Sylvester Resultant**

  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 result can be understood from

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

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

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

$$\begin{bmatrix} 1 & 1 \\ x_1 & x_2 \\ x_1^2 & x_2^2 \\ x_1^3 & x_2^3 \end{bmatrix} \begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \\ x_1^2 & x_2^2 \\ x_1^3 & x_2^3 \\ x_1^4 & x_2^4 \end{bmatrix}$$

or

$$\underline{K}.D = S_1 K D = \overline{K} = S_2 K$$

The Vandermonde 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

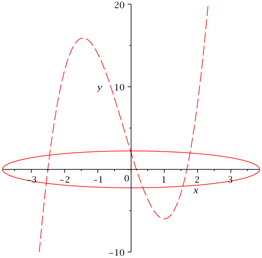
$$(S_2 Z)V = (S_1 Z)V D$$

KU LEUVEN

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

KU LEUVEN

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

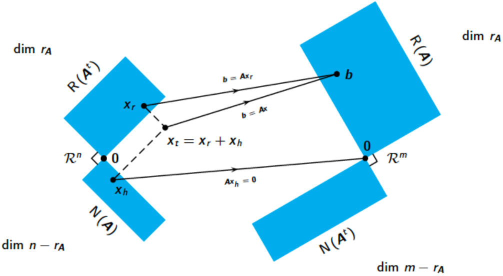
Continue to enlarge M:

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 $\Rightarrow$ overdetermined system
- If solution exists: rank deficient by construction!

- **Row space**:
  - ideal; Hilbert Basis Theorem
  - *Subspace based elimination theory*
- **Left null space**:
  - syzygies, Hilbert Syzygy Theorem
  - *Syzygy: numerical linear algebra paper bdm/kb*
- **Right null space**:
  - Variety; Hilbert Nullstellensatz (existence of solutions); Hilbert polynomial (number of solutions = nullity)
  - *Modelling the Macaulay null space with nD singular autonomous systems*
- **Column space**: *Rank tests: Affine roots, roots at* ∞

David Hilbert

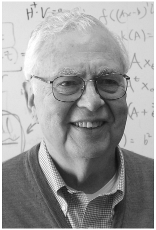


$$A = USV^t = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^t \\ V_2^t \end{pmatrix}$$

with

$$U_1^t U_1 = I_{r_A} \qquad V_1^t V_1 = I_{r_A}$$
$$U_2^t U_2 = I_{m-r_A} \qquad V_2^t V_2 = I_{n-r_A}$$
$$U_1^t U_2 = 0 \qquad V_1^t V_2 = 0$$

Geometry	Basis
$\mathrm{R}(A)$	U_1
$\mathrm{N}(A^t)$	U_2
$\mathrm{R}(A^t)$	V_1
$\mathrm{N}(A)$	V_2



Gene Howard Golub

(Dr. SVD)

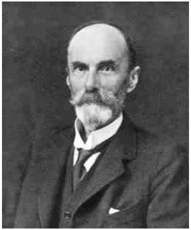
- Macaulay matrix M:

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

- Solutions generate vectors in kernel of M:

$$MK = 0$$

- Number of solutions s follows from corank



Francis Sowerby Macaulay

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

$$\begin{bmatrix} 1 & 1 & \ldots & 1 \\ \hline x_1 & x_2 & \ldots & x_s \\ y_1 & y_2 & \ldots & y_s \\ \hline x_1^2 & x_2^2 & \ldots & x_s^2 \\ x_1y_1 & x_2y_2 & \ldots & x_sy_s \\ y_1^2 & y_2^2 & \ldots & y_s^2 \\ \hline x_1^3 & x_2^3 & \ldots & x_s^3 \\ x_1^2y_1 & x_2^2y_2 & \ldots & x_s^2y_s \\ x_1y_1^2 & x_2y_2^2 & \ldots & x_sy_s^2 \\ y_1^3 & y_2^3 & \ldots & y_s^3 \\ \hline x_1^4 & x_2^4 & \ldots & x_4^4 \\ x_1^3y_1 & x_2^3y_2 & \ldots & x_s^3y_s \\ x_1^2y_1^2 & x_2^2y_2^2 & \ldots & x_s^2y_s^2 \\ x_1y_1^3 & x_2y_2^3 & \ldots & x_sy_s^3 \\ y_1^4 & y_2^4 & \ldots & y_s^4 \\ \hline \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

