The Behavioral Toolbox

Ivan Markovsky

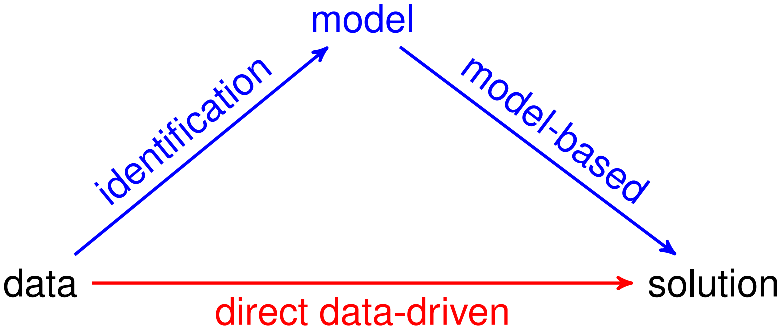
Outline

Why? Software supporting the behavioral approach

What? Analysis, signal processing, and control

How? Algorithms and implementation

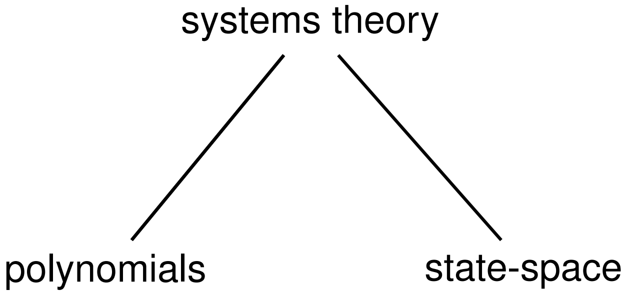
The current trend in systems & control is bypassing parametric model identification



Systems theory, signal processing, and control are going through third paradigm shift

period	paradigm	types of systems
1940–60	classical	SISO transfer funct.
1960–80	modern	MIMO state space
1980–00	behavioral	system as a set
2000–	data-driven	using directly data

However, existing computational methods are still based on model parameters



New paradigm brings new notion of system and new techniques for solving problems

system	techniques
transfer funct.	Laplace/Z, Fourier transforms
state-space	numerical linear algebra
	Lyapunov, Riccati eqn., LMIs
kernel repr.	polynomial algebra
data-driven	numerical linear algebra
	for structured matrices

The Behavioral Toolbox implements direct data-driven solution method

embraces the behavioral approach

deals seamlessly with general LTI systems

does computations without (A, B, C, D) and $H(z)$

Simple motivating example:
How to check if two LTI systems are equal?

we would like this Matlab code to give “true”

```
m = 2; p = 2; n = 3;
sys1 = drss(n, p, m);
sys2 = ss2ss(sys1, rand(n));
sys1 == sys2 % -> error
```

how are the systems specified?

what does it mean that they are equal?

In this talk, the data are trajectories of discrete-time dynamical systems

$w = \big(w(1), \dots, w(T)\big)$ — T-samples trajectory

$w(t) \in \mathbb{R}^q$ — samples are vectors

$\mathscr{B}$ — set of all possible infinite trajectories
$\mathscr{B}|_T$ — set of all possible T-samples trajectories

Key insight: viewing systems as sets of trajectories (the behavioral approach)

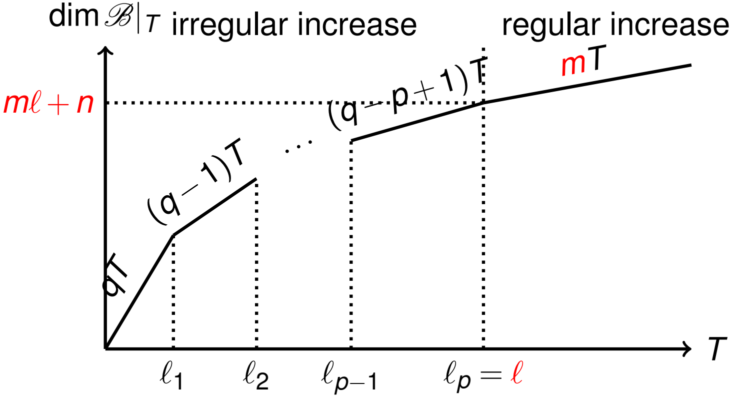
$w \in \mathscr{B}|_T \iff w$ is trajectory of $\mathscr{B}$

$\mathscr{B} \in \mathscr{L}$ — linear time-invariant system

$\mathscr{B}|_T \subset (\mathbb{R}^q)^T$ is subspace — linearity

$\sigma\mathscr{B} = \mathscr{B}$, $\sigma w(t) := w(t+1)$ — time-invariance

Finite-dimensional LTI systems have structure characterized by integers $(\ell_1, \ldots, \ell_p)$



$$\dim \mathscr{B}|_T = mT + n, \quad \text{for } T \geq \ell \qquad (\dim \mathscr{B}|_T)$$

$(\ell_1, \dots, \ell_p)$ are properties of $\mathscr{B}$ and fully specify its structure

p — number of outputs (dependent variables)
$m = q - p$ — number of inputs (free variables)

$\ell := \max\{\ell_1, \dots, \ell_p\}$ — lag
$n := \ell_1 + \cdots + \ell_p$ — order

$c := (m, \ell, n)$ — complexity of the system

$\mathscr{L}_{(m,\ell,n)}$ — bounded complexity LTI model class

Informativity of the data $w_\mathrm{d} \in \mathscr{B}|_{T_\mathrm{d}}$ is rank condition of Hankel matrix

for $w_\mathrm{d} = \big(w_\mathrm{d}(1), \dots, w_\mathrm{d}(T_\mathrm{d})\big)$ and $1 \le T \le T_\mathrm{d}$

$$\mathscr{H}_T(w_\mathrm{d}) := \begin{bmatrix} (\sigma^0 w_\mathrm{d})|_T & (\sigma^1 w_\mathrm{d})|_T & \cdots & (\sigma^{T_\mathrm{d}-T} w_\mathrm{d})|_T \end{bmatrix}$$

generalized persistence of excitation condition

$$\operatorname{rank} \mathscr{H}_T(w_\mathrm{d}) = mT + n \qquad \text{(GPE)}$$

fact: let $w_\mathrm{d} \in \mathscr{B}|_{T_\mathrm{d}}$ and (GPE), then

$$\widehat{\mathscr{B}}_T := \operatorname{image} \mathscr{H}_T(w_\mathrm{d}) = \mathscr{B}|_T, \quad \text{for } T \ge \ell \quad \text{(DDR)}$$

Sufficient conditions: the fundamental lemma

(DDR) holds if

1. $\mathscr{B}$ is controllable
2. $w_\mathrm{d} = \left[\begin{smallmatrix} u_\mathrm{d} \\ y_\mathrm{d} \end{smallmatrix}\right]$ is an input/output partitioning
3. $\mathscr{H}_{T+\mathbf{n}(\mathscr{B})}(u_\mathrm{d})$ is full row rank

J.C. Willems et al., A note on persistency of excitation
Systems & Control Letters, 54:325–329, 2005

input design result

useful byproduct: (DDR) is the cornerstone for data-driven analysis and design

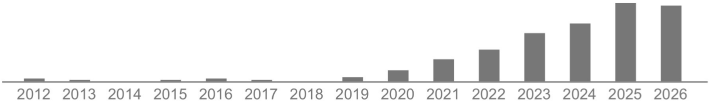
The fundamental lemma comes from SYSID and was first presented at ERNSI 2003

our motivation was subspace identification

data-driven algorithms were foreseen back then

I. Markovsky and P. Rapisarda. Data-driven simulation and control. Int. J. Contr., 81(12):1946–1959, 2008.

delay of ~15 years in uptake of the results



The Data Enabled Predictive Control popularized the fundamental lemma

J. Coulson and J. Lygeros and F. Dörfler,
Data-enabled predictive control: In the shallows of the DeePC, ECC 2019

added regularization for dealing with noise

implicit feedback be embedding it in MPC

demonstrated effectiveness in practice

(DDR) leads to computational methods for:

identification (find representations from data)

e.g., subspace methods aim at (A, B, C, D)

system analysis (direct from data)

e.g., data-driven frequency response evaluation

signal processing and control (direct from data)

e.g., simulation, smoothing, missing data recovery

The "main player" is orthonormal basis for the restricted behavior

$$\mathscr{B}|_T = \operatorname{image} B_T, \quad B_T^\top B_T = I_r \qquad (B_T)$$

$\mathscr{B} \in \mathscr{L}_c^q$ and $T \geq \boldsymbol{\ell}(\mathscr{B})$ implies that

$$r = \dim \mathscr{B}|_T = \mathbf{m}(\mathscr{B})T + \mathbf{n}(\mathscr{B}) \qquad (\dim \mathscr{B}|_T)$$

$\mathscr{B}|_{\boldsymbol{\ell}(\mathscr{B})+1}$ defines $\mathscr{B}$ $\implies$ representation of $\mathscr{B}$

Core sub-problem is obtaining B_T from other representations and data

$$\mathscr{W}_{\mathrm{d}} = \{\, w_{\mathrm{d}}^1, \dots, w_{\mathrm{d}}^N \,\} \;\mapsto\; B_T = \texttt{orth}\, \mathscr{H}_T(w_{\mathrm{d}})$$

$$(A, B, C, D) \;\mapsto\; B_T = \texttt{orth}\, \Pi \begin{bmatrix} 0_{mT \times n} & I_{mT} \\ \mathscr{O}_T(A,C) & \mathscr{T}_T(H) \end{bmatrix}$$

alternative indirect approach:

(A, B, C, D) simulation $\mathscr{W}_{\mathrm{d}}$ $\texttt{orth}\, \mathscr{H}_T(w_{\mathrm{d}})$ B_T

Dual of (DDR) is the kernel representation

$$\mathscr{B} = \ker R(\sigma), \quad \begin{bmatrix} R^1(z) \\ \vdots \\ R^g(z) \end{bmatrix} = \begin{bmatrix} R_0^1 + R_1^1 z + \cdots + R_{\ell_1}^1 z^{\ell_1} \\ \vdots \\ R_0^g + R_1^g z + \cdots + R_{\ell_g}^g z^{\ell_g} \end{bmatrix} \quad \text{(KER)}$$

multiplication matrix (generalized Sylvester)

$$\mathscr{M}_T(R) := \begin{bmatrix} \mathscr{M}_T(R^1) \\ \vdots \\ \mathscr{M}_T(R^g) \end{bmatrix}, \quad \mathscr{M}_T(R^i) := \begin{bmatrix} R_0^i & R_1^i & \cdots & R_{\ell_i}^i & & \\ & \ddots & \ddots & & \ddots & \\ & & R_0^i & R_1^i & \cdots & R_{\ell_i}^i \end{bmatrix}$$

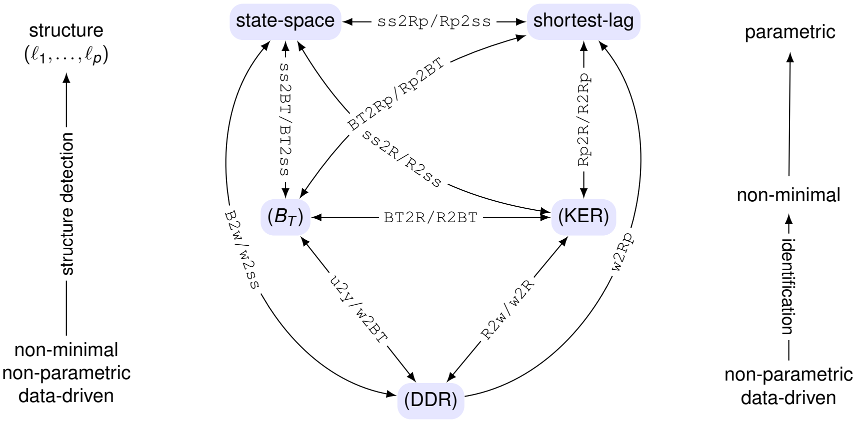
$$\mathscr{B}|_T = \ker \mathscr{M}_T(R) \quad \text{(FHK)}$$

theorem

$$\Updownarrow$$

$$\operatorname{rank} R = \mathbf{p}(\mathscr{B})T - \mathbf{n}(\mathscr{B}) \quad \text{(DGPE)}$$

Transitions among state-space, kernel, and data-driven representations



Back to the example of systems equality

equality in the behavioral setting: $\mathscr{B}^1 \stackrel{?}{=} \mathscr{B}^2$

using (B_T) with $T \geq \max\{\,\boldsymbol{\ell}(\mathscr{B}^1), \boldsymbol{\ell}(\mathscr{B}^2)\,\}$

$$\mathscr{B}^1 = \mathscr{B}^2 \quad \iff \quad \mathscr{B}^1|_T = \mathscr{B}^2|_T$$

systems equality checking

```
T = max(lag(sys1), lag(sys2));
BT1 = ss2BT(sys1, T);
BT2 = ss2BT(sys2, T);
ans = (size(BT1, 2) == size(BT2, 2)) && ...
                 subspace(BT1, BT2) < 1e-10
```

Summary

main assumptions: LTI structure + (GPE)

data-driven representation: $\widehat{\mathscr{B}} := \text{image}\, \mathscr{H}_T(w_\text{d})$

leads to new problems, methods, and insights

- distance to uncontrollability
- interpolation of LTI trajectories
- frequency response evaluation without leakage

decoupling problems from solution methods