The Behavioral Toolbox

Ivan Markovsky

ICREA

CIMNE
INTERNATIONAL CENTRE
FOR NUMERICAL METHODS
IN ENGINEERING

Outline

Historical perspective

Non-parametric repr. of the finite-horizon behavior

LTI structure and its implications

New problems and solutions

Discussion

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	directly using data

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

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

How to check if two LTI systems are equal?

we would like this 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?

Terminology and notation

$w \in (\mathbb{R}^q)^{\mathbb{N}}$, $w : \mathbb{N} \to \mathbb{R}^q$	q-variate discrete-time signal
$w\vert_T := \big(w(1), \dots, w(T)\big)$	restriction of w to $[1, T]$
$\mathscr{B} \subset (\mathbb{R}^q)^{\mathbb{N}}$	q-variables discrete-time system
$\mathscr{B}\vert_T := \{\, w\vert_T \mid w \in \mathscr{B}\,\}$	restriction of $\mathscr{B}$ to $[1, T]$
$\mathscr{L}^q$	set of q-variables LTI systems
$\mathbf{m}(\mathscr{B})$ / $\boldsymbol{\ell}(\mathscr{B})$ / $\mathbf{n}(\mathscr{B})$	number of inputs/lag/order of $\mathscr{B}$
$\mathbf{c}(\mathscr{B}) := \big(\mathbf{m}(\mathscr{B}), \boldsymbol{\ell}(\mathscr{B}), \mathbf{n}(\mathscr{B})\big)$	complexity of $\mathscr{B}$
$\mathscr{L}^q_c := \{\, \mathscr{B} \in \mathscr{L}^q \mid \mathbf{c}(\mathscr{B}) \le c\,\}$	bounded complexity LTI systems

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

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

$\mathscr{B} \in \mathscr{L}_c^q$ 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 $\mathscr{B}$ $\Longrightarrow$ representation of $\mathscr{B}$

We want to obtain B_T from other representations or data

$$(A, B, C, D) \mapsto B_T = \Pi \begin{bmatrix} 0_{mT\times n} & I_{mT} \\ \mathscr{O}_T(A,C) & \mathscr{T}_T(H) \end{bmatrix} \quad (\texttt{ss2BT})$$

$$\text{kernel: } R \xmapsto{???} B_T \quad (\texttt{R2BT})$$

$$\text{data: } \mathscr{W}_\text{d} = \{ w_\text{d}^1, \dots, w_\text{d}^N \} \xmapsto{???} B_T \quad (\texttt{w2BT})$$

$\texttt{w2BT}\colon \mathscr{W}_\mathrm{d} \xmapsto{\texttt{hank}} \mathscr{H}_T(\mathscr{W}_\mathrm{d}) \xmapsto{\texttt{orth}} B_T$

$$\mathscr{W}_\mathrm{d} = \{\, w_\mathrm{d}^1, \dots, w_\mathrm{d}^N \,\}, \quad w_\mathrm{d}^i = \big(w_\mathrm{d}^i(1), \dots, w_\mathrm{d}^i(T_\mathrm{d}^i)\big) \in \mathscr{B}|_{T_\mathrm{d}^i}$$

Hankel matrix

$$\mathscr{H}(\mathscr{W}_\mathrm{d}) := \begin{bmatrix} \mathscr{H}(w_\mathrm{d}^1) & \cdots & \mathscr{H}(w_\mathrm{d}^N) \end{bmatrix}$$

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

$$\mathscr{B}|_T = \operatorname{image} \mathscr{H}_T(\mathscr{W}_\mathrm{d}) \qquad \text{(DDR)}$$

theorem

$$\Updownarrow$$

$$\operatorname{rank} \mathscr{H}_T(\mathscr{W}_\mathrm{d}) = \mathbf{m}(\mathscr{B})T + \mathbf{n}(\mathscr{B}), \; T \geq \boldsymbol{\ell}(\mathscr{B}) \qquad \text{(GPE)}$$

$$\texttt{R2BT}: R \xrightarrow{\texttt{multmat}} \mathscr{M}_T(R) \xrightarrow{\texttt{null}} B_T$$

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

multiplication matrix

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

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

theorem

$$\Updownarrow$$

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

B_T is structured, due to the LTI dynamics

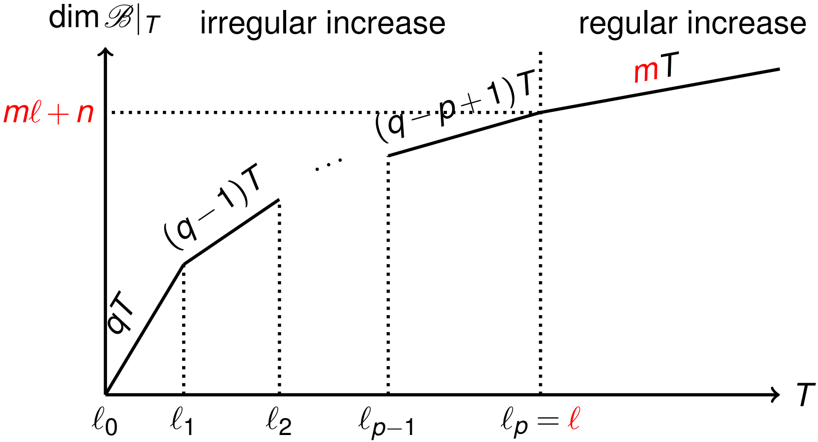
the structure is fully revealed when $T \geq \boldsymbol{\ell}(\mathscr{B})$

complexity bounded $:\iff \dim \mathscr{B}|_T < qT$

time-invariance $\iff$ shift-invariance

- autonomous case: sum-of-exponentials
- open systems: the structure is hidden
- image $\mathscr{H}_T(w_\text{d})$ and ker $\mathscr{M}_T(R)$ impose shift-invariance

$\dim \mathscr{B}|_T$ is a piecewise affine function of T



LTI system's structure is defined by the *structure indices* $(\ell_1, \ldots, \ell_p)$

the structure indices determine the complexity

- $\mathbf{m}(\mathscr{B}) = q - p$
- $\boldsymbol{\ell}(\mathscr{B}) = \ell_p$
- $\mathbf{n}(\mathscr{B}) = \ell_1 + \cdots + \ell_p$

show up in shortest-lag kernel representations

$=$ observability indices in state-space setting

Back to the example of systems equality

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

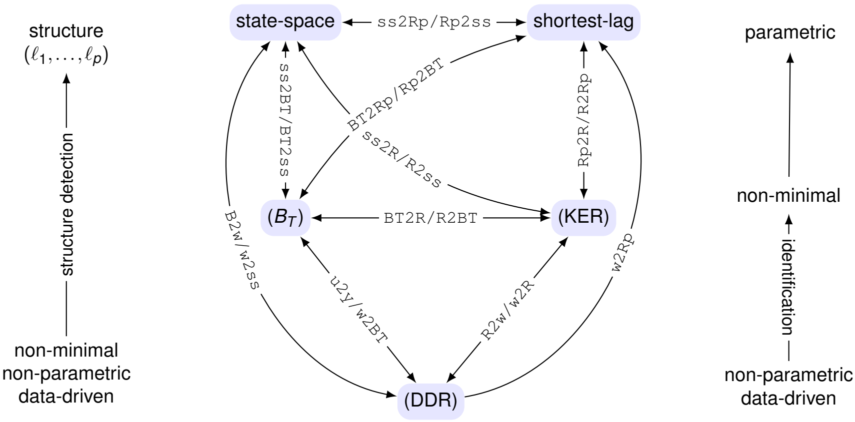
using (B_T) with $T \geq \max \{ \boldsymbol{\ell}(\mathscr{B}^1), \boldsymbol{\ell}(\mathscr{B}^2) \}$

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

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
```

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



Analysis

unified treatment of LTI systems

- static vs dynamic
- autonomous vs open
- unstable, non-minimum phase, descriptor, ...

distance to uncontrollability

- controllability as a property of the system
- data-driven tests for controllability
- meaningful distance measures

causality and optimal sensor/actuator placement

- variables selection from the behavioral perspective
- how fundamental is the notion of causality?

Interconnection and design

interconnection by variables sharing

control as interconnection

dynamic networks in the behavioral setting

Identification

structure detection

identification with pre-specified structure

identification of dynamic networks

Signal processing

input estimation

missing data estimation

data-driven smoothing

Discussion

Demo of structure detection

Hidden dangers in computing (B_T)

On scientific software development

Comparison with convolution representation

Research topics

Demo of how LTI structure is detected

1. random system and trajectory

2. transitions among representations

3. computation time and accuracy

```matlab
%% random system
q = 5; ells = [0 3 3 5]'; % structure
Rp = randB(q, ells); % shortest-lag kernel repr.

%% random trajectory
Td = 100; % number of samples
tic, wd = R2w(Rp, q, Td); t = toc % -> 0.058

%% (KER) -> (BT)
T = max(ells) + 1;
tic, BT = R2BT(Rp, q, T); t = toc % -> 0.006

%% wd -> (BT)
tic, BTh = w2BT(wd, T); t = toc   % -> 0.014

%% check
ans = (size(BT, 2) == size(BTh, 2)) && ...
            subspace(BT, BTh) < 1e-10 % -> true
```

```
%% (KER) -> (SS)
tic, B = R2ss(Rp, q); t = toc % -> 0.161

%% check
size(B)  % -> 4 outputs, 1 inputs, and 11 states
equal(B, BT) % -> true

%% wd -> (SS)
tic, Bh = w2ss(wd); t = toc   % -> 0.0718
equal(Bh, BT) % -> true

%% wd -> Rp, ells (structure detection)
tic, [Rph, ellsh] = w2Rp(wd); t = toc % -> 0.027
all(ellsh == ells) % -> true
```

Computing B_T has hidden dangers

(GPE) / (DGPE) condition

rank computation

$$\dim \mathscr{B}|_T = qT - \operatorname{rank} \mathscr{M}_T(R) = \operatorname{rank} \mathscr{H}_T(\mathscr{W}_\mathrm{d})$$

numerical rank

$$\widehat{r} := \text{\# of singular values } \geq \varepsilon \qquad (\widehat{r})$$

The SVD approximation doesn't impose time-invariant structure on B_T

two options for complexity estimation:

1. specify ε, in which case r is found from ($\widehat{r}$), or
2. $c := (m, \ell, n)$, in which case r is found from ($\dim \mathscr{B}|_T$)

SVD approximation

when "small" singular values are discarded,
B_T is not shift-invariant and therefore
does not represent $\mathscr{B}|_T$ of $\mathscr{B} \in \mathscr{L}_c^q$

The horizon T depends on the usage of (B_T)

in system analysis, where $\mathscr{B} \in \mathscr{L}_c^q$ is given

- T is of the order of $\boldsymbol{\ell}(\mathscr{B})$
- B_T is constructed from another repr. or data $w_{\rm d} \in \mathscr{B}|_{T_{\rm d}}$

$$T_{\rm d} \geq T_{\rm d,min} := \big(\mathbf{m}(\mathscr{B}) + 1\big)\, T + \mathbf{n}(\mathscr{B}) - 1$$

- the results are exact up to numerical errors ($\sim 10^{-10}$)

in SYSID/data-driven control, where $\mathscr{W}_{\rm d}$ is given

- T may be large (*e.g.*, control horizon in DeePC)
- $\mathscr{W}_{\rm d} \mapsto B_T$ is nontrivial (it is an identification problem)
  typically $T_{\rm d} \gg T$ and generally it is best to use all the data
- the results are approximate and the errors "large" ($\sim$20%)

Trefethen's "Ten Digit Algorithms"

accuracy: at least ten digits of precision

efficiency: computed in less than five seconds

conciseness: code fits one page

Software development is a key part of doing research in mathematical engineering

Knuth's story about the development of the $\TeX$

- from 2 PhD students' summer project
- to 10 year long personal project

new ideas are discovered during implementation

organically evolving rather than pre-specified

Convolution is also nonparameteric LTI system repr. but is essentially different

assumes input/output partitioning

assumes zero initial conditions

truncation to finite-horizon $\rightsquigarrow$ approximation

Research topics

systems theory

- descriptor and unstable systems
- distance to uncontrollability
- dynamic networks: interconnection of systems

signal processing

- input estimation
- missing data estimation
- data-driven smoothing

identification

- identification / model reduction with pre-specified structure
- identification of dynamic networks
- control as interconnection

Research topics

numerical linear algebra

- structure-exploiting methods
- uncertainty quantification
- polynomial computations

applications

- fault detection
- dynamic measurements
- optimal sensor / actuator placement