A software package for exact linear system identification

Ivan Markovsky, Jan C. Willems, and Bart De Moor

K.U.Leuven, ESAT-SISTA

Outline

System identification: $w_{\mathrm{d}} \mapsto \widehat{\mathscr{B}} \in \mathscr{M}$

Notation:

- $w_{\mathrm{d}} = (u_{\mathrm{d}}, y_{\mathrm{d}})$ — given data, in this talk a vector time series
- $\widehat{\mathscr{B}}$ — to be found model for w_{d}, in this talk an LTI system
- $\mathscr{M}$ — model class, in this talk the set of LTI systems $\mathscr{L}$

System identification

- defines a mapping $w_{\mathrm{d}} \mapsto \mathscr{B}$
- derives effective algorithms that realize the mapping, and
- develops efficient software that implements the algorithms

Exact identification: two points of view

Find the true data generating system

- assume that $w_{\mathrm{d}} \in \bar{\mathscr{B}} \in \mathscr{L}$
- find back $\bar{\mathscr{B}}$ from w_{d} (and an upper bound of the order)
- this is possible provided $\bar{\mathscr{B}}$ is controllable and an input component of w_{d} is persistently exciting

Find the least complex LTI system that fits w_{d}

- no assumption about w_{d}
- find $\widehat{\mathscr{B}} \in \mathscr{L}$ with minimal # of inputs and order, s.t. $w_{\mathrm{d}} \in \widehat{\mathscr{B}}$
- $\widehat{\mathscr{B}}$—most powerful unfalsified model (MPUM) for w_{d} in $\mathscr{L}$

Exact identification: not a practical SYSID problem

$w_{\rm d}$ can always be fitted exactly

- take all variables as inputs
- for finite $w_{\rm d} \in (\mathbb{R}^{\tt w})^T$, take the order sufficiently large

$\rightsquigarrow$ trivial solution

Of interest is a nontrivial solutions, *i.e.*, we want

$\mathscr{M}$ to be a set of bounded complexity LTI systems $\mathscr{L}_{m,n}$, $\leq {\tt m}$ inputs and order $\leq$ n.

However,

- $w_{\rm d} \in \bar{\mathscr{B}} \in \mathscr{L}_{m,n}$ is a too restrictive assumption
- alternatively, the MPUM generically does not exist in $\mathscr{L}_{m,n}$.

Approximate identification: optimization point of view

- the model need not fit the data exactly
- choose a distance measure $M(w_{\rm d},\mathscr{B})$ between $w_{\rm d}$ and $\mathscr{B}$
- minimize $M(w_{\rm d},\mathscr{B})$ over all models in $\mathscr{M}$

Computing $M(w_{\rm d},\mathscr{B})$ is equivalent to

- finding the "best" approximation of $w_{\rm d}$ in $\mathscr{B}$,
- smoothing or filtering (if causality is imposed) $w_{\rm d}$ by $\mathscr{B}$,
- projecting $w_{\rm d}$ on $\mathscr{B}$.

$M(w_{\rm d},\mathscr{B})$ can be computed in various ways:
smoothing, spectral factorization, Cholesky factorization, . . .

Approximate identification: suboptimal methods

exact identification is more than an academic problem

it leads to suboptimal approximate identification methods

an exact ID method can be used for approximate SYSID by

- rank computation $\leftrightarrow$ numerical rank computation
- solution of equations $\leftrightarrow$ solution of a LS problem

MATLAB does this substitution automatically where necessary

LTI model representations

- Kernel representation (parameter $R(z) := \sum_{i=0}^{\tt l} R_i z^i$)

$$R_0 w(t) + R_1 w(t+1) + \cdots + R_{\tt l} w(t+{\tt l}) = 0$$

- Impulse response represent (parameter $h : \mathbb{Z} \to \mathbb{R}^{p\times m}$)

$$w = \operatorname{col}(u,y), \qquad y(t) = \sum_{\tau=-\infty}^{t} h(\tau) u(t-\tau)$$

- Input/state/output representation (parameter (A,B,C,D))

$$w = \operatorname{col}(u,y), \qquad \begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned}$$

${\tt p} := \dim(y) = \operatorname{rowdim}(R)$ is the # of outputs
${\tt m} := \dim(u)$ is the # of inputs, ${\tt l} := \operatorname{degree}(R)$ is the lag

Algorithms for exact identification

1. $w_d \mapsto R(z)$
2. $w_d \mapsto$ impulse response h
3. $w_d \mapsto (A, B, C, D)$ (possibly balanced)

   3.1 $w_d \mapsto R(\xi) \mapsto (A, B, C, D)$ or $w_d \mapsto H \mapsto (A, B, C, D)$

   3.2 $w_d \mapsto \mathscr{O}_{l_{max}+1}(A, C) \mapsto (A, B, C, D)$

   3.3 $w_d \mapsto \big(x_d(1), \ldots, x_d(n_{max} + m + 1)\big) \mapsto (A, B, C, D)$

There are various ways to implement the mapping

$$w = \mathrm{col}(u, y) \mapsto (A, B, C, D).$$

SISTA

Building blocks for the algorithms

Function	Description
`w2r`	from data (time series w) to a kernel repr.
`r2pq`	from a kernel repr. to an LMF representation
`pq2ss`	from an LMF repr. to an I/S/O representation
`uy2h`	from data to the impulse response
`h2ss`	from the impulse resp. to an I/S/O repr.
`uy2y0`	from data to sequential free responses
`y02ox`	from free responses to an observability matrix and a state sequence
`h2ox`	from the impulse response to an observability matrix and a state sequence
`uy02ss`	from data and an observability matrix to an I/S/O representation
`uyx2ss`	from data and a state seq. to an I/S/O repr.
`hy02xbal`	from the impulse response and sequential free responses to a balanced state sequence

SISTA

Conclusions

- choice of representation
- decomposition of the identification problem into standard easy to solve subproblems
- various ways to achieve the mapping $w_d \mapsto \mathscr{B}$
- can be used as suboptimal approximate ID methods
- open question:

when w_d is not exact, which choice of representation and computational algorithm gives best approximate system?

SISTA

Thank you

SISTA