Least Squares Projection Onto the Behavior for SISO LTI Models

20th IFAC Symposium on System Identification (SYSID) - Boston (MA), USA

Sibren Lagauw*, Bart De Moor
{sibren.lagauw, bart.demoor}@esat.kuleuven.be

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Center for Dynamical Systems, Signal Processing, and Data Analytics (STADIUS),
Department of Electrical Engineering (ESAT), KU Leuven, Belgium

July 17, 2024

Fonds Wetenschappelijk Onderzoek
Vlaanderen
Opening new horizons

erc

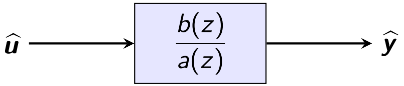
Table of contents

Table of contents

SISO LTI recurrence relation

- LTI dynamics of $(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{y}}) \in \mathbb{R}^{2\times N}$:

$$\sum_{i=0}^{n} a_i \widehat{y}_{k-i} - \sum_{i=0}^{n} b_i \widehat{u}_{k-i} = 0,\ \forall k \in \{n, \ldots, N-1\},$$



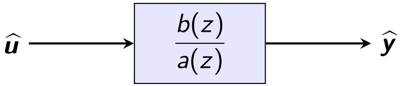
SISO LTI recurrence relation

- LTI dynamics of $(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{y}}) \in \mathbb{R}^{2\times N}$:

$$\sum_{i=0}^{n} a_i \widehat{y}_{k-i} - \sum_{i=0}^{n} b_i \widehat{u}_{k-i} = 0, \ \forall k \in \{n, \ldots, N-1\},$$

- Transfer function:

$$\widehat{H}(z) = \frac{b(z)}{a(z)} = \frac{b_0 z^n + \cdots + b_{n-1} z + b_n}{a_0 z^n + \cdots + a_{n-1} z + a_n}.$$



SISO LTI recurrence relation

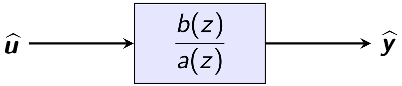
- LTI dynamics of $(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{y}}) \in \mathbb{R}^{2\times N}$:

$$\sum_{i=0}^{n} a_i \widehat{y}_{k-i} - \sum_{i=0}^{n} b_i \widehat{u}_{k-i} = 0, \ \forall k \in \{n, \dots, N-1\},$$

- Transfer function:

$$\widehat{H}(z) = \frac{b(z)}{a(z)} = \frac{b_0 z^n + \cdots + b_{n-1} z + b_n}{a_0 z^n + \cdots + a_{n-1} z + a_n}.$$

- Model params: $\boldsymbol{a} \in \mathbb{R}^{n+1}$, $\boldsymbol{b} \in \mathbb{R}^{n+1}$
  - Normalization ($a_0 = 1$)



Behavior in the null space

$$\sum_{i=0}^{n} a_i \widehat{y}_{k-i} - \sum_{i=0}^{n} b_i \widehat{u}_{k-i} = 0, \ \forall k \in \{n, \ldots, N-1\},$$

Behavior in the null space

$$\sum_{i=0}^{n} a_i \widehat{y}_{k-i} - \sum_{i=0}^{n} b_i \widehat{u}_{k-i} = 0, \ \forall k \in \{n, \dots, N-1\},$$

$$\iff \begin{bmatrix} a_n & -b_n & a_{n-1} & -b_{n-1} & \dots & \dots & a_0 & -b_0 & & & & \\ & & a_n & -b_n & \dots & \dots & a_1 & -b_1 & a_0 & -b_0 & & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & & a_n & -b_n & \dots & \dots & a_1 & -b_1 & a_0 & -b_0 \end{bmatrix} \begin{bmatrix} \hat{y}_0 \\ \hat{u}_0 \\ \vdots \\ \hat{y}_{N-1} \\ \hat{u}_{N-1} \end{bmatrix} = \widetilde{\boldsymbol{T}}^{\mathsf{T}} \widehat{\boldsymbol{w}} = \boldsymbol{0}.$$

where $\widehat{\boldsymbol{w}} \in \mathbb{R}^{2N}$ is the *model-compliant* data trajectory:

$$\widehat{\boldsymbol{w}} = \begin{bmatrix} \widehat{\boldsymbol{w}}_0^{\mathsf{T}} & \dots & \widehat{\boldsymbol{w}}_{\boldsymbol{N-1}}^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} \widehat{y}_0 & \widehat{u}_0 & \dots & \widehat{y}_{N-1} & \widehat{u}_{N-1} \end{bmatrix}^{\mathsf{T}}.$$

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Behavior in the null space

$$\sum_{i=0}^{n} a_i \widehat{y}_{k-i} - \sum_{i=0}^{n} b_i \widehat{u}_{k-i} = 0, \ \forall k \in \{n, \dots, N-1\},$$

$$\iff \begin{bmatrix} a_n & -b_n & a_{n-1} & -b_{n-1} & \dots & \dots & a_0 & -b_0 & & & & \\ & & a_n & -b_n & \dots & \dots & a_1 & -b_1 & a_0 & -b_0 & & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & & a_n & -b_n & \dots & \dots & a_1 & -b_1 & a_0 & -b_0 \end{bmatrix} \begin{bmatrix} \hat{y}_0 \\ \hat{u}_0 \\ \vdots \\ \hat{y}_{N-1} \\ \hat{u}_{N-1} \end{bmatrix} = \widetilde{\boldsymbol{T}}^{\mathsf{T}} \widehat{\boldsymbol{w}} = \boldsymbol{0}.$$

where $\widehat{\boldsymbol{w}} \in \mathbb{R}^{2N}$ is the *model-compliant* data trajectory:

$$\widehat{\boldsymbol{w}} = \begin{bmatrix} \widehat{\boldsymbol{w}}_0^{\mathsf{T}} & \dots & \widehat{\boldsymbol{w}}_{\boldsymbol{N-1}}^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} \widehat{y}_0 & \widehat{u}_0 & \dots & \widehat{y}_{N-1} & \widehat{u}_{N-1} \end{bmatrix}^{\mathsf{T}}.$$

- $\widetilde{\boldsymbol{T}} \in \mathbb{R}^{2N \times (N-n)}$ has full-column rank $(a_0 = 1)$

