Can a mathematician and an engineer be friends II

Infinite-dimensional Euclidean Distance Degrees in Model Order Reduction

Leonie Kayser & Sibren Lagauw

leokayser.github.io

November 27, 2025

MAX PLANCK INSTITUTE
FOR MATHEMATICS
IN THE SCIENCES

PREVIOUSLY ON

/ TALK / **08.05.25**, 17:00

Can a mathematician and an engineer be friends? - Critical points to low-rank Hankel matrix approximation

Leonie Kayser (MPI MiS, Leipzig)

G3 10 (Lecture hall)

A mathematician and an engineer walked into a bar

Engineer	Mathematician
Uses calculus	Teaches calculus
Can build a bridge but doesn't know why it holds	Will count the number of possible bridges
Likes two columns	Hates two columns
Has funding from industry	
Likes the smell of whiteboard markers	Crazy for specific chalk from Japan
Cares about real solutjons	Invents imaginary numbers and points at ∞ just to be right
Wants solutions quickly	Wants correct solutions

Realization of linear time-invariant difference equations

$$a_0 y_k + a_1 y_{k+1} + \cdots + a_r y_{k+r} = 0, \qquad k = 0, 1, \ldots$$

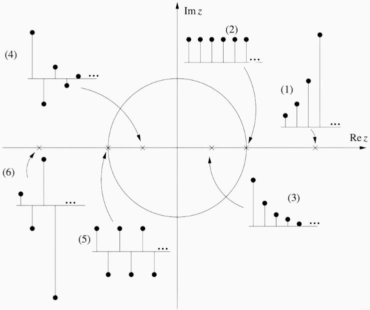
- Discrete-time physical system generating signals $y = (y_0, y_1, \ldots)$
- Explain observed data with a mathematical model
- Impose a model class: *autonomous LTI models of finite order*
  - autonomous = no input signals, no influence from outside world
  - linear = linear relation between past outputs
  - time-invariant = coefficients $a = (a_0, \ldots, a_r)$ are independent of time
  - finite order r = the relation involves at most r past outputs
- y "model compliant" data

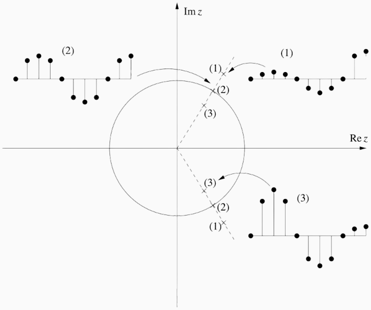
Roots of $a(z) = \sum_{i=0}^{r} a_i z^i$ determine dynamics of model

- Simple roots: Each root λ generates mode $\mathrm{vand}(\lambda) := (1, \lambda, \lambda^2, \dots)$

$$y = \sum_{\lambda} c_\lambda \cdot \mathrm{vand}(\lambda) = \left[\sum_{\lambda} c_\lambda \cdot \lambda^k\right]_{k \geq 0}$$

- Multiple roots introduce *confluent Vandermonde vectors* $\frac{\partial^j}{\partial \lambda^j} \mathrm{vand}(\lambda)$
- Magnitude of λ's determines growth or decay, argument determines phase





Today's menu: SISO linear time-independent systems

- Extend the model class to allow for an input $(u_k)_{k\in\mathbb{N}}$ at each time step k
- Consider discrete-time single-input (u_k) single-output (y_k) LTI models
- State space description: states $x_0, x_1, \ldots \in \mathbb{R}^r$ and

$$x_{k+1} = Ax_k + Bu_k, \quad y_k = C^\mathsf{T} x_k + d \cdot u_k, \qquad A \in \mathbb{R}^{r\times r},\ B, C \in \mathbb{R}^r,\ d \in \mathbb{R}$$

- A is the state transition matrix; will always assume initial state $x_0 = 0$
- The impulse response $h \coloneqq \mathrm{response}[\delta]$, $\delta = (1, 0, 0, \ldots)$ is

$$h_k = \begin{cases} d & k = 0 \\ C^\mathsf{T} A^{k-1} B & k \geq 1 \end{cases}$$

- Any response is convolution with impulse response: $\mathrm{response}[u] = h * u$
- Distance of systems is ℓ^2-norm of impulse responses $\|h - \hat{h}\|_{\ell^2}$

