Solving Applications from Systems Theory Via Efficient Numerical Linear Algebra Root-Finding Algorithms

SIAM Conference on Applied Linear Algebra (LA24)
Sorbonne Université - Paris, France

Sibren Lagauw*, Christof Vermeersch, Bart De Moor
{sibren.lagauw, christof.vermeersch, bart.demoor}@esat.kuleuven.be

Center for Dynamical Systems, Signal Processing, and Data Analytics (STADIUS),
Department of Electrical Engineering (ESAT), KU Leuven, Belgium

May 14, 2024

Fonds Wetenschappelijk Onderzoek Vlaanderen
Opening new horizons

Table of Contents

Polynomials and Macaulay matrix

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

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

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Right null space of the Macaulay matrix

$$
\begin{array}{c} \\ p_1(\boldsymbol{x}) \\ x_1 p_1(\boldsymbol{x}) \\ x_2 p_1(\boldsymbol{x}) \\ p_2(\boldsymbol{x}) \\ x_1 p_2(\boldsymbol{x}) \\ x_2 p_2(\boldsymbol{x}) \end{array}
\begin{array}{c}
\begin{array}{cccccccccc} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 & x_1^3 & x_1^2x_2 & x_1x_2^2 & x_2^3 \end{array} \\
\begin{bmatrix}
a_{00} & a_{10} & a_{01} & a_{20} & a_{11} & a_{02} & 0 & 0 & 0 & 0 \\
0 & a_{00} & 0 & a_{10} & a_{01} & 0 & a_{20} & a_{11} & a_{02} & 0 \\
0 & 0 & a_{00} & 0 & a_{10} & a_{01} & 0 & a_{20} & a_{11} & a_{02} \\
b_{00} & b_{10} & b_{01} & b_{20} & b_{11} & b_{02} & 0 & 0 & 0 & 0 \\
0 & b_{00} & 0 & b_{10} & b_{01} & 0 & b_{20} & b_{11} & b_{02} & 0 \\
0 & 0 & b_{00} & 0 & b_{10} & b_{01} & 0 & b_{20} & b_{11} & b_{02}
\end{bmatrix}
\end{array}
\begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \\ x_1^3 \\ x_1^2x_2 \\ x_1x_2^2 \\ x_2^3 \end{bmatrix} = \mathbf{0}
$$

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

- Basis matrix of the right null space can be written in terms of the solutions ($d \geq d_*$).
- from the rank-nullity theorem

$$
\begin{aligned}
n_d &= q_d - r_d \\
&= \dim \mathcal{P}_d^n / \langle p_1(\boldsymbol{x}), \ldots, p_s(\boldsymbol{x}) \rangle_d
\end{aligned}
$$

- right null space $\approx$ dual vector space of the quotient space
- requires differential functionals

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

Null space based solution approach

- Shift-trick:

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

Null space based solution approach

- Shift-trick:

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

- solution vectors $\boldsymbol{v}_d$ in $\boldsymbol{V}_d$ are not known in advanced:

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

$$\Downarrow$$

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

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

Null space based solution approach

- Shift-trick:

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

- solution vectors $\boldsymbol{v}_d$ in $\boldsymbol{V}_d$ are not known in advanced:

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

$$\Downarrow$$

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

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

- it is possible to *shift with any polynomial* in the variables:

$$\boldsymbol{T} \boldsymbol{D}_{x_2} \boldsymbol{T}^{-1} = (\boldsymbol{S}_1 \boldsymbol{Z}_d)^{-1} (\boldsymbol{S}_{x_2} \boldsymbol{Z}_d),$$

which leads to the same $\boldsymbol{T}$

What is a “large enough” degree?

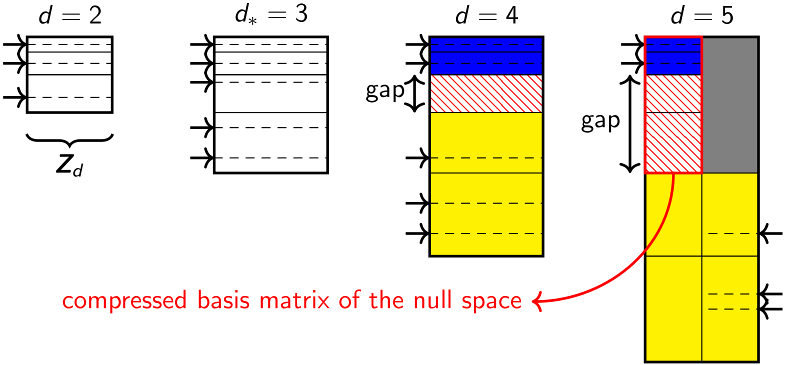


Figure: The column compression deflates the solutions at infinity from the eigenvalue problems.

MacaulayLab

For more details...