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Behavior in the null space

$$\sum_{i=0}^{n} a_i \widehat{y}_{k-i} - \sum_{i=0}^{n} b_i \widehat{u}_{k-i} = 0,\ \forall k \in \{n, \ldots, N-1\},$$

$$\iff \begin{bmatrix} a_n & -b_n & a_{n-1} & -b_{n-1} & \ldots & \ldots & a_0 & -b_0 & & & & \\ & & a_n & -b_n & \ldots & \ldots & a_1 & -b_1 & a_0 & -b_0 & & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & & a_n & -b_n & \ldots & \ldots & a_1 & -b_1 & a_0 & -b_0 \end{bmatrix} \begin{bmatrix} \hat{y}_0 \\ \hat{u}_0 \\ \vdots \\ \hat{y}_{N-1} \\ \hat{u}_{N-1} \end{bmatrix} = \widetilde{\boldsymbol{T}}^{\mathsf{T}} \widehat{\boldsymbol{w}} = \boldsymbol{0}.$$

where $\widehat{\boldsymbol{w}} \in \mathbb{R}^{2N}$ is the *model-compliant* data trajectory:

$$\widehat{\boldsymbol{w}} = \begin{bmatrix} \widehat{\boldsymbol{w}}_0^{\mathsf{T}} & \ldots & \widehat{\boldsymbol{w}}_{\boldsymbol{N-1}}^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} \widehat{y}_0 & \widehat{u}_0 & \ldots & \widehat{y}_{N-1} & \widehat{u}_{N-1} \end{bmatrix}^{\mathsf{T}}.$$

- $\widetilde{\boldsymbol{T}} \in \mathbb{R}^{2N\times(N-n)}$ has full-column rank $(a_0 = 1)$
- Behavior: $(N+n)$-dim. subspace $\text{null}\left(\widetilde{\boldsymbol{T}}^{\mathsf{T}}\right)$

Behavior in the null space

$$\sum_{i=0}^{n} a_i \widehat{y}_{k-i} - \sum_{i=0}^{n} b_i \widehat{u}_{k-i} = 0,\ \forall k \in \{n, \dots, N-1\},$$

$$\iff \begin{bmatrix} a_n & -b_n & a_{n-1} & -b_{n-1} & \dots & \dots & a_0 & -b_0 & & & & \\ & & a_n & -b_n & \dots & \dots & a_1 & -b_1 & a_0 & -b_0 & & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & & a_n & -b_n & \dots & \dots & a_1 & -b_1 & a_0 & -b_0 \end{bmatrix} \begin{bmatrix} \hat{y}_0 \\ \hat{u}_0 \\ \vdots \\ \hat{y}_{N-1} \\ \hat{u}_{N-1} \end{bmatrix} = \widetilde{\boldsymbol{T}}^{\mathsf{T}} \widehat{\boldsymbol{w}} = \boldsymbol{0}.$$

where $\widehat{\boldsymbol{w}} \in \mathbb{R}^{2N}$ is the *model-compliant* data trajectory:

$$\widehat{\boldsymbol{w}} = \left[\widehat{\boldsymbol{w}}_0^{\mathsf{T}}\ \dots\ \widehat{\boldsymbol{w}}_{\boldsymbol{N-1}}^{\mathsf{T}}\right]^{\mathsf{T}} = \left[\widehat{y}_0\ \widehat{u}_0\ \dots\ \widehat{y}_{N-1}\ \widehat{u}_{N-1}\right]^{\mathsf{T}}.$$

- $\widetilde{\boldsymbol{T}} \in \mathbb{R}^{2N \times (N-n)}$ has full-column rank $(a_0 = 1)$
- Behavior: $(N+n)$-dim. subspace $\text{null}\left(\widetilde{\boldsymbol{T}}^{\mathsf{T}}\right) \longrightarrow$ #dofs: inputs (N) + init. state (n)

Basis matrix for the null space

$$\begin{bmatrix} a_n & -b_n & a_{n-1} & -b_{n-1} & \dots & \dots & a_0 & -b_0 & & & & & \\ & & a_n & -b_n & \dots & \dots & a_1 & -b_1 & a_0 & -b_0 & & & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & & & a_n & -b_n & \dots & \dots & a_1 & -b_1 & a_0 & -b_0 \end{bmatrix} \quad \widehat{\boldsymbol{T}} \in \mathbb{R}^{2N \times (N+n)} \quad = \mathbf{0}.$$

$$\iff \widetilde{\boldsymbol{T}}^{\top} \widehat{\boldsymbol{T}} = \mathbf{0}$$

Basis matrix for the null space

$$
\underbrace{\begin{bmatrix}
a_n & -b_n & a_{n-1} & -b_{n-1} & \cdots & \cdots & a_0 & -b_0 & & & & \\
 & & a_n & -b_n & \cdots & \cdots & a_1 & -b_1 & a_0 & -b_0 & & \\
 & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\
 & & & & & a_n & -b_n & \cdots & \cdots & a_1 & -b_1 & a_0 & -b_0
\end{bmatrix}}_{\widetilde{\boldsymbol{T}}^\top \in \mathbb{R}^{(N-n)\times 2N}}
\underbrace{\left[\begin{array}{ccc|cccc|ccc}
b_n & \cdots & b_1 & b_0 & & & & & & \\
a_n & \cdots & a_1 & a_0 & & & & & & \\
 & \ddots & \vdots & b_1 & b_0 & & & & & \\
 & \ddots & \vdots & a_1 & a_0 & \ddots & & & & \\
 & & b_n & \vdots & \vdots & \ddots & b_0 & & & \\
 & & a_n & \vdots & \vdots & & a_0 & & & \\
 & & & b_n & b_{n-1} & & \vdots & b_0 & & \\
 & & & a_n & a_{n-1} & \ddots & \vdots & a_0 & & \\
 & & & & b_n & \ddots & b_{n-1} & \vdots & \ddots & \\
 & & & & a_n & \ddots & a_{n-1} & \vdots & \ddots & \\
 & & & & & \ddots & b_n & b_{n-1} & \cdots & b_0 \\
 & & & & & & a_n & a_{n-1} & \cdots & a_0
\end{array}\right]}_{\widehat{\boldsymbol{T}} \in \mathbb{R}^{2N\times(N+n)}}
= \mathbf{0}.
$$

$$
\iff \quad \widetilde{\boldsymbol{T}}^\top \widehat{\boldsymbol{T}} = \mathbf{0}
$$

