State representations from finite time series

I. Markovsky, J. C. Willems, and B. De Moor

K.U.Leuven, ESAT-SISTA

Outline

Exact identification

Let $w_\mathrm{d} = (u_\mathrm{d}, y_\mathrm{d}) \in (\mathbb{R}^\mathtt{w})^T$ be generated by an LTI system $\mathscr{B}$.

Problem: find $\mathscr{B}$ back from the data $w_\mathrm{d} = \big(w_\mathrm{d}(1), \dots, w_\mathrm{d}(T)\big)$.

Alternatively, given an arbitrary w_d, find the most powerful unfalsified model $\mathscr{B}_\mathrm{mpum}$ for w_d in the LTI model class.

We consider an input/state/output representation of $\mathscr{B}$:

$$\sigma x = Ax + Bu, \qquad y = Cx + Du$$

where σ is the forward shift operator $(\sigma x)(t) = x(t+1)$

$\mathtt{m}$ — number of inputs, $\mathtt{p} = \mathtt{w} - \mathtt{m}$ — number of outputs of $\mathscr{B}$

Approach

we consider N4SID-type algorithms: $w_\mathrm{d} \mapsto x_\mathrm{d} \mapsto (A, B, C, D)$

x_d — minimal state sequence of $\mathscr{B}$ corresponding to w_d

$(w_\mathrm{d}, x_\mathrm{d}) \mapsto (A, B, C, D)$ is solving a linear system of equations:

$$\begin{bmatrix} \sigma x_\mathrm{d} \\ y_\mathrm{d} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_\mathrm{d} \\ u_\mathrm{d} \end{bmatrix}$$

$w_\mathrm{d} \mapsto x_\mathrm{d}$ is the heart of the identification problem

Result for infinite time series

Define

$$\begin{bmatrix} \mathscr{H}_{\mathrm{p}} \\ \hline \mathscr{H}_{\mathrm{f}} \end{bmatrix} := \left[\begin{array}{ccccc} & \vdots & \vdots & \vdots & \\ \cdots & w_{\mathrm{d}}(-2) & w_{\mathrm{d}}(-1) & w_{\mathrm{d}}(0) & \cdots \\ \cdots & w_{\mathrm{d}}(-1) & w_{\mathrm{d}}(0) & w_{\mathrm{d}}(1) & \cdots \\ \hline \cdots & w_{\mathrm{d}}(0) & w_{\mathrm{d}}(1) & w_{\mathrm{d}}(2) & \cdots \\ \cdots & w_{\mathrm{d}}(1) & w_{\mathrm{d}}(2) & w_{\mathrm{d}}(3) & \cdots \\ & \vdots & \vdots & \vdots & \end{array}\right]$$

The minimal state dimension of $\mathscr{B}_{\mathrm{mpum}}(w_{\mathrm{d}})$ is equal to

$$\mathtt{n} = \operatorname{rank}(\mathscr{H}_{\mathrm{p}}) + \operatorname{rank}(\mathscr{H}_{\mathrm{f}}) - \operatorname{rank}(\mathscr{H}).$$

Moreover, a basis for $\operatorname{row\,span}(\mathscr{H}_{\mathrm{p}}) \cap \operatorname{row\,span}(\mathscr{H}_{\mathrm{f}})$ is a minimal state sequence x_{d} of $\mathscr{B}_{\mathrm{mpum}}(w_{\mathrm{d}})$, corresponding to w_{d}.

Notation

The Hankel matrix with Δ block-rows composed of w_{d} is

$$\mathscr{H}_{\Delta}(w_{\mathrm{d}}) := \begin{bmatrix} w_{\mathrm{d}}(1) & w_{\mathrm{d}}(2) & \cdots & w_{\mathrm{d}}(T-\Delta+1) \\ w_{\mathrm{d}}(2) & w_{\mathrm{d}}(3) & \cdots & w_{\mathrm{d}}(T-\Delta+2) \\ w_{\mathrm{d}}(3) & w_{\mathrm{d}}(4) & \cdots & w_{\mathrm{d}}(T-\Delta+3) \\ \vdots & \vdots & & \vdots \\ w_{\mathrm{d}}(\Delta) & w_{\mathrm{d}}(\Delta+1) & \cdots & w_{\mathrm{d}}(T) \end{bmatrix}$$

Past–future intersection for finite time series

Define the (finite) input and output, “past” and “future” matrices

$$\begin{array}{cc} \begin{bmatrix} U_{\mathrm{p}} \\ U_{\mathrm{f}} \end{bmatrix} & \begin{matrix} \mathtt{m}\Delta \\ \mathtt{m}\Delta \end{matrix} \end{array} := \mathscr{H}_{2\Delta}(u_{\mathrm{d}}), \qquad \begin{array}{cc} \begin{bmatrix} Y_{\mathrm{p}} \\ Y_{\mathrm{f}} \end{bmatrix} & \begin{matrix} \mathtt{p}\Delta \\ \mathtt{p}\Delta \end{matrix} \end{array} := \mathscr{H}_{2\Delta}(y_{\mathrm{d}}),$$

and the matrix of the “past” and “future” state sequences

$$X_{\mathrm{p}} := \begin{bmatrix} x_{\mathrm{d}}(1) & \cdots & x_{\mathrm{d}}(T-2\Delta+1) \end{bmatrix}, \; X_{\mathrm{f}} := \begin{bmatrix} x_{\mathrm{d}}(\Delta+1) & \cdots & x_{\mathrm{d}}(T-\Delta+1) \end{bmatrix}$$

In an early subspece ID paper it is shown that

$$\operatorname{row\,span}(X_{\mathrm{f}}) = \operatorname{row\,span}(W_{\mathrm{p}}) \cap \operatorname{row\,span}(W_{\mathrm{f}})$$

holds under the rank condition $\operatorname{rank}\left(\begin{bmatrix} X_{\mathrm{p}} \\ U_{\mathrm{p}} \end{bmatrix}\right) = \mathtt{n} + \Delta\mathtt{m}$.

This gives two following two-stage SVD procedure.

Input: $w_{\mathrm{d}} = (u_{\mathrm{d}}, y_{\mathrm{d}})$, and $\Delta \in \mathbb{Z}$, $\Delta \geq \mathtt{L}$ (the lag of $\mathscr{B}$).

1: Compute the SVD

$$U\Sigma V^{\top} = \mathscr{H}_{2\Delta}(w_{\mathrm{d}}) =: \begin{array}{cc} \begin{bmatrix} W_{\mathrm{p}} \\ W_{\mathrm{f}} \end{bmatrix} & \begin{matrix} \mathtt{w}\Delta \\ \mathtt{w}\Delta \end{matrix} \end{array}$$

and define the partitionings

$$\begin{array}{cc} \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} & \begin{matrix} \mathtt{w}\Delta \\ \mathtt{w}\Delta \end{matrix} \end{array} =: \begin{array}{c} \begin{matrix} \mathtt{w}\Delta & \mathtt{w}\Delta \end{matrix} \\ \begin{bmatrix} U_1 & U_2 \end{bmatrix} \end{array} := U.$$

2: Compute the matrix $\tilde{X} := U_{12}^{\top} W_{\mathrm{p}}$.

3: Compute the SVD $\bar{U}\bar{\Sigma}\bar{V}^{\top} = \tilde{X}$, and define $\mathtt{n} := \operatorname{rank}(U_{12}^{\top} W_{\mathrm{p}})$

$$\begin{array}{cc} \begin{bmatrix} \bar{U}_1 \\ \bar{U}_2 \end{bmatrix} & \begin{matrix} \mathtt{n} \\ \mathtt{w}\Delta - \mathtt{n} \end{matrix} \end{array} := \bar{U}.$$

Output: $\bar{U}_1$ — a minimal state sequence of $\mathscr{B}$.

Explanation of the algorithm

The first SVD is used to compute a basis for $\operatorname{left\,ker}\left(\mathscr{H}_{2\Delta}(w_\mathrm{d})\right)$,

$$\operatorname{col\,span}(U_2) = \operatorname{left\,ker}\left(\mathscr{H}_{2\Delta}(w_\mathrm{d})\right)$$

From

$$\tilde{X} := U_{12}^\top W_\mathrm{p} = -U_{22}^\top W_\mathrm{f},$$

it follows that

$$\operatorname{row\,span}(\tilde{X}) = \operatorname{row\,span}(W_\mathrm{p}) \cap \operatorname{row\,span}(W_\mathrm{f}),$$

so that the columns of $\tilde{X}$ form a state sequence.

The second SVD is used to compute a minimal state sequence.
There is a nonsingular $S \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$, such that

$$\bar{U}_1 = SX_\mathrm{f}, \qquad \det(S) \neq 0.$$

Kernel structure of a Hankel matrix

Let the columns of N_1 be in the left kernel of $\mathscr{H}_{\Delta_1}(w_\mathrm{d})$, *i.e.*,

$$N_1^\top \mathscr{H}_{\Delta_1}(w_\mathrm{d}) = 0$$

and $N_{1,i} \in \mathbb{R}^{\mathtt{w}\times\operatorname{col\,dim}(N_1)}$ be the ith block element of N_1. Define

$$N_2 := \begin{bmatrix} N_{1,0} & 0 & \cdots & 0 \\ \vdots & N_{1,0} & \ddots & \vdots \\ N_{1,\Delta_1-1} & \vdots & \ddots & 0 \\ 0 & N_{1,\Delta_1-1} & & N_{1,0} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & N_{1,\Delta_1-1} \end{bmatrix}$$

Then, for $\Delta_2 > \Delta_1$,

$$N_2^\top \mathscr{H}_{\Delta_2}(w_\mathrm{d}) = 0.$$

Kernel structure of a Hankel matrix

Lemma

$$\operatorname{col\,span}(N_1) = \operatorname{left\,ker}\left(\mathscr{H}_{\Delta_1}(w_\mathrm{d})\right) \implies \operatorname{col\,span}(N_2) = \operatorname{left\,ker}\left(\mathscr{H}_{\Delta_2}(w_\mathrm{d})\right)$$

holds under the conditions that

1. $\Delta_1 > \mathtt{L}$,
2. $\mathscr{B}_\mathrm{mpum}$ is controllable, and
3. $\mathscr{H}_{\mathtt{L}+1+\mathtt{n}}(u_\mathrm{d})$ is full rank (persistency of excitation).

Therefore the knowledge of $\operatorname{left\,ker}\left(\mathscr{H}_{\Delta}(w_\mathrm{d})\right)$, $\Delta > \mathtt{L}$, suffices to construct $\operatorname{left\,ker}\left(\mathscr{H}_{2\Delta}(w_\mathrm{d})\right)$.

This gives the following algorithm.

Input: $w_\mathrm{d} = (u_\mathrm{d}, y_\mathrm{d})$, and $\Delta \in \mathbb{Z}$, $\Delta \geq \mathtt{L}$ (the lag of $\mathscr{B}$).

1: Compute a basis N for $\operatorname{left\,ker}\left(\mathscr{H}_{\Delta}(w_\mathrm{d})\right)$.

2: Compute the matrix

$$\tilde{X} := \begin{bmatrix} N_0^\top & N_1^\top & \cdots & N_{\Delta-2}^\top \\ 0 & N_0^\top & \cdots & N_{\Delta-3}^\top \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & N_0^\top \end{bmatrix} \mathscr{H}_{\Delta-1}(\sigma w_\mathrm{d})$$

3: Compute the SVD $\bar{U}\bar{\Sigma}\bar{V}^\top = \tilde{X}$ and define $\mathtt{n} := \operatorname{rank}(U_{12}^\top W_\mathrm{p})$

$$\begin{array}{cc} \begin{bmatrix} \bar{U}_1 \\ \bar{U}_2 \end{bmatrix} & \begin{array}{c} \mathtt{n} \\ \mathtt{w}\Delta - \mathtt{n} \end{array} \end{array} := \bar{U}.$$

Output: $\bar{U}_1$ — a minimal state sequence of $\mathscr{B}$.

Explanation of the algorithm

Step 2 of the new algorithm is the shift-and-cut map.

If the kernel computation on step 1 is carried out by the SVD

$$U\Sigma V^\top = \mathscr{H}_\Delta(w_\mathrm{d})$$

the new algorithm matches exactly the strucutre of the past-future intersection algorithm.

Now however we use only matrices with Δ block rows, while before the matrices were with 2Δ block rows.

This leads to improved numerical efficiency.

Simulation example

$\mathscr{B}$ is a random system, $\quad \mathtt{n} = 6$, $\mathtt{m} = 3$, and $\mathtt{p} = 2$

w_d is a random trajectory of $\mathscr{B}$, $\quad T = 100$

$\hat{\mathscr{B}}_1$ — the system computed by the classical algorithm
$\hat{\mathscr{B}}_2$ — the system computed by the shift-and-cut algorithm

$$\|\mathscr{B} - \hat{\mathscr{B}}_1\|_\infty = 2.4 \times 10^{-15} \quad \text{and} \quad \|\mathscr{B} - \hat{\mathscr{B}}_2\|_\infty = 4.3 \times 10^{-15}$$

$\implies$ $\mathscr{B}$ is recovered exactly from w_d (up to the numerical errors)

computational requirements of the algorithms:

$$f_1 = 1.9 \times 10^6 \quad \text{and} \quad f_2 = 1.3 \times 10^6$$

Thank you