Multivariate Polynomials (with MacaulayLab) for Beginners

Flanders AI Research Day

Christof Vermeersch

christof.vermeersch@esat.kuleuven.be

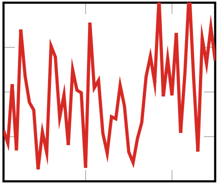
STADIUS
Center for Dynamical Systems, Signal Processing, and Data Analytics

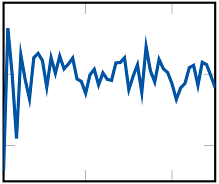
October 16, 2023

Why multivariate polynomials?

$$\underbrace{\begin{bmatrix} e_1 \\ \vdots \\ e_M \end{bmatrix}}_{\text{unknown, latent input}} \longrightarrow \boxed{\frac{c(q)}{a(q)}} \longrightarrow \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}}_{\text{known output}}$$

such that $\sum_{i=0}^{n_a} \alpha_i y_{k-i} = \sum_{i=0}^{n_c} \gamma_i e_{k-i}$





$$\min_{\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{e}} \|\boldsymbol{e}\|_2^2$$

subject to $\boldsymbol{T}_\alpha \boldsymbol{y} = \boldsymbol{T}_\gamma \boldsymbol{e}$

(Vermeersch and De Moor, 2019)

Outline

Outline

Multivariate polynomials

A **multivariate polynomial** $p(\boldsymbol{x})$, or $p(x_1, \ldots, x_n)$, in n variables is a finite linear combination of monomials $\boldsymbol{x}^{\boldsymbol{\alpha}}$ from K^n with coefficients $c_{\boldsymbol{\alpha}}$ from K:

$$p(\boldsymbol{x}) = \sum_{\mathcal{A}} c_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}},$$

where the summation runs over all the exponents in the set $\mathcal{A}$.

- K can be any field: complex numbers $\mathbb{C}$, real numbers $\mathbb{R}$, or finite numbers $\mathbb{F}_q$
- $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ indexes the monomials $\boldsymbol{\alpha}^{\boldsymbol{\alpha}}$ and coefficients $c_{\boldsymbol{\alpha}}$
- set of monomials, $\mathcal{C}_d^n$, consists out of $\binom{d+n}{n}$ elements
- example: $p(\boldsymbol{x}) = 3 + \sqrt{5}x_1 + (1+i)x_2 + \frac{3}{2}x_1^2x_2^8$

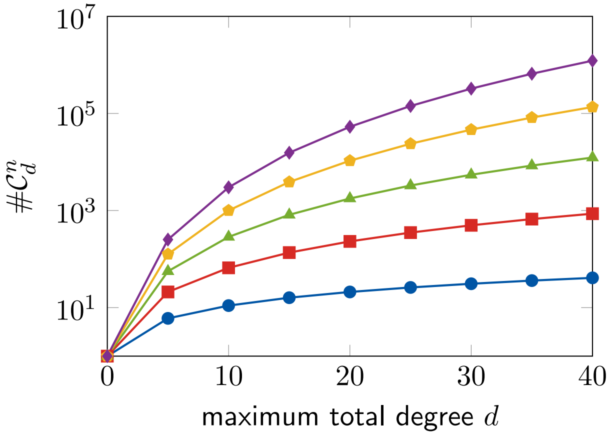
Multivariate polynomials

- Typically, multivariate polynomials appear in systems of multivariate polynomial equations:

$$\begin{cases} p_1(\boldsymbol{x}) = \sum_{\mathcal{A}} c^{(1)}_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} = 0, \\ \qquad \vdots \\ p_s(\boldsymbol{x}) = \sum_{\mathcal{A}} c^{(s)}_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} = 0, \end{cases}$$

  where we look for the solutions of s multivariate polynomials in n variables.
- Every polynomial has a total degree: $d_i = \max(|\boldsymbol{\alpha}|)$.

Combinatorial explosion



Combinatorial explosion of the number of monomials in the set $\mathcal{C}_d^n$ with respect to the maximum total degree d and number of variables n. The cardinality of $\mathcal{C}_d^n$ is given for $n = 1$ (blue circle), $n = 2$ (red square), $n = 3$ (green triangle), $n = 4$ (yellow pentagon), and $n = 5$ (purple diamond).

Two limit cases

univariate polynomial ($n = 1$)

$$p(x) = \sum_{i=0}^{d} c_i x^i$$

$$\Downarrow$$

$$\boldsymbol{C}_p = \begin{bmatrix} 0 & 0 & \cdots & 0 & -c_0/c_n \\ 1 & 0 & \cdots & 0 & -c_1/c_n \\ 0 & 1 & \cdots & 0 & -c_2/c_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1}/c_n \end{bmatrix}$$

$$\Downarrow$$

$$\boldsymbol{C}_p \boldsymbol{x} = \lambda \boldsymbol{x}$$

linear systems ($d_i = 1$)

$$\begin{cases} p_1(\boldsymbol{x}) = b_1 + \sum_{j=1}^{n} a_{1j} x_j \\ \vdots \\ p_s(\boldsymbol{x}) = b_s + \sum_{j=1}^{n} a_{sj} x_j \end{cases}$$

$$\Downarrow$$

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

These are well-known problems from linear algebra

Different polynomial basis

$$p(\boldsymbol{x}) = \sum_{\mathcal{W}} \tilde{c}_{\boldsymbol{\omega}} t_{\boldsymbol{\omega}(\boldsymbol{x})}$$

A multivariate Chebyshev polynomial of degree $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$:

$$t_{\boldsymbol{\alpha}}(\boldsymbol{x}) = \prod_{i=1}^{n} t^*_{\alpha_i}(x_i) = \prod_{i=1}^{n} \cos\big(\alpha_i \cos^{-1}(x_i)\big),$$

where $t^*_{\alpha_i}(x_i)$ corresponds to the univariate Chebyshev polynomial of degree α_i.

Examples:

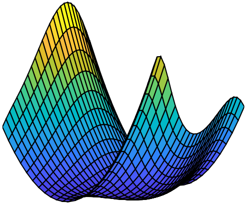
$$t_2(x) = 2x^2 - 1$$

$$t_{13}(x_1, x_2) = x_1\big(4x_2^3 - 3x_2\big)$$

$$t_{162}(x_1, x_2, x_3) = x_1\big(32x_2^6 - 48x_2^4 + 18x_2^2 - 1\big)\big(2x_3^2 - 1\big)$$

Some applications

polynomial optimization problems



$$\min_{\boldsymbol{x} \in K^n} g(\boldsymbol{x})$$

$$\text{subject to } h_1(\boldsymbol{x}) = \cdots = h_l(\boldsymbol{x}) = 0$$

$$\downarrow$$

This can be solved via the Lagrangian $\mathcal{L}(\boldsymbol{x}, \lambda_1, \ldots, \lambda_l)$:

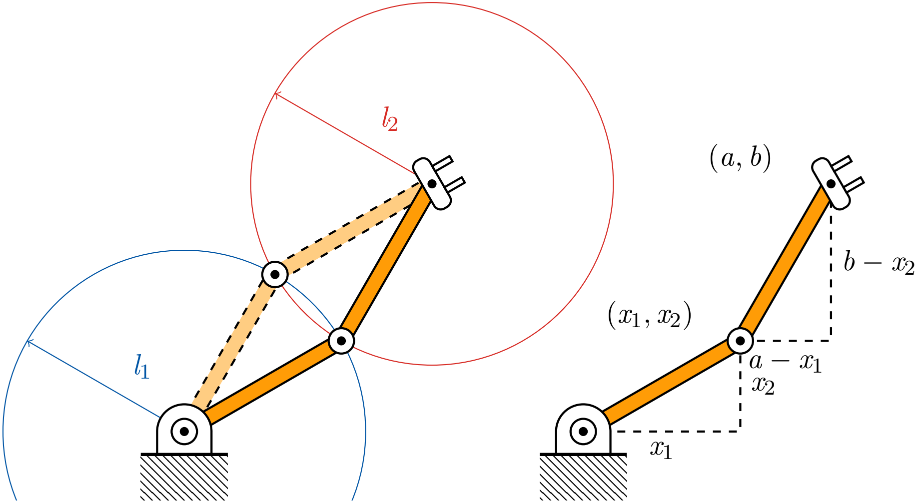
$$\mathcal{L}(\boldsymbol{x}, \lambda_1, \ldots, \lambda_l) = g(\boldsymbol{x}) - \lambda_1 h_1(\boldsymbol{x}) - \cdots - \lambda_l h_l(\boldsymbol{x})$$

The first-order necessary conditions are given by the partial derivatives:

$$\frac{\partial \mathcal{L}}{\partial x_1} = \cdots = \frac{\partial \mathcal{L}}{\partial x_n} = h_1(\boldsymbol{x}) = \cdots = h_l(\boldsymbol{x}) = 0$$

Some applications

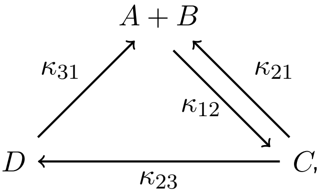
robotics and kinematics



Some applications

computational chemistry

The (bio)chemical reaction network of the T-cell signal transduction model:

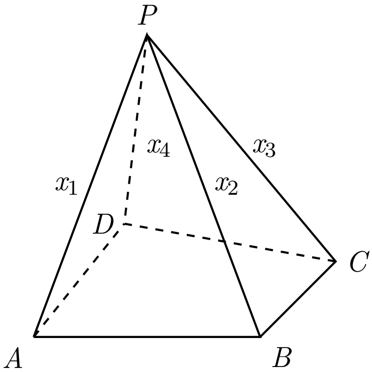


$$\Rightarrow \begin{cases} \dfrac{dx_A(t)}{dt} = -\kappa_{12} x_A(t) x_B(t) + \kappa_{21} x_C(t) + \kappa_{31} x_D(t) = 0, \\ \dfrac{dx_B(t)}{dt} = -\kappa_{12} x_A(t) x_B(t) + \kappa_{21} x_C(t) + \kappa_{31} x_D(t) = 0, \\ \dfrac{dx_C(t)}{dt} = \kappa_{12} x_A(t) x_B(t) - \kappa_{21} x_C(t) - \kappa_{23} x_C(t) = 0, \\ \dfrac{dx_D(t)}{dt} = \kappa_{23} x_C(t) - \kappa_{31} x_D(t) = 0. \end{cases}$$

Some applications

computer graphics

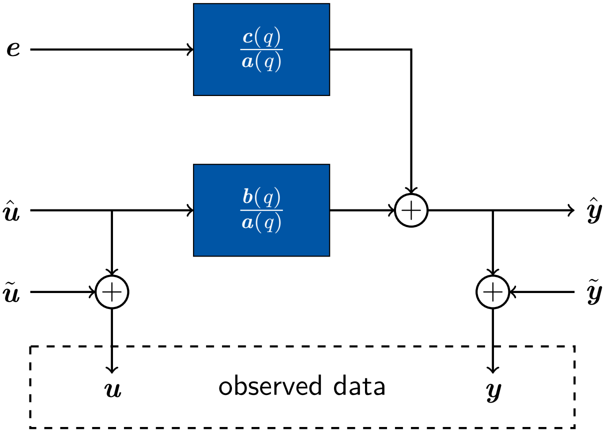
Estimating internal calibration parameters of a camera from point correspondences in a sequence of (noisy) images:

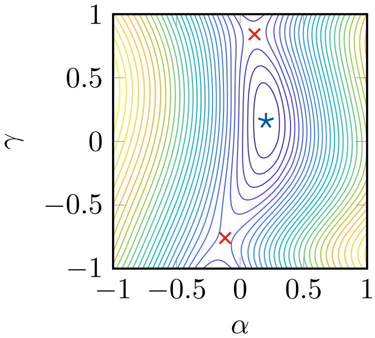


$$\Rightarrow \begin{cases} x_1^2 + x_2^2 - r x_1 x_2 - \|AB\|^2 \approx 0, \\ x_1^2 + x_3^2 - q x_1 x_3 - \|AC\|^2 \approx 0, \\ x_2^2 + x_3^2 - p x_2 x_3 - \|BC\|^2 \approx 0, \\ x_1^2 + x_4^2 - s x_1 x_4 - \|AD\|^2 \approx 0, \\ x_4^2 + x_3^2 - t x_1 x_2 - \|CD\|^2 \approx 0, \\ x_2^2 + x_4^2 - u x_2 x_4 - \|BD\|^2 \approx 0, \end{cases}$$

Some applications

system identification and time series





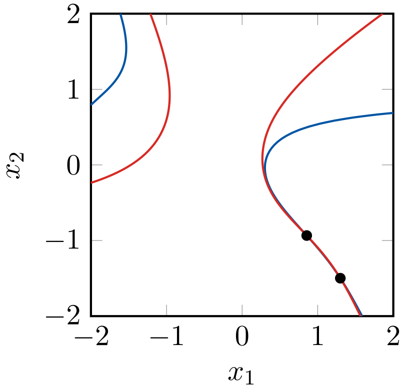
Outline

What does solving mean?

Find all the values for $\boldsymbol{x} \in \bar{K}^n$ such that $p_1(\boldsymbol{x}) = \cdots = p_s(\boldsymbol{x}) = 0$, i.e., the variety of the polynomial system

$$\mathcal{V}(p_1, \ldots, p_s) = \left\{ \boldsymbol{a} \in \bar{K}^n : p_i(\boldsymbol{a}) = 0, \forall i = 1, \ldots s \right\}$$

- Typically, we consider polynomial systems which are well-determined! This can be for both square and rectangular systems.
- Some solution approaches can deal with over-determined polynomial systems.
- Under-determined polynomial systems could be solved in a certain sense, but what does it mean?

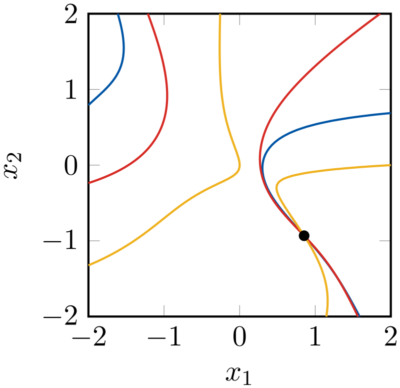


What does solving mean?

Find all the values for $\boldsymbol{x} \in \bar{K}^n$ such that $p_1(\boldsymbol{x}) = \cdots = p_s(\boldsymbol{x}) = 0$, i.e., the variety of the polynomial system

$$\mathcal{V}(p_1, \dots, p_s) = \left\{\boldsymbol{a} \in \bar{K}^n : p_i(\boldsymbol{a}) = 0, \forall i = 1, \dots s\right\}$$

- Typically, we consider polynomial systems which are well-determined! This can be for both square and rectangular systems.
- Some solution approaches can deal with over-determined polynomial systems.
- Under-determined polynomial systems could be solved in a certain sense, but what does it mean?

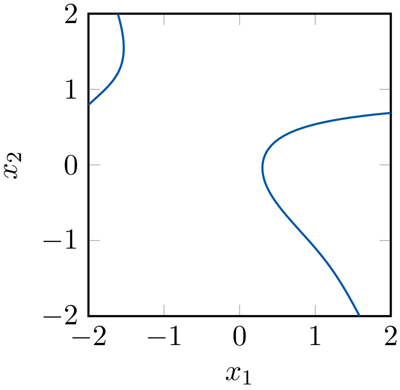


What does solving mean?

Find all the values for $\boldsymbol{x} \in \bar{K}^n$ such that $p_1(\boldsymbol{x}) = \cdots = p_s(\boldsymbol{x}) = 0$, i.e., the variety of the polynomial system

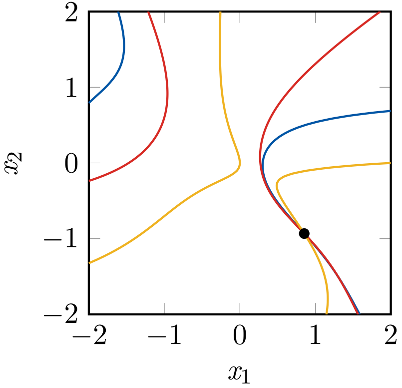
$$\mathcal{V}(p_1, \ldots, p_s) = \{\boldsymbol{a} \in \bar{K}^n : p_i(\boldsymbol{a}) = 0, \forall i = 1, \ldots s\}$$

- Typically, we consider polynomial systems which are well-determined! This can be for both square and rectangular systems.
- Some solution approaches can deal with over-determined polynomial systems.
- Under-determined polynomial systems could be solved in a certain sense, but what does it mean?



What does solving mean?

Everything depends on the ground field that you consider!



This polynomial system has 9 solutions in $\mathbb{C}^2$, 1 solution in $\mathbb{R}^2$, and 0 solutions in $\mathbb{Q}^2$

Some examples:

- $\mathbb{C}$: homotopy continuation, particle physics
- $\mathbb{R}$: optimization, chemical reaction networks, robotics
- $\mathbb{Q}$: discriminants/resultants, Grassmannians, number theory
- $\mathbb{F}_q$: cryptography
- $\mathbb{C}\{\{t\}\}, \mathbb{Q}_p$: tropical geometry

Number of solutions?

- For univariate polynomials (i.e., $n = 1$), the **fundamental theorem of algebra** states that a degree d polynomial has d roots.
- The **theorem of Bézout** is the multivariate extension of that theorem.

For any square system (i.e., $s = n$) of multivariate polynomial equations $p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x})$, the number of isolated solutions in the projective space $\mathbb{P}^n$ when the solution set is zero-dimensional, i.e., the number of isolated points in the zero-dimensional variety $\mathcal{V}(p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x})) \subset \mathbb{P}^n$, is exactly equal to

$$m_b = d_1 \cdots d_n = \prod_{i=1}^{n} d_i,$$

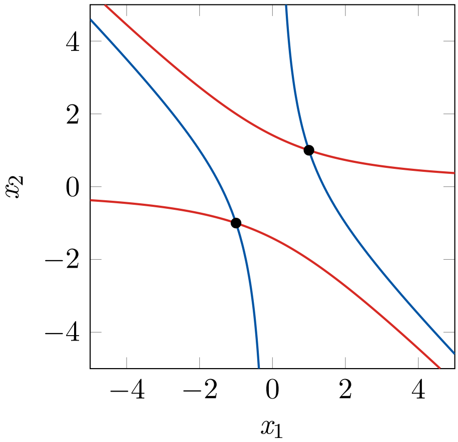
where d_i is the total degree of the polynomial $p_i(\boldsymbol{x})$.

Number of solutions?

- The theorem of Bézout counts the number of isolated solutions in the projective space:

$$m_b = m_a + m_\infty$$

- For generic systems, $m_b = m_a$, but in practice this is not the case
- There exist more refined bounds on the number of affine solutions (e.g., Kushnirenko, BKK, etc.)



Different solution approaches

When we consider an algebraically closed field, there are two main methods to solve systems of multivariate polynomial equations:

normal form methods

- reduce problem to a univariate problem
- (numerical) linear algebra
- any field K^n
- rectangular systems: $s \geq n$
- $m_b < \pm 10\,000$ solutions

homotopy continuation methods

- continuously deform a system with known solutions
- ordinary differential equations
- field of complex numbers $\mathbb{C}^n$
- square systems: $s = n$
- $m_b < \pm 1\,000\,000$ solutions

Implementations

normal form methods

- Macaulay2
- Singular
- Magma
- msolve
- Mathematica
- Oscar.jl
- Maple
- MacaulayLab
- ...

homotopy continuation methods

- HomotopyContinuation.jl
- Bertini
- NAG4M2
- Hom4PS-3
- PHCpack
- ...

Outline

Normal form methods and the Macaulay matrix

$$\begin{cases} p_1(\boldsymbol{x}) = a_{00} + a_{10}x_1 + a_{01}x_2 + a_{20}x_1^2 + a_{11}x_1x_2 + a_{02}x_2^2 = 0 \\ p_2(\boldsymbol{x}) = b_{00} + b_{10}x_1 + b_{01}x_2 + b_{20}x_1^2 + b_{11}x_1x_2 + b_{02}x_2^2 = 0 \end{cases}$$

$$\Downarrow$$

$$\begin{array}{r} \\ p_1(\boldsymbol{x}) \\ x_1p_1(\boldsymbol{x}) \\ x_2p_1(\boldsymbol{x}) \\ p_2(\boldsymbol{x}) \\ x_1p_2(\boldsymbol{x}) \\ x_2p_2(\boldsymbol{x}) \end{array}
\underbrace{\begin{array}{cccccccccc} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 & x_1^3 & x_1^2x_2 & x_1x_2^2 & x_2^3 \\ a_{00} & a_{10} & a_{01} & a_{20} & a_{11} & a_{02} & 0 & 0 & 0 & 0 \\ 0 & a_{00} & 0 & a_{10} & a_{01} & 0 & a_{20} & a_{11} & a_{02} & 0 \\ 0 & 0 & a_{00} & 0 & a_{10} & a_{01} & 0 & a_{20} & a_{11} & a_{02} \\ b_{00} & b_{10} & b_{01} & b_{20} & b_{11} & b_{02} & 0 & 0 & 0 & 0 \\ 0 & b_{00} & 0 & b_{10} & b_{01} & 0 & b_{20} & b_{11} & b_{02} & 0 \\ 0 & 0 & b_{00} & 0 & b_{10} & b_{01} & 0 & b_{20} & b_{11} & b_{02} \end{array}}_{\text{Macaulay matrix}}
\begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \\ x_1^3 \\ x_1^2x_2 \\ x_1x_2^2 \\ x_2^3 \end{bmatrix} = \mathbf{0}$$

Structured right null space

Basis matrix of the right null space can be written in terms of the solutions ($d \geq d_*$)

- solutions can be described by the dual vector space of the quotient space
- from the rank-nullity theorem

$$n_d = q_d - r_d$$
$$= \dim \mathcal{P}_d^n / \left\langle p_1^h(\tilde{\boldsymbol{x}}), \ldots, p_s^h(\tilde{\boldsymbol{x}}) \right\rangle_d$$

- requires dual vector space $\mathcal{C}_d^{n'}$ of $\mathcal{C}_d^n$ and differential functionals

$$\partial_{\boldsymbol{i}}(\cdot)\big|_{(j)} \triangleq \frac{1}{i_1! \ldots i_n!} \frac{\partial^{|\boldsymbol{i}|}(\cdot)}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}}\Bigg|_{(j)}$$

confluent multivariate Vandermonde basis matrix *(projective setting)*

$$\boldsymbol{V}_d = \left[\begin{array}{ccc} x_0^2\big|_{(1)} & 0 & x_0^2\big|_{(2)} \\ \hline x_0x_1\big|_{(1)} & x_0\big|_{(1)} & x_0x_1\big|_{(2)} \\ x_0x_2\big|_{(1)} & 0 & x_0x_2\big|_{(2)} \\ \hline x_1^2\big|_{(1)} & 2x_1\big|_{(1)} & x_1^2\big|_{(2)} \\ x_1x_2\big|_{(1)} & x_2\big|_{(1)} & x_1x_2\big|_{(2)} \\ x_2^2\big|_{(1)} & 0 & x_2^2\big|_{(2)} \end{array}\right]$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

$$\partial_{00}(\boldsymbol{v})\big|_{(1)} \quad \partial_{10}(\boldsymbol{v})\big|_{(1)} \quad \partial_{00}(\boldsymbol{v})\big|_{(2)}$$

(Batselier et al., 2014; Stetter, 2004)

Monomial multiplicative property

one projective solution

$$\underbrace{\begin{bmatrix}1&0&0&0&0&0\\0&1&0&0&0&0\\0&0&1&0&0&0\end{bmatrix}}_{\boldsymbol{S}_{x_1,x_0}}\underbrace{\begin{bmatrix}x_0^2|_{(1)}\\ \hline x_0x_1|_{(1)}\\x_0x_2|_{(1)}\\ \hline x_1^2|_{(1)}\\x_1x_2|_{(1)}\\x_2^2|_{(1)}\end{bmatrix}}_{\partial_{00}(\boldsymbol{v})|_{(1)}}x_1|_{(1)}=\underbrace{\begin{bmatrix}0&1&0&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1&0\end{bmatrix}}_{\boldsymbol{S}_{x_0,x_1}}\underbrace{\begin{bmatrix}x_0^2|_{(1)}\\ \hline x_0x_1|_{(1)}\\x_0x_2|_{(1)}\\ \hline x_1^2|_{(1)}\\x_1x_2|_{(1)}\\x_2^2|_{(1)}\end{bmatrix}}_{\partial_{00}(\boldsymbol{v})|_{(1)}}x_0|_{(1)}$$

Note that we consider differential functionals in $\mathcal{C}_2^{2\,\prime}$ in this exposition.

Monomial multiplicative property

one projective solution

$$\boldsymbol{S}_{x_1,x_0} \underbrace{\begin{bmatrix} x_0^2|_{(1)} \\ \hline x_0x_1|_{(1)} \\ x_0x_2|_{(1)} \\ \hline x_1^2|_{(1)} \\ x_1x_2|_{(1)} \\ x_2^2|_{(1)} \end{bmatrix}}_{\partial_{00}(\boldsymbol{v})|_{(1)}} x_1|_{(1)} = \boldsymbol{S}_{x_0,x_1} \underbrace{\begin{bmatrix} x_0^2|_{(1)} \\ \hline x_0x_1|_{(1)} \\ x_0x_2|_{(1)} \\ \hline x_1^2|_{(1)} \\ x_1x_2|_{(1)} \\ x_2^2|_{(1)} \end{bmatrix}}_{\partial_{00}(\boldsymbol{v})|_{(1)}} x_0|_{(1)}$$

Note that we consider differential functionals in $\mathcal{C}_2^{2'}$ in this exposition.

Monomial multiplicative property

multiple simple projective solutions

$$
\boldsymbol{S}_{x_1,x_0}\underbrace{\begin{bmatrix} x_0^2\big|_{(1)} & x_0^2\big|_{(2)} \\ \hline x_0x_1\big|_{(1)} & x_0x_1\big|_{(2)} \\ x_0x_2\big|_{(1)} & x_0x_2\big|_{(2)} \\ \hline x_1^2\big|_{(1)} & x_1^2\big|_{(2)} \\ x_1x_2\big|_{(1)} & x_1x_2\big|_{(2)} \\ x_2^2\big|_{(1)} & x_2^2\big|_{(2)} \end{bmatrix}}_{\boldsymbol{V}_d}\boldsymbol{D}_{x_1} = \boldsymbol{S}_{x_0,x_1}\underbrace{\begin{bmatrix} x_0^2\big|_{(1)} & x_0^2\big|_{(2)} \\ \hline x_0x_1\big|_{(1)} & x_0x_1\big|_{(2)} \\ x_0x_2\big|_{(1)} & x_0x_2\big|_{(2)} \\ \hline x_1^2\big|_{(1)} & x_1^2\big|_{(2)} \\ x_1x_2\big|_{(1)} & x_1x_2\big|_{(2)} \\ x_2^2\big|_{(1)} & x_2^2\big|_{(2)} \end{bmatrix}}_{\boldsymbol{V}_d}\boldsymbol{D}_{x_0}
$$

$$
\boldsymbol{D}_{x_1} = \begin{bmatrix} x_1\big|_{(1)} & 0 \\ 0 & x_1\big|_{(2)} \end{bmatrix}
$$

Note that we consider differential functionals in $\mathcal{C}_2^{2'}$ in this exposition.

Monomial multiplicative property

multiple projective solutions with multiplicity greater than one

$$\boldsymbol{S}_{x_1,x_0} \underbrace{\left[\begin{array}{ccc} x_0^2\big|_{(1)} & 0 & x_0^2\big|_{(2)} \\ \hline x_0x_1\big|_{(1)} & x_0\big|_{(1)} & x_0x_1\big|_{(2)} \\ x_0x_2\big|_{(1)} & 0 & x_0x_2\big|_{(2)} \\ \hline x_1^2\big|_{(1)} & 2x_1\big|_{(1)} & x_1^2\big|_{(2)} \\ x_1x_2\big|_{(1)} & x_2\big|_{(1)} & x_1x_2\big|_{(2)} \\ x_2^2\big|_{(1)} & 0 & x_2^2\big|_{(2)} \end{array}\right]}_{\boldsymbol{V}_d} \boldsymbol{D}_{x_1} = \boldsymbol{S}_{x_0,x_1} \underbrace{\left[\begin{array}{ccc} x_0^2\big|_{(1)} & 0 & x_0^2\big|_{(2)} \\ \hline x_0x_1\big|_{(1)} & x_0\big|_{(1)} & x_0x_1\big|_{(2)} \\ x_0x_2\big|_{(1)} & 0 & x_0x_2\big|_{(2)} \\ \hline x_1^2\big|_{(1)} & 2x_1\big|_{(1)} & x_1^2\big|_{(2)} \\ x_1x_2\big|_{(1)} & x_2\big|_{(1)} & x_1x_2\big|_{(2)} \\ x_2^2\big|_{(1)} & 0 & x_2^2\big|_{(2)} \end{array}\right]}_{\boldsymbol{V}_d} \boldsymbol{D}_{x_0}$$

$$\boldsymbol{D}_{x_1} = \begin{bmatrix} x_1\big|_{(1)} & \times & \times \\ 0 & x_1\big|_{(1)} & \times \\ 0 & 0 & x_1\big|_{(2)} \end{bmatrix}$$

Note that we consider differential functionals in $\mathcal{C}_2^{2'}$ in this exposition.

Monomial multiplicative property

multiple affine solutions with multiplicity greater than one

$$\boldsymbol{S}_1 \underbrace{\left[\begin{array}{ccc} 1 & 0 & 1 \\ \hline x_1|_{(1)} & 1 & x_1|_{(2)} \\ x_2|_{(1)} & 0 & x_2|_{(2)} \\ \hline x_1^2|_{(1)} & 2x_1|_{(1)} & x_1^2|_{(2)} \\ x_1x_2|_{(1)} & x_2|_{(1)} & x_1x_2|_{(2)} \\ x_2^2|_{(1)} & 0 & x_2^2|_{(2)} \end{array}\right]}_{\boldsymbol{V}_d} \boldsymbol{D}_{x_1} = \boldsymbol{S}_{x_1} \underbrace{\left[\begin{array}{ccc} 1 & 0 & 1 \\ \hline x_1|_{(1)} & 1 & x_1|_{(2)} \\ x_2|_{(1)} & 0 & x_2|_{(2)} \\ \hline x_1^2|_{(1)} & 2x_1|_{(1)} & x_1^2|_{(2)} \\ x_1x_2|_{(1)} & x_2|_{(1)} & x_1x_2|_{(2)} \\ x_2^2|_{(1)} & 0 & x_2^2|_{(2)} \end{array}\right]}_{\boldsymbol{V}_d} \boldsymbol{I}$$

select the linearly independent rows of $\boldsymbol{V}_d$

Note that we consider differential functionals in $\mathcal{C}_2^{2'}$ in this exposition.

Monomial multiplicative property

multiple affine solutions with multiplicity greater than one

$$\boldsymbol{S}_1 \underbrace{\begin{bmatrix} 1 & 0 & 1 \\ \hline x_1|_{(1)} & 1 & x_1|_{(2)} \\ x_2|_{(1)} & 0 & x_2|_{(2)} \\ \hline x_1^2|_{(1)} & 2x_1|_{(1)} & x_1^2|_{(2)} \\ x_1x_2|_{(1)} & x_2|_{(1)} & x_1x_2|_{(2)} \\ x_2^2|_{(1)} & 0 & x_2^2|_{(2)} \end{bmatrix}}_{\boldsymbol{V}_d} \boldsymbol{D}_{x_1} = \boldsymbol{S}_{x_1} \underbrace{\begin{bmatrix} 1 & 0 & 1 \\ \hline x_1|_{(1)} & 1 & x_1|_{(2)} \\ x_2|_{(1)} & 0 & x_2|_{(2)} \\ \hline x_1^2|_{(1)} & 2x_1|_{(1)} & x_1^2|_{(2)} \\ x_1x_2|_{(1)} & x_2|_{(1)} & x_1x_2|_{(2)} \\ x_2^2|_{(1)} & 0 & x_2^2|_{(2)} \end{bmatrix}}_{\boldsymbol{V}_d} \boldsymbol{I}$$

eigenvalue problem!

$$\boldsymbol{S}_1 \boldsymbol{V}_d \boldsymbol{D}_{x_1} = \boldsymbol{S}_{x_1} \boldsymbol{V}_d$$

Note that we consider differential functionals in $\mathcal{C}_2^{2'}$ in this exposition.

Three difficulties

$$\boldsymbol{S}_1 \boldsymbol{V}_d \boldsymbol{D}_{x_1} = \boldsymbol{S}_{x_1} \boldsymbol{V}_d$$

- Solutions/confluent Vandermonde basis vectors are not known in advanced:

numerical basis matrix $\boldsymbol{Z}_d$ of the right null space $(\boldsymbol{V}_d = \boldsymbol{Z}_d \boldsymbol{T})$

$$\Downarrow$$

$$(\boldsymbol{S}_1 \boldsymbol{Z}_d) \boldsymbol{T} \boldsymbol{D}_{x_1} = (\boldsymbol{S}_{x_1} \boldsymbol{Z}_d) \boldsymbol{T}$$

$$\boldsymbol{T} \boldsymbol{D}_{x_1} \boldsymbol{T}^{-1} = (\boldsymbol{S}_1 \boldsymbol{Z}_d)^{-1} (\boldsymbol{S}_{x_1} \boldsymbol{Z}_d)$$

- Not possible to numerically stable compute Jordan normal form:

numerically stable **Schur decomposition**

$$\Downarrow$$

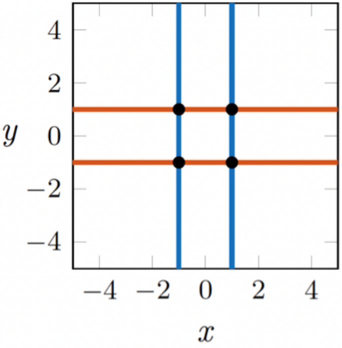
$$\boldsymbol{Q} \boldsymbol{U}_{x_1} \boldsymbol{Q}^{-1} = (\boldsymbol{S}_1 \boldsymbol{Z}_d)^{-1} (\boldsymbol{S}_{x_1} \boldsymbol{Z}_d)$$

- Solutions at infinity need to be deflated via a **column compression**

(Dreesen, 2013; Corless et al., 1997; Vermeersch and De Moor, 2021)

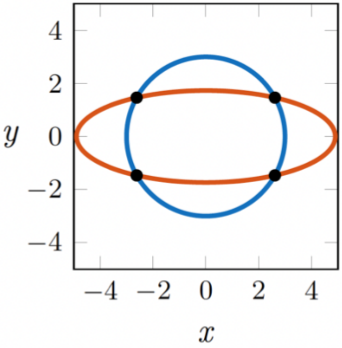
Outline

Homotopy continuation methods



$$\begin{cases} x^2 - 1 = 0 \\ y^2 - 1 = 0 \end{cases}$$

$\longrightarrow$

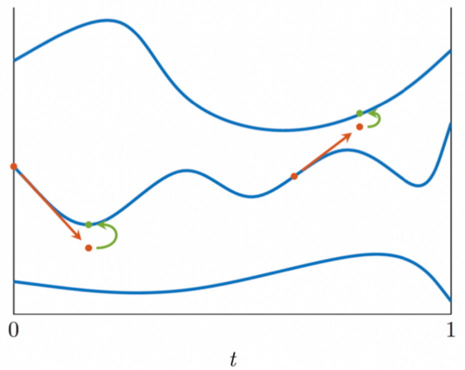


$$\begin{cases} x^2 + y^2 - 9 = 0 \\ 0.25x^2 + 2y^2 - 6 = 0 \end{cases}$$

Homotopy continuation methods

$$Q(\boldsymbol{x}) = (q_1(\boldsymbol{x}), \ldots, q_n(\boldsymbol{x})) \to P(\boldsymbol{x}) = (p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x}))$$

$$H(\boldsymbol{x}, t) = (1-t)Q(\boldsymbol{x}) + tP(\boldsymbol{x}), \quad t = 0 \to 1$$



Outline

MacaulayLab toolbox

`www.macaulaylab.net`

Command Window

```
>> disp(noon3)
   system of multivariate polynomial equations:
         - polynomial basis = monomial
         - number of equations = 3
         - number of variables = 3
         - maximum total degree = 3

   number of solutions:
         - total number of solutions = 27
         - number of affine solutions = 21

   other information:
         - time to solve via MacaulayLab = 0.0118s
         - reference = noonburg-1989-neural
```

Outline

Multiparameter eigenvalue problem

The **multiparameter eigenvalue problem (MEP)** $\mathcal{M}(\boldsymbol{\lambda})\boldsymbol{z} = \mathbf{0}$ consists in finding all n-tuples $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^n$ and corresponding vectors $\boldsymbol{z} \in \mathbb{C}^{l \times 1} \setminus \{\mathbf{0}\}$, so that

$$\mathcal{M}(\boldsymbol{\lambda})\boldsymbol{z} = \left(\sum_{\{\boldsymbol{\omega}\}} \boldsymbol{A}_{\boldsymbol{\omega}} \boldsymbol{\lambda}^{\boldsymbol{\omega}} \right) \boldsymbol{z} = \mathbf{0},$$

with $\|\boldsymbol{z}\|_2 = 1$.

- $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)$ indexes the monomials $\boldsymbol{\lambda}^{\boldsymbol{\omega}}$ and coefficient matrices $\boldsymbol{A}_{\boldsymbol{\omega}}$
- rectangular[†] coefficient matrices $\boldsymbol{A}_{\boldsymbol{\omega}} = \boldsymbol{A}_{(\omega_1, \ldots, \omega_n)} \in \mathbb{R}^{k \times l}$ with $k \geq l + n - 1$
- example: $\left(\boldsymbol{A}_{000} + \boldsymbol{A}_{200}\lambda_1^2 + \boldsymbol{A}_{013}\lambda_2\lambda_3^3\right)\boldsymbol{z} = \mathbf{0}$

[†] We consider only rectangular problems in this presentation, so we no longer mention the term “rectangular” explicitly.

Unifying framework for (multiparameter) eigenvalue problems

Different types of (multiparameter) eigenvalue problems in a two-dimensional grid†.

Spectral parameter(s)	**Linear**	**Polynomial**
Eigenvalues ($n = 1$)	Type I $\{1, \lambda\}$ $\boldsymbol{A} - \boldsymbol{B}\lambda$ SEP/GEP	Type II λ^{ω} $\boldsymbol{A}_0 + \boldsymbol{A}_1\lambda + \cdots + \boldsymbol{A}_d\lambda^d$ PEP
Eigentuples ($n > 1$) ($i = 1, \ldots, n$)	Type III λ_i $\boldsymbol{A}_{00} + \boldsymbol{A}_{10}\lambda_1 + \boldsymbol{A}_{01}\lambda_2$ linear MEP	Type IV $\boldsymbol{\lambda}^{\boldsymbol{\omega}} = \prod_{i=1}^{n} \lambda_i^{\omega_i}$ $\boldsymbol{A}_{00} + \boldsymbol{A}_{11}\lambda_1\lambda_2 + \boldsymbol{A}_{03}\lambda_2^3$ polynomial MEP

† SEP = standard eigenvalue problem, GEP = generalized eigenvalue problem, and PEP = polynomial eigenvalue problem

Examples and applications

rectangular 3×2 coefficient matrices

$$\left(\underbrace{\begin{bmatrix}1 & 2\\3 & 4\\3 & 4\end{bmatrix}}_{\boldsymbol{A}_{00}} + \underbrace{\begin{bmatrix}2 & 1\\0 & 1\\1 & 3\end{bmatrix}}_{\boldsymbol{A}_{10}}\lambda_1 + \underbrace{\begin{bmatrix}3 & 4\\2 & 1\\0 & 1\end{bmatrix}}_{\boldsymbol{A}_{11}}\lambda_1\lambda_2 + \underbrace{\begin{bmatrix}1 & 2\\4 & 2\\2 & 1\end{bmatrix}}_{\boldsymbol{A}_{02}}\lambda_2^2\right)\boldsymbol{z} = \boldsymbol{0}$$

(ω_1, ω_2) labels $\boldsymbol{\lambda}^{\boldsymbol{\omega}}$ and indexes $\boldsymbol{A}_{\boldsymbol{\omega}}$

- Identification of misfit-versus-latency models
- Model order reduction in the $\mathcal{H}_2$-norm

Extension to the block Macaulay matrix

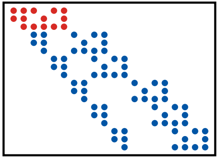
$$\boldsymbol{\mathcal{M}}(\boldsymbol{\lambda})\boldsymbol{z} = (\boldsymbol{A}_{00} + \boldsymbol{A}_{10}\lambda_1 + \boldsymbol{A}_{01}\lambda_2)\boldsymbol{z} = \boldsymbol{0}$$

$$\Downarrow$$

$$\begin{array}{r} \\ \boldsymbol{\mathcal{M}}(\lambda) \\ \lambda_1\boldsymbol{\mathcal{M}}(\lambda) \\ \lambda_2\boldsymbol{\mathcal{M}}(\lambda) \\ \lambda_1^2\boldsymbol{\mathcal{M}}(\lambda) \\ \lambda_1\lambda_2\boldsymbol{\mathcal{M}}(\lambda) \\ \lambda_2^2\boldsymbol{\mathcal{M}}(\lambda) \end{array}
\begin{array}{c} \begin{array}{cccccccccc} \boldsymbol{z} & \lambda_1\boldsymbol{z} & \lambda_2\boldsymbol{z} & \lambda_1^2\boldsymbol{z} & \lambda_1\lambda_2\boldsymbol{z} & \lambda_2^2\boldsymbol{z} & \lambda_1^3\boldsymbol{z} & \lambda_1^2\lambda_2\boldsymbol{z} & \lambda_1\lambda_2^2\boldsymbol{z} & \lambda_2^3\boldsymbol{z} \end{array} \\
\underbrace{\begin{bmatrix} \boldsymbol{A}_{00} & \boldsymbol{A}_{10} & \boldsymbol{A}_{01} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\
\boldsymbol{0} & \boldsymbol{A}_{00} & \boldsymbol{0} & \boldsymbol{A}_{10} & \boldsymbol{A}_{01} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\
\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_{00} & \boldsymbol{0} & \boldsymbol{A}_{10} & \boldsymbol{A}_{01} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\
\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_{00} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_{10} & \boldsymbol{A}_{01} & \boldsymbol{0} & \boldsymbol{0} \\
\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_{00} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_{10} & \boldsymbol{A}_{01} & \boldsymbol{0} \\
\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_{00} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_{10} & \boldsymbol{A}_{01} \end{bmatrix}}_{\text{block Macaulay matrix}} \end{array}
\begin{bmatrix} \boldsymbol{z} \\ \lambda_1\boldsymbol{z} \\ \lambda_2\boldsymbol{z} \\ \lambda_1^2\boldsymbol{z} \\ \lambda_1\lambda_2\boldsymbol{z} \\ \lambda_1^3\boldsymbol{z} \\ \lambda_1^2\lambda_2\boldsymbol{z} \\ \lambda_1\lambda_2^2\boldsymbol{z} \\ \lambda_2^3\boldsymbol{z} \end{bmatrix} = \boldsymbol{0}$$

(Vermeersch and De Moor, 2022)

Two solution approaches



null space based approach

$$\begin{bmatrix} \boldsymbol{z} \\ \lambda_1 \boldsymbol{z} \\ \lambda_2 \boldsymbol{z} \end{bmatrix} \lambda_1 \rightarrow \begin{bmatrix} \lambda_1 \boldsymbol{z} \\ \lambda_1^2 \boldsymbol{z} \\ \lambda_1 \lambda_2 \boldsymbol{z} \end{bmatrix}$$

column space based approach

$$\boldsymbol{Q} \begin{bmatrix} \boldsymbol{R}_{14} & \boldsymbol{R}_{13} & \boldsymbol{R}_{12} & \boldsymbol{R}_{11} \\ \boldsymbol{R}_{24} & \boldsymbol{R}_{23} & \boldsymbol{R}_{22} & \boldsymbol{0} \\ \boldsymbol{R}_{34} & \boldsymbol{R}_{33} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{R}_{44} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} = \boldsymbol{N}$$

(Vermeersch and De Moor, 2022)

Outline

Conclusion

- Multivariate polynomials are omnipresent in science and engineering
- There exist many different solution approaches
- Depending on the application, there could be a lot of structure available

Any questions?

KU Leuven, Dept. of Electrical Engineering (ESAT), Center for Dynamical Systems, Signal Processing, and Data Analytics (STADIUS), Kasteelpark Arenberg 10, 3001 Leuven, Belgium (christof.vermeersch@esat.kuleuven.be)

Christof Vermeersch was supported in part by the FWO Strategic Basic Research fellowship under grant SB/1SA1319N

References

Kim Batselier, Philippe Dreesen, and Bart De Moor. On the null spaces of the Macaulay matrix. *Linear Algebra and its Applications*, 460:259–289, 2014.

Robert M. Corless, Gianni M. Patrizia, and Barry M. Trager. A reordered Schur factorization method for zero-dimensional polynomial systems with multiple roots. In *Proc. of the 1997 International Symposium on Symbolic and Algebraic Computation (ISSAC)*, pages 133–140, Maui, HI, USA, 1997.

Philippe Dreesen. *Back to the Roots: Polynomial System Solving Using Linear Algebra*. PhD thesis, KU Leuven, Leuven, Belgium, 2013.

Hans J. Stetter. *Numerical Polynomial Algebra*. SIAM, Philadelphia, PA, USA, 2004.

Christof Vermeersch and Bart De Moor. Globally optimal least-squares ARMA model identification is an eigenvalue problem. *IEEE Control Systems Letters*, 3(4): 1062–1067, 2019.

References

Christof Vermeersch and Bart De Moor. A column space based approach to solve systems of multivariate polynomial equations. *IFAC-PapersOnLine*, 54(9):137–144, 2021. Part of special issue: 24th International Symposium on Mathematical Theory of Networks and Systems (MTNS).

Christof Vermeersch and Bart De Moor. Two complementary block Macaulay matrix algorithms to solve multiparameter eigenvalue problems. *Linear Algebra and its Applications*, 654:177–209, 2022.