- Stetter, 2004
- Batselier et al., 2014; De Cock and De Moor, 2021; Dreesen, 2013
- Vermeersch, 2023; Vermeersch and De Moor, 2021, 2022, 2023

Software available via: `www.macaulaylab.net`

Table of Contents

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

H_2-norm model reduction of SISO LTI models

Higer order model

$$H(s) = \frac{b(s)}{a(s)} = \frac{b_{n-1}s^{n-1}+\cdots+b_1s+b_0}{s^n+a_{n-1}s^{n-1}+\cdots+a_1s+a_0}$$

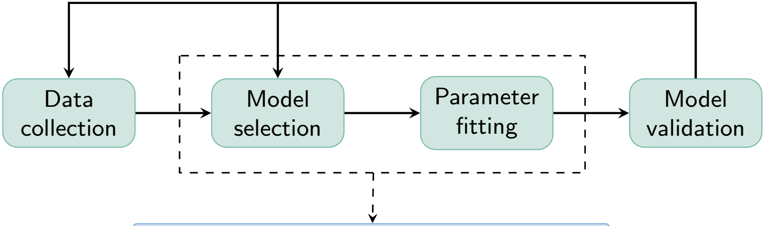
Reduced-order model $(m < n)$

$$\hat{H}(s) = \frac{\hat{b}(s)}{\hat{a}(s)} = \frac{\hat{b}_{m-1}s^{m-1}+\cdots+\hat{b}_1s+\hat{b}_0}{s^m+\hat{a}_{m-1}s^{m-1}+\cdots+\hat{a}_1s+\hat{a}_0},$$

$$\text{Min } \frac{1}{2\pi}\int_{-\infty}^{\infty} |H(j\omega) - \hat{H}(j\omega)|^2 \, \mathrm{d}\omega$$

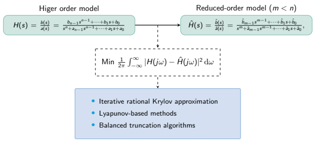
- Iterative rational Krylov approximation
- Lyapunov-based methods
- Balanced truncation algorithms

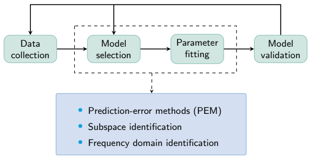
System identification for autonomous LTI models

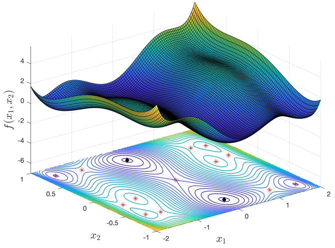


- Prediction-error methods (PEM)
- Subspace identification
- Frequency domain identification

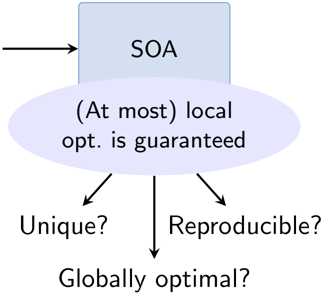
Nonlinear optimization problem(s)



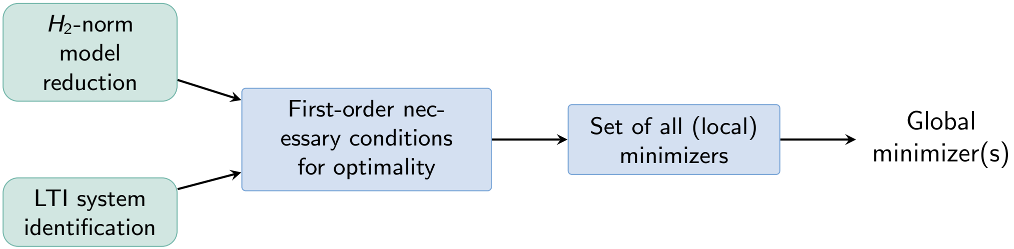




Non-convex
opt. problem



Global optimality



Global optimality

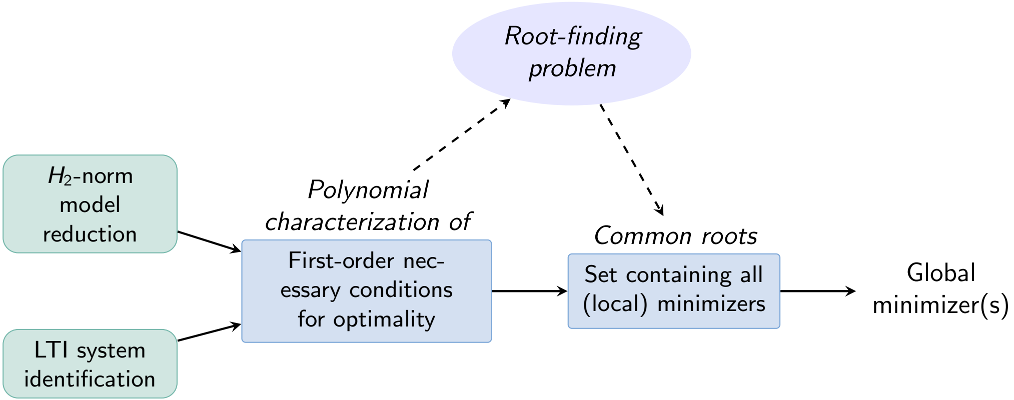


Table of Contents

Model reduction

- Class $\mathcal{M}$: minimal, stable, SISO LTI systems.

- Given $H(s) \in \mathcal{M}$ of order n,

$$H(s) = \frac{b(s)}{a(s)} = \frac{b_{n-1}s^{n-1} + \cdots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0}, \quad (1)$$

  find $\hat{H}(s) \in \mathcal{M}$ of order $m < n$,

$$\hat{H}(s) = \frac{\hat{b}(s)}{\hat{a}(s)} = \frac{\hat{b}_{m-1}s^{m-1} + \cdots + \hat{b}_1 s + \hat{b}_0}{s^m + \hat{a}_{m-1}s^{m-1} + \cdots + \hat{a}_1 s + \hat{a}_0}, \quad (2)$$

  so that $\hat{H}(s)$ is a 'good approximation' of $H(s)$.

- $2n$ unknowns: $\hat{\boldsymbol{a}} = (\hat{a}_{m-1}, \ldots, \hat{a}_0)^{\mathsf{T}} \in \mathbb{R}^n$, $\hat{\boldsymbol{b}} = (\hat{b}_{m-1}, \ldots, \hat{b}_0)^{\mathsf{T}} \in \mathbb{R}^n$.

H_2-norm model reduction of SISO LTI models

- Minimize H_2-norm of approximation error $E(s) = H(s) - \hat{H}(s)$:

$$\hat{H}(s) \in \underset{\hat{H}(s) \in \mathcal{M}}{\operatorname{argmin}} \, J^2, \tag{3}$$

where,

$$J^2 = \|E(s)\|_{H_2}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} |H(j\omega) - \hat{H}(j\omega)|^2 \, \mathrm{d}\omega$$
$$= \int_0^{\infty} \left(h(t) - \hat{h}(t)\right)^2 \mathrm{d}t,$$

with $h(t)$ and $\hat{h}(t)$ the impulse responses of $H(s)$, $\hat{H}(s)$, respectively.

Interpolatory conditions for optimality

Theorem (Meier and Luenberger, 1967)

Given a stable SISO model $H(s) \in \mathcal{M}$ of order n, let $\hat{H}(s)$ of order m ($m<n$) be a stationary point of the H_2-norm model reduction problem in (3). Then for all poles p_i of $\hat{H}(s)$,

$$H(-p_i)^{(j)} = \hat{H}(-p_i)^{(j)}, \quad j = 0, \ldots, d_i,$$

where d_i is the multiplicity of the pole p_i and the superscript j denotes the jth derivative wrt. s.

Interpolatory conditions for optimality

Theorem (Meier and Luenberger, 1967)

Given a stable SISO model $H(s) \in \mathcal{M}$ of order n, let $\hat{H}(s)$ of order m ($m<n$) be a stationary point of the H_2-norm model reduction problem in (3). Then for all poles p_i of $\hat{H}(s)$,

$$H(-p_i)^{(j)} = \hat{H}(-p_i)^{(j)}, \quad j = 0, \dots, d_i,$$

where d_i is the multiplicity of the pole p_i and the superscript j denotes the jth derivative wrt. s.

Theorem (Regalia, 1995)

Given a stable SISO model $H(s) \in \mathcal{M}$ of order n, let $\hat{H}(s)$ of order m with $m<n$ be a stationary point of the model reduction problem (3). Then for all $s \in \mathbb{C}$:

$$H(s) - \hat{H}(s) = \frac{b(s)}{a(s)} - \frac{\hat{b}(s)}{\hat{a}(s)} = \left[\frac{\hat{a}(-s)}{\hat{a}(s)}\right]^2 G(s),$$

with $G(s)$ the Laplace-transform of some real-valued, stable and causal signal.

Interpolatory conditions for optimality

Corollary (Lagauw, Agudelo, et al., 2023)

For any given stable SISO model $H(s)$ of order n and mth order approximant $\hat{H}(s)$ with $m<n$ as defined in (1)–(2), define the polynomial,

$$I(s) = b(s)\hat{a}(s) - a(s)\hat{b}(s) - [\hat{a}(-s)]^2\tilde{G}(s), \tag{4}$$

where $\tilde{G}(s)$ is a polynomial parametrized in the coefficients $\boldsymbol{g} = (g_0, \ldots, g_{n-m-1})^\mathsf{T} \in \mathbb{R}^{n-m}$:

$$\tilde{G}(s) = g_{n-m-1}s^{n-m-1} + \cdots + g_1 s + g_0.$$

Then, $\hat{H}(s)$ is a stationary point of (3) if and only if,

$$\exists\, \boldsymbol{g} \quad s.t. \quad I(s) = 0,\ \forall s \in \mathbb{C}.$$

System of multivariate polynomial equations

- Let f_k be the coefficient corresponding to s^k in the polynomial,

$$l(s) = b(s)\hat{a}(s) - a(s)\hat{b}(s) - [\hat{a}(-s)]^2\,\tilde{G}(s),$$

and define the algebraic variety,

$$\mathcal{V}_{\mathbb{R}} = \left\{(\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}}, \boldsymbol{g}) \in \mathbb{R}^{m+n} \;:\; f_k(\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}}, \boldsymbol{g}) = 0, \quad \forall k = 0, \ldots, m+n-1\right\}.$$

- If $\mathcal{V}'_{\mathbb{R}} \subseteq \mathcal{V}_{\mathbb{R}}$ contains the $(\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}}, \boldsymbol{g}) \in \mathcal{V}_{\mathbb{R}}$ for which $\hat{H}(s)$ is stable, then $\mathcal{V}'_{\mathbb{R}}$ contains all minimizers $\hat{H}(s)$ ✓

- square system of $m+n$ polynomial equations
- degree of the polynomials is at most *cubic*

Numerical example I

- Consider the third order model ($n = 3$) used in Agudelo et al., 2021:

$$H(s) = \frac{s^2 + 9s - 10}{s^3 + 12s^2 + 49s + 78}.$$

- For $m = 1$, the polynomial $I(s)$ from (4) is given as,

$$\begin{aligned} I(s) = {} & (\underbrace{1 - g_1 - \hat{b}_0}_{f_3})s^3 + (\underbrace{\hat{a}_0 - 12\hat{b}_0 - g_0 + 2\hat{a}_0 g_1 + 9}_{f_2})s^2 \\ & + (\underbrace{9\hat{a}_0 - 49\hat{b}_0 + 2\hat{a}_0 g_0 - \hat{a}_0^2 g_1 - 10}_{f_1})s \\ & + (\underbrace{-g_0\hat{a}_0^2 - 10\hat{a}_0 - 78\hat{b}_0}_{f_0})1. \end{aligned}$$

Numerical example I

Strategy: find all $(\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}}, \boldsymbol{g})$ for which $I(s) = 0, \ \forall s \in \mathbb{C}$"

$$\iff \begin{cases} 0 = f_3 = 1 - g_1 - \hat{b}_0, \\ 0 = f_2 = \hat{a}_0 - 12\hat{b}_0 - g_0 + 2\hat{a}_0 g_1 + 9, \\ 0 = f_1 = 9\hat{a}_0 - 49\hat{b}_0 + 2\hat{a}_0 g_0 - \hat{a}_0^2 g_1 - 10, \\ 0 = f_0 = -g_0 \hat{a}_0^2 - 10\hat{a}_0 - 78\hat{b}_0. \end{cases} \tag{5}$$

- Common roots of (5) contain all stationary points $\hat{H}(s)$ of (3):

J	$\hat{a}_0$	$\hat{b}_0$	g_0	g_1	stable
8.9403	$-4.1639 + 0.9026j$	$24.930 - 6.5393j$	$-106.84 - 18.287j$	$-23.930 + 6.5393j$	✗
8.9403	$-4.1639 - 0.9026j$	$24.930 + 6.5393j$	$-106.84 + 18.287j$	$-23.930 - 6.5393j$	✗
0.3982	0.2671	-0.0437	10.349	1.0437	✓
0.2784	0.6914	9.6796	1.2799	-2.0986	✓
0.5232	-16.618	1.9264	0.0576	-0.9264	✗

Numerical example II

We search for the globally optimal 4th order reduced model ($m = 4$) of the state-space model[1] ($n = 17$) considered in Žigić et al., 1992.

- System of 21 polynomial equations with $d_{\max} = 3$
- $\mathcal{V}_{\mathbb{R}}$ contains 290 tuples, 69 remain in $\mathcal{V}'_{\mathbb{R}}$
- Four best-performing stationary points $\hat{H}(s)$:

	J	$p_{1,2}$	$p_{3,4}$
$*$	9.14×10^{-3}	$-0.032 \pm 78.54j$	$-0.111 \pm 15.43j$
l_1	9.22×10^{-3}	$-0.032 \pm 78.54j$	$-5.713 \pm 52.57j$
l_2	1.03×10^{-2}	$-0.032 \pm 78.54j$	$-0.023 \pm 3.842j$
l_3	1.09×10^{-2}	$-0.032 \pm 78.54j$	$-4.663 \pm 15.88j$

[1]The model describes the interaction between a torque activator and a torsional rate sensor for the ACES structure Collins et al., 1991.

Numerical example II

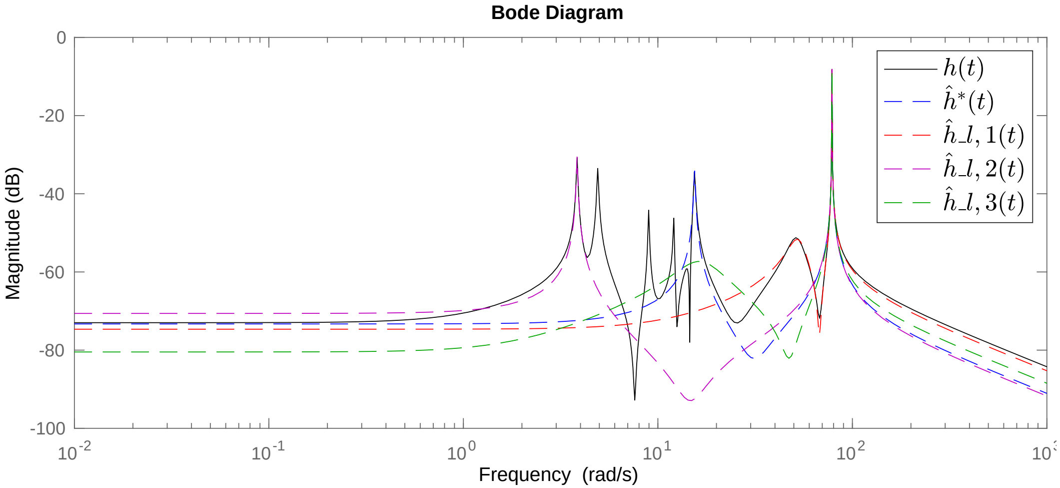
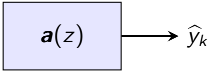


Table of Contents

Automous LTI model



- LTI dynamics of model-compliant data $\widehat{\boldsymbol{y}} \in \mathbb{R}^N$:

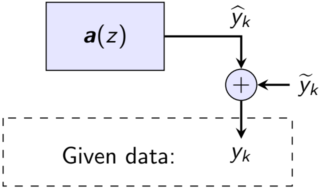
$$\widehat{y}_{k+n} + a_1 \widehat{y}_{k+n-1} + \cdots + a_n \widehat{y}_k = 0, \ \forall k{=}0, \ldots, N{-}n{-}1$$

- Kernel representation of *behavior*

$$\underbrace{\begin{bmatrix} a_n & \ldots & \ldots & a_1 & 1 & 0 & \ldots & 0 \\ 0 & a_n & \ldots & \ldots & a_1 & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & a_n & \ldots & \ldots & a_1 & 1 \end{bmatrix}}_{\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}} \begin{bmatrix} \widehat{y}_0 \\ \widehat{y}_1 \\ \vdots \\ \widehat{y}_{N-1} \end{bmatrix} = \mathbf{0}$$

- n unknown model parameters $\boldsymbol{a} \in \mathbb{R}^n$

Automous LTI model



Least-squares realization:

$$\min_{\boldsymbol{a},\widehat{\boldsymbol{y}}} \quad \frac{1}{2}\|\widetilde{\boldsymbol{y}}\|_2^2 = \frac{1}{2}\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|_2^2,$$
$$\text{s.t.} \quad \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}} = \boldsymbol{0}.$$

- LTI dynamics of model-compliant data $\widehat{\boldsymbol{y}} \in \mathbb{R}^N$:

$$\widehat{y}_{k+n} + a_1\widehat{y}_{k+n-1} + \cdots + a_n\widehat{y}_k = 0, \ \forall k{=}0,\ldots,N{-}n{-}1$$

- Kernel representation of *behavior*

$$\underbrace{\begin{bmatrix} a_n & \ldots & \ldots & a_1 & 1 & 0 & \ldots & 0 \\ 0 & a_n & \ldots & \ldots & a_1 & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & a_n & \ldots & \ldots & a_1 & 1 \end{bmatrix}}_{\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}} \begin{bmatrix} \widehat{y}_0 \\ \widehat{y}_1 \\ \vdots \\ \widehat{y}_{N-1} \end{bmatrix} = \boldsymbol{0}$$

- n unknown model parameters $\boldsymbol{a} \in \mathbb{R}^n$

Least-squares realization

$$\min_{\boldsymbol{a},\widehat{\boldsymbol{y}}} \quad \frac{1}{2}\|\widetilde{\boldsymbol{y}}\|_2^2 = \frac{1}{2}\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|_2^2,$$
$$\text{s.t.} \quad \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}} = \boldsymbol{0}.$$

↓

$$\mathcal{L}(\boldsymbol{a}, \widehat{\boldsymbol{y}}, \boldsymbol{l}) = \tfrac{1}{2}\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|_2^2 + \boldsymbol{l}^\top \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}},$$

↓

$$\partial\mathcal{L}/\partial\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{y}} - \boldsymbol{y} + (\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^\top \boldsymbol{l} = \boldsymbol{0},$$
$$\partial\mathcal{L}/\partial\boldsymbol{a} = (\widehat{\boldsymbol{Y}}_{\boldsymbol{N-n}})^\top \boldsymbol{l} - \boldsymbol{e}\lambda = \boldsymbol{0},$$
$$\partial\mathcal{L}/\partial\boldsymbol{l} = \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{Y}}_{\boldsymbol{N-n}}\boldsymbol{a} = \boldsymbol{0}$$

Least-squares realization

$$\min_{\boldsymbol{a},\widehat{\boldsymbol{y}}} \quad \frac{1}{2}\|\widetilde{\boldsymbol{y}}\|_2^2 = \frac{1}{2}\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|_2^2,$$
$$\text{s.t.} \quad \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}} = \boldsymbol{0}.$$

$\downarrow$

$$\mathcal{L}(\boldsymbol{a}, \widehat{\boldsymbol{y}}, \boldsymbol{l}) = \tfrac{1}{2}\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|_2^2 + \boldsymbol{l}^{\mathsf{T}}\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}},$$

$\downarrow$

$$\begin{aligned}
\partial\mathcal{L}/\partial\widehat{\boldsymbol{y}} &= \widehat{\boldsymbol{y}} - \boldsymbol{y} + (\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathsf{T}}\boldsymbol{l} = \boldsymbol{0},\\
\partial\mathcal{L}/\partial\boldsymbol{a} &= (\widehat{\boldsymbol{Y}}_{\boldsymbol{N-n}})^{\mathsf{T}}\boldsymbol{l} - \boldsymbol{e}\lambda = \boldsymbol{0},\\
\partial\mathcal{L}/\partial\boldsymbol{l} &= \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{Y}}_{\boldsymbol{N-n}}\boldsymbol{a} = \boldsymbol{0}
\end{aligned}$$

Theorem (De Moor, 2020)

The minimal norm misfit $\widetilde{\boldsymbol{y}} = \boldsymbol{y} - \widehat{\boldsymbol{y}} \in \mathbb{R}^N$ *corresponds to the orth. projection of* $\boldsymbol{y}$ *onto* $\mathrm{row}\left(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\right)$,

$$\widetilde{\boldsymbol{y}} = (\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathsf{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathsf{T}})^{-1}\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\boldsymbol{y}.$$

Least-squares realization

$$\min_{\boldsymbol{a},\widehat{\boldsymbol{y}}} \quad \frac{1}{2}\|\widetilde{\boldsymbol{y}}\|_2^2 = \frac{1}{2}\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|_2^2,$$
$$\text{s.t.} \quad \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}} = \boldsymbol{0}.$$

↓

$$\mathcal{L}(\boldsymbol{a}, \widehat{\boldsymbol{y}}, \boldsymbol{l}) = \tfrac{1}{2}\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|_2^2 + \boldsymbol{l}^{\mathrm{T}}\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}},$$

↓

$$\partial\mathcal{L}/\partial\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{y}} - \boldsymbol{y} + (\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathrm{T}}\boldsymbol{l} = \boldsymbol{0},$$
$$\partial\mathcal{L}/\partial\boldsymbol{a} = (\widehat{\boldsymbol{Y}}_{\boldsymbol{N-n}})^{\mathrm{T}}\boldsymbol{l} - \boldsymbol{e}\lambda = \boldsymbol{0},$$
$$\partial\mathcal{L}/\partial\boldsymbol{l} = \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{Y}}_{\boldsymbol{N-n}}\boldsymbol{a} = \boldsymbol{0}$$

Theorem (De Moor, 2020)

The minimal norm misfit $\widetilde{\boldsymbol{y}} = \boldsymbol{y} - \widehat{\boldsymbol{y}} \in \mathbb{R}^N$ *corresponds to the orth. projection of* $\boldsymbol{y}$ *onto* $\mathrm{row}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})$,

$$\widetilde{\boldsymbol{y}} = (\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathrm{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathrm{T}})^{-1}\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\boldsymbol{y}.$$

Theorem (Lagauw, Vanpoucke, et al., 2024)

If $\boldsymbol{a}$ *is a (local) minimizer, then* $\exists\, \boldsymbol{g} \in \mathbb{R}^{N-2n}$ *s.t.,*

$$\widetilde{\boldsymbol{y}} = (\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathrm{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-2n}})^{\mathrm{T}}\boldsymbol{g},$$

where $\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-2n}} \in \mathbb{R}^{(N-2n)\times(N-n)}$ *is a banded Toeplitz matrix defined similarly to the matrix* $\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}} \in \mathbb{R}^{(N-n)\times N}$.

