System identification with missing data

Ivan Markovsky and Konstantin Usevich

UNIVERSITY OF Southampton

Motivation

- Partial realization problem:

$$\underset{\text{over ?'s}}{\text{minimize}} \quad \operatorname{rank} \underbrace{\begin{bmatrix} h(1) & h(2) & h(3) & \cdots & h(T) \\ h(2) & h(3) & \iddots & h(T) & ? \\ h(3) & \iddots & \iddots & \iddots & ? \\ \vdots & h(T) & \iddots & \iddots & \vdots \\ h(T) & ? & ? & \cdots & ? \end{bmatrix}}_{\mathscr{H}_T(h)}$$

  $\rightsquigarrow$?'s (missing values) are in the "future"

  $\rightsquigarrow$ extrapolation of sequence by LTI model

- Generalization: arbitrary missing values

  $\rightsquigarrow$ interpolation/extrapolation of a sequence by (multivariable, input/output) LTI model

  $\rightsquigarrow$ Hankel low-rank matrix completion

$$\text{minimize} \quad \text{over } \widehat{h} \quad \operatorname{rank}\big(\mathscr{H}_L(\widehat{h})\big)$$
$$\text{subject to} \quad \widehat{h}(\mathscr{I}_{\text{data}}) = h(\mathscr{I}_{\text{data}})$$

Solution methods

- nuclear norm heuristic
- local optimization methods

  I. Markovsky and K. Usevich. Structured low-rank approximation with missing values.
  `http://eprints.soton.ac.uk/340718`
- subspace-type method (this poster)

Subspace-type method

- let $\operatorname{rank}\big(\mathscr{H}_{r+1}(h)\big) = r$
- there is $p \neq 0$, $\begin{bmatrix} p_0 & p_1 & \cdots & p_r \end{bmatrix} \mathscr{H}_{r+1}(h) = 0$
- consider the matrix $\mathscr{H}_{r+2}(h)$; we have

$$\underbrace{\begin{bmatrix} p_0 & p_1 & \cdots & p_r & 0 \\ 0 & p_0 & p_1 & \cdots & p_r \end{bmatrix}}_{\widetilde{P}} \mathscr{H}_{r+2}(h) = 0$$

- $\widetilde{P}$ is full row rank, so that for any i there is $\widetilde{p}^i \neq 0$,

$$\widetilde{p}^i \mathscr{H}_{r+2}(h) = 0 \qquad \text{and} \qquad \widetilde{p}^i_i = 0$$

- suppose that $\mathscr{H}_{r+2}(h)$ has at least $r+1$ columns with single missing value in the ith position and let $\widetilde{H}^i$ be the corresponding submatrix of $\mathscr{H}_{r+2}(h)$
- $\operatorname{left\,ker}(\widetilde{H}^i) = \alpha \widetilde{p}^i$, for some $\alpha \neq 0$

  $\rightsquigarrow$ $\widetilde{p}^i$ can be identified (up to scaling factor)
- suppose also that $\mathscr{H}_{r+2}(h)$ has at least $r+1$ columns with single missing value in the jth position, where $j \neq i$

  $\rightsquigarrow$ $\widetilde{p}^j$ can be identified (up to scaling factor)
- $\operatorname{GCD}(p^i, p^j) = p \quad \rightsquigarrow \quad$ identification algorithm with missing values

Applications

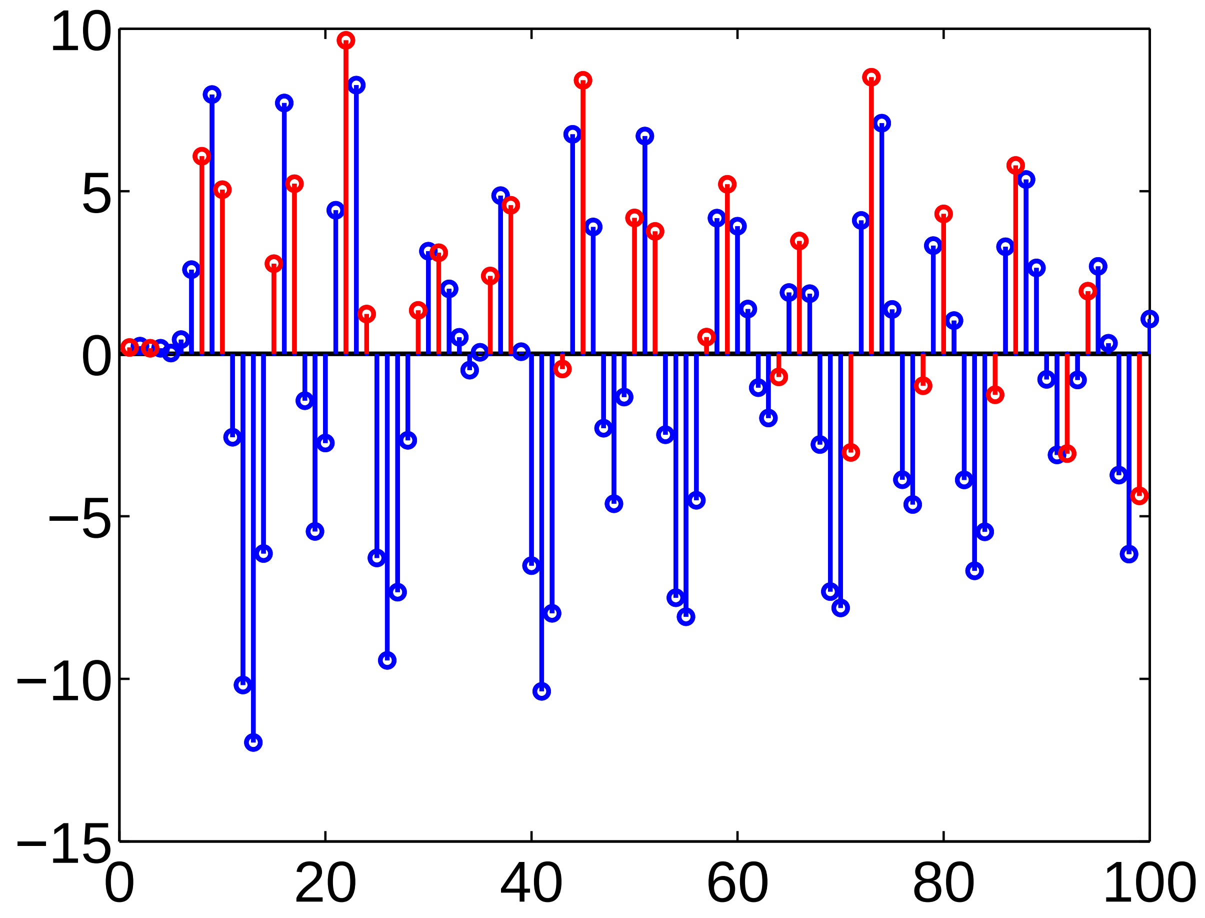
- system identification with missing data
- data-driven simulation
- data-driven tracking

Extensions

- $\Delta > 1$ missing values per frame
- multivariable time series
- input/output systems
- multiple time series
- noisy data — require approx. kernel and GCD computations

$$\min \begin{bmatrix} \operatorname{rank}\big(\mathscr{H}_L(\widehat{h})\big) \\ \|\widehat{h}(\mathscr{I}_{\text{data}}) - h(\mathscr{I}_{\text{data}})\| \end{bmatrix} \begin{matrix} \leftarrow \text{complexity} \\ \leftarrow \text{approx. error} \end{matrix}$$

Example: exact data, autonomous 6th order system



data points   interpolated values

the data is available from

`users.ecs.soton.ac.uk/im/d.mat`

Research supported by ERC