Basis matrix for the null space

$$
\begin{bmatrix}
a_n & -b_n & a_{n-1} & -b_{n-1} & \cdots & \cdots & a_0 & -b_0 & & & & \\
 & & a_n & -b_n & \cdots & \cdots & a_1 & -b_1 & a_0 & -b_0 & & \\
 & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\
 & & & & a_n & -b_n & \cdots & \cdots & a_1 & -b_1 & a_0 & -b_0
\end{bmatrix}
\left[\begin{array}{ccc|cccc|ccc}
b_n & \cdots & b_1 & b_0 & & & & & & \\
a_n & \cdots & a_1 & a_0 & & & & & & \\
 & \ddots & \vdots & b_1 & b_0 & & & & & \\
 & \ddots & \vdots & a_1 & a_0 & \ddots & & & & \\
 & & b_n & \vdots & \vdots & \ddots & b_0 & & & \\
 & & a_n & \vdots & \vdots & & a_0 & & & \\
 & & & b_n & b_{n-1} & & \vdots & b_0 & & \\
 & & & a_n & a_{n-1} & \ddots & \vdots & a_0 & & \\
 & & & & b_n & \ddots & b_{n-1} & \vdots & \ddots & \\
 & & & & a_n & \ddots & a_{n-1} & \vdots & \ddots & \\
 & & & & & \ddots & b_n & b_{n-1} & \cdots & b_0 \\
 & & & & & & a_n & a_{n-1} & \cdots & a_0
\end{array}\right] = \mathbf{0}.
$$

$$
\iff \quad \widetilde{\boldsymbol{T}}^{\top}\widehat{\boldsymbol{T}} = \mathbf{0}
$$

- for $a(z)$, $b(z)$ coprime $\Rightarrow$ $\widehat{\boldsymbol{T}}$ full column rank (Legat et al., 2023)

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Basis matrix for the null space

$$
\begin{bmatrix}
a_n & -b_n & a_{n-1} & -b_{n-1} & \cdots & \cdots & a_0 & -b_0 & & & & \\
& & a_n & -b_n & \cdots & \cdots & a_1 & -b_1 & a_0 & -b_0 & & \\
& & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\
& & & & a_n & -b_n & \cdots & \cdots & a_1 & -b_1 & a_0 & -b_0
\end{bmatrix}
\left[\begin{array}{ccc|cccc|ccc}
b_n & \cdots & b_1 & b_0 & & & & & & \\
a_n & \cdots & a_1 & a_0 & & & & & & \\
& \ddots & \vdots & b_1 & b_0 & & & & & \\
& \ddots & \vdots & a_1 & a_0 & \ddots & & & & \\
& & b_n & \vdots & \vdots & \ddots & b_0 & & & \\
& & a_n & \vdots & \vdots & & a_0 & & & \\
& & & b_n & b_{n-1} & & \vdots & b_0 & & \\
& & & a_n & a_{n-1} & \ddots & \vdots & a_0 & & \\
& & & & b_n & \ddots & b_{n-1} & \vdots & \ddots & \\
& & & & a_n & \ddots & a_{n-1} & \vdots & \ddots & \\
& & & & & \ddots & b_n & b_{n-1} & \cdots & b_0 \\
& & & & & & a_n & a_{n-1} & \cdots & a_0
\end{array}\right] = \mathbf{0}.
$$

$$
\iff \widetilde{\boldsymbol{T}}^\top \widehat{\boldsymbol{T}} = \mathbf{0}
$$

- for $a(z)$, $b(z)$ coprime $\Rightarrow$ $\widehat{\boldsymbol{T}}$ full column rank (Legat et al., 2023)
- Model *behavior*: $\widehat{\boldsymbol{w}} \in \operatorname{null}\left(\widetilde{\boldsymbol{T}}^\top\right) = \operatorname{range}\left(\widehat{\boldsymbol{T}}\right)$

Basis matrix for the null space

$$
\begin{bmatrix}
a_n & -b_n & a_{n-1} & -b_{n-1} & \cdots & \cdots & a_0 & -b_0 & & & & \\
& & a_n & -b_n & \cdots & \cdots & a_1 & -b_1 & a_0 & -b_0 & & \\
& & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\
& & & & a_n & -b_n & \cdots & \cdots & a_1 & -b_1 & a_0 & -b_0
\end{bmatrix}
\left[\begin{array}{ccc|cccc|ccc}
b_n & \cdots & b_1 & b_0 & & & & & & \\
a_n & \cdots & a_1 & a_0 & & & & & & \\
& \ddots & \vdots & b_1 & b_0 & & & & & \\
& \ddots & \vdots & a_1 & a_0 & \ddots & & & & \\
& & b_n & \vdots & \vdots & \ddots & b_0 & & & \\
& & a_n & \vdots & \vdots & & a_0 & & & \\
& & & b_n & b_{n-1} & & \vdots & b_0 & & \\
& & & a_n & a_{n-1} & \ddots & \vdots & a_0 & & \\
& & & & b_n & \ddots & b_{n-1} & \vdots & \ddots & \\
& & & & a_n & \ddots & a_{n-1} & \vdots & \ddots & \\
& & & & & \ddots & b_n & b_{n-1} & \cdots & b_0 \\
& & & & & & a_n & a_{n-1} & \cdots & a_0
\end{array}\right] = \mathbf{0}.
$$

$$
\iff \widetilde{\boldsymbol{T}}^{\top}\widehat{\boldsymbol{T}} = \mathbf{0}
$$

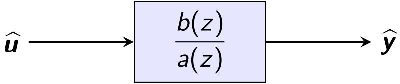
- for $a(z)$, $b(z)$ coprime $\Rightarrow$ $\widehat{\boldsymbol{T}}$ full column rank (Legat et al., 2023)
- Model *behavior*: $\widehat{\boldsymbol{w}} \in \text{null}\left(\widetilde{\boldsymbol{T}}^{\top}\right) = \text{range}\left(\widehat{\boldsymbol{T}}\right)$
- Orthogonal decomposition of ambient space: $\mathbb{R}^{2N} = \underbrace{\text{range}\left(\widehat{\boldsymbol{T}}\right)}_{\text{behavior}} \oplus \text{range}\left(\widetilde{\boldsymbol{T}}\right)$.