System of multivariate polynomial equations

- If $\boldsymbol{a}$ is a (local) minimizer, then $\exists\ \boldsymbol{g} \in \mathbb{R}^{N-2n}$ s.t.,

$$\begin{aligned} \widetilde{\boldsymbol{y}} = (\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{N-n}(\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}})^{-1}\boldsymbol{T}^{\boldsymbol{a}}_{N-n}\boldsymbol{y} = (\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{N-2n})^{\mathsf{T}}\boldsymbol{g}. \\ \iff (\boldsymbol{T}^{\boldsymbol{a}}_{N-n}(\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}})^{-1}\boldsymbol{T}^{\boldsymbol{a}}_{N-n}\boldsymbol{y} - (\boldsymbol{T}^{\boldsymbol{a}}_{N-2n})^{\mathsf{T}}\boldsymbol{g} = \boldsymbol{0}, \\ \iff \boldsymbol{T}^{\boldsymbol{a}}_{N-n}\boldsymbol{y} - \boldsymbol{T}^{\boldsymbol{a}}_{N-n}(\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{N-2n})^{\mathsf{T}}\boldsymbol{g} = \boldsymbol{0}. \end{aligned}$$

- Define the algebraic variety,

$$\mathcal{V}_{\mathbb{R}} = \left\{(\boldsymbol{a}, \boldsymbol{g}) \in \mathbb{R}^{N-n}\ :\ \boldsymbol{T}^{\boldsymbol{a}}_{N-n}\boldsymbol{y} - \boldsymbol{T}^{\boldsymbol{a}}_{N-n}(\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{N-2n})^{\mathsf{T}}\boldsymbol{g} = \boldsymbol{0}\right\}.$$

$\longrightarrow$ $\mathcal{V}_{\mathbb{R}}$ contains all minimizers $\boldsymbol{a}$ of the identification problem ✓

- square system of $N-n$ polynomial equations
- degree of the polynomials is at most *quartic*

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Numerical example III

Toy problem from De Moor, 2020

- Find globally optimal first-order ($n = 1$) autonomous LTI realization
- Given output data $\boldsymbol{y} = [4, 3, 2, 1]^\mathsf{T}$ ($N = 4$)

$$\boldsymbol{T}^{\boldsymbol{a}}_{N-n}\boldsymbol{y} - \boldsymbol{T}^{\boldsymbol{a}}_{N-n}(\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^\mathsf{T}(\boldsymbol{T}^{\boldsymbol{a}}_{N-2n})^\mathsf{T}\boldsymbol{g} = \boldsymbol{0},$$

$$\iff \begin{cases} 0 = 4a_1 - 2a_1g_1 - a_1^2g_2 - a_1^3g_1 + 3, \\ 0 = 3a_1 - g_1 - 2a_1g_2 - 2a_1^2g_1 - a_1^3g_2 + 2, \\ 0 = 2a_1 - g_2 - a_1g_1 - 2a_1^2g_2 + 1. \end{cases}$$

- 7 common-roots, one of which in $\mathcal{V}_\mathbb{R}$

$\lVert\widetilde{\boldsymbol{y}}\rVert_2^2$	a_1	g_1	g_2
0.1486	-0.6764	-0.2525	-0.2734
/	$-0.1589 \mp 0.808j$	$1.3577 \pm 3.8194j$	$1.8359 \pm 3.3491j$
/	$0.4209 \pm 0.6233j$	$3.0425 \mp 2.9959j$	$-0.0785 \pm 1.2013j$
/	$1.3261 \pm 2.0058j$	$-0.2739 \mp 0.6279j$	$0.3793 \mp 0.3849j$

Numerical Example IV

- Find globally optimal first-order ($n = 2$) autonomous LTI realization
- Given output data[2] $\boldsymbol{y}$ ($N = 16$),

$$\boldsymbol{y} = \boldsymbol{y}_{3\text{rd}} + 0.05 * \texttt{randn(N,1)},$$

where $\boldsymbol{y}_{3\text{rd}}$ is generated by a third-order autonomous LTI model with poles $(0.2, 0.7 \pm 0.4j)$

- 739 affine common-roots, 9 of which are real-valued.

$\lVert\tilde{\boldsymbol{y}}\rVert_2^2$	a_1	a_2	p_1	p_2
0.1327	-1.6255	0.7167	$0.8127 + 0.2369j$	$0.8127 - 0.2369j$
0.1514	-0.0752	-0.5850	0.8033	-0.7282
0.1606	-14.076	10.433	13.291	0.7849
0.5386	-0.7127	1.8381	$0.3564 + 1.3081j$	$0.3564 - 1.3081j$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
0.5492	-1.3053	1.0564	$0.6527 + 0.7940j$	$0.6527 - 0.7940j$

[2] The considered instance of $\boldsymbol{y}$ has $\lVert\boldsymbol{y}\rVert_2^2 = 0.5509$.

Numerical Example IV

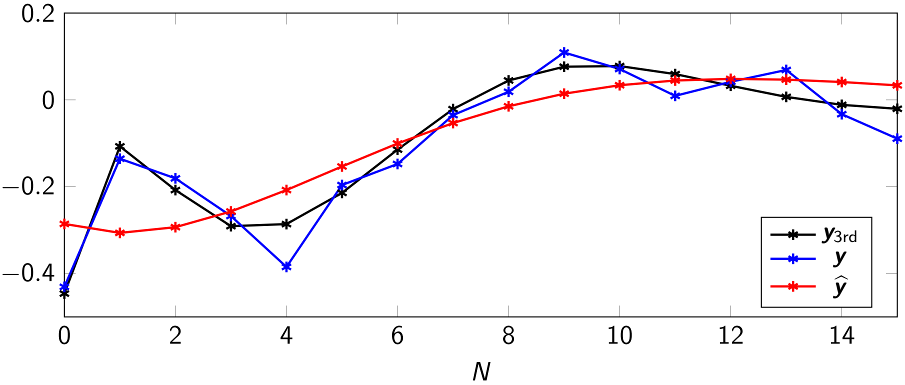


Table of Contents

Conclusions

What should we learn from these examples?

