Sep 2 – Sep 15, 2015
Shanghai University, Shanghai, China

Introductory Course on Model Reduction of Linear Time Invariant Systems

Jan Heiland

Max Planck Institute for Dynamics of Complex Technical Systems
Computational Methods in Systems and Control Theory
Magdeburg, Germany

heiland@mpi-magdeburg.mpg.de

Outline

Typical Situation

- Fry a steak
- The cook controls the heat at the fireplace
- and observes the process, e.g. via measuring the temperature in the inner

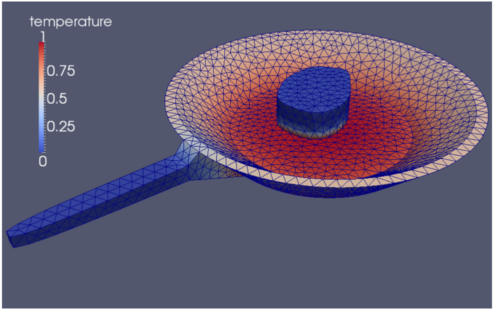
Typical Situation

- The model

$$
\begin{aligned}
\dot\theta &= \nabla \cdot (\nu \nabla \theta) && \text{in } (0,\infty)\times\Omega,\\
\theta &= u, && \text{at the plate,}\\
\theta(0) &= 0.
\end{aligned}
$$

- The cook controls the heat at the fireplace, which we denote by u
- and observes the process, e.g. he measures the temperature y in the center: $y = f(\theta)$.

Simulation



- The model:

$$\begin{aligned} \dot{\theta} &= \nabla \cdot (\nu \nabla \theta), \\ \theta &= u, \\ \theta(0) &= 0. \end{aligned}$$

- The cook controls the heat u
- and observes the process via $y = f(\theta)$.

- A *Finite Element* discretization of the problem leads to the finite dimensional model

$$E\dot{\theta}(t) = A\theta(t) + Bu(t), \quad \theta(0) = 0, \tag{1}$$
$$y(t) = C\theta(t), \tag{2}$$

a linear time invariant system.

Linear State Space System

$$
\begin{aligned}
E\dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = x_0, && \text{(3a)} \\
y(t) &= Cx(t) + Du(t), && \text{(3b)}
\end{aligned}
$$

with

- $E \in \mathbb{R}^{n\times n}$: the identity or the mass matrix
- $A \in \mathbb{R}^{n\times n}$: the system matrix
- $B \in \mathbb{R}^{n\times m}$: the input matrix
- $C \in \mathbb{R}^{q\times n}$: the output matrix
- $D \in \mathbb{R}^{q\times n}$: the throughput

- $x(t) \in \mathbb{R}^{n}$: the system's state
- $u(t) \in \mathbb{R}^{m}$: the input or control
- $y(t) \in \mathbb{R}^{q}$: the output or measurements
- n, m, $q \in \mathbb{N}$: the system dimensions

We will assume that $E = I$ and denote the LTI (3) by (A, B, C, D).

Linear State Space System

$$E\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, \tag{3a}$$
$$y(t) = Cx(t) + Du(t), \tag{3b}$$

with

- $E \in \mathbb{R}^{n\times n}$: the identity or the mass matrix
- $A \in \mathbb{R}^{n\times n}$: the system matrix
- $B \in \mathbb{R}^{n\times m}$: the input matrix
- $C \in \mathbb{R}^{q\times n}$: the output matrix
- $D \in \mathbb{R}^{q\times n}$: the throughput

- $x(t) \in \mathbb{R}^{n}$: the system's state
- $u(t) \in \mathbb{R}^{m}$: the input or control
- $y(t) \in \mathbb{R}^{q}$: the output or measurements
- $n, m, q \in \mathbb{N}$: the system dimensions

We will assume that $E = I$ and denote the LTI (3) by (A, B, C, D).

Linear State Space System

$$
\begin{aligned}
E\dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = x_0, && \text{(3a)}\\
y(t) &= Cx(t) + Du(t), && \text{(3b)}
\end{aligned}
$$

with

- $E \in \mathbb{R}^{n\times n}$: the identity or the mass matrix
- $A \in \mathbb{R}^{n\times n}$: the system matrix
- $B \in \mathbb{R}^{n\times m}$: the input matrix
- $C \in \mathbb{R}^{q\times n}$: the output matrix
- $D \in \mathbb{R}^{q\times n}$: the throughput

- $x(t) \in \mathbb{R}^{n}$: the system's state
- $u(t) \in \mathbb{R}^{m}$: the input or control
- $y(t) \in \mathbb{R}^{q}$: the output or measurements
- $n,\ m,\ q \in \mathbb{N}$: the system dimensions

We will assume that $E = I$ and denote the LTI (3) by (A, B, C, D).

Some Preliminary Thoughts

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t) \end{aligned}$$

A simple question...

What is x?

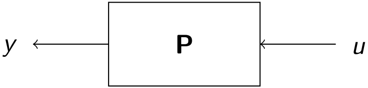
- it is a physical state in the model – like the temperature
- in practise, we may not access it – only the measurement $y = Cx$
- it is but a mathematical object as a part of a model
- furthermore, as we will see later, the state x can be severely changed e.g. in the course of model reduction

The state x can be seen...

...as nothing but an artificial object of the model for the input to output behavior

$$\mathbf{G} \colon u \mapsto y$$

of an abstract system $\mathbf{P}$:



that maps an input u to the corresponding output y.

The state x can be seen...
...as nothing but an artificial object of the model for the input to output behavior

$$\mathbf{G}\colon u \mapsto y$$

of an abstract system $\mathbf{P}$:

$$y \longleftarrow \boxed{\begin{array}{c} y(t) = Cx(t) \\ \scriptstyle \dot{x}(t) = Ax(t) + Bu(t) \end{array}} \longleftarrow u$$

that maps an input u to the corresponding output y.

Transfer Function in Time-Domain

If **P** is modelled trough an (A, B, C, D) system, then the function **G** can be defined via

$$\mathbf{G}\colon u \mapsto y\colon y(t) = C\left[e^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s)\,\mathrm{d}s\right] + Du(t),$$

known as the formula of *variation of constants*.

This is in **time-domain**: A function u depending on time $t \in [0, \infty)$ is mapped onto a function y depending on time $t \in [0, \infty)$.

Introducing Frequency-Domain

Through the Laplace transform $\mathcal{L}$ and its inverse $\mathcal{L}^{-1}$, we can switch between time-domain and frequency-domain representations of the input and output signals:

$$U(s) := \mathcal{L}\{u\}(s) := \int_0^\infty e^{-st} u(t) \, \mathrm{d}t,$$

where $s \in \mathbb{C}$ is the *frequency* and

$$y(t) := \mathcal{L}^{-1}\{Y\}(t) := \lim_{T \to \infty} \frac{1}{2\pi i} \int_{\gamma - iT}^{\gamma + iT} e^{s} Y(s) \, \mathrm{d}s$$

where $\gamma \in \mathbb{R}$ is chosen such that the contour path of the integration is the domain of convergence of Y.

Laplace Transform of an LTI

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t)\\ y(t) &= Cx(t) + Du(t)\end{aligned}$$

With the basic properties of the Laplace transform

- $\dot{X}(s) := \mathcal{L}\{\dot{x}\}(s) - x(0) = s\mathcal{L}\{x\}(s) = sX(s) - x(0)$
- and linearity $\mathcal{L}\{Ax\}(s) = AX(s)$

with zero initial value $x(0) = 0$, the (A, B, C, D) system defines the transfer function

$$G(s) := C(sI - A)^{-1}B + D$$

in frequency domain.

Realizations

Fact

An LTI (A, B, C, D) always defines a transfer function

$$G(s) = C(sI - A)^{-1}B + D$$

which is a matrix $G \in \mathbb{R}^{q \times m}$ with coefficients that are rational functions of s.

Question

Given a rational matrix function $s \mapsto G(s) \in \mathbb{R}^{q \times m}$, is there an

$$(A, B, C, D)$$

system, so that $G(s) = C(sI - A)^{-1}B + D$?

Realizations

given G, find (A, B, C, D),
$G(s) = C(sI - A)^{-1}B + D$

If there is one such (A, B, C, D), then there are infinitely many:

- For $T \in \mathbb{R}^{n\times n}$ invertible, also $(TAT^{-1}, TB, CT^{-1}, D)$ is a realization:

$$C(sI - A)^{-1}B + D = CT^{-1}(sI - TAT^{-1})^{-1}TB + D.$$

- Moreover, also

$$(\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} B \\ 0 \end{bmatrix}, \begin{bmatrix} C & 0 \end{bmatrix}, D)$$

  is a realization of G.

Realizations

Facts and Thoughts on Realizations

- If G is *proper*, then there is a realization (A, B, C, D) as a state space system.
- This realization is by no means unique.
- The dimension of the state can be arbitrary large. What is the smallest possible dimension? (cf. *model reduction*)
- What is a good choice for the state?

Remark: A transfer function $G\colon s \mapsto G(s) \in \mathbb{R}^{q\times m}$ with coefficients that are rational functions in s, is *proper*, if in each coefficient the polynomial degree of the numerators does not exceed the degree of denominators.

Controllability and Observability

Based on the previous considerations, we can say that

- The states of an LTI system (A, B, C, D) are just a part of a model that realizes a transfer function G
- The transfer function G describes how controls u lead to outputs y
- As seen above in the example, there can be states that are neither affected (*controlled*) by the inputs nor seen (*observed*) by the outputs
- These states are obviously not needed to realize the input to output behavior of G.

We will give a thorough characterization of the *controllable* and *observable* states of an LTI.

Controllability

Definition

The LTI (A, B, C, D) or the pair (A, B) is said to be *controllable* if, for any initial state $x(0) = x_0$, $t_1 > 0$ and final state x_1, there exists a (piecewise continuous) input u such that the solution of (3) satisfies $x(t_1) = x_1$. Otherwise, the system (A, B, C, D) or the pair (A, B) is said to be *uncontrollable*.

Theorem

The following statements are equivalent:

(i.) *The pair (A, B) is controllable.*

(ii.) *The controllability matrix $\mathcal{C} := \begin{bmatrix} B & AB & A^2B & \cdots & A^{n-1}B \end{bmatrix}$ has full rank.*

(iii.) *The matrix $\begin{bmatrix} A - \lambda I & B \end{bmatrix}$ has full rank for all $\lambda \in \mathbb{C}$.*

Observability

Definition

The LTI (A, B, C, D) or the pair (C, A) is said to be *observable* if, for any $t_1 > 0$, the initial state $x(0) = x_0$ can be determined from the time history of the input u and the output y in the interval of $[0, t_1]$. Otherwise, the system (A, B, C, D), or (C, A), is said to be *unobservable*.

Observability is the dual concept of controllability:

Theorem

The pair (C, A) is observable if and only if the pair $(A^\mathsf{T}, C^\mathsf{T})$ is controllable.

Invariance Under State Space Transformation

Theorem

The LTI (A, B, C, D) is controllable (observable) if, and only if, the transformed LTI $(TAT^{-1}, TB, CT^{-1}, D)$ is controllable (observable), where T is a regular matrix.

- Recall that also a transfer function is invariant with respect to state space transformations on its realization.
- Next, we find the states that are at least necessary for the realization of a transfer function...

Theorem (Kalman Canonical Decomposition)

Given an LTI (A, B, C, D), there is a state space transformation T such that the transformed system $(TAT^{-1}, TB, CT^{-1}, D)$ has the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_{co} \\ x_{c\bar{o}} \\ x_{\bar{c}o} \\ x_{\bar{c}\bar{o}} \end{bmatrix} = \begin{bmatrix} A_{co} & 0 & A_{13} & 0 \\ A_{21} & A_{c\bar{o}} & A_{23} & A_{24} \\ 0 & 0 & A_{\bar{c}o} & 0 \\ 0 & 0 & A_{43} & A_{\bar{c}\bar{o}} \end{bmatrix} \begin{bmatrix} x_{co} \\ x_{c\bar{o}} \\ x_{\bar{c}o} \\ x_{\bar{c}\bar{o}} \end{bmatrix} + \begin{bmatrix} B_{co} \\ B_{c\bar{o}} \\ 0 \\ 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} C_{co} & 0 & C_{\bar{c}o} & 0 \end{bmatrix} \begin{bmatrix} x_{co} \\ x_{c\bar{o}} \\ x_{\bar{c}o} \\ x_{\bar{c}\bar{o}} \end{bmatrix} + Du,$$

with the subsystem $(A_{co}, B_{co}, C_{co}, D)$ being controllable and observable, while the remaining states $x_{\bar{c}o}$, $x_{c\bar{o}}$, or $x_{\bar{c}\bar{o}}$ are not controllable, not observable, or neither of them.

For a constructive proof of the Theorem, see Ch. 3.3 of [Zhou, Doyle, Glover '96]

Outcomes of the Kalman Decomposition

For any state space system (A, B, C, D), there is a transformation T so that the transformed states $T^{-1}x$ decompose into

- x_{co} - controllable and observable
- $x_{c\bar{o}}$ - controllable but not observable
- $x_{\bar{c}o}$ - observable but not controllable
- $x_{\bar{c}\bar{o}}$ - not observable and not controllable

Moreover, for the transfer function, it holds that

$$G(s) = C(sI - A)^{-1}B = C_{co}(sI - A_{co})^{-1}B_{co}.$$

Conclusion from the Kalman Decomposition

What does this mean for us and a transfer function $G(s)$?

- The minimal dimension of a realization is the dimension of x_{co} in the *Kalman Canonical Decomposition*
- Such a realization is called minimal realization
- It is the starting point for further model reduction. (Throwing out $x_{\bar{c}o}$ etc. does not effect $G(s)$ and is typically not considered a model reduction)
- There are algorithm to reduce a realization to a minimal one, cf. [VARGA '90].
- In practice, the uncontrolled and unobserved states play a role and they may cause troubles. (check the literature for zero dynamics)

Linear Time Invariant Systems

Stability

- A system G is stable if all poles of G are located in the left half-plane $\mathbb{C}^-$.
- If (A, B, C, D) is a minimal realization of a stable system G, then the poles of G are the eigenvalues of A.
- In this case, the system is stable if

$$\lambda \text{ is an eigenvalue of } A, \text{ then } \lambda \in \mathbb{C}^-.$$

- Such an A is called *stable* or *Hurwitz*.
- A stable system can have a stable realization.

If $m = q = 1$, then $G(s) = \frac{N(s)}{D(s)}$, where $N(s)$ and $D(s)$ are polynomials and the *poles* are the roots of $D(s)$, i.e. those $s \in \mathbb{C}$ for which $D(s) = 0$.
If $m, q > 1$, then one can use the *McMillan* form of G to determine the poles.

Linear Time Invariant Systems

Stability

- A system G is stable if all poles of G are located in the left half-plane $\mathbb{C}^-$.
- If (A, B, C, D) is a minimal realization of a stable system G, then the poles of G are the eigenvalues of A.
- In this case, the system is stable if

$$\lambda \text{ is an eigenvalue of } A, \text{ then } \lambda \in \mathbb{C}^-.$$

- Such an A is called *stable* or *Hurwitz*.
- A stable system can have a stable realization.

If $m = q = 1$, then $G(s) = \frac{N(s)}{D(s)}$, where $N(s)$ and $D(s)$ are polynomials and the *poles* are the roots of $D(s)$, i.e. those $s \in \mathbb{C}$ for which $D(s) = 0$.
If $m, q > 1$, then one can use the *McMillan* form of G to determine the poles.

Linear Time Invariant Systems
Stability

- A system G is stable if all poles of G are located in the left half-plane $\mathbb{C}^-$.
- If (A, B, C, D) is a minimal realization of a stable system G, then the poles of G are the eigenvalues of A.
- In this case, the system is stable if

$$\lambda \text{ is an eigenvalue of } A, \text{ then } \lambda \in \mathbb{C}^-.$$

- Such an A is called *stable* or *Hurwitz*.
- A stable system can have a stable realization.

If $m = q = 1$, then $G(s) = \frac{N(s)}{D(s)}$, where $N(s)$ and $D(s)$ are polynomials and the *poles* are the roots of $D(s)$, i.e. those $s \in \mathbb{C}$ for which $D(s) = 0$.
If $m, q > 1$, then one can use the *McMillan* form of G to determine the poles.

Linear Time Invariant Systems

Stability

- A system G is stable if all poles of G are located in the left half-plane $\mathbb{C}^-$.
- If (A, B, C, D) is a minimal realization of a stable system G, then the poles of G are the eigenvalues of A.
- In this case, the system is stable if

$$\lambda \text{ is an eigenvalue of } A, \text{ then } \lambda \in \mathbb{C}^-.$$

- Such an A is called *stable* or *Hurwitz*.
- A stable system can have a stable realization.

If $m = q = 1$, then $G(s) = \frac{N(s)}{D(s)}$, where $N(s)$ and $D(s)$ are polynomials and the *poles* are the roots of $D(s)$, i.e. those $s \in \mathbb{C}$ for which $D(s) = 0$.
If $m, q > 1$, then one can use the *McMillan* form of G to determine the poles.

Linear Time Invariant Systems
Stability

- A system G is stable if all poles of G are located in the left half-plane $\mathbb{C}^-$.
- If (A, B, C, D) is a minimal realization of a stable system G, then the poles of G are the eigenvalues of A.
- In this case, the system is stable if

$$\lambda \text{ is an eigenvalue of } A, \text{ then } \lambda \in \mathbb{C}^-.$$

- Such an A is called *stable* or *Hurwitz*.
- A stable system can have a stable realization.

If $m = q = 1$, then $G(s) = \frac{N(s)}{D(s)}$, where $N(s)$ and $D(s)$ are polynomials and the *poles* are the roots of $D(s)$, i.e. those $s \in \mathbb{C}$ for which $D(s) = 0$.

If m, $q > 1$, then one can use the *McMillan* form of G to determine the poles.

Linear Time Invariant Systems

Gramians and Balanced Realizations

If A is stable, then the *Lyapunov* equations

$$A^*P + PA + BB^* = 0$$

and

$$AQ + Q^*A + C^*C = 0$$

have a unique positive definite solutions P and Q.

- The matrix P is called the the controllability Gramian
- and Q is called the observability Gramian
- and one can show that P and Q fulfill

$$P = \int_0^\infty e^{A\tau}BB^*e^{A^*\tau}\,\mathrm{d}\tau \quad \text{and} \quad Q = \int_0^\infty e^{A^*\tau}C^*Ce^{A\tau}\,\mathrm{d}\tau.$$

Linear Time Invariant Systems

Gramians and Balanced Realizations

If A is stable, then the *Lyapunov* equations

$$A^*P + PA + BB^* = 0$$

and

$$AQ + Q^*A + C^*C = 0$$

have a unique positive definite solutions P and Q.

- The matrix P is called the the controllability Gramian
- and Q is called the observability Gramian
- and one can show that P and Q fulfill

$$P = \int_0^\infty e^{A\tau} BB^* e^{A^*\tau} \, \mathrm{d}\tau \quad \text{and} \quad Q = \int_0^\infty e^{A^*\tau} C^* C e^{A\tau} \, \mathrm{d}\tau.$$

Linear Time Invariant Systems

Gramians and Balanced Realizations

$$
\begin{aligned}
A^*P + PA + BB^* &= 0 \\
AQ + Q^*A + C^*C &= 0
\end{aligned}
$$

- If P and Q are the Gramians of a stable realization (A, B, C, D),
- then the transformed system $(\hat{A}, \hat{B}, \hat{C}, \hat{D}) = (TAT^{-1}, TB, CT^{-1}, D)$ has the Gramians

$$
\hat{P} = TPT^* \quad \text{and} \quad \hat{Q} = (T^{-1})^*QT^{-1}
$$

for any regular transformation T.

Linear Time Invariant Systems

Gramians and Balanced Realizations

- For any minimal and stable system (A, B, C, D),
- there are particular transformations T,
- so that the transformed system has Gramians that are equal and diagonal:

$$\hat{P} = \hat{Q} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$.

These realizations are called Balanced Realizations.

Summary

- LTI as model for physical processes (e.g. heat transfer)
- The input/output behavior is often more important than the state
- Moreover, the state need not have a meaning
- State space systems (A, B, C, D) can be seen as realizations of transfer functions
- A transfer function has multiple realizations
- The minimal realizations are of our interest
- A stable system can have stable realization
- Minimal and stable realization can be balanced

More on the LTI topics

K. Zhou, J. C. Doyle, and K. Glover.
Robust and Optimal Control. (Chapter 3 for LTI)
Prentice-Hall, Upper Saddle River, NJ, 1996.

A. Varga.
Computation of irreducible generalized state-space realizations.
Kybernetika, 26(2):89–106, 1990.

A. Gaul.
Leckerbraten – a lightweight Python toolbox to solve the heat equation on arbitrary domains
https://github.com/andrenarchy/leckerbraten, 2013.

J. Heiland.
The slides, additional material, and information on this course
https://github.com/highlando/mor-shortcourse-SH, 2015.

Outline

Linear Time Invariant Systems

Basic Notions of Norms

Ingredients of a normed space $(V, \|\cdot\|)$:

- A linear space V over $\mathbb{C}$ (or $\mathbb{R}$)
- and a functional

  $$\|\cdot\|\colon V \to \mathbb{R}$$

  that has the following properties:

  i) $\|\alpha v\| = |\alpha| \, \|v\|$,
  ii) $\|v + w\| \leq \|v\| + \|w\|$, and
  iii) $\|v\| \geq 0$ and $\|v\| = 0$ if, and only if, $v = 0$,

  for any v, $w \in V$ and any $\alpha \in \mathbb{C}$ (or $\mathbb{R}$).

Section
Norms of Linear Operators

If $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$, then for the space of linear maps $\big(V \to W\big)$ a norm is defined via

$$\|G\|_* := \sup_{v \in V, v \neq 0} \frac{\|Gv\|_W}{\|v\|_V}.$$

This is the norm for $G \colon V \to W$ that is induced by $\|\cdot\|_V$ and $\|\cdot\|_W$.
There can be other norms that are not induced.