Table of contents

Observed data is not model-compliant

For a given model $(\boldsymbol{a}, \boldsymbol{b})$:

- In practice: e.g., measurement inaccuracies, missing data, and model mismatch



STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

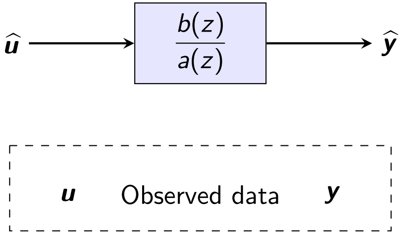
Observed data is not model-compliant

For a given model $(\boldsymbol{a}, \boldsymbol{b})$:

- In practice: e.g., measurement inaccuracies, missing data, and model mismatch

$\rightarrow$ observed data $(\boldsymbol{y}, \boldsymbol{u})$ as such are almost never *model-compliant*:

$$\boldsymbol{w} \notin \text{range}\left(\widehat{\boldsymbol{T}}\right)$$



STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Observed data is not model-compliant

For a given model $(\boldsymbol{a}, \boldsymbol{b})$:

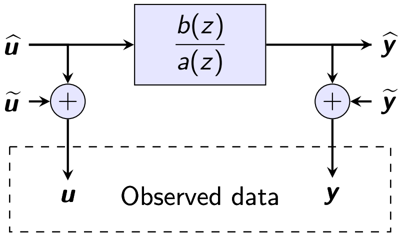
- In practice: e.g., measurement inaccuracies, missing data, and model mismatch

→ observed data $(\boldsymbol{y}, \boldsymbol{u})$ as such are almost never *model-compliant*:

$$\boldsymbol{w} \notin \operatorname{range}\left(\widehat{\boldsymbol{T}}\right)$$

- Modify observed data using *misfits* $(\widetilde{\boldsymbol{y}}, \widetilde{\boldsymbol{u}})$

$$\widehat{\boldsymbol{y}} = \boldsymbol{y} - \widetilde{\boldsymbol{y}}, \quad \widehat{\boldsymbol{u}} = \boldsymbol{u} - \widetilde{\boldsymbol{u}}$$



Observed data is not model-compliant

For a given model $(\boldsymbol{a}, \boldsymbol{b})$:

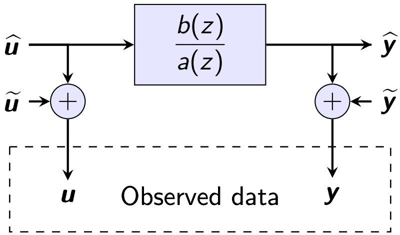
- In practice: e.g., measurement inaccuracies, missing data, and model mismatch

→ observed data $(\boldsymbol{y}, \boldsymbol{u})$ as such are almost never *model-compliant*:

$$\boldsymbol{w} \notin \text{range}\left(\widehat{\boldsymbol{T}}\right)$$

- Modify observed data using *misfits* $(\widetilde{\boldsymbol{y}}, \widetilde{\boldsymbol{u}})$

$$\widehat{\boldsymbol{y}} = \boldsymbol{y} - \widetilde{\boldsymbol{y}}, \quad \widehat{\boldsymbol{u}} = \boldsymbol{u} - \widetilde{\boldsymbol{u}}$$



$$\begin{aligned} \min_{\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{y}}} \quad & J = \left\| \begin{bmatrix} \boldsymbol{y} - \widehat{\boldsymbol{y}} \\ \boldsymbol{u} - \widehat{\boldsymbol{u}} \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \widetilde{\boldsymbol{y}} \\ \widetilde{\boldsymbol{u}} \end{bmatrix} \right\|_2^2, \\ \text{s.t.} \quad & \widetilde{\boldsymbol{T}}^{\mathsf{T}} \widehat{\boldsymbol{w}} = \boldsymbol{0} \text{ and } a_0 = 1. \end{aligned}$$

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Orthogonality between model-compliant data and misfits

With $\boldsymbol{l} = [l_0, \ldots, l_{N-n-1}]^\mathsf{T} \in \mathbb{R}^{N-n}$ Lagrange multipliers, define

$$\mathcal{L}(\widehat{\boldsymbol{w}}, \boldsymbol{l}) = \sum_{k=0}^{N-1} \|\boldsymbol{w}_k - \widehat{\boldsymbol{w}}_k\|_2^2 + \sum_{k=n}^{N-1} l_{k-n}\left(\sum_{i=0}^{n} a_i \widehat{y}_{k-i} - \sum_{i=0}^{n} b_i \widehat{u}_{k-i}\right).$$

Orthogonality between model-compliant data and misfits

With $\boldsymbol{l} = [l_0, \ldots, l_{N-n-1}]^\mathsf{T} \in \mathbb{R}^{N-n}$ Lagrange multipliers, define

$$\mathcal{L}(\widehat{\boldsymbol{w}}, \boldsymbol{l}) = \sum_{k=0}^{N-1} \|\boldsymbol{w}_k - \widehat{\boldsymbol{w}}_k\|_2^2 + \sum_{k=n}^{N-1} l_{k-n} \left(\sum_{i=0}^{n} a_i \widehat{y}_{k-i} - \sum_{i=0}^{n} b_i \widehat{u}_{k-i} \right).$$

First-order necessary conditions for optimality:

- $\partial \mathcal{L}(\ldots)/l_k = 0, \ \forall k \in \{0, \ldots, N-1\} \Rightarrow \widehat{\boldsymbol{w}} \in \mathrm{null}\left(\widetilde{\boldsymbol{T}}^\mathsf{T} \right)$

Orthogonality between model-compliant data and misfits

With $\boldsymbol{l} = [l_0, \ldots, l_{N-n-1}]^\mathsf{T} \in \mathbb{R}^{N-n}$ Lagrange multipliers, define

$$\mathcal{L}(\widehat{\boldsymbol{w}}, \boldsymbol{l}) = \sum_{k=0}^{N-1} \|\boldsymbol{w}_k - \widehat{\boldsymbol{w}}_k\|_2^2 + \sum_{k=n}^{N-1} l_{k-n} \left(\sum_{i=0}^{n} a_i \widehat{y}_{k-i} - \sum_{i=0}^{n} b_i \widehat{u}_{k-i} \right).$$

First-order necessary conditions for optimality:

- $\partial \mathcal{L}(\ldots)/l_k = 0,\ \forall k \in \{0, \ldots, N-1\} \Rightarrow \widehat{\boldsymbol{w}} \in \text{null}\left(\widetilde{\boldsymbol{T}}^\mathsf{T}\right)$
- $\partial \mathcal{L}(\ldots)/\widehat{\boldsymbol{w}}_k = 0,\ \forall k \in \{0, \ldots, N-1\} \Rightarrow \widetilde{\boldsymbol{w}} = \widetilde{\boldsymbol{T}} \boldsymbol{l} \Rightarrow \widetilde{\boldsymbol{w}} \in \text{range}\left(\widetilde{\boldsymbol{T}}\right)$,

Orthogonality between model-compliant data and misfits

With $\boldsymbol{l} = [l_0, \ldots, l_{N-n-1}]^\mathsf{T} \in \mathbb{R}^{N-n}$ Lagrange multipliers, define

$$\mathcal{L}(\widehat{\boldsymbol{w}}, \boldsymbol{l}) = \sum_{k=0}^{N-1} \|\boldsymbol{w}_k - \widehat{\boldsymbol{w}}_k\|_2^2 + \sum_{k=n}^{N-1} l_{k-n} \left(\sum_{i=0}^{n} a_i \widehat{y}_{k-i} - \sum_{i=0}^{n} b_i \widehat{u}_{k-i} \right).$$

First-order necessary conditions for optimality:

- $\partial\mathcal{L}(\ldots)/l_k = 0,\ \forall k \in \{0, \ldots, N-1\} \Rightarrow \widehat{\boldsymbol{w}} \in \operatorname{null}\left(\widetilde{\boldsymbol{T}}^\mathsf{T}\right)$
- $\partial\mathcal{L}(\ldots)/\widehat{\boldsymbol{w}}_k = 0,\ \forall k \in \{0, \ldots, N-1\} \Rightarrow \widetilde{\boldsymbol{w}} = \widetilde{\boldsymbol{T}}\boldsymbol{l} \Rightarrow \widetilde{\boldsymbol{w}} \in \operatorname{range}\left(\widetilde{\boldsymbol{T}}\right),$

$$\longrightarrow \mathbb{R}^{2N} = \underbrace{\operatorname{range}\left(\widehat{\boldsymbol{T}}\right)}_{\text{behavior}} \oplus \underbrace{\operatorname{range}\left(\widetilde{\boldsymbol{T}}\right)}_{\text{misfits}}.$$

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Optimal solution via orthogonal projection(s)

- The minimal norm misfits $\widetilde{\boldsymbol{w}}$: orth. projection of $\boldsymbol{w}$ onto range$\left(\widetilde{\boldsymbol{T}}\right)$,

$$\widetilde{\boldsymbol{w}} = (\widetilde{\boldsymbol{T}}^{\top})^{\dagger}\widetilde{\boldsymbol{T}}^{\top}\boldsymbol{w} = \widetilde{\boldsymbol{T}}(\widetilde{\boldsymbol{T}}^{\top}\widetilde{\boldsymbol{T}})^{-1}\widetilde{\boldsymbol{T}}^{\top}\boldsymbol{w},$$

Optimal solution via orthogonal projection(s)

- The minimal norm misfits $\widetilde{\boldsymbol{w}}$: orth. projection of $\boldsymbol{w}$ onto $\text{range}\left(\widetilde{\boldsymbol{T}}\right)$,

$$\widetilde{\boldsymbol{w}} = (\widetilde{\boldsymbol{T}}^{\top})^{\dagger}\widetilde{\boldsymbol{T}}^{\top}\boldsymbol{w} = \widetilde{\boldsymbol{T}}(\widetilde{\boldsymbol{T}}^{\top}\widetilde{\boldsymbol{T}})^{-1}\widetilde{\boldsymbol{T}}^{\top}\boldsymbol{w},$$

- or equivalently, the optimal model-compliant data $\widehat{\boldsymbol{w}}$: orth. projection of $\boldsymbol{w}$ onto $\text{range}\left(\widehat{\boldsymbol{T}}\right)$,

$$\begin{aligned}\widehat{\boldsymbol{w}} &= \widehat{\boldsymbol{T}}(\widehat{\boldsymbol{T}}^{\top}\widehat{\boldsymbol{T}})^{-1}\widehat{\boldsymbol{T}}^{\top}\boldsymbol{w},\\ &= (\boldsymbol{I} - (\widetilde{\boldsymbol{T}}^{\top})^{\dagger}\widetilde{\boldsymbol{T}}^{\top})\boldsymbol{w}.\end{aligned}$$

Optimal solution via orthogonal projection(s)

- The minimal norm misfits $\widetilde{\boldsymbol{w}}$: orth. projection of $\boldsymbol{w}$ onto range$\left(\widetilde{\boldsymbol{T}}\right)$,

$$\widetilde{\boldsymbol{w}} = (\widetilde{\boldsymbol{T}}^{\top})^{\dagger} \widetilde{\boldsymbol{T}}^{\top} \boldsymbol{w} = \widetilde{\boldsymbol{T}}(\widetilde{\boldsymbol{T}}^{\top} \widetilde{\boldsymbol{T}})^{-1} \widetilde{\boldsymbol{T}}^{\top} \boldsymbol{w},$$

- or equivalently, the optimal model-compliant data $\widehat{\boldsymbol{w}}$: orth. projection of $\boldsymbol{w}$ onto range$\left(\widehat{\boldsymbol{T}}\right)$,

$$\begin{aligned} \widehat{\boldsymbol{w}} &= \widehat{\boldsymbol{T}}(\widehat{\boldsymbol{T}}^{\top} \widehat{\boldsymbol{T}})^{-1} \widehat{\boldsymbol{T}}^{\top} \boldsymbol{w}, \\ &= (\boldsymbol{I} - (\widetilde{\boldsymbol{T}}^{\top})^{\dagger} \widetilde{\boldsymbol{T}}^{\top}) \boldsymbol{w}. \end{aligned}$$

- Generalizes results earlier results on autonomous models to the SISO case:

B. De Moor (2020). "Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem". In: *Communications in Information and Systems* 20.2, pp. 163–207

Table of contents

Misfit model

The optimal misfits $\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{y}}$, are heavily structured:

$$\widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{T}} = \mathbf{0} \iff \widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{T}} \boldsymbol{I} = \mathbf{0} \iff \widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{w}} = \mathbf{0}$$

Misfit model

The optimal misfits $\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{y}}$, are heavily structured:

$$\widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{T}} = \boldsymbol{0} \iff \widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{T}} \boldsymbol{I} = \boldsymbol{0} \iff \widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{w}} = \boldsymbol{0}$$

- the $(n+1)$th equation up until the Nth equation in $\widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{w}} = \boldsymbol{0}$ show that

$$\sum_{i=0}^{n} b_{n-i} \widetilde{y}_{k-i} + \sum_{i=0}^{n} a_{n-i} \widetilde{u}_{k-i} = 0, \ \forall k \in \{n, \ldots, N-1\}.$$

Misfit model

The optimal misfits $\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{y}}$, are heavily structured:

$$\widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{T}} = \boldsymbol{0} \iff \widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{T}} \boldsymbol{l} = \boldsymbol{0} \iff \widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{w}} = \boldsymbol{0}$$

- the $(n+1)$th equation up until the Nth equation in $\widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{w}} = \boldsymbol{0}$ show that

$$\sum_{i=0}^{n} b_{n-i} \widetilde{y}_{k-i} + \sum_{i=0}^{n} a_{n-i} \widetilde{u}_{k-i} = 0 \Rightarrow \widetilde{H}(z) = -\frac{a_r(z)}{b_r(z)} = \frac{a_n z^n + \cdots + a_1 z + a_0}{b_n z^n + \cdots + b_1 z + b_0}.$$

with the 'reversed-coefficients'-polynomials: $a_r(z) = z^n a(z^{-1})$ and $b_r(z) = z^n b(z^{-1})$.

Misfit model

The optimal misfits $\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{y}}$, are heavily structured:

$$\widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{T}} = \boldsymbol{0} \iff \widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{T}} \boldsymbol{l} = \boldsymbol{0} \iff \widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{w}} = \boldsymbol{0}$$

- the $(n+1)$th equation up until the Nth equation in $\widehat{\boldsymbol{T}}^\top \widetilde{\boldsymbol{w}} = \boldsymbol{0}$ show that

$$\sum_{i=0}^{n} b_{n-i} \widetilde{y}_{k-i} + \sum_{i=0}^{n} a_{n-i} \widetilde{u}_{k-i} = 0 \Rightarrow \widetilde{H}(z) = -\frac{a_r(z)}{b_r(z)} = \frac{a_n z^n + \cdots + a_1 z + a_0}{b_n z^n + \cdots + b_1 z + b_0}.$$

with the 'reversed-coefficients'-polynomials: $a_r(z) = z^n a(z^{-1})$ and $b_r(z) = z^n b(z^{-1})$.

- alternatively, express $\widehat{\boldsymbol{w}} \perp \widetilde{\boldsymbol{w}}$ in the z-domain,

$$\langle \frac{b(z)}{a(z)} \widehat{U}(z), \widetilde{Y}(z) \rangle + \langle \widehat{U}(z), \widetilde{U}(z) \rangle = \frac{1}{2\pi i} \oint_{|z|=1} \left[\widetilde{Y}(z) \frac{b(z^{-1})}{a(z^{-1})} + \widetilde{U}(z) \right] \widehat{U}(z^{-1}) \mathrm{d}z = 0$$

$$\iff \widetilde{Y}(z) = -\frac{a(z^{-1})}{b(z^{-1})} \widetilde{U}(z) = -\frac{a_r(z)}{b_r(z)} \widetilde{U}(z) = \widetilde{H}(z) \widetilde{U}(z).$$

Table of contents

Isometric state-space models

- Behavioral state-space recurrence relations,

$$\begin{aligned} \widehat{\boldsymbol{x}}_{k+1} &= \boldsymbol{A}\widehat{\boldsymbol{x}}_k + \boldsymbol{B}\widehat{v}_k \\ \widehat{\boldsymbol{w}}_k &= \boldsymbol{C}\widehat{\boldsymbol{x}}_k + \boldsymbol{D}\widehat{v}_k, \end{aligned} \quad \text{for } k = 0, \ldots, N-1,$$

with $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, $\boldsymbol{B} \in \mathbb{R}^{n}$, $\boldsymbol{C} \in \mathbb{R}^{2 \times n}$, $\boldsymbol{D} \in \mathbb{R}^{2}$.

Isometric state-space models

- **Isometric** behavioral state-space recurrence relations:

$$\begin{bmatrix} \widehat{\boldsymbol{x}}_{k+1} \\ \widehat{\boldsymbol{w}}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{x}}_k \\ \widehat{v}_k \end{bmatrix} \quad \text{for } k = 0, \ldots, N-1, \text{ with } \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} = \boldsymbol{I}_{n+1},$$

and $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, $\boldsymbol{B} \in \mathbb{R}^{n}$, $\boldsymbol{C} \in \mathbb{R}^{2 \times n}$, $\boldsymbol{D} \in \mathbb{R}^{2}$.

Isometric state-space models

- **Isometric** behavioral state-space recurrence relations:

$$\begin{bmatrix} \widehat{\boldsymbol{x}}_{k+1} \\ \widehat{\boldsymbol{w}}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{x}}_k \\ \widehat{v}_k \end{bmatrix} \quad \text{for } k = 0, \ldots, N-1, \text{ with } \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} = \boldsymbol{I}_{n+1},$$

and $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n}$, $\boldsymbol{C} \in \mathbb{R}^{2\times n}$, $\boldsymbol{D} \in \mathbb{R}^{2}$.

- (Isometric) misfit model:

$$\begin{bmatrix} \widetilde{\boldsymbol{x}}_{\boldsymbol{k+1}} \\ \widetilde{\boldsymbol{w}}_{\boldsymbol{k}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & \widetilde{\boldsymbol{B}} \\ \boldsymbol{C} & \widetilde{\boldsymbol{D}} \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{x}}_{\boldsymbol{k}} \\ \widetilde{v}_k \end{bmatrix} \quad \text{for } k = 0, \ldots, N-1, \text{ with } \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} & \widetilde{\boldsymbol{B}} \\ \boldsymbol{C} & \boldsymbol{D} & \widetilde{\boldsymbol{D}} \end{bmatrix} \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} & \widetilde{\boldsymbol{B}} \\ \boldsymbol{C} & \boldsymbol{D} & \widetilde{\boldsymbol{D}} \end{bmatrix}^{\mathsf{T}} = \boldsymbol{I}_{n+2}.$$

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Isometric state-space models

- **Isometric** behavioral state-space recurrence relations:

$$\begin{bmatrix} \widehat{\boldsymbol{x}}_{k+1} \\ \widehat{\boldsymbol{w}}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{x}}_k \\ \widehat{v}_k \end{bmatrix} \quad \text{for } k = 0, \ldots, N-1, \text{ with } \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} = \boldsymbol{I}_{n+1},$$

and $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, $\boldsymbol{B} \in \mathbb{R}^{n}$, $\boldsymbol{C} \in \mathbb{R}^{2 \times n}$, $\boldsymbol{D} \in \mathbb{R}^{2}$.

- (Isometric) misfit model:

$$\begin{bmatrix} \widetilde{\boldsymbol{x}}_{\boldsymbol{k}+\boldsymbol{1}} \\ \widetilde{\boldsymbol{w}}_{\boldsymbol{k}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & \widetilde{\boldsymbol{B}} \\ \boldsymbol{C} & \widetilde{\boldsymbol{D}} \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{x}}_{\boldsymbol{k}} \\ \widetilde{v}_k \end{bmatrix} \quad \text{for } k = 0, \ldots, N-1, \text{ with } \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} & \widetilde{\boldsymbol{B}} \\ \boldsymbol{C} & \boldsymbol{D} & \widetilde{\boldsymbol{D}} \end{bmatrix} \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} & \widetilde{\boldsymbol{B}} \\ \boldsymbol{C} & \boldsymbol{D} & \widetilde{\boldsymbol{D}} \end{bmatrix}^{\mathsf{T}} = \boldsymbol{I}_{n+2}.$$

- Isometry $\longrightarrow$ orthogonality

Ambient space

$$\widehat{\boldsymbol{H}} = \begin{bmatrix} \boldsymbol{D} & 0 & \dots & \dots & \dots & 0 \\ \boldsymbol{CB} & \boldsymbol{D} & 0 & \dots & \dots & 0 \\ \boldsymbol{CAB} & \boldsymbol{CB} & \boldsymbol{D} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ \boldsymbol{CA}^{N-2}\boldsymbol{B} & \boldsymbol{CA}^{N-3}\boldsymbol{B} & \dots & \dots & \boldsymbol{CB} & \boldsymbol{D} \end{bmatrix} \in \mathbb{R}^{2N \times N}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} \boldsymbol{C}^\mathsf{T} & (\boldsymbol{CA})^\mathsf{T} & \dots & (\boldsymbol{CA}^{N-1})^\mathsf{T} \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{2N \times n}$$

Ambient space

$$\underbrace{\begin{bmatrix} \boldsymbol{D} & 0 & \dots & \dots & \dots & 0 \\ \boldsymbol{CB} & \boldsymbol{D} & 0 & \dots & \dots & 0 \\ \boldsymbol{CAB} & \boldsymbol{CB} & \boldsymbol{D} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ \boldsymbol{CA}^{N-2}\boldsymbol{B} & \boldsymbol{CA}^{N-3}\boldsymbol{B} & \dots & \dots & \boldsymbol{CB} & \boldsymbol{D} \end{bmatrix}}_{\widehat{\boldsymbol{H}} \in \mathbb{R}^{2N \times N}} \text{ and } \underbrace{\begin{bmatrix} \boldsymbol{C}^\top & (\boldsymbol{CA})^\top & \dots & (\boldsymbol{CA}^{N-1})^\top \end{bmatrix}^\top}_{\boldsymbol{\Gamma} \in \mathbb{R}^{2N \times n}}$$

- Behavior: $(N + n)$-dimensional subspace

$$\widehat{\boldsymbol{w}} = \widehat{\boldsymbol{H}}\widehat{\boldsymbol{v}} + \boldsymbol{\Gamma}\widehat{\boldsymbol{x}}_0, \Rightarrow \widehat{\boldsymbol{w}} \in \left[\text{range}\left(\widehat{\boldsymbol{H}}\right) \oplus \text{range}(\boldsymbol{\Gamma})\right].$$

Ambient space

$$\underbrace{\begin{bmatrix} \boldsymbol{D} & 0 & \dots & \dots & \dots & 0 \\ \boldsymbol{CB} & \boldsymbol{D} & 0 & \dots & \dots & 0 \\ \boldsymbol{CAB} & \boldsymbol{CB} & \boldsymbol{D} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ \boldsymbol{CA}^{N-2}\boldsymbol{B} & \boldsymbol{CA}^{N-3}\boldsymbol{B} & \dots & \dots & \boldsymbol{CB} & \boldsymbol{D} \end{bmatrix}}_{\widehat{\boldsymbol{H}} \in \mathbb{R}^{2N\times N}} \text{ and } \underbrace{\begin{bmatrix} \boldsymbol{C}^\top & (\boldsymbol{CA})^\top & \dots & (\boldsymbol{CA}^{N-1})^\top \end{bmatrix}^\top}_{\boldsymbol{\Gamma} \in \mathbb{R}^{2N\times n}}$$

- Behavior: $(N+n)$-dimensional subspace

$$\widehat{\boldsymbol{w}} = \widehat{\boldsymbol{H}}\widehat{\boldsymbol{v}} + \boldsymbol{\Gamma}\widehat{\boldsymbol{x}}_0, \Rightarrow \widehat{\boldsymbol{w}} \in \left[\text{range}\left(\widehat{\boldsymbol{H}}\right) \oplus \text{range}(\boldsymbol{\Gamma})\right].$$

- $(N-n)$-dimensional subspace:

$$\mathbb{R}^{2N} = \underbrace{(\text{range}(\boldsymbol{H}) \oplus \text{range}(\boldsymbol{\Gamma}))}_{\text{behavior}} \oplus \underbrace{\text{range}\left(\widetilde{\boldsymbol{H}}\boldsymbol{Z}\right)}_{\text{misfits}}.$$

(FONC $\longrightarrow \widetilde{\boldsymbol{x}}_0 = \widetilde{\boldsymbol{x}}_N = \boldsymbol{0}$)

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Table of contents

Summary

- LS-criterion induces orthogonal decomposition of ambient space
- Minimal misfits can be expressed in terms of the model parameters via an orth. projection
- Misfits are heavily structured
- Two alternative yet equivalent frameworks: state-space vs. input-output

Future work: (globally optimal) SISO LTI misfit modeling

Misfit modeling (SYSID!) is double optimization problem:

$$\min_{\widehat{\boldsymbol{u}},\widehat{\boldsymbol{y}},\boldsymbol{a},\boldsymbol{b}} J = \left\| \begin{bmatrix} \boldsymbol{y} - \widehat{\boldsymbol{y}} \\ \boldsymbol{u} - \widehat{\boldsymbol{u}} \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \widetilde{\boldsymbol{y}} \\ \widetilde{\boldsymbol{u}} \end{bmatrix} \right\|_2^2,$$

$$\text{s.t.} \quad \widetilde{\boldsymbol{T}}^{\mathsf{T}} \widehat{\boldsymbol{w}} = \boldsymbol{0} \text{ and } a_0 = 1,$$

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Future work: (globally optimal) SISO LTI misfit modeling

Misfit modeling (SYSID!) is double optimization problem:

$$\min_{\widehat{\boldsymbol{u}},\widehat{\boldsymbol{y}},\boldsymbol{a},\boldsymbol{b}} J = \left\| \begin{bmatrix} \boldsymbol{y} - \widehat{\boldsymbol{y}} \\ \boldsymbol{u} - \widehat{\boldsymbol{u}} \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \widetilde{\boldsymbol{y}} \\ \widetilde{\boldsymbol{u}} \end{bmatrix} \right\|_2^2, \quad \text{s.t.} \quad \widetilde{\boldsymbol{T}}^\mathsf{T} \widehat{\boldsymbol{w}} = \boldsymbol{0} \text{ and } a_0 = 1, \quad \Rightarrow \quad \min_{\boldsymbol{a},\boldsymbol{b}} J = \boldsymbol{w} \widetilde{\boldsymbol{T}} (\widetilde{\boldsymbol{T}}^\mathsf{T} \widetilde{\boldsymbol{T}})^{-1} \widetilde{\boldsymbol{T}}^\mathsf{T} \boldsymbol{w}, \quad \text{s.t.} \quad \widetilde{\boldsymbol{T}}^\mathsf{T} \widehat{\boldsymbol{w}} = \boldsymbol{0} \text{ and } a_0 = 1.$$

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Future work: (globally optimal) SISO LTI misfit modeling

Misfit modeling (SYSID!) is double optimization problem:

$$\min_{\widehat{\boldsymbol{u}},\widehat{\boldsymbol{y}},\boldsymbol{a},\boldsymbol{b}} J = \left\| \begin{bmatrix} \boldsymbol{y} - \widehat{\boldsymbol{y}} \\ \boldsymbol{u} - \widehat{\boldsymbol{u}} \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \widetilde{\boldsymbol{y}} \\ \widetilde{\boldsymbol{u}} \end{bmatrix} \right\|_2^2, \quad \text{s.t.} \quad \widetilde{\boldsymbol{T}}^\mathsf{T} \widehat{\boldsymbol{w}} = \boldsymbol{0} \text{ and } a_0 = 1, \quad \Rightarrow \quad \min_{\boldsymbol{a},\boldsymbol{b}} J = \boldsymbol{w} \widetilde{\boldsymbol{T}} (\widetilde{\boldsymbol{T}}^\mathsf{T} \widetilde{\boldsymbol{T}})^{-1} \widetilde{\boldsymbol{T}}^\mathsf{T} \boldsymbol{w}, \quad \text{s.t.} \quad \widetilde{\boldsymbol{T}}^\mathsf{T} \widehat{\boldsymbol{w}} = \boldsymbol{0} \text{ and } a_0 = 1.$$

Earlier work: globally optimal system identification

- B. De Moor (2020). "Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem". In: *Communications in Information and Systems* 20.2, pp. 163–207
- S. Lagauw et al. (June 2024). "Exact Characterization of the Global Optima of Least Squares Realization of Autonomous LTI Models as a Multiparameter Eigenvalue Problem". In: *Proc. of the 22nd European Control Conference (ECC)*. Stockholm, Sweden, pp. 3439–3444

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Future work: (globally optimal) SISO LTI misfit modeling

Misfit modeling (SYSID!) is double optimization problem:

$$\min_{\widehat{\boldsymbol{u}},\widehat{\boldsymbol{y}},\boldsymbol{a},\boldsymbol{b}} J = \left\| \begin{bmatrix} \boldsymbol{y} - \widehat{\boldsymbol{y}} \\ \boldsymbol{u} - \widehat{\boldsymbol{u}} \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \widetilde{\boldsymbol{y}} \\ \widetilde{\boldsymbol{u}} \end{bmatrix} \right\|_2^2, \quad \text{s.t.} \quad \widetilde{\boldsymbol{T}}^\mathsf{T} \widehat{\boldsymbol{w}} = \boldsymbol{0} \text{ and } a_0 = 1, \quad \Rightarrow \quad \min_{\boldsymbol{a},\boldsymbol{b}} J = \boldsymbol{w} \widetilde{\boldsymbol{T}} (\widetilde{\boldsymbol{T}}^\mathsf{T} \widetilde{\boldsymbol{T}})^{-1} \widetilde{\boldsymbol{T}}^\mathsf{T} \boldsymbol{w}, \quad \text{s.t.} \quad \widetilde{\boldsymbol{T}}^\mathsf{T} \widehat{\boldsymbol{w}} = \boldsymbol{0} \text{ and } a_0 = 1.$$

Earlier work: globally optimal system identification

- B. De Moor (2020). "Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem". In: *Communications in Information and Systems* 20.2, pp. 163–207
- S. Lagauw et al. (June 2024). "Exact Characterization of the Global Optima of Least Squares Realization of Autonomous LTI Models as a Multiparameter Eigenvalue Problem". In: *Proc. of the 22nd European Control Conference (ECC)*. Stockholm, Sweden, pp. 3439–3444

How to generalize these globally optimal approaches from autonomous to SISO models?

Questions?

The work of Sibren Lagauw was supported by the
FWO Fundamental Research PhD Fellowship under Grant 11K5623N.

References I

- De Moor, B. (2020). "Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem". In: *Communications in Information and Systems* 20.2, pp. 163–207.
- De Moor, B. and B. Roorda (1994). "L_2-optimal linear system identification: Structured Total Least Squares for SISO Systems". In: *Proc. of the 33rd IEEE Conference on Decision and Control (CDC)*. Lake Buena Vista, Florida, USA, pp. 2874–2879.
- Golub, G. H. and V. Pereyra (2003). "Separable nonlinear least squares: the variable projection method and its applications". In: *Inverse Problems* 19.2, R1–R26.
- Lagauw, S., L. Vanpoucke, and B. De Moor (June 2024). "Exact Characterization of the Global Optima of Least Squares Realization of Autonomous LTI Models as a Multiparameter Eigenvalue Problem". In: *Proc. of the 22nd European Control Conference (ECC)*. Stockholm, Sweden, pp. 3439–3444.
- Legat, B., C. Yuan, and P. Parrilo (2023). "Low-Rank Univariate Sum of Squares Has No Spurious Local Minima". In: *SIAM Journal on Optimization* 33.3, pp. 2041–2061.
- Lemmerling, P. and B. De Moor (2001). "Misfit versus latency". In: *Automatica* 37, pp. 2057–2067.
- Roorda, B. and C. Heij (1995). "Global Total Least Squares Modeling of Multivariable Time Series". In: *IEEE Transactions on Automatic Control* 40.1, pp. 50–63.
- Willems, J. C. (1986, 1987). "From time series to linear system, Part I: Finite dimensional linear time invariant systems; Part II: Exact modelling; Part III: Approximate modelling". In: *Automatica* 22/23.5,6,1, pp. 561–580, 675–694, 87–115.

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics