MAX PLANCK INSTITUTE
FOR DYNAMICS OF COMPLEX
TECHNICAL SYSTEMS
MAGDEBURG

COMPUTATIONAL METHODS IN
SYSTEMS AND CONTROL THEORY

A Rank Minimization Approach to Learning Dynamical Systems from Frequency Response Data

Pawan Goyal

Joint work with
Peter Benner (MPI, Magdeburg)
Benjamin Peherstorfer (CIMS, NY University, NY)

Workshop on Mathematics of Reduced Order Models,

ICERM, Brown University, Providence, USA

February 17-21, 2020

Table of Contents

Dynamical systems

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t))), \quad \mathbf{x}(0) = 0,$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t),$$

where

- (generalized) states $\mathbf{x}(t) \in \mathbb{R}^n$ (invertible $\mathbf{E} \in \mathbb{R}^{n\times n}$),
- inputs (controls) $\mathbf{u}(t) \in \mathbb{R}^m$,
- outputs (measurements, quantity of interest) $\mathbf{y}(t) \in \mathbb{R}^q$.

Dynamical systems

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t))), \quad \mathbf{x}(0) = 0,$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t),$$

where

- (generalized) states $\mathbf{x}(t) \in \mathbb{R}^n$ (invertible $\mathbf{E} \in \mathbb{R}^{n\times n}$),
- inputs (controls) $\mathbf{u}(t) \in \mathbb{R}^m$,
- outputs (measurements, quantity of interest) $\mathbf{y}(t) \in \mathbb{R}^q$.

System Classes

Classical linear systems:	$\mathbf{f}(\mathbf{x}) := \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t),$
Delay systems:	$\mathbf{f}(\mathbf{x}) := \mathbf{A}\mathbf{x}(t) + \mathbf{A}_\tau\mathbf{x}(t-\tau) + \mathbf{B}\mathbf{u}(t),$
Second-order system	$\mathbf{f}(\mathbf{x}) := \mathbf{A}\mathbf{x}(t) + \mathbf{A}_1\int_0^t \mathbf{x}(\tau)d\tau + \int_0^t \mathbf{B}\mathbf{u}(\tau)\tau, \ldots,$

Dynamical systems

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{x}(0) = 0,$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t),$$

where

- (generalized) states $\mathbf{x}(t) \in \mathbb{R}^n$ (invertible $\mathbf{E} \in \mathbb{R}^{n \times n}$),
- inputs (controls) $\mathbf{u}(t) \in \mathbb{R}^m$,
- outputs (measurements, quantity of interest) $\mathbf{y}(t) \in \mathbb{R}^q$.

Frequency domain representation

- Using Laplace transform, we can convert time-domain representation into the frequency domain.

Dynamical systems

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t))), \quad \mathbf{x}(0) = 0,$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t),$$

where

- (generalized) states $\mathbf{x}(t) \in \mathbb{R}^n$ (invertible $\mathbf{E} \in \mathbb{R}^{n\times n}$),
- inputs (controls) $\mathbf{u}(t) \in \mathbb{R}^m$,
- outputs (measurements, quantity of interest) $\mathbf{y}(t) \in \mathbb{R}^q$.

Frequency domain representation

- Using Laplace transform, we can convert time-domain representation into the frequency domain.
- This yields $\mathbf{Y}(s) = \mathbf{H}(s)\mathbf{U}(s)$.

Dynamical systems

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{x}(0) = 0,$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t),$$

where

- (generalized) states $\mathbf{x}(t) \in \mathbb{R}^n$ (invertible $\mathbf{E} \in \mathbb{R}^{n \times n}$),
- inputs (controls) $\mathbf{u}(t) \in \mathbb{R}^m$,
- outputs (measurements, quantity of interest) $\mathbf{y}(t) \in \mathbb{R}^q$.

Frequency domain representation

- Using Laplace transform, we can convert time-domain representation into the frequency domain.
- This yields $\mathbf{Y}(s) = \mathbf{H}(s)\mathbf{U}(s)$.
- Hence, $\mathbf{H}(s)$, called as transfer function is known, we can write the output of a system for any given input.

Dynamical systems

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t))), \quad \mathbf{x}(0) = 0,$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t),$$

where

- (generalized) states $\mathbf{x}(t) \in \mathbb{R}^n$ (invertible $\mathbf{E} \in \mathbb{R}^{n\times n}$),
- inputs (controls) $\mathbf{u}(t) \in \mathbb{R}^m$,
- outputs (measurements, quantity of interest) $\mathbf{y}(t) \in \mathbb{R}^q$.

Frequency domain representation

- Using Laplace transform, we can convert time-domain representation into the frequency domain.
- This yields $\mathbf{Y}(s) = \mathbf{H}(s)\mathbf{U}(s)$.
- Hence, $\mathbf{H}(s)$, called as transfer function is known, we can write the output of a system for any given input.
- Moreover, $\mathbf{H}(s)$, the transfer function of a system, completely characterize the dynamics.

Linear system (standard)

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t),$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$

$\xrightarrow{\text{Time} \mapsto \text{Frequency}}$

Linear system (standard)

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}.$$

Linear system (standard)

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t),$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$

Time $\mapsto$ Frequency

Linear system (standard)

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}.$$

Second-order system

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t),$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$

Time $\mapsto$ Frequency

Second-order system

$$\mathbf{G}(s) = \mathbf{C}(s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}\mathbf{B}.$$

Linear system (standard)

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t),$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$

Time $\mapsto$ Frequency

Linear system (standard)

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}.$$

Second-order system

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t),$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$

Time $\mapsto$ Frequency

Second-order system

$$\mathbf{G}(s) = \mathbf{C}(s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}\mathbf{B}.$$

Delay system

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_\tau\mathbf{x}(t-\tau) + \mathbf{B}\mathbf{u}(t),$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$

Time $\mapsto$ Frequency

Delay system

$$\mathbf{G}(s) = C\left(s\mathbf{E} - \mathbf{A} - \mathbf{A}_\tau e^{-s\tau}\right)^{-1}\mathbf{B}.$$

Linear system (standard)

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t),$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$

Time $\mapsto$ Frequency

Linear system (standard)

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}.$$

Second-order system

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t),$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$

Time $\mapsto$ Frequency

Second-order system

$$\mathbf{G}(s) = \mathbf{C}(s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}\mathbf{B}.$$

Delay system

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_\tau\mathbf{x}(t-\tau) + \mathbf{B}\mathbf{u}(t),$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$

Time $\mapsto$ Frequency

Delay system

$$\mathbf{G}(s) = C\left(s\mathbf{E} - \mathbf{A} - \mathbf{A}_\tau e^{-s\tau}\right)^{-1}\mathbf{B}.$$

Integro system

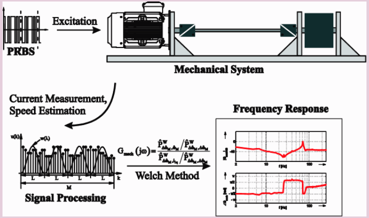
$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_\tau\int_0^t \mathbf{x}(\tau)d\tau + \mathbf{B}\mathbf{u}(t),$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$

Time $\mapsto$ Frequency

Integro system

$$\mathbf{G}(s) = \mathbf{C}\left(s\mathbf{E} - \mathbf{A} - \frac{1}{s}\mathbf{A}_\tau\right)^{-1}\mathbf{B}.$$

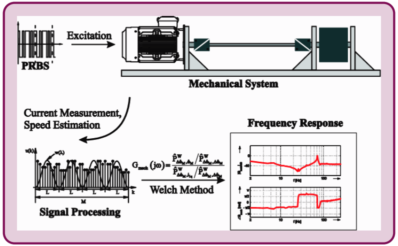
- Excite the system



- Very useful when system parameters are not known.

Introduction
How to measure the transfer function

- Excite the system



- Very useful when system parameters are not known.