- Choose s linear independent rows in K

$$S_1 K$$

- This corresponds to finding linear dependent columns in M

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ \hline x_1 & x_2 & \dots & x_s \\ y_1 & y_2 & \dots & y_s \\ \hline x_1^2 & x_2^2 & \dots & x_s^2 \\ x_1 y_1 & x_2 y_2 & \dots & x_s y_s \\ y_1^2 & y_2^2 & \dots & y_s^2 \\ \hline x_1^3 & x_2^3 & \dots & x_s^3 \\ x_1^2 y_1 & x_2^2 y_2 & \dots & x_s^2 y_s \\ x_1 y_1^2 & x_2 y_2^2 & \dots & x_s y_s^2 \\ y_1^3 & y_2^3 & \dots & y_s^3 \\ \hline x_1^4 & x_2^4 & \dots & x_4^4 \\ x_1^3 y_1 & x_2^3 y_2 & \dots & x_s^3 y_s \\ x_1^2 y_1^2 & x_2^2 y_2^2 & \dots & x_s^2 y_s^2 \\ x_1 y_1^3 & x_2 y_2^3 & \dots & x_s y_s^3 \\ y_1^4 & y_2^4 & \dots & y_s^4 \\ \hline \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

Shifting the selected rows gives (shown for 3 columns)

$$
\begin{bmatrix}
1 & 1 & 1 \\
x_1 & x_2 & x_3 \\
y_1 & y_2 & y_3 \\
x_1^2 & x_2^2 & x_3^2 \\
x_1y_1 & x_2y_2 & x_3y_3 \\
y_1^2 & y_2^2 & y_3^2 \\
x_1^3 & x_2^3 & x_3^3 \\
x_1^2y_1 & x_2^2y_2 & x_3^2y_3 \\
x_1y_1^2 & x_2y_2^2 & x_3y_3^2 \\
y_1^3 & y_2^3 & y_3^3 \\
x_1^4 & x_2^4 & x_4^4 \\
x_1^3y_1 & x_2^3y_2 & x_3^3y_3 \\
x_1^2y_1^2 & x_2^2y_2^2 & x_3^2y_3^2 \\
x_1y_1^3 & x_2y_2^3 & x_3y_3^3 \\
y_1^4 & y_2^4 & y_3^4 \\
\vdots & \vdots & \vdots
\end{bmatrix}
\longrightarrow \text{“shift with } x\text{”} \longrightarrow
\begin{bmatrix}
1 & 1 & 1 \\
x_1 & x_2 & x_3 \\
y_1 & y_2 & y_3 \\
x_1^2 & x_2^2 & x_3^2 \\
x_1y_1 & x_2y_2 & x_3y_3 \\
y_1^2 & y_2^2 & y_3^2 \\
x_1^3 & x_2^3 & x_3^3 \\
x_1^2y_1 & x_2^2y_2 & x_3^2y_3 \\
x_1y_1^2 & x_2y_2^2 & x_3y_3^2 \\
y_1^3 & y_2^3 & y_3^3 \\
x_1^4 & x_2^4 & x_4^4 \\
x_1^3y_1 & x_2^3y_2 & x_3^3y_3 \\
x_1^2y_1^2 & x_2^2y_2^2 & x_3^2y_3^2 \\
x_1y_1^3 & x_2y_2^3 & x_3y_3^3 \\
y_1^4 & y_2^4 & y_3^4 \\
\vdots & \vdots & \vdots
\end{bmatrix}
$$

simplified:

$$
\begin{bmatrix}
1 & 1 & 1 \\
x_1 & x_2 & x_3 \\
y_1 & y_2 & y_3 \\
x_1y_1 & x_2y_2 & x_3y_3 \\
x_1^3 & x_2^3 & x_3^3 \\
x_1^2y_1 & x_2^2y_2 & x_3^2y_3
\end{bmatrix}
\begin{bmatrix}
x_1 & & \\
 & x_2 & \\
 & & x_3
\end{bmatrix}
=
\begin{bmatrix}
x_1 & x_2 & x_3 \\
x_1^2 & x_2^2 & x_3^2 \\
x_1y_1 & x_2y_2 & x_3y_3 \\
x_1^2y_1 & x_2^2y_2 & x_3^2y_3 \\
x_1^4 & x_2^4 & x_4^4 \\
x_1^3y_1 & x_2^3y_2 & x_3^3y_3
\end{bmatrix}
$$

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

$$S_1 K D_x = S_x K,$$

where S_1 and S_x select rows from K w.r.t. shift property

- **Realization Theory** for the unknown x

We have

$$S_1 K D_x = S_x K$$

Generalized Vandermonde K is not known as such, instead a null space basis Z is calculated, which is a linear transformation of K:

$$ZV = K$$

which leads to

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

Here, V is the matrix with eigenvectors, D_x contains the roots x as eigenvalues.

KU LEUVEN

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!
- $(S_x Z)^{-1}(S_1 Z)$ and $(S_y Z)^{-1}(S_1 Z)$ commute
  $\Longrightarrow$ ‘commutative algebra’

Basic Algorithm outline

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

$$S_1 K D = S_{\text{shift}} K$$

with K generalized Vandermonde, not known as such. Instead a basis Z is computed as

$$ZV = K$$

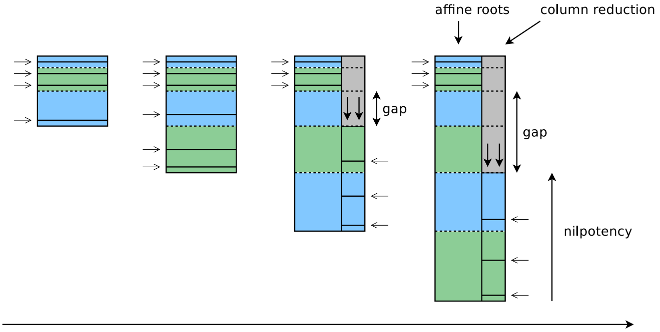
which leads to

$$(S_{\text{shift}} Z) V = (S_1 Z) V D$$

S_1 selects linear independent rows; S_{shift} selects rows 'hit' by the shift.

'Mind the Gap' and 'A Bout de Souffle'

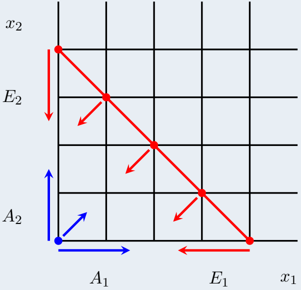
- Dynamics in the null space of $M(d)$ for increasing degree d: The index of some of the linear independent rows stabilizes (=affine zeros); The index of other ones keeps increasing (=zeros at ∞).
- 'Mind-the-gap': As a function of d, certain degree blocks become and stay linear dependent on all preceeding rows: allows to count and seperate affine zeros from zeros at ∞
- 'A bout de souffle': Effect of zeros at ∞ 'dies' out (nilpotency).



- Weierstrass Canonical Form decoupling affine and infinity roots

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

- Action of A_i and E_i represented in grid of monomials



Roots at Infinity: nD Descriptor Systems

Weierstrass Canonical Form decouples affine/infinity

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

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

$$\begin{array}{rcl} v(k+1,l) & = & A_x v(k,l) \\ v(k,l+1) & = & A_y v(k,l) \\ \\ w(k-1,l) & = & E_x w(k,l) \\ w(k,l-1) & = & E_y w(k,l) \end{array}$$

with E_x and E_y nilpotent matrices.

Summary

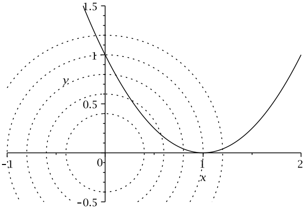
- Rooting multivariate polynomials
  - = (numerical) linear algebra
  - = (fund. thm. of algebra) $\bigcap$ (fund. thm. of linear algebra)
  - = nD realization theory in null space of Macaulay matrix
- Decisions based upon (numerical) rank
  - Dimension of variety = degree of Hilbert polynomial: follows from corank (nullity);
  - For 0-dimensional varieties ('isolated' roots): corank stabilizes = # roots (nullity)
  - 'Mind-the-gap' splits affine zeros from zeros at ∞
  - # affine roots (dimension column compression)
- not discussed
  - Multiplicity of roots ('confluent' generalized Vandermonde matrices)
  - Macaulay matrix columnspace based methods ('data driven')

Outline

KU LEUVEN

Polynomial Optimization Problems

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



Lagrange multipliers: necessary conditions for optimality:

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

$$
\begin{array}{lcl}
\partial L / \partial x = 0 & \rightarrow & 2x - 2xz + 2z = 0 \\
\partial L / \partial y = 0 & \rightarrow & 2y + z = 0 \\
\partial L / \partial z = 0 & \rightarrow & y - x^2 + 2x - 1 = 0
\end{array}
$$

Observations:

- all equations remain polynomial
- all ‘stationary’ points (local minima/maxima, saddle points) are roots of a system of polynomial equations
- shift with objective function to find minimum: only minimizing roots are needed !

Let

$$A_x V = V D_x$$

and

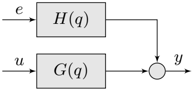
$$A_y V = V D_y$$

then find minimum eigenvalue of

$$(A_x^2 + A_y^2) V = V (D_x^2 + D_y^2)$$

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

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

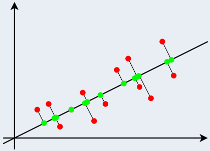
subject to the model equations

- Example

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

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

Static Linear Modeling



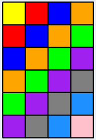
- Rank deficiency
- minimization problem:

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

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

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

Dynamical Linear Modeling



- Rank deficiency
- minimization problem:

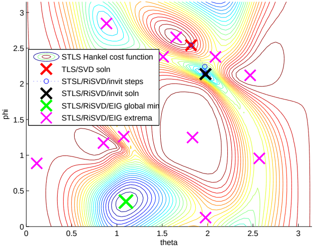
$$\begin{array}{ll} \min & \| [\Delta A \quad \Delta b] \|_F^2 , \\ \text{s. t.} & (A + \Delta A)v = b + \Delta b, \\ & v^T v = 1 \\ & [\Delta A \quad \Delta b] \text{ structured} \end{array}$$

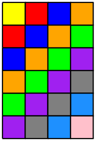
- Riemannian SVD: find (u, τ, v) which minimizes τ^2

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

KU LEUVEN

$$\begin{array}{ll} \min\limits_{v} & \tau^2 = v^T M^T D_v^{-1} M v \\ \text{s. t.} & v^T v = 1. \end{array}$$





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

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

- 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

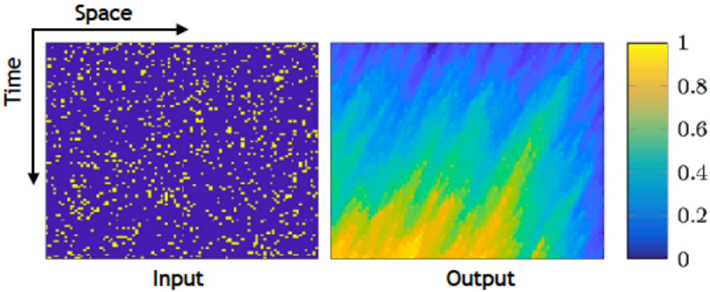
- Polynomial Optimization Problems
- Structured Total Least Squares
- H_2 Model order reduction
- Analyzing identifiability of nonlinear model structures (differential algebra)
- Robotics: kinematic problems
- Computational Biology: conformation of molecules
- Algebraic Statistics
- Signal Processing
- nD dynamical systems; Partial difference equations
- ...

Outline

- Finding roots: linear algebra and realization theory!
- Polynomial optimization: extremal eigenvalue problems
- (Numerical) linear algebra/systems theory translation of algebraic geometry/symbolic algebra
- Many problems are in fact eigenvalue 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)
- Projecting up to higher dimensional space (difficult in low number of dimensions; ‘easy’ (=large EVP) in high number of dimensions)

Current work:

- Subspace identification for spatially-temporarilly correlated signals (partial difference equations)
- Modelling in the era of IoT (Internet-of-Things) with its tsunami of data: in space and time (e.g. trajectories over time); or e.g. in MSI (mass spectrometry imaging): spectrum (1D) per space-voxel (3D) over time (1D) = 5D-tensor. How to model ?
- Example: Advection - diffusion equation space-time with input-output data:



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)

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

Anders 'free will' Lindquist

Sculpture by Joos Vandewalle

Ad multos annos !!

A variety in algebraic geometry

KU LEUVEN