Transfer functions

- $\mathcal{Z}$-transform: If $y \in \ell^2$, then

$$\mathcal{Z}\{y\}(z) := \sum_{k \geq 0} y_k z^{-k}$$

- Multiplicative $\mathcal{Z}(y * y') = \mathcal{Z}(y) \cdot \mathcal{Z}(y')$, therefore

$$\mathcal{Z}\{y\}(z) = H(z) \cdot \mathcal{Z}\{u\}(z) \qquad H := \mathcal{Z}\{h\} = C^\mathsf{T}(Iz - A)^{-1}B + d$$

- Parseval identity:

$$\|h\|_{\ell^2}^2 = \|H\|_{\mathcal{H}_2}^2 := \frac{1}{2\pi \mathrm{i}} \oint_{\mathbb{S}^1} |H(z)|^2 \frac{\mathrm{d}z}{z}$$

  where $\mathbb{S}^1 \subset \mathbb{C}$ is the contour around the unit circle
- State-space models can be “faithfully” represented by the rational function H

Difference equation revisited

- $H = C^{\mathsf{T}}(Iz - A)^{-1}B + d$ is a rational function

Theorem (Kronecker)

The following are equivalent for $\hat{h} = (\hat{h}_1, \hat{h}_2, \dots) \in \prod_{k\geq 1}\mathbb{C}$*:*

1. $\hat{h}$ *satisfies a linear recurrence relation of minimal order* r*.*
2. *The formal power series* $\widehat{H}(z) = \sum_{k=1}^{\infty} \hat{h}_k z^{-k} \in \mathbb{C}[[z^{-1}]]$ *is a rational function:* $\widehat{H}(z) = b(z)/a(z)$, $\gcd(a, b) = 1$, *and* $r = \deg a > \deg b$*.*

The polynomial $a(z)$ *in 2. is the char.polynomial of the minimal recurrence relation.*

- If the state space model is minimal, then $a(z) = \mathrm{charpol}_A(z)$
- $a(z)$ governs the behavior/dynamics of the system

Stable polynomials and the main optimization problem

- A polynomial $a(z)$ is (Schur-)stable if all roots in open unit disc

Corollary

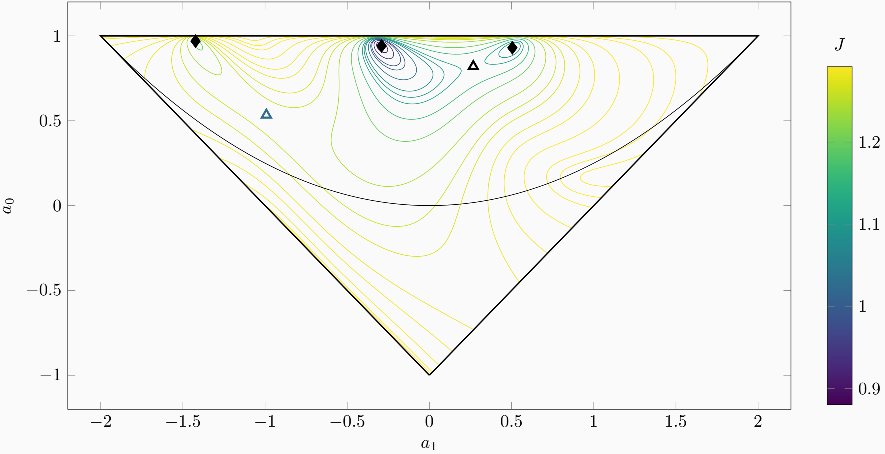
The following are equivalent for $\hat{h} = (\hat{h}_1, \hat{h}_2, \dots)$:

1. *$\hat{h}$ satisfies a linear recurrence relation of minimal order r and $\hat{h} \in \ell^2$;*
2. *The $\mathcal{Z}$-transform $\widehat{H}(z)$ is a rational function: $\widehat{H}(z) = b(z)/a(z)$, $\gcd(a, b) = 1$, and $r = \deg a > \deg b$, and $a(z)$ stable.*

- Model order reduction problem: Given model via transfer function $H = d/c$ of order R, minimize the distance $\|H - \widehat{H}\|_{\mathcal{H}_2}$ to models $\widehat{H} = b/a$ of order r
- Suffices to assume $H, \widehat{H}$ to be strictly proper (if not, match $d = \hat{d}$)
- ∞-dimensional analogue to Euclidean distance minimization to the r-th secant variety of $\nu_d(\mathbb{P}^1) \subseteq \mathbb{P}^d$ $(d \to \infty)$

Example: Model order reduction from order 6 to 2

$$H(z) = \frac{0.0448z^5 + 0.2368z^4 + 0.0013z^3 + 0.0211z^2 + 0.2250z + 0.0219}{z^6 - 1.2024z^5 + 2.3675z^4 - 2.0039z^3 + 2.2337z^2 - 1.0420z + 0.8513}$$



Checkpoint: The MOR problem, mathematically

- Given $H(z) = \frac{d(z)}{c(z)}$, $\deg d < \deg c = R$, $c(z)$ stable,

$$\underset{\widehat{H}=\frac{b(z)}{a(z)}}{\text{minimize}}\ J(\widehat{H}) \qquad \text{subject to } \deg b < \deg a \leq r,\ a(z) \text{ stable}$$

$$J(\widehat{H})^2 = \|\widehat{H} - H\|_{\mathcal{H}_2}^2 = \frac{1}{2\pi \mathrm{i}} \oint_{\mathbb{S}^1} |\widehat{H}(z) - H(z)|^2 \frac{\mathrm{d}z}{z}$$

- Given data: $(c, d) \in \mathbb{R}[z]_{<R} \times \mathbb{R}[z]_{\leq R}$
- Unknowns: $(a, b) \in \mathbb{R}[z]_{<r} \times \mathbb{R}[z]_{\leq r}$, $2r + 1$ coefficients (-1 up to scaling)
- Questions:
  1. Describe critical equations in the coefficients of a, b
  2. Is the critical locus finite, non-degenerate for *generic* c, d?
  3. What is the number of (complex) critical points in terms of R, r?

Walsh's result on optimality

- If $\widehat{H}$ is a proper rational function, then $\operatorname{rank} \widehat{H} := \deg a$, $\widehat{H} = \frac{b}{a}$ in reduced form

Lemma

If $\widehat{H}$ is a local minimum of J, then $\widehat{H}$ has rank r: $\deg a = r$ and $\gcd(a, b) = 1$.

Theorem (Walsh 1932)

Let $\widehat{H}$ be of rank exactly r, then $\widehat{H}$ is a critical point to $J(\widehat{H})$ if and only if

$$(zH(z))^{(j)}(\omega^{-1}) \overset{!}{=} (z\widehat{H}(z))^{(j)}(\omega^{-1}) \qquad \text{for all poles } \omega \text{ of } \widehat{H},\ j = 0, \ldots, \operatorname{ord}_\omega a(z),$$

where $(-)^{(j)}$ denotes the j-th derivative.

- $\widehat{H}$ simple poles $\omega_1, \ldots, \omega_r \neq 0$: $H({\omega_i}^{-1}) \overset{!}{=} \widehat{H}({\omega_i}^{-1})$ and $H'({\omega_i}^{-1}) \overset{!}{=} \widehat{H}'({\omega_i}^{-1})$

Proof idea

- Normalize $a_r = 1$, then $\widehat{H} \mathrel{\hat{=}} (a, b) \in \mathbb{R}^{2r}$. Need $\nabla_{\widehat{H}} J(\widehat{H}) \overset{!}{=} 0$
- Partial derivative in direction $v \in \mathbb{R}^{2r}$

$$\frac{\partial J(\widehat{H})^2}{\partial v} = \frac{2}{2\pi \mathrm{i}} \oint (\widehat{H} - H)(z^{-1}) \cdot \frac{\partial \widehat{H}(z)}{\partial v} z^{-1} \mathrm{d}z$$

- Partial fraction expansion yields $2r$ independent directions $\omega_1, \ldots, \omega_m$, r_{ij}

$$\widehat{H}(z) = \frac{b(z)}{a(z)} = \sum_{i=1}^{m} \sum_{j=1}^{e_i} \frac{r_{ij}}{(z - \omega_i)^j}, \qquad r_{ij} \in \mathbb{C}$$

- Derivatives $\frac{\partial \widehat{H}(z)}{\partial r_{ij}} = \frac{1}{(z - \omega_i)^j}$, similarly $\frac{\partial \widehat{H}(z)}{\partial \omega_i} = \frac{\text{const}}{(z - \omega_i)^{e_i + 1}} + \ldots$, lead to

$$\oint \frac{\widehat{H}(z^{-1}) z^{-1}}{(z - \omega_i)^j} \mathrm{d}z \overset{!}{=} \oint \frac{H(z^{-1}) z^{-1}}{(z - \omega_i)^j} \mathrm{d}z, \qquad i = 1, \ldots, m, \ j = 1, \ldots, e_i + 1$$

- Conclude by Cauchy's integral formula $(\square)$

Walsh polynomial system

Corollary

Let $\widehat{H}(z) = b(z)/a(z)$ have simple poles, then $\widehat{H}$ is a critical point if and only if there exists a polynomial $g \in \mathbb{R}[z]_{\leq R-r-1}$ such that

$$a \cdot d - b \cdot c = \tilde{a}^2 \cdot g, \qquad \tilde{a}(z) := z^r a(1/z) \in \mathbb{R}[z]_{\leq r}.$$

- *Proof:* By Walsh's theorem,

$$F(z) := \frac{zb(z)}{a(z)} - \frac{zd(z)}{c(z)}$$

  has zeros of order ≥ 2 at ω^{-1} for roots ω of $a(z)$
- Equivalently, $\tilde{a}^2 \mid F$, meaning $F = \tilde{a}^2 G$, where G has no poles outside unit disc
- Carefully clear denominators to obtain equations ($g = acG/z$) $\square$

The Walsh variety

- From now on, make restriction to $\widehat{H}$ having simple poles
- Complexify and homogenize (for example $\tilde{a}(z_0, z_1)$ becomes $a(z_1, z_0)$)

$$a \in A := \mathbb{C}[z_0, z_1]_r, \quad b \in B := \mathbb{C}[z_0, z_1]_{r-1},$$

$$c \in C := \mathbb{C}[z_0, z_1]_R, \quad d \in D := \mathbb{C}[z_0, z_1]_{R-1}, \quad g \in G := \mathbb{C}[z_0, z_1]_{R-r-1}$$

- Shorthand notation $A^\circ := A \setminus 0$, $A^\circ BCDG := A^\circ \times B \times C \times D \times G$

Definition (Walsh variety)

The *Walsh variety* is the variety (subscheme)

$$\mathcal{W} := \left\{ (a, b, c, d, g) \;\middle|\; ad - bc = \tilde{a}^2 g \right\} \subseteq A^\circ BCDG.$$

- $R + r$ equations in $3R + r + 2$ variables $\rightsquigarrow$ dimension $2R + 2$?

Irreducibility

Theorem (K.–Lagauw 2025+)

The Walsh variety $\mathcal{W} = \{ad - bc = \tilde{a}^2 g\}$ *is a reduced and irreducible ideal-theoretic complete intersection in* $A^\circ BCDG$ *of dimension* $2R + 2$*.*

- Main idea: Stratify $\mathcal{W} = \bigcup_{k=0}^{r-1} \mathcal{W}_k$, $\mathcal{W}_k = \{ (a, b, \dots) \in \mathcal{W} \mid \deg \gcd(a, b) = k \}$
- Each $\mathcal{W}_k$ is an affine bundle of rank $(R - r + 1 + k)$ over

$$\mathcal{T}_k \coloneqq \left\{ (a, b, g) \;\middle|\; \deg \gcd(a, b) = k,\ \gcd(a, b) \mid \tilde{a}^2 g \right\} \subseteq A^\circ BG$$

- $\mathcal{T}_k$ is a subset of dimension $\begin{cases} = R + r + 1, & k = 0 \\ < R + r + 1 - k, & k \geq 1 \end{cases}$
- Deduce that $\mathcal{W}$ is a set-theoretic complete intersection by dimension count
- $\mathcal{W}_0$ is smooth, irreducible and dense in $\mathcal{W}$; finally apply unmixedness theorem $\square$

Finiteness of the critical locus

- Consider the projection $\tau\colon \mathcal{W} \to CD$
- $\tau^{-1}(c, d)$ is the solution set to the Walsh polynomial system

Corollary

If $(c, d) \in CD \cong \mathbb{C}^{2R+1}$ is general, then every component of $\tau^{-1}(c, d)$ is

1. *reduced, of dimension 1 (invariant under $\lambda \cdot (a, b, g) = (\lambda a, \lambda b, \lambda^{-1} g)$) and*
2. *consists of tuples (a, b, g) such that $a(z)$ has r distinct roots and $\gcd(a, b) = 1$.*

- Either τ is not dominant, then vacously true (not the case, will see later),
- or the general fiber is reduced of dimension $\dim \mathcal{W} - \dim CD = 1$
- Gives alternative, conceptual proof of finiteness of the critical locus

The multi-parameter eigenvalue problem (MEP)

- Let $M(f)\colon \mathbb{C}[z_0,z_1]_{R+r-1-\deg(f)} \to \mathbb{C}[z_0,z_1]_{R+r-1}$ be the linear map $q \mapsto q \cdot f$
- Walsh polynomial system is affine-linear in b, g:

$$ad - bc = \tilde{a}^2 g \quad \Longleftrightarrow \quad \begin{bmatrix} M(ad) & M(c) & M(\tilde{a}^2) \end{bmatrix} \cdot \begin{pmatrix} -1 \\ b \\ g \end{pmatrix} = 0$$

- To eliminate b, g, we need to ensure that $\begin{bmatrix} M(c) & M(\tilde{a}^2) \end{bmatrix}$ never drops rank

Lemma (Emptyness of the bad locus)

$$\dim_{\mathbb{C}} \operatorname{Ker} \begin{bmatrix} M(c) & M(\tilde{a}^2) \end{bmatrix} = \max\{0, \deg\gcd(c, \tilde{a}^2) - r\}.$$

The matrix has full rank if either c has distinct roots or both a and c are stable.

Walsh meets Porteous

$$M(a,c,d) \coloneqq \begin{bmatrix} M(ad) & M(c) & M(\tilde{a}^2) \end{bmatrix} : \mathcal{O}(-1) \oplus \mathcal{O}^r \oplus \mathcal{O}(-2)^{R-r} \to \mathcal{O}^{R+r}$$

Theorem (K.–Lagauw 2025+, main result)

For given (c,d) with c distinct roots, the projection

$$\{\, (a,b,g) \mid (a,b,c,d,g) \in \mathcal{W} \,\} \to A^\circ$$

is an isomorphism onto the degeneracy locus of $M(a,c,d) \in \mathbb{C}[\underline{a}]^{(R+r)\times(R+1)}$.
If (c,d) is general, then the degeneracy locus $D_R(M(a,c,d)) \subseteq \mathbb{P}A$ is a finite set of points of degree

$$\sum_{j=0}^{r} \binom{R-r-1+j}{j} \cdot 2^j.$$

Degree of the MOR problem for various $R > r$

$R \setminus r$	1	2	3	4	5	6	7	8
2	3							
3	5	7						
4	7	17	15					
5	9	31	49	31				
6	11	49	111	129	63			
7	13	71	209	351	321	127		
8	15	97	351	769	1 023	769	255	
9	17	127	545	1 471	2 561	2 815	1 793	511

- In the example $(R, r) = (6, 2)$, the authors report 49 complex solutions, 11 real, 5 real stable.

Towards multiple poles

- Previous results also apply for specific c with distinct roots and general d
- If c has exactly one double root but the solution set of the Walsh system is still finite, then the number of solutions drops to

$$\#\text{solutions} \leq \sum_{j=0}^{r} \binom{R-r-1+j}{j} 2^j - \binom{R-2}{r-1}$$

What about lifting the assumption that a has r distinct root?

- Leads to study equations on multiple root loci
- Known examples where multiple root solutions can be local minimizer
- Rigorous global understanding of MOR requires to also consider these solutions

Thank you! Questions?

Image credit

- Slide 0: *Previously on Game of Thrones* screen, HBO
- Slide 3: “With permission” from Sibren’s lecture on systems theory
- Slide 8: Created by Sibren using MATLAB and `pgfplot`