- Modeling is done using a proprietary software

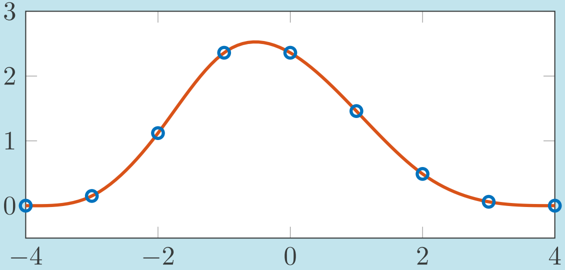
  $\rightsquigarrow$ not so easy to get system matrices

  However, we can obtain transfer function evaluation much easier

Goal of the talk

Build a linear model $\mathcal{M}$ such that

(a) it interpolates given transfer function measurements, i.e., $\mathbf{H}_{\mathcal{M}}(j\omega_i) = v_i$,

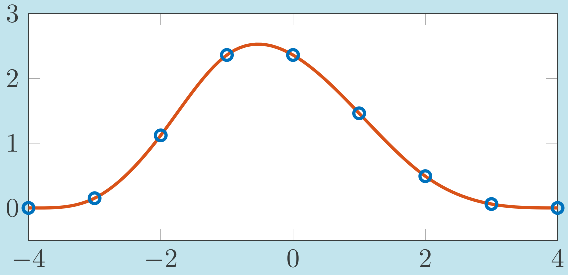


Goal of the talk

Build a linear model $\mathcal{M}$ such that

(a) it interpolates given transfer function measurements, i.e., $\mathbf{H}_{\mathcal{M}}(j\omega_i) = v_i$,

(b) the model has the structure, given by engineering experts, e.g. second-order, delay, fractional, etc.



The Loewner framework for model reduction of large-scale systems
Athanasios Antoulas, Rice University

After reviewing the basics of rational Krylov projections and of the Loewner framework, we will present an explicit generalized eigenvalue decomposition of the Loewner pencil. This brings into the picture the sensitivity of the resulting eigenvalues with respect to the choice of the data. This gives a basis for addressing the issue of "good" choices of data, which has been elusive. Several numerical examples will illustrate these sensitivity issues.

Data-driven Identification

Already talks along these lines

The Loewner framework for model reduction of large-scale systems
Athanasios Antoulas, Rice University

Embedding properties of data-driven dissipative reduced order Models
Vladimir Druskin, WPI

Realizations of reduced order models of passive SISO or MIMO LTE problems can be transformed to tridiagonal and block-tridiagonal forms, respectively, via different modifications of the Lanczos algorithm. Generally, such realizations can be interpreted as ladder resistor-capacitor-inductor (RCL) networks. They gave rise to network syntheses in the first half of the 20th century that was at the base of modern electronics design and consecutively to MOR that tremendously impacted many areas of engineering (electrical, mechanical, aerospace, etc.) by enabling efficient compression of the underlining dynamical systems. In his seminal 1950s works Krein realized that in addition to their compressing properties, network realizations can be used to embed the data back into the state space of the underlying continuum problems.

Data-driven Identification

Already talks along these lines

The Loewner framework for model reduction of large-scale systems
Athanasios Antoulas, Rice University

Af
w
Tl
cl
w

Embedding properties of data-driven dissipative reduced order Models
Vladimir Druskin, WPI

Realiz
to trid
algori
(RCL)
the ba
areas
the ur
their c
state s

The Loewner Framework for Model Reduction of Flow Equations
Matthias Heinkenschloss, Rice University

The Loewner framework is an interpolatory model reduction approach which, in contrast to other approaches, computes a reduced order model (ROM) directly from data. This talk discusses an extension of the Loewner framework to semi-discretizations of fluid flow problems such as Burgers' equation or the Navier-Stokes equations. The extension addresses behavior of the transfer function at infinity, quadratic nonlinearity of the flow equations, and stability of the ROM. Numerical results illustrate the potential of the Loewner framework, but also expose additional issues that need to be addressed to make it fully applicable. Possible approaches to deal with some of these issues are outlined.

Rational interpolation problem

Given interpolation points $\{\sigma_1, \ldots, \sigma_{2l}\} \subset \mathbb{C}$ and sample values $\{\gamma_1, \ldots, \gamma_{2l}\} \subset \mathbb{C}$, **construct** a rational function $\mathbf{H}(s) = \mathbf{C}\left(s\mathbf{E} - \mathbf{A}\right)^{-1}\mathbf{B}$, satisfying

$$\mathbf{H}(\sigma_j) = \gamma_j, \; j = 1, \ldots, 2l$$

- Let us recall the Loewner framework in the single-input single-output case.

- Let us recall the Loewner framework in the single-input single-output case.
- Data set as $\begin{cases} \text{interpolation points} \quad \sigma_k \in \mathbb{C}, \\ \text{sample values} \quad \gamma_k \in \mathbb{C}, \end{cases}$ for $k = 1, \ldots, 2l$.

- Let us recall the Loewner framework in the single-input single-output case.
- Data set as $\begin{cases} \text{interpolation points} & \sigma_k \in \mathbb{C}, \\ \text{sample values} & \gamma_k \in \mathbb{C}, \end{cases}$ for $k = 1, \ldots, 2l$.
- Partition the data into the left & right sets:

$$\{(\sigma_k, \gamma_k)\} = \{(\mu_i, \mathbf{v}_i) \cup (\lambda_i, \mathbf{w}_i)\}, \quad k = 1, \ldots, 2l, i = 1, \ldots, l.$$

- Let us recall the Loewner framework in the single-input single-output case.
- Data set as $\begin{cases} \text{interpolation points} & \sigma_k \in \mathbb{C}, \\ \text{sample values} & \gamma_k \in \mathbb{C}, \end{cases}$ for $k = 1, \ldots, 2l$.
- Partition the data into the left & right sets:

$$\{(\sigma_k, \gamma_k)\} = \{(\mu_i, \mathbf{v}_i) \cup (\lambda_i, \mathbf{w}_i)\}, \quad k = 1, \ldots, 2l, i = 1, \ldots, l.$$

Objective

Find a rational function $\mathbf{H}(s) = \mathbf{C}\left(s\mathbf{E} - \mathbf{A}\right)^{-1}\mathbf{B}$ such that

$$\mathbf{H}(\mu_j) = \mathbf{v}_j, \quad \mathbf{H}(\lambda_i) = \mathbf{w}_i.$$

Data-driven Identification
Loewner framework

- Let us recall the Loewner framework in the single-input single-output case.
- Data set as $\begin{cases} \text{interpolation points} & \sigma_k \in \mathbb{C}, \\ \text{sample values} & \gamma_k \in \mathbb{C}, \end{cases}$ for $k = 1, \ldots, 2l$.
- Partition the data into the left & right sets:

$$\{(\sigma_k, \gamma_k)\} = \{(\mu_i, \mathbf{v}_i) \cup (\lambda_i, \mathbf{w}_i)\}, \quad k = 1, \ldots, 2l, i = 1, \ldots, l.$$

Objective

Find a rational function $\mathbf{H}(s) = \mathbf{C}\left(s\mathbf{E} - \mathbf{A}\right)^{-1}\mathbf{B}$ such that

$$\mathbf{H}(\mu_j) = \mathbf{v}_j, \quad \mathbf{H}(\lambda_i) = \mathbf{w}_i.$$

Let us organize the data as follows:

$$\text{Interpolation points}: \quad \Lambda = \operatorname{diag}\left(\lambda_1, \ldots, \lambda_l\right), \qquad \Omega = \operatorname{diag}\left(\mu_1, \ldots, \mu_l\right),$$

$$\text{Sample values}: \quad \mathbf{V} = \left[v_1, \ldots, v_l\right]^T, \qquad \mathbf{W} = \left[w_1, \ldots, w_l\right]^T.$$

Loewner Approach (Matrix form)

- Let $\mathbb{L}$ and $\mathbb{L}_\sigma$ satisfy:

$$\boxed{\begin{aligned} -\mathbb{L}\Lambda + \mathbb{L}_\sigma &= \mathbf{V}1^T, \\ -\mathbb{L}^T\Omega + {\mathbb{L}_\sigma}^T &= \mathbf{W}1^T, \end{aligned}}$$

- The rational function $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$ interpolates the data, where

$$\boxed{\mathbf{E} = -\mathbb{L}, \quad \mathbf{A} = -\mathbb{L}_\sigma, \quad \mathbf{B} = \mathbf{V}, \quad \text{and} \quad \mathbf{C} = \mathbf{W},}$$

and the pencil $(\mathbb{L}, \mathbb{L}_\sigma)$ is regular.

Loewner Approach (Matrix form)

- Let $\mathbb{L}$ and $\mathbb{L}_\sigma$ satisfy:

$$\boxed{\begin{aligned} -\mathbb{L}\Lambda + \mathbb{L}_\sigma &= \mathbf{V}\mathbb{1}^T, \\ -\mathbb{L}^T\Omega + \mathbb{L}_\sigma{}^T &= \mathbf{W}\mathbb{1}^T, \end{aligned}}$$

- The rational function $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$ interpolates the data, where

$$\boxed{\mathbf{E} = -\mathbb{L}, \quad \mathbf{A} = -\mathbb{L}_\sigma, \quad \mathbf{B} = \mathbf{V}, \quad \text{and} \quad \mathbf{C} = \mathbf{W},}$$

  and the pencil $(\mathbb{L}, \mathbb{L}_\sigma)$ is regular.

Remarks

- No need to solve Sylvester equations $\Rightarrow$ matrices $\mathbb{L}$ and $\mathbb{L}_\sigma$ have analytic expressions.

Loewner Approach (Matrix form)

- Let $\mathbb{L}$ and $\mathbb{L}_\sigma$ satisfy:

$$\mathbb{L} = \begin{pmatrix} \frac{\mathbf{v}_1 - \mathbf{w}_1}{\mu_1 - \lambda_1} & \cdots & \frac{\mathbf{v}_1 - \mathbf{w}_l}{\mu_1 - \lambda_l} \\ \vdots & \ddots & \vdots \\ \frac{\mathbf{v}_l - \mathbf{w}_1}{\mu_l - \lambda_1} & \cdots & \frac{\mathbf{v}_l - \mathbf{w}_l}{\mu_l - \lambda_l} \end{pmatrix} \mathbb{L}_\sigma = \begin{pmatrix} \frac{\mu_1 \mathbf{v}_1 - \lambda_1 \mathbf{w}_1}{\mu_1 - \lambda_1} & \cdots & \frac{\mu_1 \mathbf{v}_1 - \lambda_l \mathbf{w}_l}{\mu_1 - \lambda_l} \\ \vdots & \ddots & \vdots \\ \frac{\mu_l \mathbf{v}_l - \lambda_1 \mathbf{w}_1}{\mu_l - \lambda_1} & \cdots & \frac{\mu_l \mathbf{v}_l - \lambda_l \mathbf{w}_l}{\mu_l - \lambda_l} \end{pmatrix}$$

and the pencil $(\mathbb{L}, \mathbb{L}_\sigma)$ is regular.

Remarks

- No need to solve Sylvester equations $\Rightarrow$ matrices $\mathbb{L}$ and $\mathbb{L}_\sigma$ have analytic expressions.

Loewner Approach (Matrix form)

- Let $\mathbb{L}$ and $\mathbb{L}_\sigma$ satisfy:

$$\boxed{\begin{aligned} -\mathbb{L}\Lambda + \mathbb{L}_\sigma &= \mathbf{V}\mathbb{1}^T, \\ -\mathbb{L}^T\Omega + {\mathbb{L}_\sigma}^T &= \mathbf{W}\mathbb{1}^T, \end{aligned}}$$

- The rational function $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$ interpolates the data, where

$$\boxed{\mathbf{E} = -\mathbb{L}, \quad \mathbf{A} = -\mathbb{L}_\sigma, \quad \mathbf{B} = \mathbf{V}, \quad \text{and} \quad \mathbf{C} = \mathbf{W},}$$

  and the pencil $(\mathbb{L}, \mathbb{L}_\sigma)$ is regular.

Remarks

- No need to solve Sylvester equations $\Rightarrow$ matrices $\mathbb{L}$ and $\mathbb{L}_\sigma$ have analytic expressions.
- $\mathrm{rank}\,(\mathbb{L}) =$ order of minimal realization $= r$.

Loewner Approach (Matrix form)

- Let $\mathbb{L}$ and $\mathbb{L}_\sigma$ satisfy:

$$\boxed{\begin{aligned} -\mathbb{L}\Lambda + \mathbb{L}_\sigma &= \mathbf{V}\mathbb{1}^T, \\ -\mathbb{L}^T\Omega + \mathbb{L}_\sigma{}^T &= \mathbf{W}\mathbb{1}^T, \end{aligned}}$$

- The rational function $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$ interpolates the data, where

$$\boxed{\mathbf{E} = -\mathbb{L}, \quad \mathbf{A} = -\mathbb{L}_\sigma, \quad \mathbf{B} = \mathbf{V}, \quad \text{and} \quad \mathbf{C} = \mathbf{W},}$$

  and the pencil $(\mathbb{L}, \mathbb{L}_\sigma)$ is regular.

Remarks

- No need to solve Sylvester equations $\Rightarrow$ matrices $\mathbb{L}$ and $\mathbb{L}_\sigma$ have analytic expressions.
- $\operatorname{rank}(\mathbb{L})$ = order of minimal realization $= r$.
- Hence, a compression step using SVD of $\mathbb{L}$ and $\mathbb{L}_\sigma$ can be performed to obtain a minimal or approximate.

Objective: rational functions

Find a rational function $\mathbf{H}(s) = \mathbf{C}\left(s\mathbf{E} - \mathbf{A}\right)^{-1}\mathbf{B}$ such that

$$\mathbf{H}(\mu_j) = \mathbf{v}_j, \quad \mathbf{H}(\lambda_i) = \mathbf{w}_i.$$

Objective: structured (non-)rational functions

Find a (non-)rational function $\mathbf{H}_{\text{nr}}(s) = \mathbf{C}\left(f_1(s)\mathbf{A}_1 + f_2(s)\mathbf{A}_2 + f_3(s)\mathbf{A}_3\right)^{-1}\mathbf{B}$ such that

$$\mathbf{H}_{\text{nr}}(\mu_j) = \mathbf{v}_j, \quad \mathbf{H}_{\text{nr}}(\lambda_i) = \mathbf{w}_i.$$

Objective: rational functions

Find a rational function $\mathbf{H}(s) = \mathbf{C}\left(s\mathbf{E} - \mathbf{A}\right)^{-1}\mathbf{B}$ such that

$$\mathbf{H}(\mu_j) = \mathbf{v}_j, \quad \mathbf{H}(\lambda_i) = \mathbf{w}_i.$$

Objective: structured (non-)rational functions

Find a (non-)rational function $\mathbf{H}_{\text{nr}}(s) = \mathbf{C}\left(f_1(s)\mathbf{A}_1 + f_2(s)\mathbf{A}_2 + f_3(s)\mathbf{A}_3\right)^{-1}\mathbf{B}$ such that

$$\mathbf{H}_{\text{nr}}(\mu_j) = \mathbf{v}_j, \quad \mathbf{H}_{\text{nr}}(\lambda_i) = \mathbf{w}_i.$$

Fit into second-order systems, delay systems, fractional systems, etc.

Objective: rational functions

Find a rational function $\mathbf{H}(s) = \mathbf{C}\left(s\mathbf{E} - \mathbf{A}\right)^{-1}\mathbf{B}$ such that

$$\mathbf{H}(\mu_j) = \mathbf{v}_j, \quad \mathbf{H}(\lambda_i) = \mathbf{w}_i.$$

Objective: structured (non-)rational functions

Find a (non-)rational function $\mathbf{H}_{\mathrm{nr}}(s) = \mathbf{C}\left(f_1(s)\mathbf{A}_1 + f_2(s)\mathbf{A}_2 + f_3(s)\mathbf{A}_3\right)^{-1}\mathbf{B}$ such that

$$\mathbf{H}_{\mathrm{nr}}(\mu_j) = \mathbf{v}_j, \quad \mathbf{H}_{\mathrm{nr}}(\lambda_i) = \mathbf{w}_i.$$

Fit into second-order systems, delay systems, fractional systems, etc.
For instance, second-order systems: $f_1(s) = s^2$, $f_2(s) = s$, and $f_3(s) = 1$.

Objective: rational functions

Find a rational function $\mathbf{H}(s) = \mathbf{C}\left(s\mathbf{E} - \mathbf{A}\right)^{-1}\mathbf{B}$ such that

$$\mathbf{H}(\mu_j) = \mathbf{v}_j, \quad \mathbf{H}(\lambda_i) = \mathbf{w}_i.$$

Objective: structured (non-)rational functions

Find a (non-)rational function $\mathbf{H}_{\text{nr}}(s) = \mathbf{C}\left(f_1(s)\mathbf{A}_1 + f_2(s)\mathbf{A}_2 + f_3(s)\mathbf{A}_3\right)^{-1}\mathbf{B}$ such that

$$\mathbf{H}_{\text{nr}}(\mu_j) = \mathbf{v}_j, \quad \mathbf{H}_{\text{nr}}(\lambda_i) = \mathbf{w}_i.$$

Fit into second-order systems, delay systems, fractional systems, etc.
Once again, we organize the data as follows:

$$\text{interpolation points}: \quad \Lambda = \operatorname{diag}\left(\lambda_1, \ldots, \lambda_l\right), \qquad \Omega = \operatorname{diag}\left(\mu_1, \ldots, \mu_l\right),$$

$$\text{Sample values}: \quad \mathbf{V} = \left[\mathbf{v}_1, \ldots, \mathbf{v}_l\right]^T, \qquad \mathbf{W} = \left[\mathbf{w}_1, \ldots, \mathbf{w}_l\right]^T.$$

Identification of structured systems

[Unger/Schulze/Beattie/Gugercin '16]

- Let us say the matrices $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$ satisfy:

$$\boxed{\begin{aligned} \mathbf{A}_1\mathbf{F}_1^{\Lambda} + \mathbf{A}_2\mathbf{F}_2^{\Lambda} + \mathbf{A}_3\mathbf{F}_3^{\Lambda} &= \mathbf{V}1^T, \\ \mathbf{A}_1{}^T\mathbf{F}_1^{\Omega} + \mathbf{A}_2{}^T\mathbf{F}_2^{\Omega} + \mathbf{A}_3{}^T\mathbf{F}_3^{\Omega} &= \mathbf{W}1^T, \end{aligned}}$$

where $\mathbf{F}_i^{\Lambda} = \mathrm{diag}\left(f_i(\lambda_1), \ldots, f_i(\lambda_l)\right)$, $\mathbf{F}_i^{\Omega} = \mathrm{diag}\left(f_i(\mu_1), \ldots, f_i(\mu_l)\right)$, and $\mathbf{V}$ and $\mathbf{W}$ are vectors, containing measurements.

Identification of structured systems

[Unger/Schulze/Beattie/Gugercin '16]

- Let us say the matrices $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$ satisfy:

$$\boxed{\begin{aligned} \mathbf{A}_1\mathbf{F}_1^{\Lambda} + \mathbf{A}_2\mathbf{F}_2^{\Lambda} + \mathbf{A}_3\mathbf{F}_3^{\Lambda} &= \mathbf{V}1^T, \\ \mathbf{A}_1{}^T\mathbf{F}_1^{\Omega} + \mathbf{A}_2{}^T\mathbf{F}_2^{\Omega} + \mathbf{A}_3{}^T\mathbf{F}_3^{\Omega} &= \mathbf{W}1^T, \end{aligned}}$$

where $\mathbf{F}_i^{\Lambda} = \operatorname{diag}\left(f_i(\lambda_1), \ldots, f_i(\lambda_l)\right)$, $\mathbf{F}_i^{\Omega} = \operatorname{diag}\left(f_i(\mu_1), \ldots, f_i(\mu_l)\right)$, and $\mathbf{V}$ and $\mathbf{W}$ are vectors, containing measurements.

- The function $\mathbf{H}_{\mathrm{nr}}(s) = \mathbf{W}^T(f_1(s)\mathbf{A}_1 + f_2(s)\mathbf{A}_2 + f_3(s)\mathbf{A}_3)^{-1}\mathbf{V}$ interpolates the data, i.e.,

$$\mathbf{H}_{\mathrm{nr}}(\lambda_i) = \mathbf{w}_i, \quad \mathbf{H}_{\mathrm{nr}}(\mu_i) = \mathbf{v}_i,$$

assuming $[\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3]$ is of row full-rank. If it is not full-rank, a compression step can be performed.

Realizing
$\mathbf{H} = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$

Realizing
$\mathbf{H} = \mathbf{C}(f_1(s)\mathbf{A}_1 + f_2(s)\mathbf{A}_2 + f_3(s)\mathbf{A}_3)^{-1}\mathbf{B}$

Realizing
$\mathbf{H} = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$

- Number of equations: $2l^2$

Realizing
$\mathbf{H} = \mathbf{C}(f_1(s)\mathbf{A}_1 + f_2(s)\mathbf{A}_2 + f_3(s)\mathbf{A}_3)^{-1}\mathbf{B}$

- Number of equations: $2l^2$

Realizing
$\mathbf{H} = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$

- Number of equations: $2l^2$
- Number of variables: $2l^2$

Realizing
$\mathbf{H} = \mathbf{C}(f_1(s)\mathbf{A}_1 + f_2(s)\mathbf{A}_2 + f_3(s)\mathbf{A}_3)^{-1}\mathbf{B}$

- Number of equations: $2l^2$
- Number of variables: $3l^2$

Realizing
$\mathbf{H} = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$

- Number of equations: $2l^2$
- Number of variables: $2l^2$
- Unique solution

Realizing
$\mathbf{H} = \mathbf{C}(f_1(s)\mathbf{A}_1 + f_2(s)\mathbf{A}_2 + f_3(s)\mathbf{A}_3)^{-1}\mathbf{B}$

- Number of equations: $2l^2$
- Number of variables: $3l^2$
- Infinitely many solutions

Remark: [Unger '16] have tried to enforce additional constraints/conditions to utilize extra variables.

Realizing $\mathbf{H} = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$	Realizing $\mathbf{H} = \mathbf{C}(f_1(s)\mathbf{A}_1 + f_2(s)\mathbf{A}_2 + f_3(s)\mathbf{A}_3)^{-1}\mathbf{B}$
• Number of equations: $2l^2$ • Number of variables: $2l^2$ • Unique solution • Analytical expression (divide and difference)	• Number of equations: $2l^2$ • Number of variables: $3l^2$ • Infinitely many solutions • No such thing

Realizing
$\mathbf{H} = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$

- Number of equations: $2l^2$
- Number of variables: $2l^2$
- Unique solution
- Analytical expression (divide and difference)

Realizing
$\mathbf{H} = \mathbf{C}(f_1(s)\mathbf{A}_1 + f_2(s)\mathbf{A}_2 + f_3(s)\mathbf{A}_3)^{-1}\mathbf{B}$

- Number of equations: $2l^2$
- Number of variables: $3l^2$
- Infinitely many solutions
- No such thing

> The simplest answer is often the right one.
>
> Occam's Razor

Realizing
$\mathbf{H} = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$

- Number of equations: $2l^2$
- Number of variables: $2l^2$
- Unique solution
- Analytical expression (divide and difference)

Realizing
$\mathbf{H} = \mathbf{C}(f_1(s)\mathbf{A}_1 + f_2(s)\mathbf{A}_2 + f_3(s)\mathbf{A}_3)^{-1}\mathbf{B}$

- Number of equations: $2l^2$
- Number of variables: $3l^2$
- Infinitely many solutions
- No such thing

> The simplest answer is often the right one.
>
> Occam's Razor

In dynamical systems, simplicity can be defined as "*minimal order systems, describing the dynamics, or interpolating the data*".

Identification of structured systems [Unger/Schulze/Beattie/Gugercin '16]

- Let us say the matrices $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$ satisfy:

$$\boxed{\begin{aligned} \mathbf{A}_1\mathbf{F}_1^{\Lambda} + \mathbf{A}_2\mathbf{F}_2^{\Lambda} + \mathbf{A}_3\mathbf{F}_3^{\Lambda} &= \mathbf{V}1^T, \\ {\mathbf{A}_1}^T\mathbf{F}_1^{\Omega} + {\mathbf{A}_2}^T\mathbf{F}_2^{\Omega} + {\mathbf{A}_3}^T\mathbf{F}_3^{\Omega} &= \mathbf{W}1^T, \end{aligned}}$$

where $\mathbf{F}_i^{\Lambda} = \operatorname{diag}\left(f_i(\lambda_1), \ldots, f_i(\lambda_l)\right), \quad \mathbf{F}_i^{\Omega} = \operatorname{diag}\left(f_i(\mu_1), \ldots, f_i(\mu_l)\right),$

A result [Benner/G./Pontes '19]

$\operatorname{rank}\left(\left[\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3\right]\right) =$ minimum order of a realization that interpolates the data.

A rank-minimization Problem Formulation

$$\min_{\mathbf{A}_1,\mathbf{A}_2,\mathbf{A}_3} \operatorname{rank}\left(\left[\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3\right]\right)$$

subject to

$$\boxed{\begin{aligned} \mathbf{A}_1\mathbf{F}_1^{\Lambda} + \mathbf{A}_2\mathbf{F}_2^{\Lambda} + \mathbf{A}_3\mathbf{F}_3^{\Lambda} &= \mathbf{V}1^T, \\ \mathbf{A}_1{}^T\mathbf{F}_1^{\Omega} + \mathbf{A}_2{}^T\mathbf{F}_2^{\Omega} + \mathbf{A}_3{}^T\mathbf{F}_3^{\Omega} &= \mathbf{W}1^T, \end{aligned}}$$

where $\mathbf{F}_i^{\Lambda} = \operatorname{diag}\left(f_i(\lambda_1), \ldots, f_i(\lambda_l)\right)$, $\quad \mathbf{F}_i^{\Omega} = \operatorname{diag}\left(f_i(\mu_1), \ldots, f_i(\mu_l)\right)$, and $\mathbf{V}$ and $\mathbf{W}$ are vectors, containing measurements.

A delay example

- Consider a delay system whose transfer function is: $\mathbf{H}(s) = (s + 1 - 0.25e^{-s})^{-1}$.
- Take four distinct measurements: $\mathbf{H}(\sigma_1), \mathbf{H}(\sigma_2), \mathbf{H}(\mu_1), \mathbf{H}(\mu_2)$

A delay example

- ~~Consider a delay system whose transfer function:~~ $\cancel{\mathbf{H}(s) = (s + 1 - 0.25e^{-s})^{-1}}$.
- Take four distinct measurements: $\mathbf{H}(\sigma_1), \mathbf{H}(\sigma_2), \mathbf{H}(\mu_1), \mathbf{H}(\mu_2)$

Identification (inverse) problem

Given measurements, identify a delay model that interpolates the measurement.
In other words, construct a state-space model:

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_\tau\mathbf{x}(t-1) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$$

such that $\mathbf{H}_{\mathtt{Iden}}(\sigma_i) = \mathbf{H}(\sigma_i)$ and $\mathbf{H}_{\mathtt{Iden}}(\mu_i) = \mathbf{H}(\mu_i)$, $i \in \{1, 2\}$.

Identification (inverse) problem

Given measurements, identify a delay model that interpolates the measurement.
In other words, construct a state-space model:

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_\tau \mathbf{x}(t-1) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$$

such that $\mathbf{H}_{\texttt{Iden}}(\sigma_i) = \mathbf{H}(\sigma_i)$ and $\mathbf{H}_{\texttt{Iden}}(\mu_i) = \mathbf{H}(\mu_i)$, $i \in \{1,2\}$.

Necessary Conditions for interpolation [Unger et. al '16]

$$\mathbf{E}\begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix} + \mathbf{A}\begin{bmatrix} 1 & \\ & 1 \end{bmatrix} + \mathbf{A}_\tau \begin{bmatrix} e^{-\sigma_1} & \\ & e^{-\sigma_2} \end{bmatrix} = \begin{bmatrix} \mathbf{H}(\mu_1) \\ \mathbf{H}(\mu_2) \end{bmatrix} 1^T =: \mathbf{B}1^T,$$

$$\mathbf{E}^T\begin{bmatrix} \mu_1 & \\ & \mu_2 \end{bmatrix} + \mathbf{A}^T\begin{bmatrix} 1 & \\ & 1 \end{bmatrix} + \mathbf{A}_\tau^T \begin{bmatrix} e^{-\mu_1} & \\ & e^{-\mu_2} \end{bmatrix} = \begin{bmatrix} \mathbf{H}(\sigma_1) \\ \mathbf{H}(\sigma_2) \end{bmatrix} 1^T =: \mathbf{C}^T1^T,$$

- Every triplet $(\mathbf{E}, \mathbf{A}, \mathbf{A}_\tau)$, satisfying the above equations, interpolates the data.
- Infinite possibility since $3 \cdot 4$ variables and $2 \cdot 4$ equations.

Identification (inverse) problem

Given measurements, identify a delay model that interpolates the measurement.
In other words, construct a state-space model:

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_\tau \mathbf{x}(t-1) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$$

such that $\mathbf{H}_{\mathtt{Iden}}(\sigma_i) = \mathbf{H}(\sigma_i)$ and $\mathbf{H}_{\mathtt{Iden}}(\mu_i) = \mathbf{H}(\mu_i)$, $i \in \{1, 2\}$.

Necessary Conditions for interpolation [Unger et. al '16]

$$\mathbf{E}\begin{bmatrix}\sigma_1 & \\ & \sigma_2\end{bmatrix} + \mathbf{A}\begin{bmatrix}1 & \\ & 1\end{bmatrix} + \mathbf{A}_\tau\begin{bmatrix}e^{-\sigma_1} & \\ & e^{-\sigma_2}\end{bmatrix} = \begin{bmatrix}\mathbf{H}(\mu_1) \\ \mathbf{H}(\mu_2)\end{bmatrix} 1^T =: \mathbf{B}1^T,$$

$$\mathbf{E}^T\begin{bmatrix}\mu_1 & \\ & \mu_2\end{bmatrix} + \mathbf{A}^T\begin{bmatrix}1 & \\ & 1\end{bmatrix} + \mathbf{A}_\tau^T\begin{bmatrix}e^{-\mu_1} & \\ & e^{-\mu_2}\end{bmatrix} = \begin{bmatrix}\mathbf{H}(\sigma_1) \\ \mathbf{H}(\sigma_2)\end{bmatrix} 1^T =: \mathbf{C}^T1^T,$$

- Every triplet $(\mathbf{E}, \mathbf{A}, \mathbf{A}_\tau)$, satisfying the above equations, interpolates the data.
- Infinite possibility since $3 \cdot 4$ variables and $2 \cdot 4$ equations.
- For $\mathbf{A}_\tau = 0$, it yields a rational function, obtained by the Loewner approach, of order $r = 2$.

Rank Minimization Problems
Using a rank-minimization approach

- We seek among infinite triplets that minimize as follows:

$$\operatorname{rank}\left(\left[\mathbf{E}, \mathbf{A}, \mathbf{A}_\tau\right]\right),$$

satisfying

$$\mathbf{E}\begin{bmatrix}\sigma_1 & \\ & \sigma_2\end{bmatrix} + \mathbf{A}\begin{bmatrix}1 & \\ & 1\end{bmatrix} + \mathbf{A}_\tau\begin{bmatrix}e^{-\sigma_1} & \\ & e^{-\sigma_2}\end{bmatrix} = \begin{bmatrix}\mathbf{H}(\mu_1)\\ \mathbf{H}(\mu_2)\end{bmatrix}1^T =: \mathbf{B}1^T,$$

$$\mathbf{E}^T\begin{bmatrix}\mu_1 & \\ & \mu_2\end{bmatrix} + \mathbf{A}^T\begin{bmatrix}1 & \\ & 1\end{bmatrix} + \mathbf{A}_\tau^T\begin{bmatrix}e^{-\mu_1} & \\ & e^{-\mu_2}\end{bmatrix} = \begin{bmatrix}\mathbf{H}(\sigma_1)\\ \mathbf{H}(\sigma_2)\end{bmatrix}1^T =: \mathbf{C}^T1^T.$$

- We seek among infinite triplets that minimize as follows:

$$\operatorname{rank}\left(\left[\mathbf{E}, \mathbf{A}, \mathbf{A}_\tau\right]\right),$$

- If we solve, then we get

$$\mathbf{E} = \begin{bmatrix} \mathbf{H}(\mu_1) & \\ & \mathbf{H}(\mu_2) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{H}(\sigma_1) & \\ & \mathbf{H}(\sigma_2) \end{bmatrix},$$

$$\mathbf{A} = -\begin{bmatrix} \mathbf{H}(\mu_1) & \\ & \mathbf{H}(\mu_2) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{H}(\sigma_1) & \\ & \mathbf{H}(\sigma_2) \end{bmatrix},$$

$$\mathbf{A}_\tau = 0.25 \begin{bmatrix} \mathbf{H}(\mu_1) & \\ & \mathbf{H}(\mu_2) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{H}(\sigma_1) & \\ & \mathbf{H}(\sigma_2) \end{bmatrix}.$$

- We seek among infinite triplets that minimize as follows:

$$\operatorname{rank}\left([\mathbf{E}, \mathbf{A}, \mathbf{A}_\tau]\right),$$

- If we solve, then we get

$$\begin{aligned}
\mathbf{E} &= \begin{bmatrix} \mathbf{H}(\mu_1) & \\ & \mathbf{H}(\mu_2) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{H}(\sigma_1) & \\ & \mathbf{H}(\sigma_2) \end{bmatrix}, \\
\mathbf{A} &= -\begin{bmatrix} \mathbf{H}(\mu_1) & \\ & \mathbf{H}(\mu_2) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{H}(\sigma_1) & \\ & \mathbf{H}(\sigma_2) \end{bmatrix}, \\
\mathbf{A}_\tau &= 0.25 \begin{bmatrix} \mathbf{H}(\mu_1) & \\ & \mathbf{H}(\mu_2) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{H}(\sigma_1) & \\ & \mathbf{H}(\sigma_2) \end{bmatrix}.
\end{aligned}$$

- Note that the rank of $[\mathbf{E}, \mathbf{A}, \mathbf{A}_\tau] = 1$.

- We seek among infinite triplets that minimize as follows:

$$\operatorname{rank}\left(\left[\mathbf{E}, \mathbf{A}, \mathbf{A}_\tau\right]\right),$$

- If we solve, then we get

$$\begin{aligned}
\mathbf{E} &= \begin{bmatrix} \mathbf{H}(\mu_1) & \\ & \mathbf{H}(\mu_2) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{H}(\sigma_1) & \\ & \mathbf{H}(\sigma_2) \end{bmatrix}, \\
\mathbf{A} &= -\begin{bmatrix} \mathbf{H}(\mu_1) & \\ & \mathbf{H}(\mu_2) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{H}(\sigma_1) & \\ & \mathbf{H}(\sigma_2) \end{bmatrix}, \\
\mathbf{A}_\tau &= 0.25 \begin{bmatrix} \mathbf{H}(\mu_1) & \\ & \mathbf{H}(\mu_2) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{H}(\sigma_1) & \\ & \mathbf{H}(\sigma_2) \end{bmatrix}.
\end{aligned}$$

- Note that the rank of $\left[\mathbf{E}, \mathbf{A}, \mathbf{A}_\tau\right] = 1$.
- Hence, using a compression step, we can obtain the same transfer function.

- We have set-up a possible ideal problem to identify low-order models.

- We have set-up a possible ideal problem to identify low-order models.
- But rank-minimization problems are NP hard ones, and there is little hope of finding the global minimum efficiently in all instances.

- We have set-up a possible ideal problem to identify low-order models.
- But rank-minimization problems are NP hard ones, and there is little hope of finding the global minimum efficiently in all instances.

A general rank-minimization formulation

$$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X}) \quad \text{subject to} \quad \mathcal{A}\operatorname{vec}(\mathbf{X}) = \mathbf{b}$$

Rank Minimization Problems

Challenges

- We have set-up a possible ideal problem to identify low-order models.
- But rank-minimization problems are NP hard ones, and there is little hope of finding the global minimum efficiently in all instances.

A general rank-minimization formulation

$$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X}) \quad \text{subject to} \quad \mathcal{A}\operatorname{vec}(\mathbf{X}) = \mathbf{b}$$

A rank-minimization Problem Formulation

$$\min_{\mathbf{A}_1,\mathbf{A}_2,\mathbf{A}_3} \operatorname{rank}\left(\left[\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3\right]\right)$$

subject to

$$\boxed{\begin{aligned} \mathbf{A}_1\mathbf{F}_1^{\Lambda} + \mathbf{A}_2\mathbf{F}_2^{\Lambda} + \mathbf{A}_3\mathbf{F}_3^{\Lambda} &= \mathbf{V}1^T, \\ \mathbf{A}_1{}^T\mathbf{F}_1^{\Omega} + \mathbf{A}_2{}^T\mathbf{F}_2^{\Omega} + \mathbf{A}_3{}^T\mathbf{F}_3^{\Omega} &= \mathbf{W}1^T, \end{aligned}}$$

$$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X}) \quad \text{subject to} \quad \mathcal{A}\operatorname{vec}(X) = \mathbf{b} \tag{2}$$

$$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X}) \quad \text{subject to} \quad \mathcal{A}\operatorname{vec}(X) = \mathbf{b} \tag{2}$$

- Observe: $\operatorname{rank}(\mathbf{X}) = \|\sigma(\mathbf{X})\|_{l_0}$, where $\sigma(\mathbf{X}) = [\sigma_1, \ldots, \sigma_n]$.

$$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X}) \quad \text{subject to} \quad \mathcal{A}\operatorname{vec}(X) = \mathbf{b} \tag{2}$$

Relaxed: $\min_{\mathbf{X}} \sum_i g(\sigma_i) \quad \text{subject to} \quad \mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

- Observe: $\operatorname{rank}(\mathbf{X}) = \|\sigma(\mathbf{X})\|_{l_0}$, where $\sigma(\mathbf{X}) = \left[\sigma_1, \ldots, \sigma_n\right]$.
- What we are going to look at, instead, efficient heuristics.

$$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X}) \quad \text{subject to} \quad \mathcal{A}\operatorname{vec}(X) = \mathbf{b} \tag{2}$$

Relaxed: $\min_{\mathbf{X}} \sum_i g(\sigma_i)$ subject to $\mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

- Observe: $\operatorname{rank}(\mathbf{X}) = \|\sigma(\mathbf{X})\|_{l_0}$, where $\sigma(\mathbf{X}) = [\sigma_1, \ldots, \sigma_n]$.
- What we are going to look at, instead, efficient heuristics.
  - $\|\sigma(\mathbf{X})\|_{l_0} \to \|\sigma(\mathbf{X})\|_{l_1} := \|\mathbf{X}\|_*$ (nuclear norm of $\mathbf{X}$).
    – the best convex relaxation. [Fazel '02]

$$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X}) \quad \text{subject to} \quad \mathcal{A}\operatorname{vec}(X) = \mathbf{b} \tag{2}$$

Relaxed: $\min_{\mathbf{X}} \sum_i g(\sigma_i)$ subject to $\mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

- Observe: $\operatorname{rank}(\mathbf{X}) = \|\sigma(\mathbf{X})\|_{l_0}$, where $\sigma(\mathbf{X}) = [\sigma_1, \ldots, \sigma_n]$.
- What we are going to look at, instead, efficient heuristics.
  - $\|\sigma(\mathbf{X})\|_{l_0} \rightarrow \|\sigma(\mathbf{X})\|_{l_1} := \|\mathbf{X}\|_*$ (nuclear norm of $\mathbf{X}$).
    – the best convex relaxation. [Fazel '02]
  - $\|\sigma(\mathbf{X})\|_{l_0} \rightarrow \sum_i (\sigma_i)^p$
    – concave function but better approximation of cardinality.

An ideal problem

$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X})$ subject to $\mathcal{A} \operatorname{vec}(X) = \mathbf{b}$

An ideal problem

$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X})$ subject to $\mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

Relaxation

$\min_{\mathbf{X}} \sum_i g(\sigma_i)$ subject to $\mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

An ideal problem

$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X})$ subject to $\mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

Relaxation

$\min_{\mathbf{X}} \sum_i g(\sigma_i)$ subject to $\mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

Shrinkage Operator

Let $\mathbf{M}$ be a matrix and its SVD be $\mathbf{M} = \mathbf{U}\Sigma\mathbf{V}^*$ with $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$. The shrinkage operator $\mathcal{D}_\tau$ is defined as

$$\mathcal{D}_\tau(\mathbf{M}) = \mathbf{U}\mathcal{D}_\tau(\Sigma)\mathbf{V}^*, \quad \mathcal{D}_\tau(\Sigma) = \operatorname{diag}\left((\sigma_1 - \tau)_+, \ldots, (\sigma_n - \tau)_+\right),$$

where $t_+ = \max(t, 0)$.

Rank Minimization Problems

Singular value thresholding

An ideal problem

$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X})$ subject to $\mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

Relaxation

$\min_{\mathbf{X}} \sum_i g(\sigma_i)$ subject to $\mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

If $g(\sigma_i) = \sigma_i$:

$\min_{\mathbf{X}} \|\mathbf{X}\|_*$ subject to $\mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

Shrinkage Operator

Let $\mathbf{M}$ be a matrix and its SVD be $\mathbf{M} = \mathbf{U}\Sigma\mathbf{V}^*$ with $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$. The shrinkage operator $\mathcal{D}_\tau$ is defined as

$$\mathcal{D}_\tau(\mathbf{M}) = \mathbf{U}\mathcal{D}_\tau(\Sigma)\mathbf{V}^*, \quad \mathcal{D}_\tau(\Sigma) = \operatorname{diag}\left((\sigma_1 - \tau)_+, \ldots, (\sigma_n - \tau)_+\right),$$

where $t_+ = \max(t, 0)$.

An ideal problem

$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X})$ subject to $\mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

Relaxation

$\min_{\mathbf{X}} \sum_i g(\sigma_i)$ subject to $\mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

If $g(\sigma_i) = \sigma_i$:

$\min_{\mathbf{X}} \|\mathbf{X}\|_*$ subject to $\mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

Uzawa's iterations [Cai/Candes/Shen '10]

Optimal solution is given by

$$\begin{cases} z &= \mathcal{A}^T(y^{k-1}); \\ Z &= \mathtt{reshape}(z, n, 3n) \\ \mathbf{X}^k &= \mathcal{D}_\tau(\mathcal{A}^T(y^{k-1})) \\ y^k &= y^{k-1} + \delta_k\left(b - \mathcal{A}\operatorname{vec}\left(\mathbf{X}^K\right)\right) \end{cases}$$

n : number of data points

Rank Minimization Problems
Singular value thresholding

An ideal problem

$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X}) \quad$ subject to $\quad \mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

Relaxation

$\min_{\mathbf{X}} \sum_i g(\sigma_i) \quad$ subject to $\quad \mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

If $g(\sigma_i) = \sigma_i$:

$\min_{\mathbf{X}} \|\mathbf{X}\|_* \quad$ subject to $\quad \mathcal{A}\operatorname{vec}(X) = \mathbf{b}$

Uzawa's iterations [Cai/Candes/Shen '10]

Optimal solution is given by

$$\begin{cases} z &= \mathcal{A}^T(y^{k-1}); \\ Z &= \texttt{reshape}(z, n, 3n) \\ \mathbf{X}^k &= \mathcal{D}_\tau(\mathcal{A}^T(y^{k-1})) \\ y^k &= y^{k-1} + \delta_k\left(b - \mathcal{A}\operatorname{vec}\left(\mathbf{X}^K\right)\right) \end{cases}$$

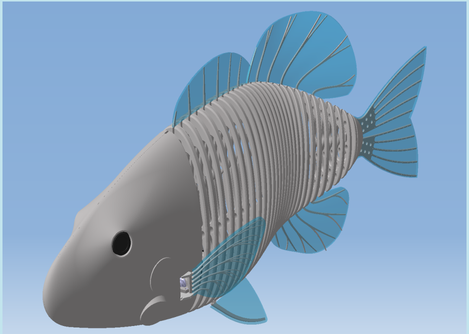
n : number of data points

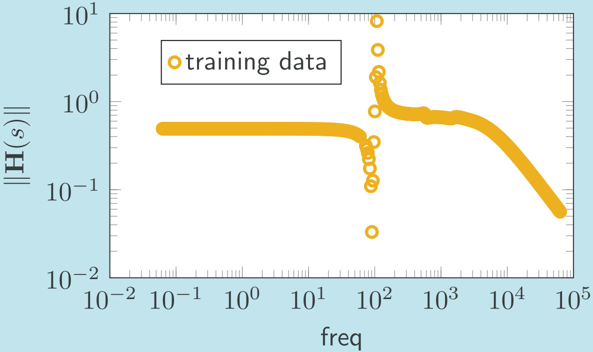
- Note that similar iterations are possible for a concave heuristics of the rank function, i.e., $g(\sigma_i) = \sigma_i^{0.5}$.

A second-order example: Fishtail example ($n = 779, 232$)

- Consider 296 points and look for a minimal order second-order system that fits the data.

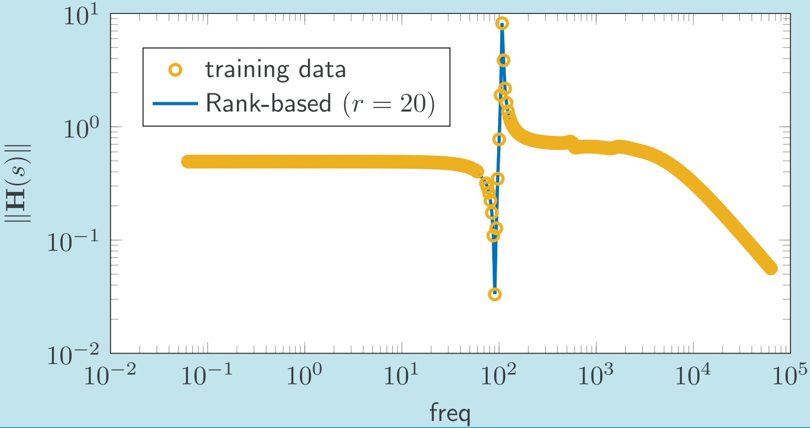
$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$





A second-order example: Fishtail example ($n = 779,232$)

- Consider 296 points and look for a minimal order second-order system that fits the data.



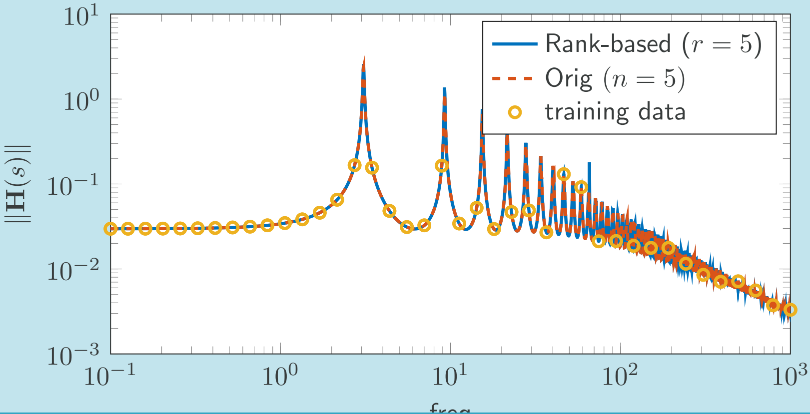
A delay example

- Consider 40 points and look for a minimal order delay system that fits the data.

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_\tau \mathbf{x}(t-1) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$

A delay example

- Consider 40 points and look for a minimal order delay system that fits the data.




NOISE
PROBLEM

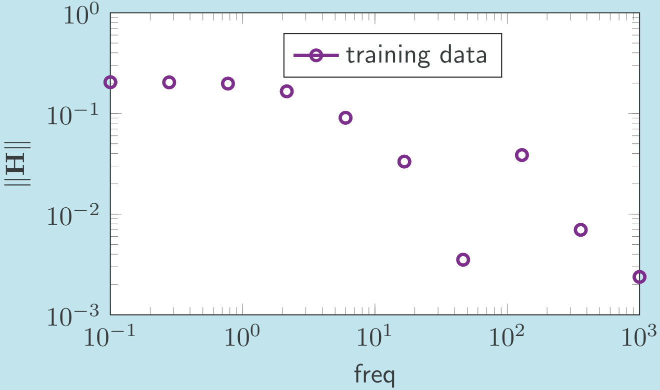
- Data obtained using sensors or in a lab are corrupted with noise.

Measurement Noise
Illustration of Loewner with noise

- Data obtained using sensors or in a lab are corrupted with noise.
- Hence, data-based algorithm should be robust with respect to noise.

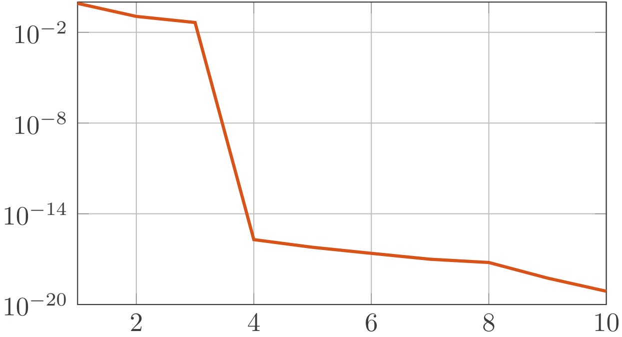
- Data obtained using sensors or in a lab are corrupted with noise.
- Hence, data-based algorithm should be robust with respect to noise.

Consider 10 data points

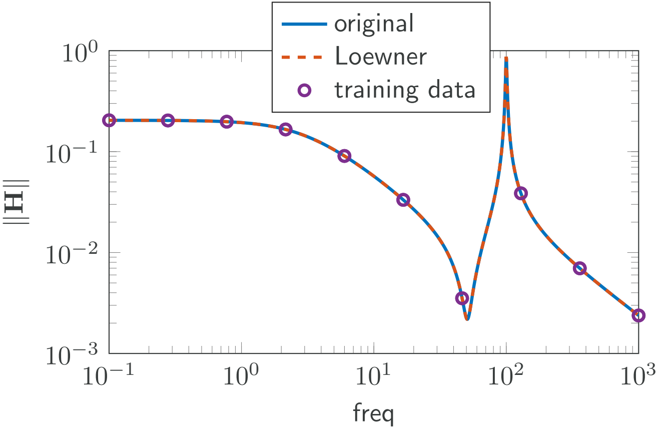


- Let us apply the Loewner method to have a rational function $(\mathbf{E}, \mathbf{A}, \mathbf{B}, \mathbf{C})$.

- Observe the decay of singular values of the Loewner pencil.



- Observe the decay of singular values of the Loewner pencil.
- We obtain a 3rd order system, interpolating the data.



$$\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

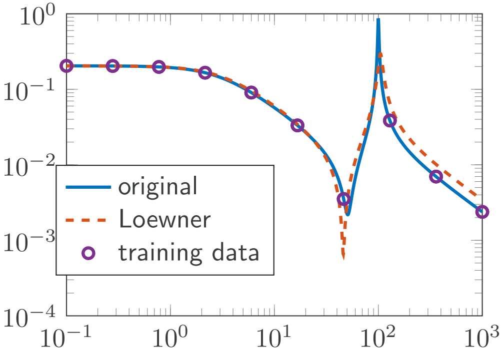
$$\mathbf{A} = \begin{bmatrix} -1 & 100 & 0 \\ -100 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

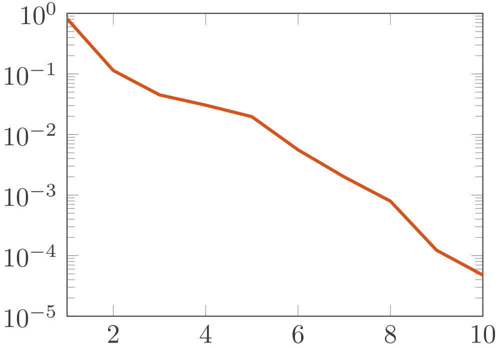
$$\mathbf{B} = \begin{bmatrix} -1.5461 \\ -0.7585 \\ -1.1096 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} -0.8456 \\ -0.5727 \\ -0.5587 \end{bmatrix}.$$

- Add 1% noise in the data and construct $3rd$-order dynamical system using the Loewner method.

- Add 1% noise in the data and construct $3rd$-order dynamical system using the Loewner method.





A rank-minimization Problem Formulation + Noise (for the simplest linear system)

$$\min_{\mathbf{E},\mathbf{A}} \operatorname{rank}\left(\left[\mathbf{E},\mathbf{A}\right]\right) \tag{3}$$

subject to

$$\boxed{\begin{aligned} \|\mathbf{E}\Lambda + \mathbf{A} - \mathbf{V}1^T\|_F \leq \epsilon, \\ \|\mathbf{E}^T\Omega + \mathbf{A}^T - \mathbf{W}1^T\|_F \leq \epsilon, \end{aligned}}$$

where $\Lambda = \operatorname{diag}\left(\lambda_1, \ldots, \lambda_l\right), \quad \Omega = \operatorname{diag}\left(\mu_1, \ldots, \mu_l\right),$

A rank-minimization Problem Formulation + Noise (for the simplest linear system)

$$\min_{\mathbf{E},\mathbf{A}} \operatorname{rank}\left(\left[\mathbf{E}, \mathbf{A}\right]\right) \tag{3}$$

subject to

$$\boxed{\begin{aligned} \|\mathbf{E}\Lambda + \mathbf{A} - \mathbf{V}1^T\|_F \leq \epsilon, \\ \|\mathbf{E}^T\Omega + \mathbf{A}^T - \mathbf{W}1^T\|_F \leq \epsilon, \end{aligned}}$$

where $\Lambda = \operatorname{diag}\left(\lambda_1, \ldots, \lambda_l\right), \quad \Omega = \operatorname{diag}\left(\mu_1, \ldots, \mu_l\right),$

- Intuition: Looking for a minimal-order systems at approximately interpolate data but not exactly.

A rank-minimization Problem Formulation + Noise (for the simplest linear system)

$$\min_{\mathbf{E},\mathbf{A}} \operatorname{rank}\left(\left[\mathbf{E}, \mathbf{A}\right]\right) \tag{3}$$

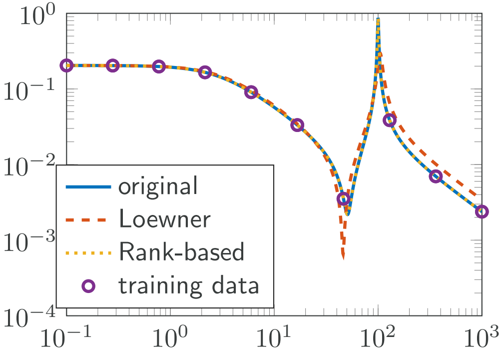
subject to

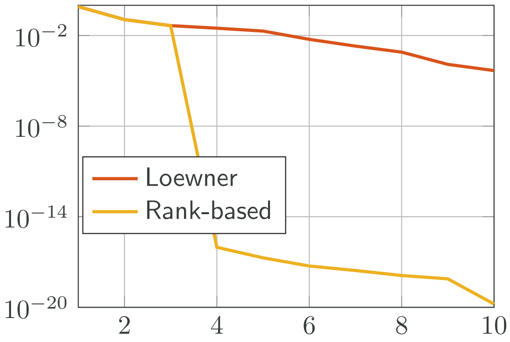
$$\boxed{\begin{aligned} \|\mathbf{E}\Lambda + \mathbf{A} - \mathbf{V}1^T\|_F \leq \epsilon, \\ \|\mathbf{E}^T\Omega + \mathbf{A}^T - \mathbf{W}1^T\|_F \leq \epsilon, \end{aligned}}$$

where $\Lambda = \operatorname{diag}\left(\lambda_1, \ldots, \lambda_l\right), \quad \Omega = \operatorname{diag}\left(\mu_1, \ldots, \mu_l\right),$

- Intuition: Looking for a minimal-order systems at approximately interpolate data but not exactly.
- A similar formulation was proposed in [Fazel/Hindi/Byod '04] in the context of system identification using time-domain noisy data.

Measurement Noise
Rank-based problem formulation





- We could recover the original system.

Conclusions

Contribution of this talk

- Identification of linear systems from frequency response
- Can impose structure obtained using prior engineering knowledge
- An efficient algorithm to solve the optimization problem
- Proof of concepts by mean of numerical examples

Conclusions

Contribution of this talk

- Identification of linear systems from frequency response
- Can impose structure obtained using prior engineering knowledge
- An efficient algorithm to solve the optimization problem
- Proof of concepts by mean of numerical examples

Open questions and future work

- Working with data obtained in a lab
- What if the structure is not known: structure discovery
- Need to do analysis for noisy case

Conclusions

Contribution of this talk

- Identification of linear systems from frequency response
- Can impose structure obtained using prior engineering knowledge
- An efficient algorithm to solve the optimization problem
- Proof of concepts by mean of numerical examples

Thank you for your attention!!

Open questions and future work

- Working with data obtained in a lab
- What if the structure is not known: structure discovery
- Need to do analysis for noisy case