- Systems of polynomial equations from applications are *sparse*
- Systems of polynomial equations from applications are *structured*
- Particular interest in the *real-valued* common-roots
- Solving systems of polynomial equations is useful!

⟶ S. Lagauw, O. M. Agudelo, et al. (2023). "Globally Optimal SISO H_2-Norm Model Reduction Using Walsh's Theorem". In: *IEEE Control Systems Letters* 7, pp. 1670–1675

⟶ S. Lagauw, L. Vanpoucke, et al. (2024). *Exact Characterization of the Global Optima of Least Squares Realization of Autonomous LTI Models as a Multiparameter Eigenvalue Problem*. Tech. rep. Accepted for publication in the Proc. of the 22nd European Control Conference (ECC), Stockholm, Sweden. ESAT-STADIUS, KU Leuven

Questions?

The work of Sibren Lagauw was supported by the
FWO Fundamental Research Fellowship under Grant 11K5623N.

References I

- Agudelo, O. M., C. Vermeersch, and B. De Moor (2021). "Globally Optimal H_2-Norm Model Reduction: A Numerical Linear Algebra Approach". In: *IFAC PapersOnLine (Special Issue: Proceedings of MTNS 2020)* 54.9, pp. 564–571.
- Batselier, Kim, Philippe Dreesen, and B. De Moor (2014). "A fast recursive orthogonalization scheme for the Macaulay matrix". In: *Journal of Computational and Applied Mathematics*.
- Collins, E.G., D.J. Phillips, and D.C. Hyland (1991). "Robust decentralized control laws for the ACES structure". In: *IEEE Control Systems Magazine* 11.3, pp. 62–70.
- De Cock, K. and B. De Moor (2021). "Multiparameter Eigenvalue Problems and Shift-Invariance". In: *IFAC-PapersOnLine (Special Issue: Proceedings of MTNS 2020)* 54.9, pp. 159–165.
- De Moor, B. (2020). "Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem". In: *Communications in Information and Systems* 20.2, pp. 163–207.
- Dreesen, P. (2013). "Back to the Roots: Polynomial System Solving Using Linear Algebra". PhD thesis. KU Leuven.
- Lagauw, S., O. M. Agudelo, and B. De Moor (2023). "Globally Optimal SISO H_2-Norm Model Reduction Using Walsh's Theorem". In: *IEEE Control Systems Letters* 7, pp. 1670–1675.
- Lagauw, S., L. Vanpoucke, and B. De Moor (2024). *Exact Characterization of the Global Optima of Least Squares Realization of Autonomous LTI Models as a Multiparameter Eigenvalue Problem*. Tech. rep. Accepted for publication in the Proc. of the 22nd European Control Conference (ECC), Stockholm, Sweden. ESAT-STADIUS, KU Leuven.
- Meier, L. and D.G. Luenberger (Oct. 1967). "Approximation of Linear Constant Systems". In: *IEEE Transactions on Automatic Control* 12.5, pp. 585–588.

References II

- Regalia, Phillip A. (1995). *Adaptive IIR Filtering in Signal Processing and Control*. Marcel Dekker Inc., New York, NY.
- Stetter, H.J. (2004). *Numerical Polynomial Algebra*. SIAM, Philadelphia, PA.
- Vermeersch, C. (2023). "The (Block) Macaulay Matrix: Solving Systems of Multivariate Polynomial Equations and Multiparameter Eigenvalue Problems". PhD thesis. Leuven, Belgium: Faculty of Engineering, KU Leuven.
- Vermeersch, C. and B. De Moor (2021). "A Column Space Based Approach to Solve Systems of Multivariate Polynomial Equations". In: *IFAC-PapersOnLine* 54.9. Part of Special Issue: 24th International Symposium on Mathematical Theory of Networks and Systems (MTNS), pp. 137–144.
- — (2022). "Two complementary block Macaulay matrix algorithms to solve multiparameter eigenvalue problems". In: *Linear Algebra and its Applications* 654, pp. 177–209.
- — (2023). "Two Double Recursive Block Macaulay Matrix Algorithms to Solve Multiparameter Eigenvalue Problems". In: *IEEE Control Systems Letters* 7, pp. 319–324.
- Žigić, D. et al. (1992). "Homotopy methods for solving the optimal projection equations for the H_2 reduced order model problem". In: *International Journal of Control* 56.1, pp. 173–191.

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics