Back to the Roots

Polynomial System Solving Using Linear Algebra

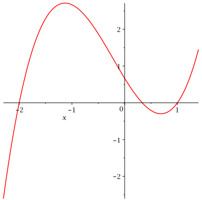
Philippe Dreesen

KU Leuven
Department of Electrical Engineering ESAT-STADIUS
Stadius Center for Dynamical Systems, Signal Processing and Data Analytics

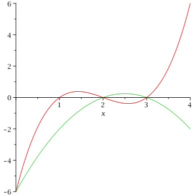
Outline

Four instances of polynomial root-finding problems

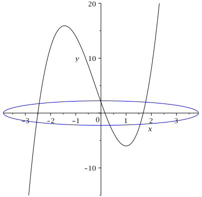
$$(x-1)(x+2)\left(x-\frac{1}{3}\right) \quad = \quad 0$$



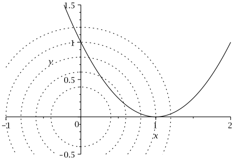
$$\begin{aligned}(x-1)(x-3)(x-2) &= 0 \\ -(x-2)(x-3) &= 0\end{aligned}$$



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

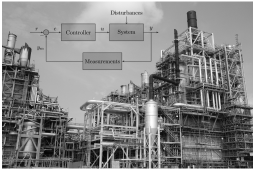


$$\begin{aligned}\min_{x,y} \quad & x^2+y^2 \\ \text{s.t.} \quad & y-x^2+2x-1=0\end{aligned}$$

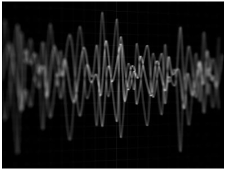


Why Study Polynomial Equations?

- fundamental mathematical objects
- powerful modelling tools
- ubiquitous in Science and Engineering (often *hidden*)



Systems and Control

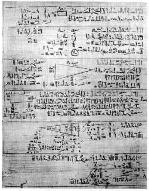


Signal Processing

Computational Biology

Kinematics/Robotics

A long and rich history. . .



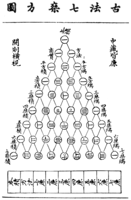
Egypt
(3000BCE-300BCE)

Babylon
(3000BCE-539BCE)

Euclid
(fl. 300BCE)

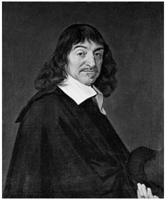
Diophantus
(c200-c284)

Al-Khwarizmi
(c780-c850)

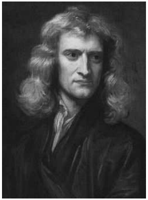


Zhu Shijie
(c1260-c1320)

Pierre de Fermat
(c1601-1665)



René Descartes
(1596-1650)



Isaac Newton
(1643-1727)

Gottfried Leibniz
(1646-1716)

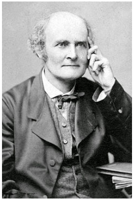
. . . leading to “Algebraic Geometry”

Etienne Bézout
(1730-1783)

Carl Friedrich Gauss
(1777-1755)

Jean-Victor Poncelet
(1788-1867)

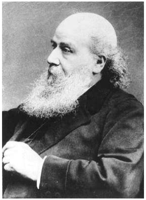
Evariste Galois
(1811-1832)



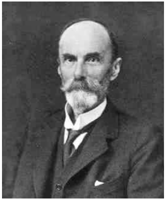
Arthur Cayley
(1821-1895)

Leopold Kronecker
(1823-1891)

Edmond Laguerre
(1834-1886)



James J. Sylvester
(1814-1897)

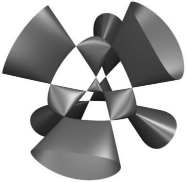


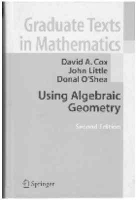
Francis S. Macaulay
(1862-1937)

David Hilbert
(1862-1943)

Algebraic Geometry and Computer Algebra

- large body of literature
- emphasis not (anymore) on *solving* equations
- computer algebra: symbolic manipulations (e.g., Gröbner Bases)
- numerical issues!

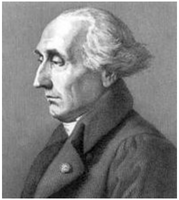




Wolfgang Gröbner
(1899-1980)

Bruno Buchberger

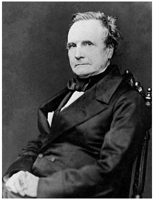
. . . and (Numerical) Linear Algebra



Joseph-Louis Lagrange
(1736-1813)

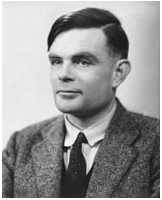
Augustin-Louis Cauchy
(1789-1857)

Hermann Grassmann
(1809-1877)



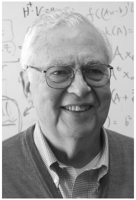
Charles Babbage
(1791-1871)

Ada Lovelace
(1815-1852)



Alan Turing
(1912-1954)

John von Neumann
(1903-1957)



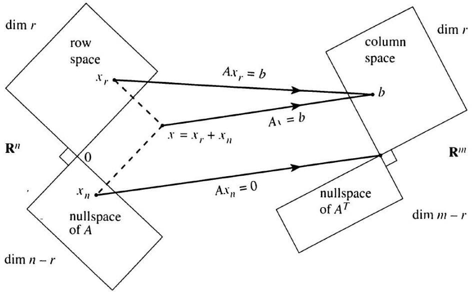
Gene Golub
(1932-2007)

Daniel Lazard

Hans J. Stetter

Why Linear Algebra?

- comprehensible and accessible language
- intuitive geometric interpretation
- computationally powerful framework
- well-established methods and stable numerics



KU LEUVEN

Eigenvalue Problems

Eigenvalue equation

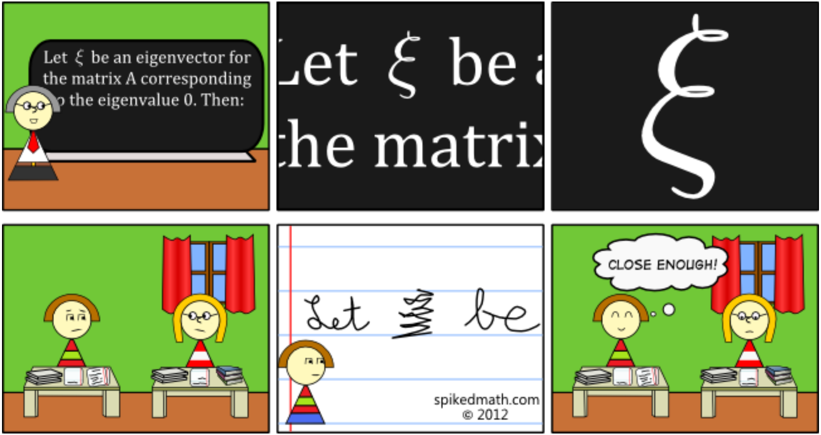
$$Av = \lambda v$$

and eigenvalue decomposition

$$A = V \Lambda V^{-1}$$

Enormous importance in (numerical) linear algebra and apps

- 'understand' the action of matrix A
- at the heart of a multitude of applications: oscillations, vibrations, quantum mechanics, data analytics, graph theory, and **many** more


Let ξ be an eigenvector for the matrix A corresponding to the eigenvalue 0. Then:
CLOSE ENOUGH!
spikedmath.com
© 2012

Outline

Univariate Polynomials and Linear Algebra: Known Facts

Characteristic Polynomial

The eigenvalues of A are the roots of

$$p(\lambda) = |A - \lambda I|$$

Companion Matrix

Solving

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

leads to

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

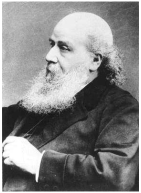
Sylvester Matrix

Consider two polynomial equations

$$\begin{array}{rcl} f(x) & = & x^3 - 6x^2 + 11x - 6 & = & (x-1)(x-2)(x-3) \\ g(x) & = & -x^2 + 5x - 6 & = & -(x-2)(x-3) \end{array}$$

Common roots if $|S(f,g)| = 0$

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



James Joseph Sylvester

Sylvester's construction can be understood from

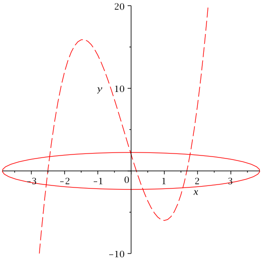
$$
\begin{array}{c}
f(x)=0 \\
x\cdot f(x)=0 \\
g(x)=0 \\
x\cdot g(x)=0 \\
x^2\cdot g(x)=0
\end{array}
\quad
\overset{\begin{array}{ccccc} 1 & x & x^2 & x^3 & x^4 \end{array}}{
\begin{bmatrix}
-6 & 11 & -6 & 1 & 0 \\
 & -6 & 11 & -6 & 1 \\
-6 & 5 & -1 & & \\
 & -6 & 5 & -1 & \\
 & & -6 & 5 & -1
\end{bmatrix}}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

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

Outline

Consider the system

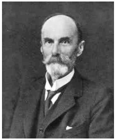
$$\begin{aligned} p(x,y) &= x^2 + 3y^2 - 15 &= 0 \\ q(x,y) &= y - 3x^3 - 2x^2 + 13x - 2 &= 0 \end{aligned}$$



Matrix representation of the system: Macaulay matrix M

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

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



Continue to enlarge the Macaulay matrix M:

		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	$\rightarrow$
$d=3$	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				
$\downarrow$								$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$

- Macaulay 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 null space

$$MK = 0$$

- number of solutions m = nullity

Multivariate Vandermonde basis for the null space:

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

Select the ‘top’ m linear independent rows of K

$$S_1 K$$

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

Shifting the selected rows gives (shown for 3 columns)

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

simplified:

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

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

$$S_1 K D_x = S_x K,$$

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

- reminiscent of **Realization Theory**

We have

$$S_1 \, K D_x = S_x \, K$$

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

$$ZV = K$$

Which leads to

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

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

The null space of the Macaulay matrix is the interface between polynomial system and nD state space description

- Attasi model (for $n = 2$)

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

- null space of Macaulay matrix: nD state sequence

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

$$\left(\begin{array}{c|cc|c|cccc} | & | & | & & | & | & | & | \\ v_{00} & A_x v_{00} & A_y v_{00} & \cdots & A_x^3 v_{00} & A_x^2 A_y v_{00} & A_x A_y^2 v_{00} & A_y^3 v_{00} \\ | & | & | & & | & | & | & | \end{array} \right)^T$$

- shift-invariance property, e.g., for y:

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

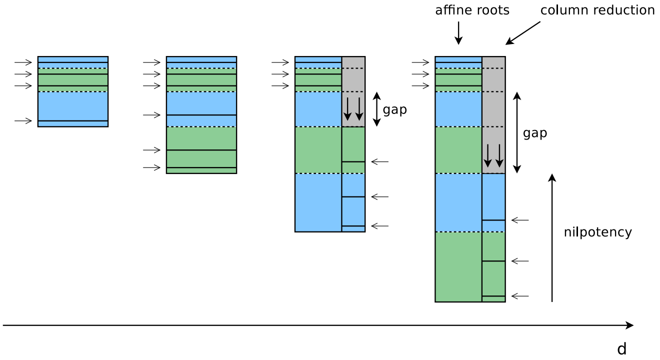
- corresponding nD system realization

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

- choice of basis null space leads to different system realizations
- eigenvalues of A_x and A_y invariant: x and y components of roots

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



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.

KU LEUVEN

Two extensions of the root-finding method:

Column-space based root-finding method

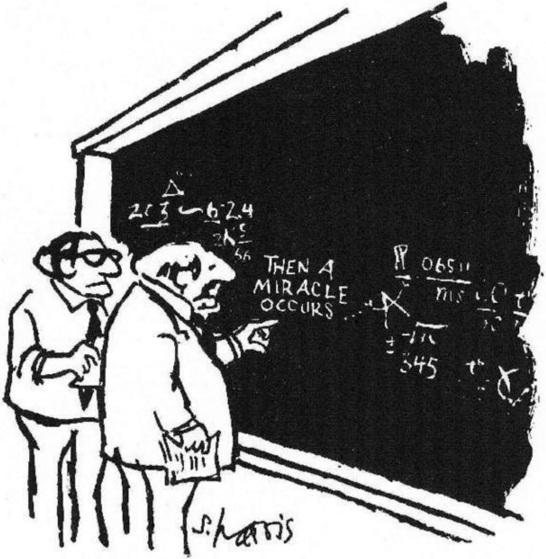
- dual method operating on column space instead of null space
- leads again to eigenvalue problems
- employs (Q)R-decomposition

Finding approximate solutions of over-constrained systems

- generalization to over-constrained (noisy) systems
- approximate solutions detectable by computing SVD of M
- example from computer vision: camera pose determination

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

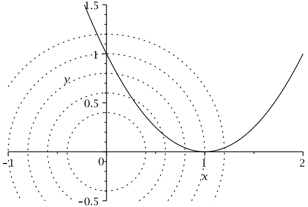


"I think you should be more explicit here in step two."

Outline

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 give conditions for optimality:

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

we find

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

Observations:

- 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

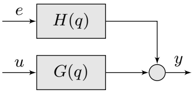
$$(A_x^2 + A_y^2)V = (x^2 + y^2)V$$

KU LEUVEN

Polynomial Optimization Problems: Applications

- PEM System identification = EVP !!
- 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

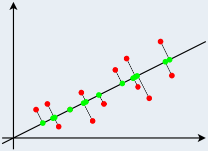
$$\hat{y}_k = H^{-1}(q)G(q)u_k + (1 - H^{-1})y_k,$$

subject to the model equations

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



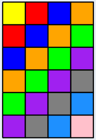
- Rank deficiency
- minimization problem:

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

- Singular Value Decomposition:
  find (u, σ, v) which minimizes σ^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



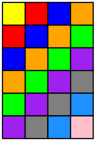
- Rank deficiency
- minimization problem:

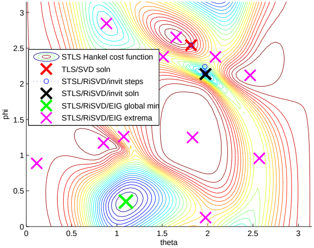
$$\begin{array}{ll} \min & \|[\Delta a \quad \Delta b]\|_F^2, \\ \text{s.t.} & (A+\Delta A)v = B + \Delta B, \\ & \Delta A = f(\Delta a) \text{ structured} \\ & \Delta B = g(\Delta b) \text{ structured} \end{array}$$

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

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

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





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**

Outline

Conclusions

- bridging the gap between algebraic geometry and engineering
- finding roots: linear algebra and realization theory!
- extension to over-constrained systems
- polynomial optimization: extremal eigenvalue problems

Open Problems

Many challenges remain

- exploiting sparsity and structure of M
- efficient (more direct) construction of the eigenvalue problem
- algorithms to find the minimizing solution efficiently (inverse power method?)
- nD version of Cayley-Hamilton theorem
- analyzing the conditioning of the root-finding problem

Thank you for listening!