Linear Time Invariant Systems

Norms of Signals

Common norms and spaces for the input or output signals

$$u\colon [0,\infty) \to \mathbb{R}^m \quad \text{or} \quad y\colon [0,\infty) \to \mathbb{R}^q$$

- All definitions work similar for finite time intervals $[0, T]$ or the whole time axis $(-\infty, \infty)$.
- Where it is clear from the context, we will drop the superscripts p and m that denote the dimension of the signals.

Linear Time Invariant Systems

Norms of Signals

Definition

The $\mathbf{L}_1^m$ norm

$$\|u\|_{\mathbf{L}_1} := \int_0^\infty \sum_{i=1}^m |u_i(t)| \, \mathrm{d}t$$

defines the $\mathbf{L}_1^m$ space of integrable (summable) functions

$$\mathbf{L}_1^m := \left\{ u \colon [0, \infty) \to \mathbb{R}^m : \|u\|_{\mathbf{L}_1} < \infty \right\}$$

on the positive time axis.

Linear Time Invariant Systems

Norms of Signals

Definition

The $\mathbf{L}_\infty^m$ norm

$$\|u\|_{\mathbf{L}_\infty} := \max_{i=\{1,\dots,m\}} \sup_{t>0} |u_i(t)|$$

defines the $\mathbf{L}_\infty^m$ space of bounded functions

$$\mathbf{L}_\infty^m := \{u\colon [0,\infty) \to \mathbb{R}^m : \|u\|_{\mathbf{L}_\infty} < \infty\}.$$

Definition

The $\mathbf{L}_2^q$ norm

$$\|y\|_{\mathbf{L}_2} := \Big(\int_0^\infty \sum_{i=1}^q |y_i(t)|^2 \;\mathrm{d}t\Big)^{\frac{1}{2}}$$

defines the $\mathbf{L}_2^q$ space of square integrable functions

$$\mathbf{L}_2^q := \{y\colon [0,\infty) \to \mathbb{R}^q : \|y\|_{\mathbf{L}_2} < \infty\}$$

Linear Time Invariant Systems

Norms of Signals

The $\mathbf{L}_2$ norm can also be evaluated in frequency domain

Theorem

For $u \in \mathbf{L}_2$ it holds that

$$\|u\|_{\mathbf{L}_2} = \Big(\frac{1}{2\pi}\int_{-\infty}^{\infty} U(i\omega)^* U(i\omega)\, \mathrm{d}\omega\Big)^{\frac{1}{2}},$$

where U is the Fourier transform of u.

The Fourier transform $\mathcal{F}$ and the Laplace transform $\mathcal{L}$ coincide for $s = i\omega$, $\omega \in \mathbb{R}$ and $u(t) = 0$ for $t \leq 0$:

$$\mathcal{F}(u)(i\omega) := \int_{-\infty}^{\infty} u(t)e^{-i\omega t}\, \mathrm{d}t = \int_{0}^{\infty} u(t)e^{-st}\, \mathrm{d}t = \mathcal{L}(u)(s)$$

Linear Time Invariant Systems
Norm of a System

A system G or (A, B, C, D) transfers inputs to outputs.

Ask yourself. . .

- What does a norm mean for a system?
- What is a large system, what is a small system?

Linear Time Invariant Systems
Norm of a System

From the definition of an operator norm:

$$\|G\| = \sup_{u \neq 0} \frac{\|Gu\|}{\|u\|}$$

we derive that for all u:

$$\|y\| = \|Gu\| \leq \|G\|\|u\|.$$

An Answer

For systems, large refers to what extend an input is amplified.
Therefore, $\|G\|$ is often called the *gain*.

Linear Time Invariant Systems

Norm of a System

From the definition of an operator norm:

$$\|G\| = \sup_{u \neq 0} \frac{\|Gu\|}{\|u\|}$$

we derive that for all u:

$$\|y\| = \|Gu\| \leq \|G\|\|u\|.$$

With a norm, one can compare two systems G_1 and G_2 via the difference in the output for the same input:

$$\|y_1 - y_2\| = \|G_1 u - G_2 u\| \leq \|G_1 - G_2\|\|u\|.$$

Linear Time Invariant Systems

Defining a Norm for Systems

We consider a SISO system $(A, B, C, -)$, i.e $m = q = 1$ and $D = 0$.

Consider $(A, B, C, -)$ a with stable and strictly proper transfer function G is stable. Then the *impulse response* of the system

$$g(t) = C\int_0^t e^{A(t-\tau)}B\delta(\tau)\,\mathrm{d}s = Ce^{At}B$$

decays exponentially and

$$\|g\|_{L_2} = \big(\frac{1}{2\pi}\int_{-\infty}^{\infty} G(i\omega)^* G(i\omega)\,\mathrm{d}\omega\big)^{\frac{1}{2}} =: \|G\|_2 < \infty.$$

A system (A, B, C, D) or A is stable, if there exists a $\lambda > 0$, such that $\|e^At\| \leq e^{-\lambda t}$, for $t > 0$. This means that all eigenvalues of A must have a negative real part.

Tmpulse response: $\delta(\tau) := \begin{cases} 0, \text{ if } t \neq 0, \\ \text{very large, if } t = 0 \end{cases}$ so that $\int_{-\infty}^{\infty} u(\tau)\delta(\tau)\,\mathrm{d}\tau = u(0)$.

Linear Time Invariant Systems

Defining a Norm for Systems

We consider a SISO system $(A, B, C, -)$, i.e $m = q = 1$ and $D = 0$.

Consider $(A, B, C, -)$ a with stable and strictly proper transfer function G is stable. Then the *impulse response* of the system

$$g(t) = C \int_0^t e^{A(t-\tau)} B\delta(\tau)\, \mathrm{d}s = Ce^{At}B$$

decays exponentially and

$$\|g\|_{\mathbf{L}_2} = \Big(\frac{1}{2\pi} \int_{-\infty}^{\infty} G(i\omega)^* G(i\omega)\, \mathrm{d}\omega\Big)^{\frac{1}{2}} =: \|G\|_2 < \infty.$$

Linear Time Invariant Systems

Defining a Norm for Systems

We consider a SISO system $(A, B, C, -)$, i.e $m = q = 1$ and $D = 0$.

Consider $(A, B, C, -)$ a with stable and strictly proper transfer function G is stable. Then the *impulse response* of the system

$$g(t) = C \int_0^t e^{A(t-\tau)} B\delta(\tau)\, \mathrm{d}s = Ce^{At}B$$

decays exponentially and

$$\|g\|_{\mathbf{L}_2} = \Big(\frac{1}{2\pi} \int_{-\infty}^{\infty} G(i\omega)^* G(i\omega)\, \mathrm{d}\omega\Big)^{\frac{1}{2}} =: \|G\|_2 < \infty.$$

This defines a norm for systems since (Exercise!)

- $G = C(sI - A)^{-1}B$ is indeed the Laplace transform of g
- the functional $\|\cdot\|_2$ for stable and strictly proper transferfunctions is a norm

Furthermore, $\|y\|_{\mathbf{L}_\infty} \leq \|G\|_2 \|u\|_{\mathbf{L}_\infty}$. (Exercise!)

Linear Time Invariant Systems

Defining a Norm for Systems

For MIMO systems $(A, B, C, -)$ with $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^q$, with a stable and strictly proper transferfunction $\mathcal{G}\colon s \to \mathbb{R}^{q\times m}$, the $\mathcal{H}_2$ norm is defined as

$$\|G\|_2 := \Big(\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{tr}\left(G(i\omega)^* G(i\omega)\right)\, \mathrm{d}\omega\Big)^{\frac{1}{2}}.$$

Fact

This is the norm of the *Hardy* space $\mathcal{H}_2$ of matrix functions that are analytic in the open right half of the complex plane. Stable and strictly proper transfer functions are in $\mathcal{H}_2$.

Linear Time Invariant Systems

Defining a Norm for Systems

For a stable and proper transfer function one can define the $\mathcal{H}_\infty$ norm:

$$\|G\|_\infty := \sup_{\omega\in\mathbb{R}} \sigma_{\max}\big(G(i\omega)\big),$$

where $\sigma_{\max}\big(G(i\omega)\big)$ is the largest singular value of $G(i\omega)$.

Fact 1

This is the norm of the *Hardy* space $\mathcal{H}_\infty$ of matrix functions that are analytic in the open right half of the complex plane and bounded on the imaginary axis. Stable and strictly proper transfer functions are in $\mathcal{H}_\infty$.

Fact 2

The $\mathcal{H}_\infty$-norm is induced by the $\mathbf{L}_2$ norm:

$$\|G\|_\infty = \sup_{u\in\mathbf{L}_2, u\neq 0} \frac{\|Gu\|_{\mathbf{L}_2}}{\|u\|_{\mathbf{L}_2}}.$$

Relation to Model Reduction
Approximation Problems - Model Reduction

Output errors in time-domain

Comparing the original system G and the reduced system $\hat{G}$:

$$
\begin{aligned}
||y-\hat{y}||_2 &\leq \left|\left|G-\hat{G}\right|\right|_\infty ||u||_2 && \Longrightarrow \left|\left|G-\hat{G}\right|\right|_\infty < \mathrm{tol} \\
||y-\hat{y}||_\infty &\leq \left|\left|G-\hat{G}\right|\right|_2 ||u||_2 && \Longrightarrow \left|\left|G-\hat{G}\right|\right|_2 < \mathrm{tol}
\end{aligned}
$$

Relation to Model Reduction

Approximation Problems - Model Reduction

Output errors in time-domain

Comparing the original system G and the reduced system $\hat{G}$:

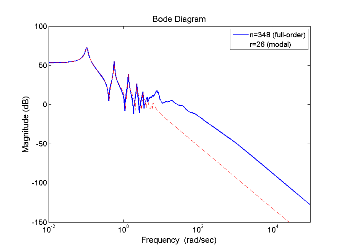
$$
\begin{aligned}
||y - \hat{y}||_2 &\leq \left|\left|G - \hat{G}\right|\right|_\infty ||u||_2 &\implies \left|\left|G - \hat{G}\right|\right|_\infty < \mathrm{tol} \\
||y - \hat{y}||_\infty &\leq \left|\left|G - \hat{G}\right|\right|_2 ||u||_2 &\implies \left|\left|G - \hat{G}\right|\right|_2 < \mathrm{tol}
\end{aligned}
$$

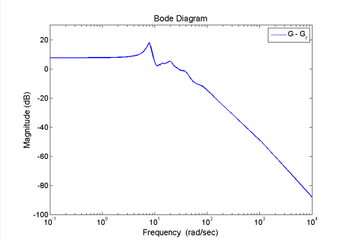
$\mathcal{H}_\infty$-norm	best approximation problem for given reduced order r in general open; balanced truncation yields suboptimal solution with computable $\mathcal{H}_\infty$-norm bound.
$\mathcal{H}_2$-norm	necessary conditions for best approximation known; (local) optimizer computable with iterative rational Krylov algorithm (IRKA)
Hankel-norm $\|\|G\|\|_H := \sigma_{\max}$	optimal Hankel norm approximation (AAK theory).

Evaluating system norms is computationally very (sometimes too) expensive.

Other measures

- absolute errors $\left|\left|G(\jmath\omega_j) - \hat{G}(\jmath\omega_j)\right|\right|_2$, $\left|\left|G(\jmath\omega_j) - \hat{G}(\jmath\omega_j)\right|\right|_\infty$ $(j = 1, \ldots, N_\omega)$;
- relative errors $\frac{\left|\left|G(\jmath\omega_j) - \hat{G}(\jmath\omega_j)\right|\right|_2}{\left|\left|G(\jmath\omega_j)\right|\right|_2}$, $\frac{\left|\left|G(\jmath\omega_j) - \hat{G}(\jmath\omega_j)\right|\right|_\infty}{\left|\left|G(\jmath\omega_j)\right|\right|_\infty}$;
- "eyeball norm", i.e. look at frequency response/Bode (magnitude) plot: for SISO system, log-log plot frequency vs. $|G(\jmath\omega)|$ (or $|G(\jmath\omega) - \hat{G}(\jmath\omega)|$) in decibels, 1 dB $\simeq 20\log_{10}(\text{value})$.





Outline

1. Linear Time Invariant Systems
2. Introduction to Model Reduction
   - Model Reduction for Dynamical Systems
   - Application Areas
   - Motivating Examples
3. Model Reduction by Projection
4. Balanced Truncation
5. Linear Time-invariant DAEs

Introduction to Model Reduction

Model Reduction — Abstract Definition

Problem

Given a model of a physical problem with dynamics described by the states $x(t) \in \mathbb{R}^n$, where n is the dimension of the state space.

The dimension n is large because $x(t)$ typically contains information that

- *is (almost) redundant,*
- *not (really) important,*
- *or not (really) of interest.*

We want to adjust the model such that the new state is of small dimension but still bears all important and interesting information.

This is the task of model reduction (also: dimension reduction, order reduction).

Introduction to Model Reduction

Model Reduction — Abstract Definition

Problem

Given a model of a physical problem with dynamics described by the states $x(t) \in \mathbb{R}^n$, where n is the dimension of the state space.

The dimension n is large because $x(t)$ typically contains information that

- *is (almost) redundant,*
- *not (really) important,*
- *or not (really) of interest.*

We want to adjust the model such that the new state is of small dimension but still bears all important and interesting information.

This is the task of model reduction (also: dimension reduction, order reduction).

Introduction to Model Reduction

Model Reduction — Abstract Definition

Problem

Given a model of a physical problem with dynamics described by the states $x(t) \in \mathbb{R}^n$, where n is the dimension of the state space.

The dimension n is large because $x(t)$ typically contains information that

- *is (almost) redundant,*
- *not (really) important,*
- *or not (really) of interest.*

We want to adjust the model such that the new state is of small dimension but still bears all important and interesting information.

This is the task of model reduction (also: dimension reduction, order reduction).

Introduction to Model Reduction

Model Reduction — Abstract Definition

Problem

Given a model of a physical problem with dynamics described by the states $x(t) \in \mathbb{R}^n$, where n is the dimension of the state space.

The dimension n is large because $x(t)$ typically contains information that

- *is (almost) redundant,*
- *not (really) important,*
- *or not (really) of interest.*

We want to adjust the model such that the new state is of small dimension but still bears all important and interesting information.

This is the task of model reduction (also: dimension reduction, order reduction).

Introduction to Model Reduction

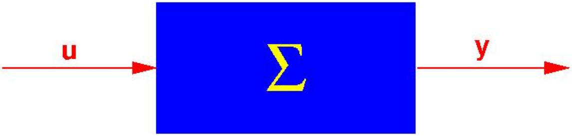
Model Reduction for Dynamical Systems

Dynamical Systems

$$\Sigma : \begin{cases} \dot{x}(t) &= f(t, x(t), u(t)), \quad x(t_0) = x_0, \\ y(t) &= g(t, x(t), u(t)) \end{cases}$$

with

- states $x(t) \in \mathbb{R}^n$,
- inputs $u(t) \in \mathbb{R}^m$,
- outputs $y(t) \in \mathbb{R}^q$.

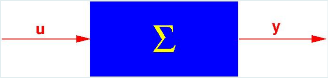


Model Reduction for Dynamical Systems

Original System

$$\Sigma : \begin{cases} \dot{x}(t) = f(t, x(t), u(t)), \\ y(t) = g(t, x(t), u(t)). \end{cases}$$

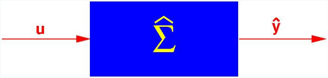
- states $x(t) \in \mathbb{R}^n$,
- inputs $u(t) \in \mathbb{R}^m$,
- outputs $y(t) \in \mathbb{R}^q$.



Reduced-Order Model (ROM)

$$\widehat{\Sigma} : \begin{cases} \dot{\hat{x}}(t) = \widehat{f}(t, \hat{x}(t), u(t)), \\ \hat{y}(t) = \widehat{g}(t, \hat{x}(t), u(t)). \end{cases}$$

- states $\hat{x}(t) \in \mathbb{R}^r$, $r \ll n$
- inputs $u(t) \in \mathbb{R}^m$,
- outputs $\hat{y}(t) \in \mathbb{R}^q$.



Goal:

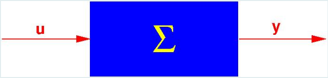
$||y - \hat{y}|| <$ tolerance $\cdot$ $||u||$ for all admissible input signals.

Model Reduction for Dynamical Systems

Original System

$$\Sigma : \begin{cases} \dot{x}(t) = f(t, x(t), u(t)), \\ y(t) = g(t, x(t), u(t)). \end{cases}$$

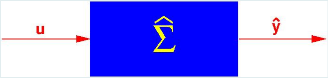
- states $x(t) \in \mathbb{R}^n$,
- inputs $u(t) \in \mathbb{R}^m$,
- outputs $y(t) \in \mathbb{R}^q$.



Reduced-Order Model (ROM)

$$\widehat{\Sigma} : \begin{cases} \dot{\hat{x}}(t) = \widehat{f}(t, \hat{x}(t), u(t)), \\ \hat{y}(t) = \widehat{g}(t, \hat{x}(t), u(t)). \end{cases}$$

- states $\hat{x}(t) \in \mathbb{R}^r$, $r \ll n$
- inputs $u(t) \in \mathbb{R}^m$,
- outputs $\hat{y}(t) \in \mathbb{R}^q$.



Goal:

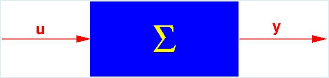
$||y - \hat{y}|| <$ tolerance $\cdot \, ||u||$ for all admissible input signals.

Model Reduction for Dynamical Systems

Original System

$$\Sigma : \begin{cases} \dot{x}(t) = f(t, x(t), u(t)), \\ y(t) = g(t, x(t), u(t)). \end{cases}$$

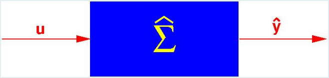
- states $x(t) \in \mathbb{R}^n$,
- inputs $u(t) \in \mathbb{R}^m$,
- outputs $y(t) \in \mathbb{R}^q$.



Reduced-Order Model (ROM)

$$\widehat{\Sigma} : \begin{cases} \dot{\hat{x}}(t) = \widehat{f}(t, \hat{x}(t), u(t)), \\ \hat{y}(t) = \widehat{g}(t, \hat{x}(t), u(t)). \end{cases}$$

- states $\hat{x}(t) \in \mathbb{R}^r$, $r \ll n$
- inputs $u(t) \in \mathbb{R}^m$,
- outputs $\hat{y}(t) \in \mathbb{R}^q$.



Goal:

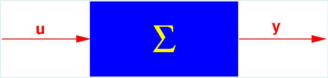
$||y - \hat{y}|| < \text{tolerance} \cdot ||u||$ for all admissible input signals.

Model Reduction for Dynamical Systems

Original System

$$\Sigma : \begin{cases} \dot{x}(t) &= f(t, x(t), u(t)), \\ y(t) &= g(t, x(t), u(t)). \end{cases}$$

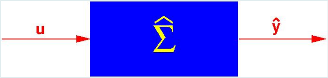
- states $x(t) \in \mathbb{R}^n$,
- inputs $u(t) \in \mathbb{R}^m$,
- outputs $y(t) \in \mathbb{R}^q$.



Reduced-Order Model (ROM)

$$\widehat{\Sigma} : \begin{cases} \dot{\hat{x}}(t) &= \widehat{f}(t, \hat{x}(t), u(t)), \\ \hat{y}(t) &= \widehat{g}(t, \hat{x}(t), u(t)). \end{cases}$$

- states $\hat{x}(t) \in \mathbb{R}^r$, $r \ll n$
- inputs $u(t) \in \mathbb{R}^m$,
- outputs $\hat{y}(t) \in \mathbb{R}^q$.



Goal:

$||y - \hat{y}|| <$ tolerance $\cdot$ $||u||$ for all admissible input signals.

Secondary goal: reconstruct approximation of x from $\hat{x}$.

Model Reduction for Dynamical Systems

Linear Systems

Linear, Time-Invariant (LTI) Systems

$$
\begin{aligned}
E\dot{x} &= f(t,x,u) &= Ax + Bu, \quad & E, A \in \mathbb{R}^{n\times n}, & B \in \mathbb{R}^{n\times m}, \\
y &= g(t,x,u) &= Cx + Du, \quad & C \in \mathbb{R}^{q\times n}, & D \in \mathbb{R}^{q\times m}.
\end{aligned}
$$

Model Reduction for Dynamical Systems

Linear Systems

Linear, Time-Invariant (LTI) Systems

$$
\begin{array}{rclclll}
E\dot{x} & = & f(t,x,u) & = & Ax + Bu, & E, A \in \mathbb{R}^{n\times n}, & B \in \mathbb{R}^{n\times m}, \\
y & = & g(t,x,u) & = & Cx + Du, & C \in \mathbb{R}^{q\times n}, & D \in \mathbb{R}^{q\times m}.
\end{array}
$$

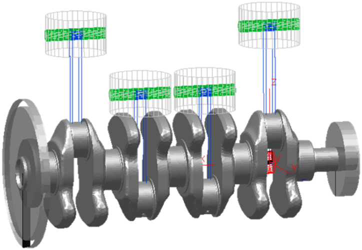
Linear, Time-Invariant Parametric Systems

$$
\begin{aligned}
E(p)\dot{x}(t;p) &= A(p)x(t;p) + B(p)u(t), \\
y(t;p) &= C(p)x(t;p) + D(p)u(t),
\end{aligned}
$$

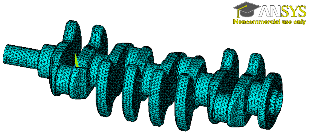
where $A(p), E(p) \in \mathbb{R}^{n\times n}, B(p) \in \mathbb{R}^{n\times m}, C(p) \in \mathbb{R}^{q\times n}, D(p) \in \mathbb{R}^{q\times m}$.

Application Areas

Structural Mechanics / Finite Element Modeling since $\sim$1960ies



$\rightsquigarrow$



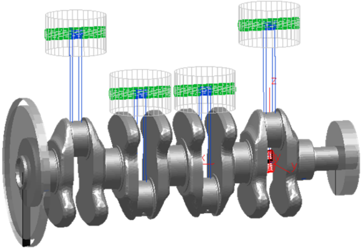
- Resolving complex 3D geometries $\Rightarrow$ millions of degrees of freedom.
- Analysis of elastic deformations requires many simulation runs for varying external forces.

Standard MOR techniques in structural mechanics: modal truncation, combined with Guyan reduction (static condensation) $\rightsquigarrow$ Craig-Bampton method.

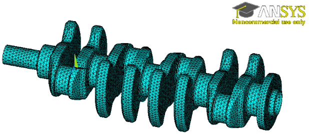
Application Areas

Structural Mechanics / Finite Element Modeling

since ∼1960ies



$\rightsquigarrow$



- Resolving complex 3D geometries $\Rightarrow$ millions of degrees of freedom.
- Analysis of elastic deformations requires many simulation runs for varying external forces.

Standard MOR techniques in structural mechanics: modal truncation, combined with Guyan reduction (static condensation) $\rightsquigarrow$ Craig-Bampton method.

Application Areas

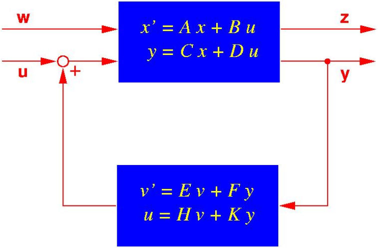
(Optimal) Control **since ∼1980ies**

Feedback Controllers

A feedback controller (dynamic compensator) is a linear system of order N, where

- input = output of plant,
- output = input of plant.

Modern (LQG-/$\mathcal{H}_2$-/$\mathcal{H}_\infty$-) control design: $N \geq n$.



Practical controllers require small N ($N \sim 10$, say) due to

- real-time constraints,
- increasing fragility for larger N.

$\Longrightarrow$ reduce order of plant (n) and/or controller (N).

Standard MOR techniques in systems and control: balanced truncation and related methods.

Application Areas

(Optimal) Control

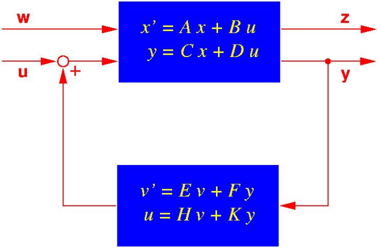
since $\sim$1980ies

Feedback Controllers

A feedback controller (dynamic compensator) is a linear system of order N, where

- input $=$ output of plant,
- output $=$ input of plant.

Modern (LQG-/$\mathcal{H}_2$-/$\mathcal{H}_\infty$-) control design: $N \geq n$.



Practical controllers require small N ($N \sim 10$, say) due to

- real-time constraints,
- increasing fragility for larger N.

$\Longrightarrow$ reduce order of plant (n) and/or controller (N).

Standard MOR techniques in systems and control: balanced truncation and related methods.

Application Areas

(Optimal) Control

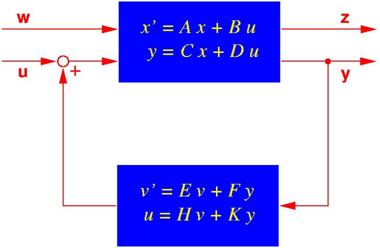
since ∼1980ies

Feedback Controllers

A feedback controller (dynamic compensator) is a linear system of order N, where

- input $=$ output of plant,
- output $=$ input of plant.

Modern (LQG-/$\mathcal{H}_2$-/$\mathcal{H}_\infty$-) control design: $N \geq n$.



Practical controllers require small N ($N \sim 10$, say) due to

- real-time constraints,
- increasing fragility for larger N.

$\Longrightarrow$ reduce order of plant (n) and/or controller (N).

Standard MOR techniques in systems and control: balanced truncation and related methods.

Application Areas

(Optimal) Control

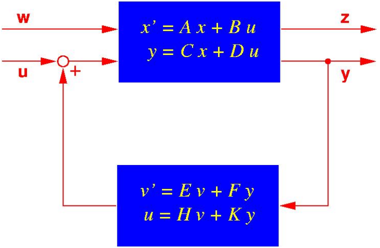
since $\sim$1980ies

Feedback Controllers

A feedback controller (dynamic compensator) is a linear system of order N, where

- input $=$ output of plant,
- output $=$ input of plant.

Modern (LQG-/$\mathcal{H}_2$-/$\mathcal{H}_\infty$-) control design: $N \geq n$.



Practical controllers require small N ($N \sim 10$, say) due to

- real-time constraints,
- increasing fragility for larger N.

$\Longrightarrow$ reduce order of plant (n) and/or controller (N).

Standard MOR techniques in systems and control: balanced truncation and related methods.

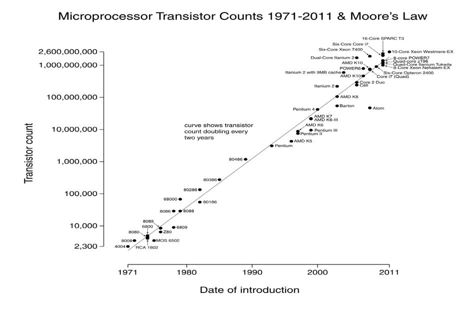
Application Areas

Micro Electronics/Circuit Simulation

since $\sim$1990ies

Progressive miniaturization

- Verification of VLSI/ULSI chip design needs a large number of simulations.
- **Moore's Law (1965/75)** states that the number of on-chip transistors doubles each 24 months.



Source: http://en.wikipedia.org/wiki/File:Transistor_Count_and_Moore'sLaw_-_2011.svg

Application Areas

Micro Electronics/Circuit Simulation since ∼1990ies

Progressive miniaturization

- Verification of VLSI/ULSI chip design needs a large number of simulations.
- **Moore's Law (1965/75)** $\rightsquigarrow$ steady increase of describing equations, i.e., network topology (Kirchhoff's laws) and characteristic element/semi-conductor equations.

Application Areas

Micro Electronics/Circuit Simulation

since ∼1990ies

Progressive miniaturization

- Verification of VLSI/ULSI chip design needs a large number of simulations.
- **Moore's Law (1965/75)** $\rightsquigarrow$ steady increase of describing equations, i.e., network topology (Kirchhoff's laws) and characteristic element/semi-conductor equations.
- Increase in packing density and multilayer technology requires modeling of interconnect to ensure that thermic/electro-magnetic effects do not disturb signal transmission.

Intel 4004 (1971)	Intel Core 2 Extreme (quad-core) (2007)
1 layer, 10μ technology	9 layers, $45nm$ technology
2,300 transistors	$> 8,200,000$ transistors
64 kHz clock speed	> 3 GHz clock speed.

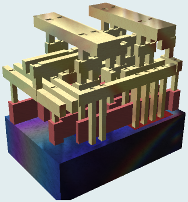
Application Areas

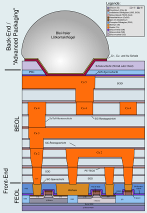
Micro Electronics/Circuit Simulation

since $\sim$1990ies

Progressive miniaturization

- Verification of VLSI/ULSI chip design needs a large number of simulations.
- **Moore's Law (1965/75)** $\rightsquigarrow$ steady increase of describing equations, i.e., network topology (Kirchhoff's laws) and characteristic element/semi-conductor equations.
- Increase in packing density and multilayer technology requires modeling of interconnect to ensure that thermic/electro-magnetic effects do not disturb signal transmission.





Source: http://en.wikipedia.org/wiki/Image:Silicon_chip_3d.png.

Application Areas

Micro Electronics/Circuit Simulation since $\sim$1990ies

Progressive miniaturization

- Verification of VLSI/ULSI chip design needs a large number of simulations.
- **Moore's Law (1965/75)** $\rightsquigarrow$ steady increase of describing equations, i.e., network topology (Kirchhoff's laws) and characteristic element/semi-conductor equations.
- Here: mostly MOR for linear systems, they occur in micro electronics through modified nodal analysis (MNA) for RLC networks. e.g., when
  - decoupling large linear subcircuits,
  - modeling transmission lines,
  - modeling pin packages in VLSI chips,
  - modeling circuit elements described by Maxwell's equation using partial element equivalent circuits (PEEC).

Application Areas

Micro Electronics/Circuit Simulation **since ∼1990ies**

Progressive miniaturization

- Verification of VLSI/ULSI chip design needs a large number of simulations.
- **Moore's Law (1965/75)** ⇝ steady increase of describing equations, i.e., network topology (Kirchhoff's laws) and characteristic element/semi-conductor equations.

⇝ Clear need for model reduction techniques in order to facilitate or even enable circuit simulation for current and future VLSI design.

Application Areas

Micro Electronics/Circuit Simulation

since $\sim$1990ies

Progressive miniaturization

- Verification of VLSI/ULSI chip design needs a large number of simulations.
- **Moore's Law (1965/75)** $\rightsquigarrow$ steady increase of describing equations, i.e., network topology (Kirchhoff's laws) and characteristic element/semi-conductor equations.

$\rightsquigarrow$ Clear need for model reduction techniques in order to facilitate or even enable circuit simulation for current and future VLSI design.

Standard MOR techniques in circuit simulation:
Krylov subspace / Padé approximation / rational interpolation methods.

Application Areas

Many other disciplines in computational sciences and engineering like

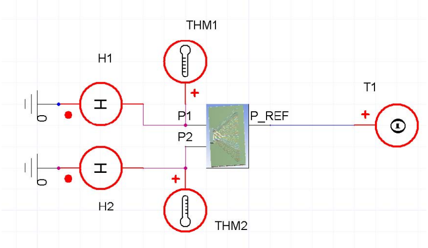
- computational fluid dynamics (CFD),
- computational electromagnetics,
- chemical process engineering,
- design of MEMS/NEMS (micro/nano-electrical-mechanical systems),
- computational acoustics,
- ...

Motivating Examples

Electro-Thermic Simulation of Integrated Circuit (IC)

[Source: Evgenii Rudnyi, CADFEM GmbH]

- SIMPLORER® test circuit with 2 transistors.



- Conservative thermic sub-system in SIMPLORER:
  voltage ⇝ temperature, current ⇝ heat flow.
- Original model: $n = 270.593$, $m = q = 2 \Rightarrow$
  Computing time (on Intel Xeon dualcore 3GHz, 1 Thread):
  - Main computational cost for set-up data $\approx 22min$.
  - Computation of reduced models from set-up data: 44–49sec. ($r = 20$–70).
  - Bode plot (MATLAB on Intel Core i7, 2,67GHz, 12GB):
    7.5h for original system, $<$ 1min for reduced system.
  - Speed-up factor: 18 including / $\geq$ 450 excluding reduced model generation!

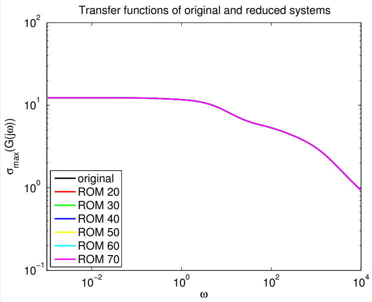
Motivating Examples

Electro-Thermic Simulation of Integrated Circuit (IC)

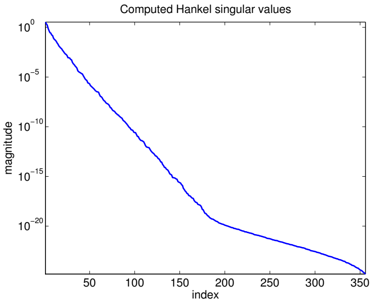
[Source: Evgenii Rudnyi, CADFEM GmbH]

- Original model: $n = 270.593$, $m = q = 2 \Rightarrow$
  Computing time (on Intel Xeon dualcore 3GHz, 1 Thread):
  - Main computational cost for set-up data $\approx 22min$.
  - Computation of reduced models from set-up data: 44–49sec. ($r = 20$–70).
  - Bode plot (MATLAB on Intel Core i7, 2,67GHz, 12GB):
    7.5h for original system, $<$ 1min for reduced system.
  - Speed-up factor: 18 including / $\geq$ 450 excluding reduced model generation!

Bode Plot (Amplitude)



Hankel Singular Values



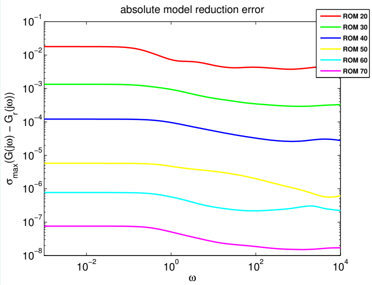
Motivating Examples

Electro-Thermic Simulation of Integrated Circuit (IC)

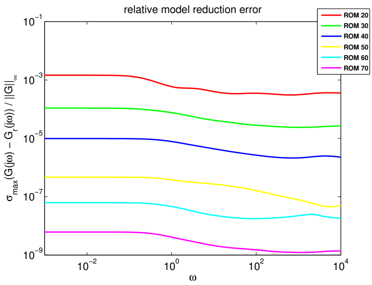
[Source: Evgenii Rudnyi, CADFEM GmbH]

- Original model: $n = 270.593,\ m = q = 2 \Rightarrow$
  Computing time (on Intel Xeon dualcore 3GHz, 1 Thread):
  - Main computational cost for set-up data $\approx 22min$.
  - Computation of reduced models from set-up data: 44–49sec. ($r = 20$–70).
  - Bode plot (MATLAB on Intel Core i7, 2,67GHz, 12GB):
    7.5h for original system, $<$ 1min for reduced system.
  - Speed-up factor: 18 including / $\geq$ 450 excluding reduced model generation!

Absolute Error



Relative Error



Motivating Examples

A Nonlinear Model from Computational Neurosciences: the FitzHugh-Nagumo System

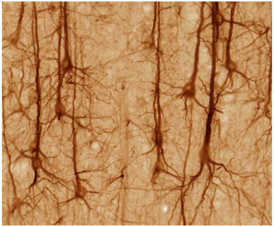
- Simple model for neuron (de-)activation [CHATURANTABUT/SORENSEN 2009]

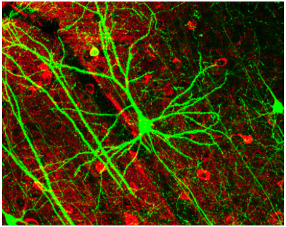
$$
\begin{aligned}
\epsilon v_t(x,t) &= \epsilon^2 v_{xx}(x,t) + f(v(x,t)) - w(x,t) + g,\\
w_t(x,t) &= hv(x,t) - \gamma w(x,t) + g,
\end{aligned}
$$

with $f(v) = v(v-0.1)(1-v)$ and initial and boundary conditions

$$
\begin{aligned}
v(x,0) &= 0, & w(x,0) &= 0, & x &\in [0,1]\\
v_x(0,t) &= -i_0(t), & v_x(1,t) &= 0, & t &\geq 0,
\end{aligned}
$$

where $\epsilon = 0.015, h = 0.5, \gamma = 2, g = 0.05, i_0(t) = 50000t^3 \exp(-15t)$.





Source: http://en.wikipedia.org/wiki/Neuron

Motivating Examples

A Nonlinear Model from Computational Neurosciences: the FitzHugh-Nagumo System

- Simple model for neuron (de-)activation [Chaturantabut/Sorensen 2009]

$$\begin{aligned} \epsilon v_t(x,t) &= \epsilon^2 v_{xx}(x,t) + f(v(x,t)) - w(x,t) + g, \\ w_t(x,t) &= hv(x,t) - \gamma w(x,t) + g, \end{aligned}$$

  with $f(v) = v(v-0.1)(1-v)$ and initial and boundary conditions

$$\begin{aligned} &v(x,0) = 0, && w(x,0) = 0, && x \in [0,1] \\ &v_x(0,t) = -i_0(t), && v_x(1,t) = 0, && t \geq 0, \end{aligned}$$

  where $\epsilon = 0.015, h = 0.5, \gamma = 2, g = 0.05, i_0(t) = 50000t^3 \exp(-15t)$.

- Parameter g handled as an additional input.
- Original state dimension $n = 2 \cdot 400$, QBDAE dimension $N = 3 \cdot 400$, reduced QBDAE dimension $r = 26$, chosen expansion point $\sigma = 1$.

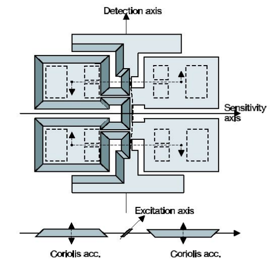
Motivating Examples

A Nonlinear Model from Computational Neurosciences: the FitzHugh-Nagumo System

Motivating Examples
Parametric MOR: Applications in Microsystems/MEMS Design

Microgyroscope (butterfly gyro)

- Application: inertial navigation.



- Voltage applied to electrodes induces vibration of wings, resulting rotation due to Coriolis force yields sensor data.
- FE model of second order: $N = 17.361 \rightsquigarrow n = 34.722,\ m = 1,\ q = 12.$
- Sensor for position control based on acceleration and rotation.

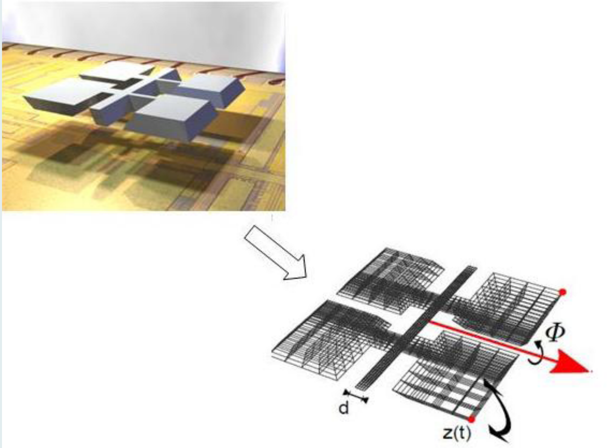
Source: The Oberwolfach Benchmark Collection http://www.imtek.de/simulation/benchmark

Motivating Examples

Parametric MOR: Applications in Microsystems/MEMS Design

Microgyroscope (butterfly gyro)

Parametric FE model: $M(d)\ddot{x}(t) + D(\Phi, d, \alpha, \beta)\dot{x}(t) + T(d)x(t) = Bu(t).$



Motivating Examples
Parametric MOR: Applications in Microsystems/MEMS Design

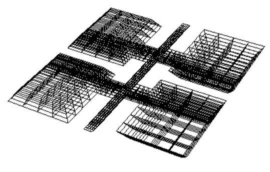
Microgyroscope (butterfly gyro)

Parametric FE model:

$$M(d)\ddot{x}(t) + D(\Phi, d, \alpha, \beta)\dot{x}(t) + T(d)x(t) = Bu(t),$$

wobei

$$\begin{aligned} M(d) &= M_1 + dM_2, \\ D(\Phi, d, \alpha, \beta) &= \Phi(D_1 + dD_2) + \alpha M(d) + \beta T(d), \\ T(d) &= T_1 + \frac{1}{d} T_2 + dT_3, \end{aligned}$$



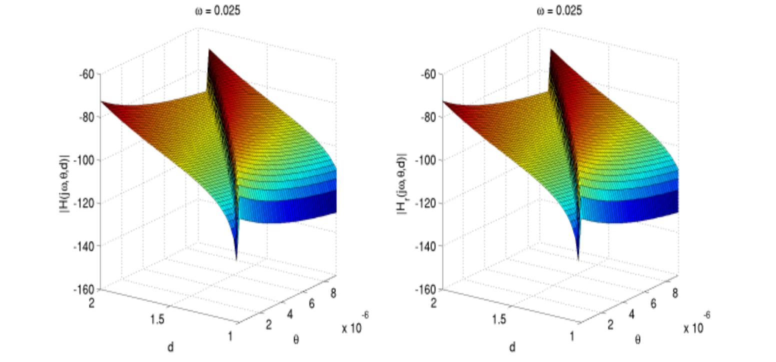
with

- width of bearing: d,
- angular velocity: Φ,
- Rayleigh damping parameters: α, β.

Motivating Examples
Parametric MOR: Applications in Microsystems/MEMS Design

Microgyroscope (butterfly gyro)

Original. . . and reduced-order model.



Outline

Model Reduction by Projection
Goals

Requirements: A Model Reduction approach should:

- Automatically generate compact models $\hat{G}$ from a given model G
- Satisfy desired error tolerance tol for all admissible input signals u

$$\|y - \hat{y}\| = \|Gu - \hat{G}u\| < \text{tol} \cdot \|u\| \quad \text{for all admissable } u.$$

$\Longrightarrow$ Provide computable error bound/estimate!

- Preserve physical properties:
  - stability
  - minimum phase
  - passivity

Model Reduction by Projection

Goals

Requirements: A Model Reduction approach should:

- Automatically generate compact models $\hat{G}$ from a given model G
- Satisfy desired error tolerance tol for all admissible input signals u

$$\|y - \hat{y}\| = \|Gu - \hat{G}u\| < \text{tol} \cdot ||u|| \quad \text{for all admissable } u.$$

$\Longrightarrow$ Provide computable error bound/estimate!

- Preserve physical properties:
  - stability
  - minimum phase
  - passivity

Model Reduction by Projection

Goals

Requirements: A Model Reduction approach should:

- Automatically generate compact models $\hat{G}$ from a given model G
- Satisfy desired error tolerance tol for all admissible input signals u

$$\|y - \hat{y}\| = \|Gu - \hat{G}u\| < \mathrm{tol} \cdot ||u|| \quad \text{for all admissable } u.$$

$\Longrightarrow$ Provide computable error bound/estimate!

- Preserve physical properties:
  - stability
  - minimum phase
  - passivity

Model Reduction by Projection

Goals

Requirements: A Model Reduction approach should:

- Automatically generate compact models $\hat{G}$ from a given model G
- Satisfy desired error tolerance tol for all admissible input signals u

$$\|y - \hat{y}\| = \|Gu - \hat{G}u\| < \text{tol} \cdot ||u|| \quad \text{for all admissable } u.$$

$\Longrightarrow$ Provide computable error bound/estimate!

- Preserve physical properties:
  - stability
  - minimum phase
  - passivity

A G is stable, if all poles of G are in $\mathbb{C}^-$. A system (A, B, C, D) or A is stable, if all eigenvalues of A have a negative real part. Compare: $G(s) = C(sI - A)^{-1}B$

Model Reduction by Projection

Goals

Requirements: A Model Reduction approach should:

- Automatically generate compact models $\hat{G}$ from a given model G
- Satisfy desired error tolerance tol for all admissible input signals u

$$\|y - \hat{y}\| = \|Gu - \hat{G}u\| < \mathrm{tol} \cdot ||u|| \quad \text{for all admissable } u.$$

$\Longrightarrow$ Provide computable error bound/estimate!

- Preserve physical properties:
  - stability
  - minimum phase
  - passivity

A system G has minimum phase if all zeros of G are in the left half-plane $\mathbb{C}^-$.

Model Reduction by Projection

Goals

Requirements: A Model Reduction approach should:

- Automatically generate compact models $\hat{G}$ from a given model G
- Satisfy desired error tolerance tol for all admissible input signals u

$$\|y - \hat{y}\| = \|Gu - \hat{G}u\| < \text{tol} \cdot ||u|| \quad \text{for all admissable } u.$$

$\Longrightarrow$ Provide computable error bound/estimate!

- Preserve physical properties:
  - stability
  - minimum phase
  - passivity

A system G is passive if, bluntly speaking, it does not generate energy. Condition for passivity:

$$\int_{-\infty}^{t} u(\tau)^T y(\tau)\, d\tau \geq 0 \quad \text{for all } t \in \mathbb{R}, \quad \text{for all } u \in L_2(\mathbb{R}, \mathbb{R}^m).$$

Model Reduction by Projection

Projection Basics

Definition

A projector $P\colon \mathcal{X} \to \mathcal{X}$ is a linear map, or a matrix, with $P^2 = P$.

Example

- $\mathcal{X} = \mathbb{R}^2$
- $P = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ is a projector in $\mathcal{X}$

Model Reduction by Projection

Notion and Properties of Projectors

- A projector is a linear map $P\colon \mathcal{X} \to \mathcal{X}$ with $P^2 = P$.
- If $\mathcal{X} = \mathbb{R}^n$, a projector is a matrix $P \in \mathbb{R}^{n\times n}$ with $P^2 = P$.
- Let $\mathcal{V} = \operatorname{range}(P)$, then P is called a projector onto $\mathcal{V}$.
- If $\{v_1, \dots, v_r\}$ is a basis of some $\mathcal{V} \in \mathcal{X}$ and $V = [\, v_1, \dots, v_r \,]$, then

  $$P := V(V^T V)^{-1} V^T$$

  defines the orthogonal projector onto $\mathcal{V}$.
- If $\mathcal{W} \subset \mathcal{X}$ is another r-dimensional subspace with a basis matrix $W = [\, w_1, \dots, w_r \,]$ so that $W^T V$ is not singular, then

  $$P = V(W^T V)^{-1} W^T$$

  defines the oblique projector onto $\mathcal{V}$ along the orthogonal complement $\mathcal{W}_\perp$ of $\mathcal{W}$.
- For a projector P, the projector $I - P$ onto $\ker P$ is the complementary projector.

Model Reduction by Projection

Notion and Properties of Projectors

- A projector is a linear map $P\colon \mathcal{X} \to \mathcal{X}$ with $P^2 = P$.
- If $\mathcal{X} = \mathbb{R}^n$, a projector is a matrix $P \in \mathbb{R}^{n\times n}$ with $P^2 = P$.
- Let $\mathcal{V} = \operatorname{range}(P)$, then P is called a projector onto $\mathcal{V}$.
- If $\{v_1, \dots, v_r\}$ is a basis of some $\mathcal{V} \in \mathcal{X}$ and $V = [\, v_1, \dots, v_r \,]$, then

$$P := V(V^T V)^{-1} V^T$$

  defines the orthogonal projector onto $\mathcal{V}$.
- If $\mathcal{W} \subset \mathcal{X}$ is another r-dimensional subspace with a basis matrix $W = [\, w_1, \dots, w_r \,]$ so that $W^T V$ is not singular, then

$$P = V(W^T V)^{-1} W^T$$

  defines the oblique projector onto $\mathcal{V}$ along the orthogonal complement $\mathcal{W}_\perp$ of $\mathcal{W}$.
- For a projector P, the projector $I - P$ onto $\ker P$ is the complementary projector.

Model Reduction by Projection

Notion and Properties of Projectors

- A projector is a linear map $P\colon \mathcal{X} \to \mathcal{X}$ with $P^2 = P$.
- If $\mathcal{X} = \mathbb{R}^n$, a projector is a matrix $P \in \mathbb{R}^{n\times n}$ with $P^2 = P$.
- Let $\mathcal{V} = \mathrm{range}\,(P)$, then P is called a projector onto $\mathcal{V}$.
- If $\{v_1, \ldots, v_r\}$ is a basis of some $\mathcal{V} \in \mathcal{X}$ and $V = [\,v_1, \ldots, v_r\,]$, then

$$P := V(V^TV)^{-1}V^T$$

  defines the orthogonal projector onto $\mathcal{V}$.
- If $\mathcal{W} \subset \mathcal{X}$ is another r-dimensional subspace with a basis matrix $W = [\,w_1, \ldots, w_r\,]$ so that W^TV is not singular, then

$$P = V(W^TV)^{-1}W^T$$

  defines the oblique projector onto $\mathcal{V}$ along the orthogonal complement $\mathcal{W}_\perp$ of $\mathcal{W}$.
- For a projector P, the projector $I - P$ onto $\ker P$ is the complementary projector.

Model Reduction by Projection

Notion and Properties of Projectors

- A projector is a linear map $P\colon \mathcal{X} \to \mathcal{X}$ with $P^2 = P$.
- If $\mathcal{X} = \mathbb{R}^n$, a projector is a matrix $P \in \mathbb{R}^{n\times n}$ with $P^2 = P$.
- Let $\mathcal{V} = \operatorname{range}(P)$, then P is called a projector onto $\mathcal{V}$.
- If $\{v_1, \dots, v_r\}$ is a basis of some $\mathcal{V} \in \mathcal{X}$ and $V = [\, v_1, \dots, v_r \,]$, then

$$P := V(V^T V)^{-1} V^T$$

defines the orthogonal projector onto $\mathcal{V}$.

- If $\mathcal{W} \subset \mathcal{X}$ is another r-dimensional subspace with a basis matrix $W = [\, w_1, \dots, w_r \,]$ so that $W^T V$ is not singular, then

$$P = V(W^T V)^{-1} W^T$$

defines the oblique projector onto $\mathcal{V}$ along the orthogonal complement $\mathcal{W}_\perp$ of $\mathcal{W}$.

- For a projector P, the projector $I - P$ onto $\ker P$ is the complementary projector.

Model Reduction by Projection

Notion and Properties of Projectors

- A projector is a linear map $P\colon \mathcal{X} \to \mathcal{X}$ with $P^2 = P$.
- If $\mathcal{X} = \mathbb{R}^n$, a projector is a matrix $P \in \mathbb{R}^{n\times n}$ with $P^2 = P$.
- Let $\mathcal{V} = \operatorname{range}(P)$, then P is called a projector onto $\mathcal{V}$.
- If $\{v_1, \dots, v_r\}$ is a basis of some $\mathcal{V} \in \mathcal{X}$ and $V = [\, v_1, \dots, v_r \,]$, then

$$P := V(V^T V)^{-1} V^T$$

  defines the orthogonal projector onto $\mathcal{V}$.
- If $\mathcal{W} \subset \mathcal{X}$ is another r-dimensional subspace with a basis matrix $W = [\, w_1, \dots, w_r \,]$ so that $W^T V$ is not singular, then

$$P = V(W^T V)^{-1} W^T$$

  defines the oblique projector onto $\mathcal{V}$ along the orthogonal complement $\mathcal{W}_\perp$ of $\mathcal{W}$.
- For a projector P, the projector $I - P$ onto $\ker P$ is the complementary projector.

Model Reduction by Projection

Notion and Properties of Projectors

- A projector is a linear map $P\colon \mathcal{X} \to \mathcal{X}$ with $P^2 = P$.
- If $\mathcal{X} = \mathbb{R}^n$, a projector is a matrix $P \in \mathbb{R}^{n\times n}$ with $P^2 = P$.
- Let $\mathcal{V} = \mathrm{range}\,(P)$, then P is called a projector onto $\mathcal{V}$.
- If $\{v_1, \dots, v_r\}$ is a basis of some $\mathcal{V} \in \mathcal{X}$ and $V = [\, v_1, \dots, v_r \,]$, then

$$P := V(V^T V)^{-1} V^T$$

  defines the orthogonal projector onto $\mathcal{V}$.
- If $\mathcal{W} \subset \mathcal{X}$ is another r-dimensional subspace with a basis matrix $W = [\, w_1, \dots, w_r \,]$ so that $W^T V$ is not singular, then

$$P = V(W^T V)^{-1} W^T$$

  defines the oblique projector onto $\mathcal{V}$ along the orthogonal complement $\mathcal{W}_\perp$ of $\mathcal{W}$.
- For a projector P, the projector $I - P$ onto $\ker P$ is the complementary projector.

Model Reduction by Projection
Projection and Interpolation

Methods:

1. Modal Truncation
2. Rational Interpolation (Padé-Approximation and (rational) Krylov Subspace Methods)
3. Balanced Truncation
4. many more...

Joint feature of these methods:
computation of reduced-order model (ROM) by projection!

Model Reduction by Projection

computation of reduced-order model (ROM) by projection!

The ideal model reduction

- There is a space $\mathcal{V} \subset \mathbb{R}^n$ with $\dim \mathcal{V} = r < n$, such that $x \in \mathcal{V}$ for all time t and input u.
- Take a space $\mathcal{W}$, so that $\mathcal{W}_\perp \oplus \mathcal{V} = \mathbb{R}^n$.
- Galerkin-type projections: $\mathcal{W} = \mathcal{V}$.
- Petrov-Galerkin projections: $\mathcal{W} \neq \mathcal{V}$.
- Take matrices V and W that form bases of $\mathcal{V}$ and $\mathcal{W}$, with

$$W^\mathsf{T} V = I_r$$

- Then $V(W^\mathsf{T} V)^{-1} W = VW^\mathsf{T}$ is a projector onto $\mathcal{V}$
- Define $\hat{x} := W^\mathsf{T} x \in \mathbb{R}^r$ and define $\tilde{x} := V\hat{x} = VW^\mathsf{T} x$
- If everything is exact, then

$$\|x - \tilde{x}\| = \|x - VW^\mathsf{T} x\| = 0$$

- and given (A, B, C, D), the reduced-order model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is

$$\hat{A} := W^T AV, \quad \hat{B} := W^T B, \quad \hat{C} := CV, \quad (\hat{D} := D).$$

Model Reduction by Projection

computation of reduced-order model (ROM) by projection!

The ideal model reduction

- There is a space $\mathcal{V} \subset \mathbb{R}^n$ with $\dim \mathcal{V} = r < n$, such that $x \in \mathcal{V}$ for all time t and input u.
- Take a space $\mathcal{W}$, so that $\mathcal{W}_\perp \oplus \mathcal{V} = \mathbb{R}^n$.
- Galerkin-type projections: $\mathcal{W} = \mathcal{V}$.
- Petrov-Galerkin projections: $\mathcal{W} \neq \mathcal{V}$.
- Take matrices V and W that form bases of $\mathcal{V}$ and $\mathcal{W}$, with

$$W^\mathsf{T} V = I_r$$

- Then $V(W^\mathsf{T} V)^{-1} W = VW^\mathsf{T}$ is a projector onto $\mathcal{V}$
- Define $\hat{x} := W^\mathsf{T} x \in \mathbb{R}^r$ and define $\tilde{x} := V\hat{x} = VW^\mathsf{T} x$
- If everything is exact, then

$$\|x - \tilde{x}\| = \|x - VW^\mathsf{T} x\| = 0$$

- and given (A, B, C, D), the reduced-order model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is

$$\hat{A} := W^T AV, \quad \hat{B} := W^T B, \quad \hat{C} := CV, \quad (\hat{D} := D).$$

Model Reduction by Projection

computation of reduced-order model (ROM) by projection!

The ideal model reduction

- There is a space $\mathcal{V} \subset \mathbb{R}^n$ with $\dim \mathcal{V} = r < n$, such that $x \in \mathcal{V}$ for all time t and input u.
- Take a space $\mathcal{W}$, so that $\mathcal{W}_\perp \oplus \mathcal{V} = \mathbb{R}^n$.
- Galerkin-type projections: $\mathcal{W} = \mathcal{V}$.
- Petrov-Galerkin projections: $\mathcal{W} \neq \mathcal{V}$.
- Take matrices V and W that form bases of $\mathcal{V}$ and $\mathcal{W}$, with

$$W^\mathsf{T} V = I_r$$

- Then $V(W^\mathsf{T}V)^{-1}W = VW^\mathsf{T}$ is a projector onto $\mathcal{V}$
- Define $\hat{x} := W^\mathsf{T}x \in \mathbb{R}^r$ and define $\tilde{x} := V\hat{x} = VW^\mathsf{T}x$
- If everything is exact, then

$$\|x - \tilde{x}\| = \|x - VW^\mathsf{T}x\| = 0$$

- and given (A, B, C, D), the reduced-order model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is

$$\hat{A} := W^T AV, \quad \hat{B} := W^T B, \quad \hat{C} := CV, \quad (\hat{D} := D).$$

Model Reduction by Projection

computation of reduced-order model (ROM) by projection!

The ideal model reduction

- There is a space $\mathcal{V} \subset \mathbb{R}^n$ with $\dim \mathcal{V} = r < n$, such that $x \in \mathcal{V}$ for all time t and input u.
- Take a space $\mathcal{W}$, so that $\mathcal{W}_\perp \oplus \mathcal{V} = \mathbb{R}^n$.
- Galerkin-type projections: $\mathcal{W} = \mathcal{V}$.
- Petrov-Galerkin projections: $\mathcal{W} \neq \mathcal{V}$.
- Take matrices V and W that form bases of $\mathcal{V}$ and $\mathcal{W}$, with

$$W^\mathsf{T} V = I_r$$

- Then $V(W^\mathsf{T} V)^{-1} W = VW^\mathsf{T}$ is a projector onto $\mathcal{V}$
- Define $\hat{x} := W^\mathsf{T} x \in \mathbb{R}^r$ and define $\tilde{x} := V\hat{x} = VW^\mathsf{T} x$
- If everything is exact, then

$$\|x - \tilde{x}\| = \|x - VW^\mathsf{T} x\| = 0$$

- and given (A, B, C, D), the reduced-order model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is

$$\hat{A} := W^T AV, \quad \hat{B} := W^T B, \quad \hat{C} := CV, \quad (\hat{D} := D).$$

Model Reduction by Projection

computation of reduced-order model (ROM) by projection!

The ideal model reduction

- There is a space $\mathcal{V} \subset \mathbb{R}^n$ with $\dim \mathcal{V} = r < n$, such that $x \in \mathcal{V}$ for all time t and input u.
- Take a space $\mathcal{W}$, so that $\mathcal{W}_\perp \oplus \mathcal{V} = \mathbb{R}^n$.
- Galerkin-type projections: $\mathcal{W} = \mathcal{V}$.
- Petrov-Galerkin projections: $\mathcal{W} \neq \mathcal{V}$.
- Take matrices V and W that form bases of $\mathcal{V}$ and $\mathcal{W}$, with

$$W^\mathsf{T} V = I_r$$

- Then $V(W^\mathsf{T} V)^{-1} W = VW^\mathsf{T}$ is a projector onto $\mathcal{V}$
- Define $\hat{x} := W^\mathsf{T} x \in \mathbb{R}^r$ and define $\tilde{x} := V\hat{x} = VW^\mathsf{T} x$
- If everything is exact, then

$$\|x - \tilde{x}\| = \|x - VW^\mathsf{T} x\| = 0$$

- and given (A, B, C, D), the reduced-order model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is

$$\hat{A} := W^T AV, \quad \hat{B} := W^T B, \quad \hat{C} := CV, \quad (\hat{D} := D).$$

Model Reduction by Projection

computation of reduced-order model (ROM) by projection!

The ideal model reduction

- There is a space $\mathcal{V} \subset \mathbb{R}^n$ with $\dim \mathcal{V} = r < n$, such that $x \in \mathcal{V}$ for all time t and input u.
- Take a space $\mathcal{W}$, so that $\mathcal{W}_\perp \oplus \mathcal{V} = \mathbb{R}^n$.
- Galerkin-type projections: $\mathcal{W} = \mathcal{V}$.
- Petrov-Galerkin projections: $\mathcal{W} \neq \mathcal{V}$.
- Take matrices V and W that form bases of $\mathcal{V}$ and $\mathcal{W}$, with

$$W^\mathsf{T} V = I_r$$

- Then $V(W^\mathsf{T} V)^{-1} W = VW^\mathsf{T}$ is a projector onto $\mathcal{V}$
- Define $\hat{x} := W^\mathsf{T} x \in \mathbb{R}^r$ and define $\tilde{x} := V\hat{x} = VW^\mathsf{T} x$
- If everything is exact, then

$$\|x - \tilde{x}\| = \|x - VW^\mathsf{T} x\| = 0$$

- and given (A, B, C, D), the reduced-order model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is

$$\hat{A} := W^T AV, \quad \hat{B} := W^T B, \quad \hat{C} := CV, \quad (\hat{D} := D).$$

Model Reduction by Projection

computation of reduced-order model (ROM) by projection!

The ideal model reduction

- There is a space $\mathcal{V} \subset \mathbb{R}^n$ with $\dim \mathcal{V} = r < n$, such that $x \in \mathcal{V}$ for all time t and input u.
- Take a space $\mathcal{W}$, so that $\mathcal{W}_\perp \oplus \mathcal{V} = \mathbb{R}^n$.
- Galerkin-type projections: $\mathcal{W} = \mathcal{V}$.
- Petrov-Galerkin projections: $\mathcal{W} \neq \mathcal{V}$.
- Take matrices V and W that form bases of $\mathcal{V}$ and $\mathcal{W}$, with

$$W^\mathsf{T} V = I_r$$

- Then $V(W^\mathsf{T} V)^{-1} W = VW^\mathsf{T}$ is a projector onto $\mathcal{V}$
- Define $\hat{x} := W^\mathsf{T} x \in \mathbb{R}^r$ and define $\tilde{x} := V\hat{x} = VW^\mathsf{T} x$
- If everything is exact, then

$$\|x - \tilde{x}\| = \|x - VW^\mathsf{T} x\| = 0$$

- and given (A, B, C, D), the reduced-order model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is

$$\hat{A} := W^T AV, \quad \hat{B} := W^T B, \quad \hat{C} := CV, \quad (\hat{D} := D).$$

Model Reduction by Projection

computation of reduced-order model (ROM) by projection!

The ideal model reduction

- There is a space $\mathcal{V} \subset \mathbb{R}^n$ with $\dim \mathcal{V} = r < n$, such that $x \in \mathcal{V}$ for all time t and input u.
- Take a space $\mathcal{W}$, so that $\mathcal{W}_\perp \oplus \mathcal{V} = \mathbb{R}^n$.
- Galerkin-type projections: $\mathcal{W} = \mathcal{V}$.
- Petrov-Galerkin projections: $\mathcal{W} \neq \mathcal{V}$.
- Take matrices V and W that form bases of $\mathcal{V}$ and $\mathcal{W}$, with

$$W^\mathsf{T} V = I_r$$

- Then $V(W^\mathsf{T} V)^{-1} W = VW^\mathsf{T}$ is a projector onto $\mathcal{V}$
- Define $\hat{x} := W^\mathsf{T} x \in \mathbb{R}^r$ and define $\tilde{x} := V\hat{x} = VW^\mathsf{T} x$
- If everything is exact, then

$$\|x - \tilde{x}\| = \|x - VW^\mathsf{T} x\| = 0$$

- and given (A, B, C, D), the reduced-order model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is

$$\hat{A} := W^T AV, \quad \hat{B} := W^T B, \quad \hat{C} := CV, \quad (\hat{D} := D).$$

Model Reduction by Projection

computation of reduced-order model (ROM) by projection!

The ideal model reduction

- There is a space $\mathcal{V} \subset \mathbb{R}^n$ with $\dim \mathcal{V} = r < n$, such that $x \in \mathcal{V}$ for all time t and input u.
- Take a space $\mathcal{W}$, so that $\mathcal{W}_\perp \oplus \mathcal{V} = \mathbb{R}^n$.
- Galerkin-type projections: $\mathcal{W} = \mathcal{V}$.
- Petrov-Galerkin projections: $\mathcal{W} \neq \mathcal{V}$.
- Take matrices V and W that form bases of $\mathcal{V}$ and $\mathcal{W}$, with

$$W^\mathsf{T} V = I_r$$

- Then $V(W^\mathsf{T} V)^{-1} W = VW^\mathsf{T}$ is a projector onto $\mathcal{V}$
- Define $\hat{x} := W^\mathsf{T} x \in \mathbb{R}^r$ and define $\tilde{x} := V\hat{x} = VW^\mathsf{T} x$
- If everything is exact, then

$$\|x - \tilde{x}\| = \|x - VW^\mathsf{T} x\| = 0$$

- and given (A, B, C, D), the reduced-order model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is

$$\hat{A} := W^T AV, \quad \hat{B} := W^T B, \quad \hat{C} := CV, \quad (\hat{D} := D).$$

Model Reduction by Projection

computation of reduced-order model (ROM) by projection!

Model reduction in practise

- Assume that there is a space $\mathcal{V} \subset \mathbb{R}^n$ with $\dim \mathcal{V} = r < n$, such that $x \in \mathcal{V}$ for all time t and input u.
- Take a space $\mathcal{W}$, so that $\mathcal{W}_\perp \oplus \mathcal{V} = \mathbb{R}^n$.
- Galerkin-type projections: $\mathcal{W} = \mathcal{V}$.
- Petrov-Galerkin projections: $\mathcal{W} \neq \mathcal{V}$.
- Find matrices V and W that approximate bases of $\mathcal{V}$ and $\mathcal{W}$, with

$$W^\mathsf{T} V = I_r$$

- Then $V(W^\mathsf{T} V)^{-1} W = VW^\mathsf{T}$ is a projector almost onto $\mathcal{V}$
- Define $\hat{x} := W^\mathsf{T} x \in \mathbb{R}^r$ and define $\tilde{x} := V\hat{x} = VW^\mathsf{T} x$
- If everything is done well, then

$$\|x - \tilde{x}\| = \|x - VW^\mathsf{T} x\| \approx 0$$

- and given (A, B, C, D), the reduced-order model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is

$$\hat{A} := W^T AV, \quad \hat{B} := W^T B, \quad \hat{C} := CV, \quad (\hat{D} := D).$$

Model Reduction by Projection

Definition of the reduced model

... and given an (A, B, C, D) system,

the reduced-order model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is

$$\hat{A} := W^T AV, \quad \hat{B} := W^T B, \quad \hat{C} := CV, \quad (\hat{D} := D).$$

Why is the ROM defined like this:
It is the (Petrov)-Galerkin condition $\dot{\tilde{x}} - A\tilde{x} - Bu \perp \mathcal{W}$:

$$W^T \left(\dot{\tilde{x}} - A\tilde{x} - Bu\right) = \qquad W^T \left(VW^T\dot{x} - AVW^T x - Bu\right)$$

is zero, if, and only if,

$$\dot{\hat{x}} - \hat{A}\hat{x} - \hat{B}u = 0.$$

Model Reduction by Projection

Definition of the reduced model

…and given an (A, B, C, D) system,

the reduced-order model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is

$$\hat{A} := W^T AV, \quad \hat{B} := W^T B, \quad \hat{C} := CV, \quad (\hat{D} := D).$$

Why is the ROM defined like this:

It is the (Petrov)-Galerkin condition $\dot{\tilde{x}} - A\tilde{x} - Bu \perp \mathcal{W}$:

$$\begin{aligned} W^T \left(\dot{\tilde{x}} - A\tilde{x} - Bu\right) &= W^T \left(VW^T \dot{x} - AVW^T x - Bu\right) \\ &= \underbrace{W^T \dot{x}}_{\dot{\hat{x}}} - \underbrace{W^T AV}_{=\hat{A}} \underbrace{W^T x}_{=\hat{x}} - \underbrace{W^T B}_{=\hat{B}} u \end{aligned}$$

is zero, if, and only if,

$$\dot{\hat{x}} - \hat{A}\hat{x} - \hat{B}u = 0.$$

Model Reduction by Projection

Definition of the reduced model

. . . and given an (A, B, C, D) system,

the reduced-order model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is

$$\hat{A} := W^T AV, \quad \hat{B} := W^T B, \quad \hat{C} := CV, \quad (\hat{D} := D).$$

Why is the ROM defined like this:

It is the (Petrov)-Galerkin condition $\dot{\tilde{x}} - A\tilde{x} - Bu \perp \mathcal{W}$:

$$\begin{aligned} W^T \left(\dot{\tilde{x}} - A\tilde{x} - Bu\right) &= W^T \left(VW^T\dot{x} - AVW^T x - Bu\right) \\ &= \underbrace{W^T\dot{x}}_{\dot{\hat{x}}} - \underbrace{W^T AV}_{=\hat{A}} \underbrace{W^T x}_{=\hat{x}} - \underbrace{W^T B}_{=\hat{B}} u \end{aligned}$$

is zero, if, and only if,

$$\dot{\hat{x}} - \hat{A}\hat{x} - \hat{B}u = 0.$$

Model Reduction by Projection

Projection $\rightsquigarrow$ Rational Interpolation

A Petrov-Galerkin projected ROM interpolates the transfer function:

Theorem 3.3 [Grimme '97, Villemagne/Skelton '87]

Given the ROM

$$\hat{A} = W^T AV, \quad \hat{B} = W^T B, \quad \hat{C} = CV, \quad (\hat{D} = D),$$

and $s_* \in \mathbb{C} \setminus (\Lambda(A) \cup \Lambda(\hat{A}))$, if either

- $(s_* I_n - A)^{-1} B \in \mathrm{range}(V)$, or
- $(s_* I_n - A)^{-*} C^T \in \mathrm{range}(W)$,

then the interpolation condition

$$G(s_*) = \hat{G}(s_*).$$

in s_* holds.

Note: extension to Hermite interpolation conditions later!

Model Reduction by Projection

Projection $\rightsquigarrow$ Rational Interpolation

Given the ROM

$$\hat{A} = W^T AV, \quad \hat{B} = W^T B, \quad \hat{C} = CV, \quad (\hat{D} = D),$$

the error transfer function can be written as

$$G(s) - \hat{G}(s) = \quad \left(C(sI_n - A)^{-1}B + D\right) - \left(\hat{C}(sI_r - \hat{A})^{-1}\hat{B} + \hat{D}\right)$$

Model Reduction by Projection

Projection $\rightsquigarrow$ Rational Interpolation

Given the ROM

$$\hat{A} = W^T AV, \quad \hat{B} = W^T B, \quad \hat{C} = CV, \quad (\hat{D} = D),$$

the error transfer function can be written as

$$\begin{aligned} G(s) - \hat{G}(s) &= \left(C(sI_n - A)^{-1}B + D\right) - \left(\hat{C}(sI_r - \hat{A})^{-1}\hat{B} + \hat{D}\right) \\ &= C\left((sI_n - A)^{-1} - V(sI_r - \hat{A})^{-1}W^T\right)B \end{aligned}$$

Model Reduction by Projection

Projection $\rightsquigarrow$ Rational Interpolation

Given the ROM

$$\hat{A} = W^T AV, \quad \hat{B} = W^T B, \quad \hat{C} = CV, \quad (\hat{D} = D),$$

the error transfer function can be written as

$$\begin{aligned} G(s) - \hat{G}(s) &= \left(C(sI_n - A)^{-1}B + D\right) - \left(\hat{C}(sI_r - \hat{A})^{-1}\hat{B} + \hat{D}\right) \\ &= C\left((sI_n - A)^{-1} - V(sI_r - \hat{A})^{-1}W^T\right)B \\ &= C\left(I_n - \underbrace{V(sI_r - \hat{A})^{-1}W^T(sI_n - A)}_{=:P(s)}\right)(sI_n - A)^{-1}B. \end{aligned}$$

Model Reduction by Projection

Projection $\rightsquigarrow$ Rational Interpolation

Given the ROM

$$\hat{A} = W^T AV, \quad \hat{B} = W^T B, \quad \hat{C} = CV, \quad (\hat{D} = D),$$

the error transfer function can be written as

$$\begin{aligned} G(s) - \hat{G}(s) &= \left(C(sI_n - A)^{-1}B + D\right) - \left(\hat{C}(sI_r - \hat{A})^{-1}\hat{B} + \hat{D}\right) \\ &= C\left(I_n - \underbrace{V(sI_r - \hat{A})^{-1}W^T(sI_n - A)}_{=:P(s)}\right)(sI_n - A)^{-1}B. \end{aligned}$$

If $s_* \in \mathbb{C} \setminus (\Lambda(A) \cup \Lambda(\hat{A}))$, then $P(s_*)$ is a projector onto $\mathcal{V}$:

$\mathrm{range}\,(P(s_*)) \subset \mathrm{range}\,(V)$, all matrices have full rank $\Rightarrow$ "=",

$$P(s_*)^2 = V(s_*I_r - \hat{A})^{-1}W^T(s_*I_n - A)V(s_*I_r - \hat{A})^{-1}W^T(s_*I_n - A)$$

Model Reduction by Projection

Projection $\rightsquigarrow$ Rational Interpolation

Given the ROM

$$\hat{A} = W^T A V, \quad \hat{B} = W^T B, \quad \hat{C} = CV, \quad (\hat{D} = D),$$

the error transfer function can be written as

$$G(s) - \hat{G}(s) = \left(C(sI_n - A)^{-1}B + D\right) - \left(\hat{C}(sI_r - \hat{A})^{-1}\hat{B} + \hat{D}\right)$$
$$= C\left(I_n - \underbrace{V(sI_r - \hat{A})^{-1}W^T(sI_n - A)}_{=:P(s)}\right)(sI_n - A)^{-1}B.$$

If $s_* \in \mathbb{C} \setminus (\Lambda(A) \cup \Lambda(\hat{A}))$, then $P(s_*)$ is a projector onto $\mathcal{V}$:

$\mathrm{range}\,(P(s_*)) \subset \mathrm{range}\,(V)$, all matrices have full rank $\Rightarrow$ "$=$",

$$P(s_*)^2 = V(s_* I_r - \hat{A})^{-1}W^T(s_* I_n - A)V(s_* I_r - \hat{A})^{-1}W^T(s_* I_n - A)$$
$$= V(s_* I_r - \hat{A})^{-1}\underbrace{(s_* I_r - \hat{A})(s_* I_r - \hat{A})^{-1}}_{=I_r}W^T(s_* I_n - A) = P(s_*).$$

Model Reduction by Projection

Projection $\rightsquigarrow$ Rational Interpolation

Given the ROM

$$\hat{A} = W^T AV, \quad \hat{B} = W^T B, \quad \hat{C} = CV, \quad (\hat{D} = D),$$

the error transfer function can be written as

$$\begin{aligned} G(s) - \hat{G}(s) = & \quad \left(C(sI_n - A)^{-1}B + D\right) - \left(\hat{C}(sI_r - \hat{A})^{-1}\hat{B} + \hat{D}\right) \\ = & \quad C\Big(I_n - \underbrace{V(sI_r - \hat{A})^{-1}W^T(sI_n - A)}_{=:P(s)}\Big)(sI_n - A)^{-1}B. \end{aligned}$$

If $s_* \in \mathbb{C} \setminus (\Lambda(A) \cup \Lambda(\hat{A}))$, then $P(s_*)$ is a projector onto $\mathcal{V}$ $\Longrightarrow$

if $(s_* I_n - A)^{-1}B \in \mathcal{V}$, *then* $(I_n - P(s_*))(s_* I_n - A)^{-1}B = 0$,

hence

$$G(s_*) - \hat{G}(s_*) = 0 \;\Rightarrow\; G(s_*) = \hat{G}(s_*), \text{ i.e., } \hat{G} \text{ interpolates } G \text{ in } s_*!$$

Model Reduction by Projection

Projection $\rightsquigarrow$ Rational Interpolation

Given the ROM

$$\hat{A} = W^T A V, \quad \hat{B} = W^T B, \quad \hat{C} = CV, \quad (\hat{D} = D),$$

the error transfer function can be written as

$$\begin{aligned} G(s) - \hat{G}(s) &= \left(C(sI_n - A)^{-1}B + D\right) - \left(\hat{C}(sI_r - \hat{A})^{-1}\hat{B} + \hat{D}\right) \\ &= C\Big(I_n - \underbrace{V(sI_r - \hat{A})^{-1}W^T(sI_n - A)}_{=:P(s)}\Big)(sI_n - A)^{-1}B. \end{aligned}$$

Analogously, $= C(sI_n - A)^{-1}\Big(I_n - \underbrace{(sI_n - A)V(sI_r - \hat{A})^{-1}W^T}_{=:Q(s)}\Big)B.$

If $s_* \in \mathbb{C} \setminus (\Lambda(A) \cup \Lambda(\hat{A}))$, then $Q(s)^H$ is a projector onto $\mathcal{W} \Longrightarrow$

if $(s_* I_n - A)^{-*}C^T \in \mathcal{W}$, *then* $C(s_* I_n - A)^{-1}(I_n - Q(s_*)) = 0,$

hence

$$G(s_*) - \hat{G}(s_*) = 0 \;\Rightarrow\; G(s_*) = \hat{G}(s_*), \text{ i.e., } \hat{G} \text{ interpolates } G \text{ in } s_*!$$

Model Reduction by Projection

Projection $\rightsquigarrow$ Rational Interpolation

A Petrov-Galerkin projected ROM interpolates the transfer function:

Theorem 3.3 [GRIMME '97, VILLEMAGNE/SKELTON '87]

Given the ROM

$$\hat{A} = W^T AV, \quad \hat{B} = W^T B, \quad \hat{C} = CV, \quad (\hat{D} = D),$$

and $s_* \in \mathbb{C} \setminus (\Lambda(A) \cup \Lambda(\hat{A}))$, if either

- $(s_* I_n - A)^{-1} B \in \mathrm{range}\,(V)$, or
- $(s_* I_n - A)^{-*} C^T \in \mathrm{range}\,(W)$,

then the interpolation condition

$$G(s_*) = \hat{G}(s_*).$$

in s_* holds.

Note: extension to Hermite interpolation conditions later!

Modal Truncation

Basic method:

Assume A is diagonalizable, $T^{-1}AT = D_A$, project state-space onto A-invariant subspace $\mathcal{V} = \mathrm{span}(t_1, \ldots, t_r)$, $t_k =$ eigenvectors corresp. to "dominant" modes / eigenvalues of A. Then with

$$V = T(:, 1:r) = [\, t_1, \ldots, t_r \,], \quad \tilde{W}^H = T^{-1}(1:r, :), \quad W = \tilde{W}(V^H \tilde{W})^{-1},$$

reduced-order model is

$$\hat{A} := W^H A V = \mathrm{diag}\,\{\lambda_1, \ldots, \lambda_r\}, \quad \hat{B} := W^H B, \quad \hat{C} = CV$$

Also computable by truncation:

$$T^{-1}AT = \begin{bmatrix} \hat{A} & \\ & A_2 \end{bmatrix}, \quad T^{-1}B = \begin{bmatrix} \hat{B} \\ B_2 \end{bmatrix}, \quad CT = [\, \hat{C}, \; C_2 \,], \quad \hat{D} = D.$$

Modal Truncation

Basic method:

Assume A is diagonalizable, $T^{-1}AT = D_A$, project state-space onto A-invariant subspace $\mathcal{V} = \operatorname{span}(t_1, \ldots, t_r)$, t_k = eigenvectors corresp. to "dominant" modes / eigenvalues of A. Then with

$$V = T(:, 1:r) = [\, t_1, \ldots, t_r \,], \quad \tilde{W}^H = T^{-1}(1:r, :), \quad W = \tilde{W}(V^H \tilde{W})^{-1},$$

reduced-order model is

$$\hat{A} := W^H A V = \operatorname{diag}\{\lambda_1, \ldots, \lambda_r\}, \quad \hat{B} := W^H B, \quad \hat{C} = CV$$

Also computable by truncation:

$$T^{-1}AT = \begin{bmatrix} \hat{A} & \\ & A_2 \end{bmatrix}, \quad T^{-1}B = \begin{bmatrix} \hat{B} \\ B_2 \end{bmatrix}, \quad CT = [\, \hat{C}, \ C_2 \,], \quad \hat{D} = D.$$

Properties:

Simple computation for large-scale systems, using, e.g., Krylov subspace methods (Lanczos, Arnoldi), Jacobi-Davidson method.

Modal Truncation

Basic method:

$$T^{-1}AT = \begin{bmatrix} \hat{A} & \\ & A_2 \end{bmatrix}, \quad T^{-1}B = \begin{bmatrix} \hat{B} \\ B_2 \end{bmatrix}, \quad CT = [\, \hat{C},\ C_2 \,], \quad \hat{D} = D.$$

Properties:

Error bound:

$$\left|\left| G - \hat{G} \right|\right|_\infty \leq ||C_2||\,||B_2||\,\frac{1}{\min_{\lambda \in \Lambda(A_2)} |\mathrm{Re}(\lambda)|}.$$

Proof:

$$\begin{aligned} G(s) &= C(sI - A)^{-1}B + D = CTT^{-1}(sI - A)^{-1}TT^{-1}B + D \\ &= CT(sI - T^{-1}AT)^{-1}T^{-1}B + D \\ &= [\, \hat{C},\ C_2 \,] \begin{bmatrix} (sI_r - \hat{A})^{-1} & \\ & (sI_{n-r} - A_2)^{-1} \end{bmatrix} \begin{bmatrix} \hat{B} \\ B_2 \end{bmatrix} + D \\ &= \hat{G}(s) + C_2(sI_{n-r} - A_2)^{-1}B_2, \end{aligned}$$

Modal Truncation

Basic method:

$$T^{-1}AT = \begin{bmatrix} \hat{A} & \\ & A_2 \end{bmatrix}, \quad T^{-1}B = \begin{bmatrix} \hat{B} \\ B_2 \end{bmatrix}, \quad CT = [\,\hat{C},\ C_2\,], \quad \hat{D} = D.$$

Properties:

Error bound:

$$\left\|G - \hat{G}\right\|_\infty \leq \|C_2\|\,\|B_2\| \frac{1}{\min_{\lambda \in \Lambda(A_2)} |\mathrm{Re}(\lambda)|}.$$

Proof:

$$G(s) \quad = \quad \hat{G}(s) + C_2(sI_{n-r} - A_2)^{-1}B_2,$$

observing that $\left\|G - \hat{G}\right\|_\infty = \sup_{\omega \in \mathbb{R}} \sigma_{\max}(C_2(\jmath\omega I_{n-r} - A_2)^{-1}B_2)$, and

$$C_2(\jmath\omega I_{n-r} - A_2)^{-1}B_2 = C_2 \mathrm{diag}\left(\frac{1}{\jmath\omega - \lambda_{r+1}}, \ldots, \frac{1}{\jmath\omega - \lambda_n}\right) B_2.$$

Modal Truncation

Basic method:

Assume A is diagonalizable, $T^{-1}AT = D_A$, project state-space onto A-invariant subspace $\mathcal{V} = \mathrm{span}(t_1, \ldots, t_r)$, t_k = eigenvectors corresp. to "dominant" modes / eigenvalues of A. Then reduced-order model is

$$\hat{A} := W^H A V = \mathrm{diag}\{\lambda_1, \ldots, \lambda_r\}, \quad \hat{B} := W^H B, \quad \hat{C} = CV$$

Also computable by truncation:

$$T^{-1}AT = \begin{bmatrix} \hat{A} & \\ & A_2 \end{bmatrix}, \quad T^{-1}B = \begin{bmatrix} \hat{B} \\ B_2 \end{bmatrix}, \quad CT = [\,\hat{C},\ C_2\,], \quad \hat{D} = D.$$

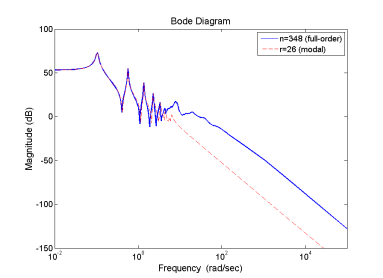
Difficulties:

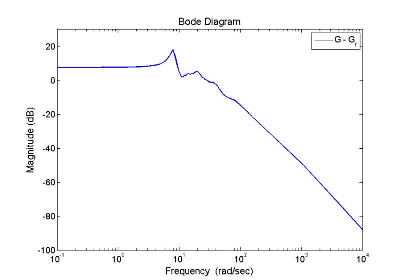
- Eigenvalues contain only limited system information.
- Dominance measures are difficult to compute.
  ([Litz '79] use Jordan canoncial form; otherwise merely heuristic criteria, e.g., [Varga '95]. Recent improvement: dominant pole algorithm.)
- Error bound not computable for really large-scale problems.

Modal Truncation
Example

BEAM, SISO system from SLICOT Benchmark Collection for Model Reduction, $n = 348$, $m = q = 1$, reduced using 13 dominant complex conjugate eigenpairs, error bound yields $\left\|G - \hat{G}\right\|_\infty \leq 1.21 \cdot 10^3$

Bode plots of transfer functions and error function





Modal Truncation

Extensions

Base enrichment

Static modes are defined by setting $\dot{x} = 0$ and assuming unit loads, i.e., $u(t) \equiv e_j$, $j = 1, \dots, m$:

$$0 = Ax(t) + Be_j \quad \Longrightarrow \quad x(t) \equiv -A^{-1} b_j.$$

Projection subspace $\mathcal{V}$ is then augmented by $A^{-1}[\, b_1, \dots, b_m \,] = A^{-1}B$.

Interpolation-projection framework $\Longrightarrow$ $G(0) = \hat{G}(0)$!

If two sided projection is used, complimentary subspace can be augmented by $A^{-T}C^T \Longrightarrow G'(0) = \hat{G}'(0)$! (If $m \neq q$, add random vectors or delete some of the columns in $A^{-T}C^T$).

Modal Truncation

Extensions

Guyan reduction (static condensation)

Partition states in masters $x_1 \in \mathbb{R}^r$ and slaves $x_2 \in \mathbb{R}^{n-r}$ (FEM terminology)
Assume stationarity, i.e., $\dot{x} = 0$ and solve for x_2 in

$$\begin{aligned} 0 &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u \\ \Rightarrow \quad x_2 &= -A_{22}^{-1} A_{21} x_1 - A_{22}^{-1} B_2 u. \end{aligned}$$

Inserting this into the first part of the dynamic system

$$\dot{x}_1 = A_{11}x_1 + A_{12}x_2 + B_1 u, \quad y = C_1 x_1 + C_2 x_2$$

then yields the reduced-order model

$$\begin{aligned} \dot{x}_1 &= (A_{11} - A_{12}A_{22}^{-1}A_{21})x_1 + (B_1 - A_{12}A_{22}^{-1}B_2)u \\ y &= (C_1 - C_2 A_{22}^{-1} A_{21})x_1 - C_2 A_{22}^{-1} B_2 u. \end{aligned}$$

Modal Truncation

Dominant Poles

Pole-Residue Form of Transfer Function

Consider partial fraction expansion of transfer function with $D = 0$:

$$G(s) = \sum_{k=1}^{n} \frac{R_k}{s - \lambda_k}$$

with the residues $R_k := (Cx_k)(y_k^H B) \in \mathbb{C}^{q \times m}$.

Modal Truncation

Dominant Poles

Pole-Residue Form of Transfer Function

Consider partial fraction expansion of transfer function with $D = 0$:

$$G(s) = \sum_{k=1}^{n} \frac{R_k}{s - \lambda_k}$$

with the residues $R_k := (Cx_k)(y_k^H B) \in \mathbb{C}^{q\times m}$.

Note: this follows using the spectral decomposition $A = XDX^{-1}$, with $X = [\,x_1, \ldots, x_n]$ the right and $X^{-1} =: Y = [y_1, \ldots, y_n]^H$ the left eigenvector matrices:

$$\begin{aligned} G(s) &= C(sI - XDX^{-1})^{-1}B = CX(sI - \operatorname{diag}\{\lambda_1, \ldots, \lambda_n\})^{-1} YB \\ &= [\, Cx_1, \ldots, Cx_n\,] \begin{bmatrix} \frac{1}{s-\lambda_1} & & \\ & \ddots & \\ & & \frac{1}{s-\lambda_n} \end{bmatrix} \begin{bmatrix} y_1^H B \\ \vdots \\ y_n^H B \end{bmatrix} \\ &= \sum_{k=1}^{n} \frac{(Cx_k)(y_k^H B)}{s - \lambda_k}. \end{aligned}$$

Modal Truncation

Dominant Poles

Pole-Residue Form of Transfer Function

Consider partial fraction expansion of transfer function with $D = 0$:

$$G(s) = \sum_{k=1}^{n} \frac{R_k}{s - \lambda_k}$$

with the residues $R_k := (Cx_k)(y_k^H B) \in \mathbb{C}^{q \times m}$.

Note: $R_k = (Cx_k)(y_k^H B)$ are the residues of G in the sense of the residue theorem of complex analysis:

$$\mathrm{res}\,(G, \lambda_\ell) = \lim_{s \to \lambda_\ell} (s - \lambda_\ell) G(s) = \sum_{k=1}^{n} \underbrace{\lim_{s \to \lambda_\ell} \frac{s - \lambda_\ell}{s - \lambda_k}}_{= \begin{cases} 0 \text{ for } k \neq \ell \\ 1 \text{ for } k = \ell \end{cases}} R_k = R_\ell.$$

Modal Truncation

Dominant Poles

Pole-Residue Form of Transfer Function

Consider partial fraction expansion of transfer function with $D = 0$:

$$G(s) = \sum_{k=1}^{n} \frac{R_k}{s - \lambda_k}$$

with the residues $R_k := (Cx_k)(y_k^H B) \in \mathbb{C}^{q \times m}$.

As projection basis use spaces spanned by right/left eigenvectors corresponding to dominant poles, i.e.. $(\lambda_j,\ x_j,\ y_j)$ with largest

$$\|R_k\| / |\operatorname{re}(\lambda_k)|.$$

Modal Truncation

Dominant Poles

Pole-Residue Form of Transfer Function

Consider partial fraction expansion of transfer function with $D = 0$:

$$G(s) = \sum_{k=1}^{n} \frac{R_k}{s - \lambda_k}$$

with the residues $R_k := (Cx_k)(y_k^H B) \in \mathbb{C}^{q \times m}$.

As projection basis use spaces spanned by right/left eigenvectors corresponding to dominant poles, i.e.. $(\lambda_j,\ x_j,\ y_j)$ with largest

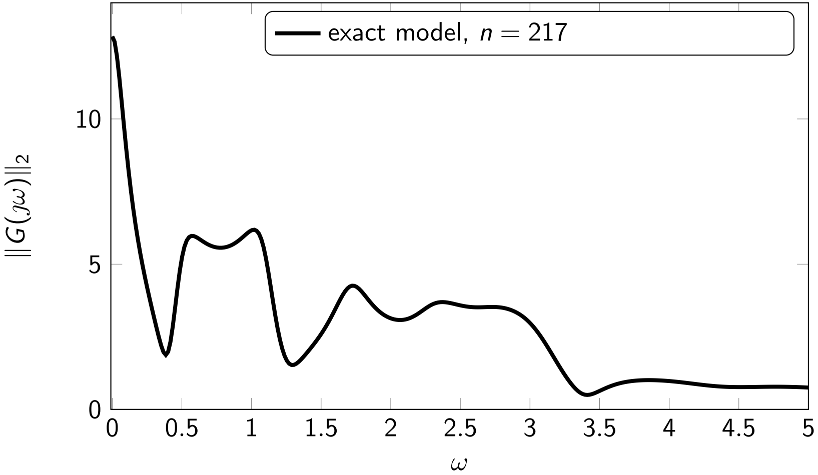
$$\|R_k\| / |\operatorname{re}(\lambda_k)|.$$

Remark

The dominant modes have most important influence on the input-output behavior of the system and are responsible for the "peaks"' in the frequency response.

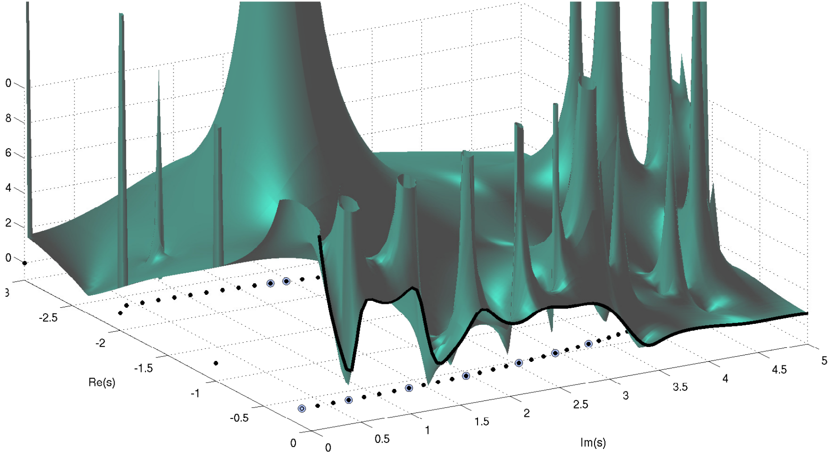
Dominant Poles

Random SISO Example ($B,\ C^T \in \mathbb{R}^n$)



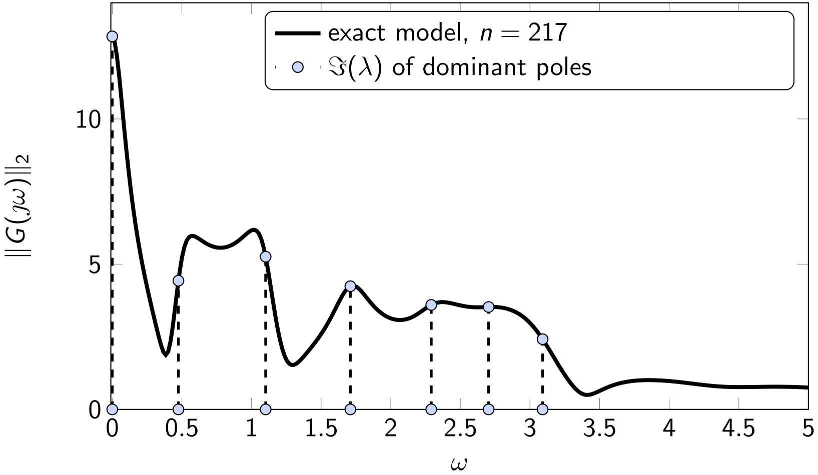
Dominant Poles

Random SISO Example ($B,\ C^T \in \mathbb{R}^n$)



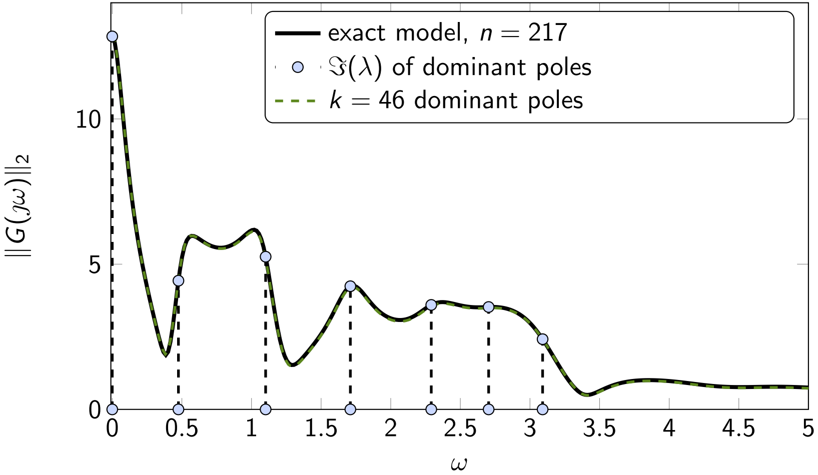
Dominant Poles

Random SISO Example ($B,\ C^T \in \mathbb{R}^n$)



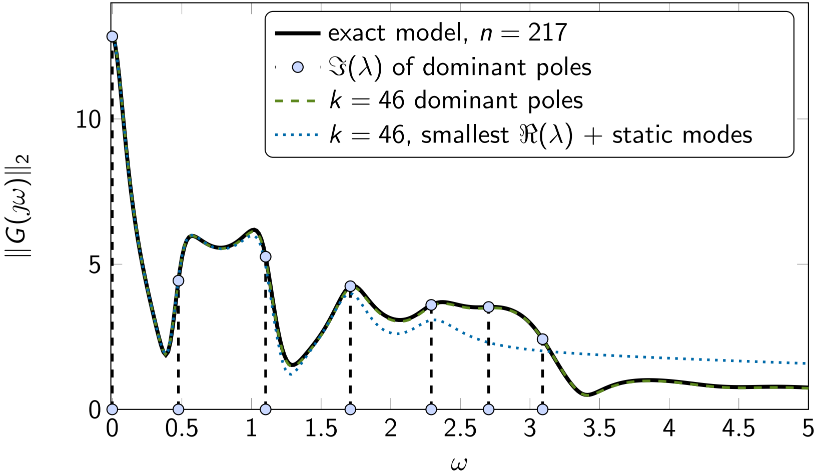
Dominant Poles

Random SISO Example ($B,\ C^T \in \mathbb{R}^n$)



Dominant Poles

Random SISO Example ($B,\ C^T \in \mathbb{R}^n$)

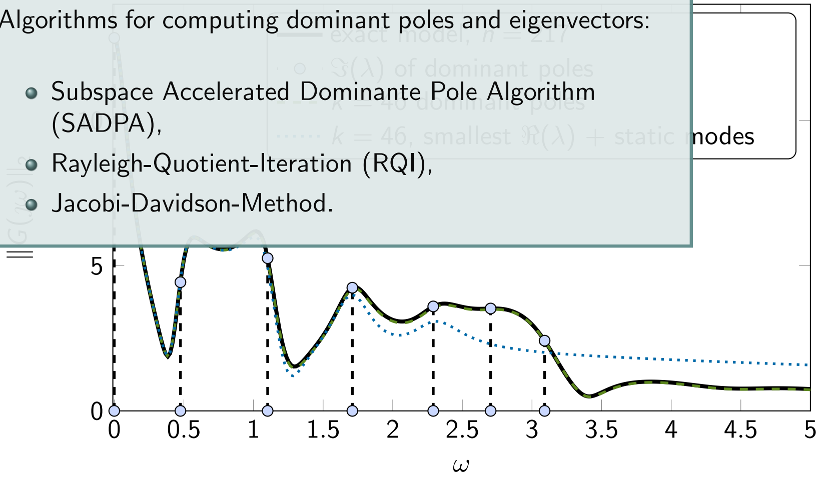


Dominant Poles

Random SISO Example ($B,\ C^T \in \mathbb{R}^n$)

Algorithms for computing dominant poles and eigenvectors:

- Subspace Accelerated Dominante Pole Algorithm (SADPA),
- Rayleigh-Quotient-Iteration (RQI),
- Jacobi-Davidson-Method.



Outline

Model Reduction by Projection

Rational Interpolation

Computation of reduced-order model by projection

Given an LTI system $\dot{x} = Ax + Bu,\ y = Cx$ with transfer function $G(s) = C(sI_n - A)^{-1}B,$ a reduced-order model is obtained using projection approach with $V, W \in \mathbb{R}^{n\times r}$ and $W^T V = I_r$ by computing

$$\hat{A} = W^T AV, \quad \hat{B} = W^T B, \quad \hat{C} = CV.$$

Petrov-Galerkin-type (two-sided) projection: $W \neq V$,

Galerkin-type (one-sided) projection: $W = V$.

Model Reduction by Projection

Rational Interpolation

Computation of reduced-order model by projection

Given an LTI system $\dot{x} = Ax + Bu,\ y = Cx$ with transfer function $G(s) = C(sI_n - A)^{-1}B,$ a reduced-order model is obtained using projection approach with $V, W \in \mathbb{R}^{n \times r}$ and $W^T V = I_r$ by computing

$$\hat{A} = W^T A V, \quad \hat{B} = W^T B, \quad \hat{C} = CV.$$

Petrov-Galerkin-type (two-sided) projection: $W \neq V$,

Galerkin-type (one-sided) projection: $W = V$.

Rational Interpolation/Moment-Matching

Choose V, W such that

$$G(s_j) = \hat{G}(s_j), \quad j = 1, \ldots, k,$$

and

$$\frac{d^i}{ds^i} G(s_j) = \frac{d^i}{ds^i} \hat{G}(s_j), \quad i = 1, \ldots, K_j, \quad j = 1, \ldots, k.$$

Model Reduction by Projection

Rational Interpolation

Theorem (simplified) [Grimme '97, Villemagne/Skelton '87]

If

$$\begin{aligned} \mathrm{span}\left\{(s_1 I_n - A)^{-1}B, \ldots, (s_k I_n - A)^{-1}B\right\} &\subset \mathrm{Ran}(V), \\ \mathrm{span}\left\{(s_1 I_n - A)^{-T}C^T, \ldots, (s_k I_n - A)^{-T}C^T\right\} &\subset \mathrm{Ran}(W), \end{aligned}$$

then

$$G(s_j) = \hat{G}(s_j), \quad \frac{d}{ds}G(s_j) = \frac{d}{ds}\hat{G}(s_j), \quad \text{for } j = 1, \ldots, k.$$

Model Reduction by Projection

Rational Interpolation

Theorem (simplified) [Grimme '97, Villemagne/Skelton '87]

If

$$\begin{aligned} \operatorname{span}\left\{(s_1 I_n - A)^{-1}B, \ldots, (s_k I_n - A)^{-1}B\right\} &\subset \operatorname{Ran}(V), \\ \operatorname{span}\left\{(s_1 I_n - A)^{-T}C^T, \ldots, (s_k I_n - A)^{-T}C^T\right\} &\subset \operatorname{Ran}(W), \end{aligned}$$

then

$$G(s_j) = \hat{G}(s_j), \quad \frac{d}{ds}G(s_j) = \frac{d}{ds}\hat{G}(s_j), \quad \text{for } j = 1, \ldots, k.$$

Remarks:

using Galerkin/one-sided projection yields $G(s_j) = \hat{G}(s_j)$, but in general

$$\frac{d}{ds}G(s_j) \neq \frac{d}{ds}\hat{G}(s_j).$$

Model Reduction by Projection

Rational Interpolation

Theorem (simplified) [Grimme '97, Villemagne/Skelton '87]

If

$$\begin{aligned} \operatorname{span}\left\{(s_1 I_n - A)^{-1}B, \ldots, (s_k I_n - A)^{-1}B\right\} &\subset \operatorname{Ran}(V), \\ \operatorname{span}\left\{(s_1 I_n - A)^{-T}C^T, \ldots, (s_k I_n - A)^{-T}C^T\right\} &\subset \operatorname{Ran}(W), \end{aligned}$$

then

$$G(s_j) = \hat{G}(s_j), \quad \frac{d}{ds}G(s_j) = \frac{d}{ds}\hat{G}(s_j), \quad \text{for } j = 1, \ldots, k.$$

Remarks:

$k = 1$, standard Krylov subspace(s) of dimension K $\rightsquigarrow$ moment-matching methods/Padé approximation,

$$\frac{d^i}{ds^i}G(s_1) = \frac{d^i}{ds^i}\hat{G}(s_1), \quad i = 0, \ldots, K-1(+K).$$

Model Reduction by Projection

Rational Interpolation

Theorem (simplified) [Grimme '97, Villemagne/Skelton '87]

If

$$\begin{aligned} \operatorname{span}\left\{(s_1 I_n - A)^{-1}B, \ldots, (s_k I_n - A)^{-1}B\right\} &\subset \operatorname{Ran}(V), \\ \operatorname{span}\left\{(s_1 I_n - A)^{-T}C^T, \ldots, (s_k I_n - A)^{-T}C^T\right\} &\subset \operatorname{Ran}(W), \end{aligned}$$

then

$$G(s_j) = \hat{G}(s_j), \quad \frac{d}{ds}G(s_j) = \frac{d}{ds}\hat{G}(s_j), \quad \text{for } j = 1, \ldots, k.$$

Remarks:

computation of V, W from rational Krylov subspaces, e.g.,

- dual rational Arnoldi/Lanczos [Grimme '97],
- Iterative Rational Krylov-Algo. [Antoulas/Beattie/Gugercin '07].

$\mathcal{H}_2$-Optimal Model Reduction

Best $\mathcal{H}_2$-norm approximation problem

$$\text{Find} \quad \arg\min_{\hat{G}\in\mathcal{H}_2 \text{ of order } \leq r} \left|\left| G - \hat{G} \right|\right|_2 .$$

$\mathcal{H}_2$-Optimal Model Reduction

Best $\mathcal{H}_2$-norm approximation problem

$$\text{Find} \quad \arg\min_{\hat{G}\in\mathcal{H}_2 \text{ of order } \leq r} \left\| G - \hat{G} \right\|_2 .$$

$\rightsquigarrow$ First-order necessary $\mathcal{H}_2$-optimality conditions:

For SISO systems

$$\begin{aligned} G(-\mu_i) &= \hat{G}(-\mu_i), \\ G'(-\mu_i) &= \hat{G}'(-\mu_i), \end{aligned}$$

where μ_i are the poles of the reduced transfer function $\hat{G}$.

$\mathcal{H}_2$-Optimal Model Reduction

Best $\mathcal{H}_2$-norm approximation problem

Find $\quad \arg\min_{\hat{G}\in\mathcal{H}_2 \text{ of order } \leq r} \left|\left|G - \hat{G}\right|\right|_2.$

$\rightsquigarrow$ First-order necessary $\mathcal{H}_2$-optimality conditions:

For MIMO systems

$$\begin{aligned} G(-\mu_i)\tilde{B}_i &= \hat{G}(-\mu_i)\tilde{B}_i, && \text{for } i = 1,\dots,r,\\ \tilde{C}_i^T G(-\mu_i) &= \tilde{C}_i^T \hat{G}(-\mu_i), && \text{for } i = 1,\dots,r,\\ \tilde{C}_i^T G'(-\mu_i)\tilde{B}_i &= \tilde{C}_i^T \hat{G}'(-\mu_i)\tilde{B}_i, && \text{for } i = 1,\dots,r, \end{aligned}$$

where $T^{-1}\hat{A}T = \operatorname{diag}\{\mu_1,\dots,\mu_r\}$ = spectral decomposition and

$$\tilde{B} = \hat{B}^T T^{-T}, \quad \tilde{C} = \hat{C}T.$$

$\rightsquigarrow$ tangential interpolation conditions.

Model Reduction by Projection

Interpolation of the Transfer Function by Projection

Construct reduced transfer function by Petrov-Galerkin projection $\mathcal{P} = VW^T$, i.e.

$$\hat{G}(s) = CV\left(sI - W^TAV\right)^{-1}W^TB,$$

where V and W are given as the rational Krylov subspaces

$$V = \left[(-\mu_1 I - A)^{-1}B, \ldots, (-\mu_r I - A)^{-1}B\right],$$
$$W = \left[(-\mu_1 I - A^T)^{-1}C^T, \ldots, (-\mu_r I - A^T)^{-1}C^T\right].$$

Then

$$G(-\mu_i) = \hat{G}(-\mu_i) \quad \text{and} \quad G'(-\mu_i) = \hat{G}'(-\mu_i),$$

for $i = 1, \ldots, r$ as desired.
$\rightsquigarrow$ iterative algorithms (IRKA/MIRIAm) that yield $\mathcal{H}_2$-optimal models.

[Gugercin et al. '06], [Bunse-Gerstner et al. '07], [Van Dooren et al. '08]

Model Reduction by Projection
Interpolation of the Transfer Function by Projection

Construct reduced transfer function by Petrov-Galerkin projection $\mathcal{P} = VW^T$, i.e.

$$\hat{G}(s) = CV\left(sI - W^TAV\right)^{-1}W^TB,$$

where V and W are given as the rational Krylov subspaces

$$V = \left[(-\mu_1 I - A)^{-1}B, \dots, (-\mu_r I - A)^{-1}B\right],$$
$$W = \left[(-\mu_1 I - A^T)^{-1}C^T, \dots, (-\mu_r I - A^T)^{-1}C^T\right].$$

Then

$$G(-\mu_i) = \hat{G}(-\mu_i) \quad \text{and} \quad G'(-\mu_i) = \hat{G}'(-\mu_i),$$

for $i = 1, \dots, r$ as desired.
$\rightsquigarrow$ iterative algorithms (IRKA/MIRIAm) that yield $\mathcal{H}_2$-optimal models.

[Gugercin et al. '06], [Bunse-Gerstner et al. '07], [Van Dooren et al. '08]

Model Reduction by Projection

Interpolation of the Transfer Function by Projection

Construct reduced transfer function by Petrov-Galerkin projection $\mathcal{P} = VW^T$, i.e.

$$\hat{G}(s) = CV\left(sI - W^TAV\right)^{-1}W^TB,$$

where V and W are given as the rational Krylov subspaces

$$\begin{aligned} V &= \left[(-\mu_1 I - A)^{-1}B, \ldots, (-\mu_r I - A)^{-1}B\right], \\ W &= \left[(-\mu_1 I - A^T)^{-1}C^T, \ldots, (-\mu_r I - A^T)^{-1}C^T\right]. \end{aligned}$$

Then

$$G(-\mu_i) = \hat{G}(-\mu_i) \quad \text{and} \quad G'(-\mu_i) = \hat{G}'(-\mu_i),$$

for $i = 1, \ldots, r$ as desired.

$\rightsquigarrow$ iterative algorithms (IRKA/MIRIAm) that yield $\mathcal{H}_2$-optimal models.

[Gugercin et al. '06], [Bunse-Gerstner et al. '07], [Van Dooren et al. '08]

Model Reduction by Projection

Interpolation of the Transfer Function by Projection

Construct reduced transfer function by Petrov-Galerkin projection $\mathcal{P} = VW^T$, i.e.

$$\hat{G}(s) = CV\left(sI - W^T AV\right)^{-1} W^T B,$$

where V and W are given as the rational Krylov subspaces

$$\begin{aligned} V &= \left[(-\mu_1 I - A)^{-1}B, \ldots, (-\mu_r I - A)^{-1}B\right], \\ W &= \left[(-\mu_1 I - A^T)^{-1}C^T, \ldots, (-\mu_r I - A^T)^{-1}C^T\right]. \end{aligned}$$

Then

$$G(-\mu_i) = \hat{G}(-\mu_i) \quad \text{and} \quad G'(-\mu_i) = \hat{G}'(-\mu_i),$$

for $i = 1, \ldots, r$ as desired.

$\rightsquigarrow$ iterative algorithms (IRKA/MIRIAm) that yield $\mathcal{H}_2$-optimal models.

[Gugercin et al. '06], [Bunse-Gerstner et al. '07], [Van Dooren et al. '08]

$\mathcal{H}_2$-Optimal Model Reduction

The Basic IRKA Algorithm

Algorithm 1 IRKA (MIMO version/MIRIAm)

Input: A stable, B, C, $\hat{A}$ stable, $\hat{B}$, $\hat{C}$, $\delta > 0$.

Output: A^{opt}, B^{opt}, C^{opt}

1: **while** $\left(\max_{j=1,\ldots,r} \left\{ \frac{|\mu_j - \mu_j^{\text{old}}|}{|\mu_j|} \right\} > \delta\right)$ **do**

2: $\quad \operatorname{diag}\{\mu_1, \ldots, \mu_r\} := T^{-1}\hat{A}T =$ spectral decomposition, $\tilde{B} = \hat{B}^H T^{-T}$, $\tilde{C} = \hat{C}T$.

3: $\quad V = \left[(-\mu_1 I - A)^{-1} B\tilde{B}_1, \ldots, (-\mu_r I - A)^{-1} B\tilde{B}_r\right]$

4: $\quad W = \left[(-\mu_1 I - A^T)^{-1} C^T\tilde{C}_1, \ldots, (-\mu_r I - A^T)^{-1} C^T\tilde{C}_r\right]$

5: $\quad V = \operatorname{orth}(V)$, $W = \operatorname{orth}(W)$, $W = W(V^H W)^{-1}$

6: $\quad \hat{A} = W^H A V$, $\hat{B} = W^H B$, $\hat{C} = CV$

7: **end while**

8: $A^{opt} = \hat{A}$, $B^{opt} = \hat{B}$, $C^{opt} = \hat{C}$

Outline

Balanced Truncation

Basic principle:

- Recall: a stable system Σ, realized by (A, B, C, D), is called balanced, if the Gramians, i.e., solutions P, Q of the Lyapunov equations

$$AP + PA^T + BB^T = 0, \qquad A^T Q + QA + C^T C = 0,$$

satisfy: $P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$.

- $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the Hankel singular values (HSVs) of Σ.

Balanced Truncation

Basic principle:

- Recall: a stable system Σ, realized by (A, B, C, D), is called balanced, if the Gramians, i.e., solutions P, Q of the Lyapunov equations

$$AP + PA^T + BB^T = 0, \qquad A^TQ + QA + C^TC = 0,$$

satisfy: $P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$.

- $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the Hankel singular values (HSVs) of Σ.

Balanced Truncation

Basic principle:

- Recall: a stable system Σ, realized by (A, B, C, D), is called balanced, if the Gramians, i.e., solutions P, Q of the Lyapunov equations

$$AP + PA^T + BB^T = 0, \qquad A^TQ + QA + C^TC = 0,$$

satisfy: $P = Q = \mathrm{diag}(\sigma_1, \dots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$.

- $\Lambda\left(PQ\right)^{\frac{1}{2}} = \{\sigma_1, \dots, \sigma_n\}$ are the Hankel singular values (HSVs) of Σ.
- Compute balanced realization of the system via state-space transformation

$$\begin{aligned} \mathcal{T} : (A, B, C, D) &\mapsto (TAT^{-1}, TB, CT^{-1}, D) \\ &= \left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \begin{bmatrix} C_1 & C_2 \end{bmatrix}, D \right) \end{aligned}$$

- Truncation $\rightsquigarrow (\hat{A}, \hat{B}, \hat{C}, \hat{D}) := (A_{11}, B_1, C_1, D)$.

Balanced Truncation

Basic principle:

- Recall: a stable system Σ, realized by (A, B, C, D), is called balanced, if the Gramians, i.e., solutions P, Q of the Lyapunov equations

$$AP + PA^T + BB^T = 0, \qquad A^T Q + QA + C^T C = 0,$$

satisfy: $P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$.

- $\Lambda\left(PQ\right)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the Hankel singular values (HSVs) of Σ.
- Compute balanced realization of the system via state-space transformation

$$\begin{aligned} \mathcal{T} : (A, B, C, D) &\mapsto (TAT^{-1}, TB, CT^{-1}, D) \\ &= \left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \begin{bmatrix} C_1 & C_2 \end{bmatrix}, D \right) \end{aligned}$$

- Truncation $\rightsquigarrow (\hat{A}, \hat{B}, \hat{C}, \hat{D}) := (A_{11}, B_1, C_1, D)$.

Balanced Truncation

Motivation:

The HSVs $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \dots, \sigma_n\}$ are system invariants: they are preserved under

$$\mathcal{T} : (A, B, C, D) \mapsto (TAT^{-1}, TB, CT^{-1}, D)$$

Balanced Truncation

Motivation:

The HSVs $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are system invariants: they are preserved under

$$\mathcal{T} : (A, B, C, D) \mapsto (TAT^{-1}, TB, CT^{-1}, D)$$

in transformed coordinates, the Gramians satisfy

$$\begin{aligned} (TAT^{-1})(TPT^T) + (TPT^T)(TAT^{-1})^T + (TB)(TB)^T &= 0, \\ (TAT^{-1})^T(T^{-T}QT^{-1}) + (T^{-T}QT^{-1})(TAT^{-1}) + (CT^{-1})^T(CT^{-1}) &= 0 \end{aligned}$$

$$\Rightarrow \; (TPT^T)(T^{-T}QT^{-1}) = TPQT^{-1},$$

hence $\Lambda(PQ) = \Lambda((TPT^T)(T^{-T}QT^{-1}))$.

Balanced Truncation

Implementation: SR Method

1. Compute (Cholesky) factors of the Gramians, $P = S^T S, \; Q = R^T R.$
2. Compute SVD $SR^T = [\, U_1, \; U_2 \,] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}.$
3. ROM is $(W^T AV, W^T B, CV, D)$, where

$$W = R^T V_1 \Sigma_1^{-\frac{1}{2}}, \qquad V = S^T U_1 \Sigma_1^{-\frac{1}{2}}.$$

Balanced Truncation

Implementation: SR Method

1. Compute (Cholesky) factors of the Gramians, $P = S^T S, \; Q = R^T R.$
2. Compute SVD $SR^T = [\, U_1, \; U_2 \,] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}.$
3. ROM is $(W^T AV, W^T B, CV, D)$, where

$$W = R^T V_1 \Sigma_1^{-\frac{1}{2}}, \qquad V = S^T U_1 \Sigma_1^{-\frac{1}{2}}.$$

Balanced Truncation

Implementation: SR Method

1. Compute (Cholesky) factors of the Gramians, $P = S^T S, \; Q = R^T R.$
2. Compute SVD $SR^T = [\, U_1, \; U_2 \,] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}.$
3. ROM is $(W^T AV, W^T B, CV, D)$, where

$$W = R^T V_1 \Sigma_1^{-\frac{1}{2}}, \qquad V = S^T U_1 \Sigma_1^{-\frac{1}{2}}.$$

Balanced Truncation

Implementation: SR Method

1. Compute (Cholesky) factors of the Gramians, $P = S^T S$, $Q = R^T R$.
2. Compute SVD $SR^T = [\, U_1,\ U_2 \,] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$.
3. ROM is $(W^T AV, W^T B, CV, D)$, where

$$W = R^T V_1 \Sigma_1^{-\frac{1}{2}}, \qquad V = S^T U_1 \Sigma_1^{-\frac{1}{2}}.$$

Note:

$$V^T W \quad = \quad (\Sigma_1^{-\frac{1}{2}} U_1^T S)(R^T V_1 \Sigma_1^{-\frac{1}{2}})$$

Balanced Truncation

Implementation: SR Method

1. Compute (Cholesky) factors of the Gramians, $P = S^T S,\ Q = R^T R.$
2. Compute SVD $SR^T = [\, U_1,\ U_2 \,] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}.$
3. ROM is $(W^T AV, W^T B, CV, D)$, where

$$W = R^T V_1 \Sigma_1^{-\frac{1}{2}}, \qquad V = S^T U_1 \Sigma_1^{-\frac{1}{2}}.$$

Note:

$$V^T W \quad = \quad (\Sigma_1^{-\frac{1}{2}} U_1^T S)(R^T V_1 \Sigma_1^{-\frac{1}{2}}) \ = \ \Sigma_1^{-\frac{1}{2}} U_1^T U \Sigma V^T V_1 \Sigma_1^{-\frac{1}{2}}$$

Balanced Truncation

Implementation: SR Method

1. Compute (Cholesky) factors of the Gramians, $P = S^T S, \; Q = R^T R$.
2. Compute SVD $SR^T = [\, U_1, \; U_2 \,] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$.
3. ROM is $(W^T AV, W^T B, CV, D)$, where

$$W = R^T V_1 \Sigma_1^{-\frac{1}{2}}, \qquad V = S^T U_1 \Sigma_1^{-\frac{1}{2}}.$$

Note:

$$\begin{aligned} V^T W &= (\Sigma_1^{-\frac{1}{2}} U_1^T S)(R^T V_1 \Sigma_1^{-\frac{1}{2}}) = \Sigma_1^{-\frac{1}{2}} U_1^T U \Sigma V^T V_1 \Sigma_1^{-\frac{1}{2}} \\ &= \Sigma_1^{-\frac{1}{2}} [\, I_r, \, 0 \,] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} I_r \\ 0 \end{bmatrix} \Sigma_1^{-\frac{1}{2}} \end{aligned}$$

Balanced Truncation

Implementation: SR Method

1. Compute (Cholesky) factors of the Gramians, $P = S^T S$, $Q = R^T R$.
2. Compute SVD $SR^T = [\, U_1,\ U_2 \,] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$.
3. ROM is $(W^T AV, W^T B, CV, D)$, where

$$W = R^T V_1 \Sigma_1^{-\frac{1}{2}}, \qquad V = S^T U_1 \Sigma_1^{-\frac{1}{2}}.$$

Note:

$$\begin{aligned} V^T W &= (\Sigma_1^{-\frac{1}{2}} U_1^T S)(R^T V_1 \Sigma_1^{-\frac{1}{2}}) = \Sigma_1^{-\frac{1}{2}} U_1^T U \Sigma V^T V_1 \Sigma_1^{-\frac{1}{2}} \\ &= \Sigma_1^{-\frac{1}{2}} [\, I_r,\ 0 \,] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} I_r \\ 0 \end{bmatrix} \Sigma_1^{-\frac{1}{2}} = \Sigma_1^{-\frac{1}{2}} \Sigma_1 \Sigma_1^{-\frac{1}{2}} = I_r \end{aligned}$$

$\Longrightarrow$ VW^T is a projector, hence BT is a projection method.

Balanced Truncation

Properties:

- Reduced-order model is stable with HSVs $\sigma_1, \ldots, \sigma_r$.
- Adaptive choice of r via computable error bound:

$$||y - \hat{y}||_2 \leq \left(2 \sum_{k=r+1}^{n} \sigma_k\right) ||u||_2 .$$

Balanced Truncation

Properties:

- Reduced-order model is stable with HSVs $\sigma_1, \ldots, \sigma_r$.
- Adaptive choice of r via computable error bound:

$$||y - \hat{y}||_2 \leq \left(2 \sum\nolimits_{k=r+1}^{n} \sigma_k\right) ||u||_2 .$$

Balanced Truncation

Theoretical Background

Linear, Time-Invariant (LTI) Systems

$$\begin{aligned} \dot{x} &= Ax + Bu, \qquad A \in \mathbb{R}^{n\times n}, \quad B \in \mathbb{R}^{n\times m}, \\ y &= Cx, \qquad C \in \mathbb{R}^{q\times n}, \quad x(-\infty) = 0. \end{aligned}$$

Alternative to State-Space Operator: Hankel Operator

Instead of

$$\mathcal{S}\colon u \mapsto y, \quad y(t) = \int_{-\infty}^{t} Ce^{A(t-\tau)}Bu(\tau)\,d\tau \quad \text{for all } t \in \mathbb{R}.$$

use the Hankel operator: (the future response of the past inputs)

$$\mathcal{H}\colon u_- \mapsto y_+, \quad y_+(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu(\tau)\,\mathrm{d}\tau \quad \text{for } t > 0,$$

Balanced Truncation

Theoretical Background

Linear, Time-Invariant (LTI) Systems

$$\begin{aligned} \dot{x} &= Ax + Bu, \qquad A \in \mathbb{R}^{n\times n}, \quad B \in \mathbb{R}^{n\times m}, \\ y &= Cx, \qquad C \in \mathbb{R}^{q\times n}, \quad x(-\infty) = 0. \end{aligned}$$

Alternative to State-Space Operator: Hankel Operator

Instead of

$$\mathcal{S}\colon u \mapsto y, \quad y(t) = \int_{-\infty}^{t} Ce^{A(t-\tau)}Bu(\tau)\,d\tau \quad \text{for all } t \in \mathbb{R}.$$

use the Hankel operator: (the future response of the past inputs)

$$\begin{aligned} \mathcal{H}\colon u_- \mapsto y_+, \quad y_+(t) &= \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu(\tau)\,\mathrm{d}\tau \quad \text{for } t > 0, \\ &= Ce^{At}\int_{-\infty}^{0} e^{-A\tau}Bu(\tau)\,\mathrm{d}\tau \quad \text{for } t > 0, \\ &= Ce^{At}x(0; u_-) \quad \text{for } t > 0. \end{aligned}$$

Balanced Truncation

Theoretical Background

Linear, Time-Invariant (LTI) Systems

$$
\begin{aligned}
\dot{x} &= Ax + Bu, \qquad && A \in \mathbb{R}^{n\times n}, \quad B \in \mathbb{R}^{n\times m},\\
y &= Cx, && C \in \mathbb{R}^{q\times n}, \quad x(-\infty) = 0.
\end{aligned}
$$

Alternative to State-Space Operator: Hankel Operator

use the Hankel operator: (the future response of the past inputs)

$$
\mathcal{H} \colon u_- \mapsto y_+, \quad y_+(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu(\tau)\,\mathrm{d}\tau \quad \text{for } t > 0,
$$

- The operator $\mathcal{H}$ is compact $\Rightarrow$ $\mathcal{H}$ has discrete SVD
  - $\rightarrow$ The Hankel singular values: $\{\sigma_j\}_{j=1}^{\infty}: \ \sigma_1 \geq \sigma_2 \geq \ldots \geq 0$
  - $\rightarrow$ An SVD-type approximation of the linear map $\mathcal{H}$ is possible!

Balanced Truncation

Theoretical Background

Linear, Time-Invariant (LTI) Systems

$$\begin{aligned} \dot{x} &= Ax + Bu, \qquad A \in \mathbb{R}^{n \times n}, \quad B \in \mathbb{R}^{n \times m}, \\ y &= Cx, \qquad C \in \mathbb{R}^{q \times n}, \quad x(-\infty) = 0. \end{aligned}$$

Alternative to State-Space Operator: Hankel Operator

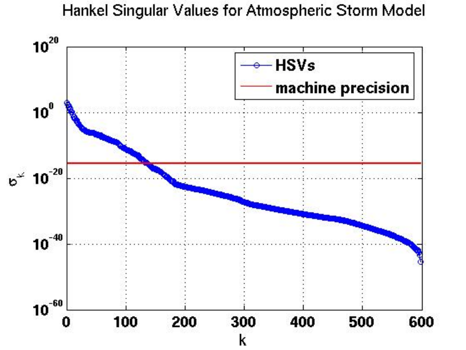
$\mathcal{H}$ compact

$\Downarrow$

$\mathcal{H}$ has a discrete SVD

$\Downarrow$

Hankel singular values



Balanced Truncation

Theoretical Background

Linear, Time-Invariant (LTI) Systems

$$\begin{array}{rcllll} \dot{x} &=& Ax + Bu, & A \in \mathbb{R}^{n\times n}, & B \in \mathbb{R}^{n\times m}, \\ y &=& Cx, & C \in \mathbb{R}^{q\times n}, & x(-\infty) = 0. \end{array}$$

Alternative to State-Space Operator: Hankel Operator

$$\mathcal{H}\colon u_- \mapsto y_+, \quad y_+(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu(\tau)\,d\tau \quad \text{for all } t > 0.$$

$\mathcal{H}$ compact $\Rightarrow$ $\mathcal{H}$ has discrete SVD

$\Rightarrow$ Best approximation problem w.r.t. 2-induced operator norm well-posed

$\Rightarrow$ solution: Adamjan-Arov-Krein (AAK Theory, 1971/78).

But: computationally unfeasible for large-scale systems.

Balanced Truncation

Theoretical Background

Linear, Time-Invariant (LTI) Systems

$$\begin{aligned} \dot{x} &= Ax + Bu, \qquad A \in \mathbb{R}^{n \times n}, \quad B \in \mathbb{R}^{n \times m}, \\ y &= Cx, \qquad C \in \mathbb{R}^{q \times n}, \quad x(-\infty) = 0. \end{aligned}$$

Alternative to State-Space Operator: Hankel Operator

$$\mathcal{H} \colon u_- \mapsto y_+, \quad y_+(t) = \int_{-\infty}^{0} C e^{A(t-\tau)} B u(\tau)\, d\tau \quad \text{for all } t > 0.$$

$\mathcal{H}$ compact $\Rightarrow$ $\mathcal{H}$ has discrete SVD

$\Rightarrow$ Best approximation problem w.r.t. 2-induced operator norm well-posed

$\Rightarrow$ solution: Adamjan-Arov-Krein (AAK Theory, 1971/78).

But: computationally unfeasible for large-scale systems.

Balanced Truncation

Theoretical Background

Linear, Time-Invariant (LTI) Systems

$$\begin{aligned} \dot{x} &= Ax + Bu, \qquad && A \in \mathbb{R}^{n \times n}, \quad B \in \mathbb{R}^{n \times m}, \\ y &= Cx, && C \in \mathbb{R}^{q \times n}, \quad x(-\infty) = 0. \end{aligned}$$

Alternative to State-Space Operator: Hankel Operator

$$\mathcal{H}\colon u_- \mapsto y_+, \quad y_+(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu(\tau)\,d\tau \quad \text{for all } t > 0.$$

$\mathcal{H}$ compact $\Rightarrow$ $\mathcal{H}$ has discrete SVD

$\Rightarrow$ Best approximation problem w.r.t. 2-induced operator norm well-posed

$\Rightarrow$ solution: Adamjan-Arov-Krein (AAK Theory, 1971/78).

But: computationally unfeasible for large-scale systems.

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \dots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Hankel operator

$$y_+(t) = \mathcal{H}u_-(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu_-(\tau)\, d\tau$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda\left(PQ\right)^{\frac{1}{2}} = \{\sigma_1, \dots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Hankel operator

$$y_+(t) = \mathcal{H}u_-(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu_-(\tau)\,d\tau =: Ce^{At}\underbrace{\int_{-\infty}^{0} e^{-A\tau}Bu_-(\tau)\,d\tau}_{=:z}$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Hankel operator

$$y_+(t) = \mathcal{H}u_-(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu_-(\tau)\,d\tau =: Ce^{At}\underbrace{\int_{-\infty}^{0} e^{-A\tau}Bu_-(\tau)\,d\tau}_{=:z} = Ce^{At}z.$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda\left(PQ\right)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Hankel operator

$$y_+(t) = \mathcal{H}u_-(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu_-(\tau)\, d\tau = Ce^{At}z.$$

Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \dots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Hankel operator

$$y_+(t) = \mathcal{H}u_-(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu_-(\tau)\,d\tau = Ce^{At}z.$$

Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

$$\mathcal{H}^* y_+(t) = \int_0^{\infty} B^T e^{A^T(\tau - t)} C^T y_+(\tau)\,d\tau$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda\left(PQ\right)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Hankel operator

$$y_+(t) = \mathcal{H}u_-(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu_-(\tau)\,d\tau = Ce^{At}z.$$

Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

$$\mathcal{H}^*y_+(t) = \int_0^\infty B^Te^{A^T(\tau-t)}C^Ty_+(\tau)\,d\tau = B^Te^{-A^Tt}\int_0^\infty e^{A^T\tau}C^Ty_+(\tau)\,d\tau.$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda\left(PQ\right)^{\frac{1}{2}} = \{\sigma_1, \dots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Hankel operator

$$y_+(t) = \mathcal{H}u_-(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu_-(\tau)\,d\tau = Ce^{At}z.$$

Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

$$\mathcal{H}^*y_+(t) = \; = B^Te^{-A^Tt}\int_0^\infty e^{A^T\tau}C^Ty_+(\tau)\,d\tau.$$

$$\mathcal{H}^*\mathcal{H}u_-(t) \quad = \quad B^Te^{-A^Tt}\int_0^\infty e^{A^T\tau}C^TCe^{A\tau}z\,d\tau$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Hankel operator

$$y_+(t) = \mathcal{H}u_-(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu_-(\tau)\,d\tau = Ce^{At}z.$$

Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

$$\mathcal{H}^*y_+(t) = \;= B^Te^{-A^Tt}\int_0^\infty e^{A^T\tau}C^Ty_+(\tau)\,d\tau.$$

Hence,

$$\begin{aligned}
\mathcal{H}^*\mathcal{H}u_-(t) &= B^Te^{-A^Tt}\int_0^\infty e^{A^T\tau}C^TCe^{A\tau}z\,d\tau \\
&= B^Te^{-A^Tt}\underbrace{\int_0^\infty e^{A^T\tau}C^TCe^{A\tau}\,d\tau}_{\equiv Q}\,z
\end{aligned}$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Hankel operator

$$y_+(t) = \mathcal{H}u_-(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu_-(\tau)\,d\tau = Ce^{At}z.$$

Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

$$\mathcal{H}^* y_+(t) = \; = B^T e^{-A^T t} \int_0^\infty e^{A^T\tau} C^T y_+(\tau)\,d\tau.$$

Hence,

$$\begin{aligned} \mathcal{H}^*\mathcal{H}u_-(t) &= B^T e^{-A^T t} \int_0^\infty e^{A^T\tau} C^T Ce^{A\tau} z\,d\tau \\ &= B^T e^{-A^T t} Qz \end{aligned}$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Hankel operator

$$y_+(t) = \mathcal{H}u_-(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu_-(\tau)\,d\tau = Ce^{At}z.$$

Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

$$\mathcal{H}^*y_+(t) = \; = B^T e^{-A^T t} \int_0^{\infty} e^{A^T\tau} C^T y_+(\tau)\,d\tau.$$

Hence,

$$\mathcal{H}^*\mathcal{H}u_-(t) \quad = \quad B^T e^{-A^T t} Qz$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Hankel operator

$$y_+(t) = \mathcal{H}u_-(t) = \int_{-\infty}^{0} Ce^{A(t-\tau)}Bu_-(\tau)\,d\tau = Ce^{At}z.$$

Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

$$\mathcal{H}^*y_+(t) = \ = B^Te^{-A^Tt}\int_0^{\infty} e^{A^T\tau}C^Ty_+(\tau)\,d\tau.$$

Hence,

$$\mathcal{H}^*\mathcal{H}u_-(t) \quad = \quad B^Te^{-A^Tt}Qz \ \doteq \ \sigma^2u_-(t).$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$, Hence,

$$\mathcal{H}^*\mathcal{H}u_-(t) = B^T e^{-A^T t} Q z \doteq \sigma^2 u_-(t).$$

$\implies u_-(t) = \frac{1}{\sigma^2} B^T e^{-A^T t} Q z$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \dots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

$$\mathcal{H}^*\mathcal{H}u_-(t) \quad = \quad B^T e^{-A^T t} Q z \doteq \sigma^2 u_-(t).$$

$\implies u_-(t) = \frac{1}{\sigma^2} B^T e^{-A^T t} Q z \implies$ (recalling $z = \int_{-\infty}^{0} e^{-A\tau} B u_-(\tau)\, d\tau$)

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

$$\mathcal{H}^*\mathcal{H}u_-(t) = B^T e^{-A^T t} Qz \doteq \sigma^2 u_-(t).$$

$\implies u_-(t) = \frac{1}{\sigma^2} B^T e^{-A^T t} Qz \implies$ (recalling $z = \int_{-\infty}^0 e^{-A\tau} B u_-(\tau)\, d\tau$)

$$z = \int_{-\infty}^0 e^{-A\tau} B \frac{1}{\sigma^2} B^T e^{-A^T \tau} Qz \, d\tau$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda\left(PQ\right)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

$$\mathcal{H}^*\mathcal{H}u_-(t) \quad = \quad B^Te^{-A^Tt}Qz \doteq \sigma^2 u_-(t).$$

$\implies u_-(t) = \frac{1}{\sigma^2}B^Te^{-A^Tt}Qz \implies$ (recalling $z = \int_{-\infty}^0 e^{-A\tau}Bu_-(\tau)\,d\tau$)

$$\begin{aligned} z &= \int_{-\infty}^0 e^{-A\tau}B\frac{1}{\sigma^2}B^Te^{-A^T\tau}Qz\,d\tau \\ &= \frac{1}{\sigma^2}\int_{-\infty}^0 e^{-A\tau}BB^Te^{-A^T\tau}\,d\tau\,Qz \end{aligned}$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda\left(PQ\right)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

$$\mathcal{H}^*\mathcal{H}u_-(t) \quad = \quad B^T e^{-A^T t} Qz \doteq \sigma^2 u_-(t).$$

$\implies u_-(t) = \frac{1}{\sigma^2} B^T e^{-A^T t} Qz \implies$ (recalling $z = \int_{-\infty}^0 e^{-A\tau} B u_-(\tau)\, d\tau$)

$$\begin{aligned} z &= \int_{-\infty}^0 e^{-A\tau} B \frac{1}{\sigma^2} B^T e^{-A^T\tau} Qz \, d\tau \\ &= \frac{1}{\sigma^2} \int_{-\infty}^0 e^{-A\tau} BB^T e^{-A^T\tau}\, d\tau\, Qz \\ &= \frac{1}{\sigma^2} \underbrace{\int_0^\infty e^{A\tau} BB^T e^{A^T\tau}\, \mathrm{d}\tau}_{\equiv P}\; Qz \end{aligned}$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda\left(PQ\right)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

$$\mathcal{H}^*\mathcal{H}u_-(t) \quad = \quad B^T e^{-A^T t} Qz \doteq \sigma^2 u_-(t).$$

$\implies u_-(t) = \frac{1}{\sigma^2} B^T e^{-A^T t} Qz \implies$ (recalling $z = \int_{-\infty}^0 e^{-A\tau} B u_-(\tau)\, d\tau$)

$$\begin{aligned} z &= \int_{-\infty}^0 e^{-A\tau} B \frac{1}{\sigma^2} B^T e^{-A^T\tau} Qz\, d\tau \\ &= \frac{1}{\sigma^2} \underbrace{\int_0^\infty e^{A\tau} BB^T e^{A^T\tau}\, \mathrm{d}\tau}_{\equiv P}\, Qz \\ &= \frac{1}{\sigma^2} PQz \end{aligned}$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda\left(PQ\right)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Proof: Singular values of $\mathcal{H}$ = square roots of eigenvalues of $\mathcal{H}^*\mathcal{H}$,

$$\mathcal{H}^*\mathcal{H}u_-(t) = B^T e^{-A^T t} Qz \doteq \sigma^2 u_-(t).$$

$$\implies u_-(t) = \frac{1}{\sigma^2} B^T e^{-A^T t} Qz \implies \text{(recalling } z = \int_{-\infty}^{0} e^{-A\tau} B u_-(\tau)\, d\tau\text{)}$$

$$\begin{aligned} z &= \int_{-\infty}^{0} e^{-A\tau} B \frac{1}{\sigma^2} B^T e^{-A^T\tau} Qz\, d\tau \\ &= \frac{1}{\sigma^2} \underbrace{\int_0^\infty e^{A\tau} BB^T e^{A^T\tau}\, \mathrm{d}\tau}_{\equiv P}\, Qz \\ &= \frac{1}{\sigma^2} PQz \end{aligned}$$

$$\iff \quad PQz = \sigma^2 z. \quad \square$$

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let P, Q be the controllability and observability Gramians of an LTI system Σ. Then the Hankel singular values $\Lambda(PQ)^{\frac{1}{2}} = \{\sigma_1, \ldots, \sigma_n\}$ are the singular values of the Hankel operator associated to Σ.

Theorem

Let the reduced-order system $\hat{\Sigma} : (\hat{A}, \hat{B}, \hat{C}, \hat{D})$ with $r \leq \hat{n}$ be computed by balanced truncation. Then the reduced-order model $\hat{\Sigma}$ is balanced, stable, minimal, and its HSVs are $\sigma_1, \ldots, \sigma_r$.

Balanced Truncation

The Hankel Singular Values are Singular Values!

Theorem

Let the reduced-order system $\hat{\Sigma} : (\hat{A}, \hat{B}, \hat{C}, \hat{D})$ with $r \le \hat{n}$ be computed by balanced truncation. Then the reduced-order model $\hat{\Sigma}$ is balanced, stable, minimal, and its HSVs are $\sigma_1, \dots, \sigma_r$.

Proof: Note that in balanced coordinates, the Gramians are diagonal and equal to

$$\mathrm{diag}(\Sigma_1, \Sigma_2) = \mathrm{diag}(\sigma_1, \dots, \sigma_r, \sigma_{r+1}, \dots, \sigma_n).$$

Hence, the Gramian satisfies

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} + \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^T + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}^T = 0,$$

whence we obtain the "controllability Lyapunov equation" of the reduced-order system,

$$A_{11}\Sigma_1 + \Sigma_1 A_{11}^T + B_1 B_1^T = 0.$$

The result follows from $\hat{A} = A_{11}, \hat{B} = B_1$, $\Sigma_1 > 0$, the solution theory of Lyapunov equations and the analogous considerations for the observability Gramian. (Minimality is a simple consequence of $\hat{P} = \Sigma_1 = \hat{Q} > 0$.)

Singular Perturbation Approximation (aka Balanced Residualization)

Assume the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u, \quad y = [\, C_1,\ C_2 \,] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + Du$$

is in balanced coordinates.

Singular Perturbation Approximation (aka Balanced Residualization)

Assume the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u, \quad y = [\, C_1, \; C_2 \,] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + Du$$

is in balanced coordinates.

Balanced truncation would set $x_2 = 0$ and use (A_{11}, B_1, C_1, D) as reduced-order model, thereby the information present in the remaining model is ignored!

Singular Perturbation Approximation (aka Balanced Residualization)

Assume the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u, \quad y = [\, C_1, \; C_2 \,] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + Du$$

is in balanced coordinates.

Balanced truncation would set $x_2 = 0$ and use (A_{11}, B_1, C_1, D) as reduced-order model, thereby the information present in the remaining model is ignored!

Particularly, if $G(0) = \hat{G}(0)$ ("zero steady-state error") is required, one can apply the same condensation technique as in Guyan reduction: instead of $x_2 = 0$, set $\dot{x}_2 = 0$. This yields the reduced-order model

$$\begin{aligned} \dot{x}_1 &= (A_{11} - A_{12}A_{22}^{-1}A_{21})x_1 + (B_1 - A_{12}A_{22}^{-1}B_2)u, \\ y &= (C_1 - C_2A_{22}^{-1}A_{21})x_1 + (D - C_2A_{22}^{-1}B_2)u, \end{aligned}$$

with

- the same properties as the reduced-order model w.r.t. stability, minimality, error bound, but $\hat{D} \neq D$;
- zero steady-state error, $G(0) = \hat{G}(0)$ as desired.

Singular Perturbation Approximation (aka Balanced Residualization)

Particularly, if $G(0) = \hat{G}(0)$ ("zero steady-state error") is required, one can apply the same condensation technique as in Guyan reduction: instead of $x_2 = 0$, set $\dot{x}_2 = 0$. This yields the reduced-order model

$$\begin{aligned}\dot{x}_1 &= (A_{11} - A_{12}A_{22}^{-1}A_{21})x_1 + (B_1 - A_{12}A_{22}^{-1}B_2)u,\\ y &= (C_1 - C_2A_{22}^{-1}A_{21})x_1 + (D - C_2A_{22}^{-1}B_2)u,\end{aligned}$$

with

- the same properties as the reduced-order model w.r.t. stability, minimality, error bound, but $\hat{D} \neq D$;
- zero steady-state error, $G(0) = \hat{G}(0)$ as desired.

Note:

- A_{22} invertible as in balanced coordinates, $A_{22}\Sigma_2 + \Sigma_2 A_{22}^T + B_2 B_2^T = 0$ and (A_{22}, B_2) controllable, $\Sigma_2 > 0 \;\Rightarrow\; A_{22}$ stable.
- If the original system is not balanced, first compute a minimal realization by applying balanced truncation with $r = \hat{n}$.

Balancing-Related Methods

Basic Principle

Given positive semidefinite matrices $P = S^T S$, $Q = R^T R$, compute balancing state-space transformation so that

$$P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) = \Sigma, \quad \sigma_1 \geq \ldots \geq \sigma_n > 0,$$

and truncate corresponding realization at size r with $\sigma_r > \sigma_{r+1}$.

Balancing-Related Methods

Basic Principle

Given positive semidefinite matrices $P = S^T S$, $Q = R^T R$, compute balancing state-space transformation so that

$$P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) = \Sigma, \quad \sigma_1 \geq \ldots \geq \sigma_n > 0,$$

and truncate corresponding realization at size r with $\sigma_r > \sigma_{r+1}$.

Classical Balanced Truncation (BT) [Mullis/Roberts '76, Moore '81]

- $P =$ controllability Gramian of system given by (A, B, C, D).
- $Q =$ observability Gramian of system given by (A, B, C, D).
- P, Q solve dual Lyapunov equations

$$AP + PA^T + BB^T \ = \ 0, \qquad A^T Q + QA + C^T C \ = \ 0.$$

Balancing-Related Methods

Basic Principle

Given positive semidefinite matrices $P = S^T S$, $Q = R^T R$, compute balancing state-space transformation so that

$$P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) = \Sigma, \quad \sigma_1 \geq \ldots \geq \sigma_n > 0,$$

and truncate corresponding realization at size r with $\sigma_r > \sigma_{r+1}$.

LQG Balanced Truncation (LQGBT) [Jonckheere/Silverman '83]

- $P/Q =$ controllability/observability Gramian of closed-loop system based on LQG compensator.
- P, Q solve dual algebraic Riccati equations (AREs)

$$\begin{aligned} 0 &= AP + PA^T - PC^T CP + B^T B, \\ 0 &= A^T Q + QA - QBB^T Q + C^T C. \end{aligned}$$

Balancing-Related Methods

Basic Principle

Given positive semidefinite matrices $P = S^T S$, $Q = R^T R$, compute balancing state-space transformation so that

$$P = Q = \mathrm{diag}(\sigma_1, \dots, \sigma_n) = \Sigma, \quad \sigma_1 \geq \dots \geq \sigma_n > 0,$$

and truncate corresponding realization at size r with $\sigma_r > \sigma_{r+1}$.

Balanced Stochastic Truncation (BST) [Desai/Pal '84, Green '88]

- $P =$ controllability Gramian of system given by (A, B, C, D), i.e., solution of Lyapunov equation $AP + PA^T + BB^T = 0$.
- $Q =$ observability Gramian of right spectral factor of power spectrum of system given by (A, B, C, D), i.e., solution of ARE

$$\hat{A}^T Q + Q\hat{A} + QB_W(DD^T)^{-1}B_W^T Q + C^T(DD^T)^{-1}C = 0,$$

where $\hat{A} := A - B_W(DD^T)^{-1}C$, $B_W := BD^T + PC^T$.

Balancing-Related Methods

Basic Principle

Given positive semidefinite matrices $P = S^T S$, $Q = R^T R$, compute balancing state-space transformation so that

$$P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) = \Sigma, \quad \sigma_1 \geq \ldots \geq \sigma_n > 0,$$

and truncate corresponding realization at size r with $\sigma_r > \sigma_{r+1}$.

Positive-Real Balanced Truncation (PRBT) [Green '88]

- Based on positive-real equations, related to positive real (Kalman-Yakubovich-Popov-Anderson) lemma.
- P, Q solve dual AREs

$$\begin{aligned} 0 &= \bar{A}P + P\bar{A}^T + PC^T\bar{R}^{-1}CP + B\bar{R}^{-1}B^T, \\ 0 &= \bar{A}^T Q + Q\bar{A} + QB\bar{R}^{-1}B^T Q + C^T\bar{R}^{-1}C, \end{aligned}$$

where $\bar{R} = D + D^T$, $\bar{A} = A - B\bar{R}^{-1}C$.

Balancing-Related Methods

Basic Principle

Given positive semidefinite matrices $P = S^T S$, $Q = R^T R$, compute balancing state-space transformation so that

$$P = Q = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) = \Sigma, \quad \sigma_1 \geq \ldots \geq \sigma_n > 0,$$

and truncate corresponding realization at size r with $\sigma_r > \sigma_{r+1}$.

Other Balancing-Based Methods

- Bounded-real balanced truncation (BRBT) – based on bounded real lemma [Opdenacker/Jonckheere '88];
- H_∞ balanced truncation (HinfBT) – closed-loop balancing based on H_∞ compensator [Mustafa/Glover '91].

Both approaches require solution of dual AREs.

- Frequency-weighted versions of the above approaches.

Balancing-Related Methods

Properties

- Guaranteed preservation of physical properties like
  - stability (all),
  - passivity (PRBT),
  - minimum phase (BST).
- Computable error bounds, e.g.,

$$\text{BT:} \quad \|G - G_r\|_\infty \leq 2 \sum_{j=r+1}^{n} \sigma_j^{BT},$$

$$\text{LQGBT:} \quad \|G - G_r\|_\infty \leq 2 \sum_{j=r+1}^{n} \frac{\sigma_j^{LQG}}{\sqrt{1+(\sigma_j^{LQG})^2}}$$

$$\text{BST:} \quad \|G - G_r\|_\infty \leq \left(\prod_{j=r+1}^{n} \frac{1+\sigma_j^{BST}}{1-\sigma_j^{BST}} - 1 \right) \|G\|_\infty,$$

- Can be combined with singular perturbation approximation for steady-state performance.
- Computations can be modularized.

References for BT I

U. B. Desai and D. Pal.
A transformation approach to stochastic model reduction.
IEEE Trans. Autom. Control, AC–29:1097–1100, 1984.

M. Green.
Balanced stochastic realization.
Linear Algebra Appl., 98:211–247, 1988.

E. A. Jonckheere and L. M. Silverman.
A new set of invariants for linear systems – application to reduced order compensator design.
IEEE Trans. Autom. Control, 28:953–964, 1983.

C. Mullis and R. A. Roberts.
Synthesis of minimum roundoff noise fixed point digital filters.
IEEE Trans. Circuits and Systems, CAS-23(9):551–562, 1976.

References for BT II

D. Mustafa and K. Glover.
Controller design by $\mathcal{H}_\infty$-balanced truncation.
IEEE Trans. Autom. Control, 36(6):668–682, 1991.

P. C. Opdenacker and E. A. Jonckheere.
A contraction mapping preserving balanced reduction scheme and its infinity norm error bounds.
IEEE Trans. Circuits Syst., 35(2):184–189, 1988.

Outline

Linear Time-invariant DAEs

System Theoretic Aspects of DAEs

Consider

$$Ex(t) = Ax(t) + Bu(t), \quad x(0) = x_0,$$
$$y(t) = Cx(t),$$

where

- $x(t) \in \mathbb{R}^n$: the system's state
- $u(t) \in \mathbb{R}^m$: the input or control
- $y(t) \in \mathbb{R}^q$: the output or measurements

- $E \in \mathbb{R}^{n \times n}$ is *singular*
- $A \in \mathbb{R}^{n \times n}$: the system matrix
- $B \in \mathbb{R}^{n \times m}$: the input matrix
- $C \in \mathbb{R}^{q \times n}$: the output matrix

- We will denote the system by $(E; A, B, C, D)$.
- $(E; A, B, C, D)$ are referred to as descriptor or singular systems.

Linear Time-invariant DAEs

System Theoretic Aspects of DAEs

Consider

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = x_0, \\ y(t) &= Cx(t), \end{aligned}$$

where

- $x(t) \in \mathbb{R}^n$: the system's state
- $u(t) \in \mathbb{R}^m$: the input or control
- $y(t) \in \mathbb{R}^q$: the output or measurements

- $E \in \mathbb{R}^{n\times n}$ is *singular*
- $A \in \mathbb{R}^{n\times n}$: the system matrix
- $B \in \mathbb{R}^{n\times m}$: the input matrix
- $C \in \mathbb{R}^{q\times n}$: the output matrix

- We will denote the system by $(E; A, B, C, D)$.
- $(E; A, B, C, D)$ are referred to as descriptor or singular systems.

Linear Time-invariant DAEs

System Theoretic Aspects of DAEs

Consider

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = x_0, \\ y(t) &= Cx(t), \end{aligned}$$

where

- $x(t) \in \mathbb{R}^n$: the system's state
- $u(t) \in \mathbb{R}^m$: the input or control
- $y(t) \in \mathbb{R}^q$: the output or measurements

- $E \in \mathbb{R}^{n\times n}$ is *singular*
- $A \in \mathbb{R}^{n\times n}$: the system matrix
- $B \in \mathbb{R}^{n\times m}$: the input matrix
- $C \in \mathbb{R}^{q\times n}$: the output matrix

- We will denote the system by $(E; A, B, C, D)$.
- $(E; A, B, C, D)$ are referred to as descriptor or singular systems.

Linear Time-invariant DAEs

System Theoretic Aspects of DAEs

The transfer function of an $(E; A, B, C, D)$ system in time domain:

$\mathbf{G}\colon u \mapsto y\colon$

$$y(t) = C\Big[e^{E^{\mathrm{D}}At}x_0 + \int_0^t e^{E^{\mathrm{D}}A(t-\tau)}E^{\mathrm{D}}Bu(\tau)\,\mathrm{d}\tau - \\ - (I - E^{\mathrm{D}}E)\sum_{i=0}^{\nu-1}(EA^{\mathrm{D}})^i A^{\mathrm{D}}Bu^{(i)}(t)\Big] + Du(t),$$

where

- E^{D} is the Drazin inverse of E
- ν is the differentiation index of the DAE $E\dot{x} = Ax$
- $u^{(i)}$ denotes the i-th derivative of u

Note that if $E = I$, then $E^{\mathrm{D}} = I$ and the transfer function is well-known:

$$\mathbf{G}\colon u \mapsto y\colon y(t) = C\Big[e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)\,\mathrm{d}\tau\Big] + Du(t).$$

Linear Time-invariant DAEs
System Theoretic Aspects of DAEs

The transfer function of an $(E; A, B, C, D)$ system in time domain:

$\mathbf{G}\colon u \mapsto y\colon$

$$y(t) = C\Big[e^{E^{\mathrm{D}}At}x_0 + \int_0^t e^{E^{\mathrm{D}}A(t-\tau)}E^{\mathrm{D}}Bu(\tau)\,\mathrm{d}\tau - \\ - (I - E^{\mathrm{D}}E)\sum_{i=0}^{\nu-1}(EA^{\mathrm{D}})^i A^{\mathrm{D}}Bu^{(i)}(t)\Big] + Du(t),$$

where

- E^{D} is the Drazin inverse of E
- ν is the differentiation index of the DAE $E\dot{x} = Ax$
- $u^{(i)}$ denotes the i-th derivative of u

Note that if $E = I$, then $E^{\mathrm{D}} = I$ and the transfer function is well-known:

$$\mathbf{G}\colon u \mapsto y\colon y(t) = C\Big[e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)\,\mathrm{d}\tau\Big] + Du(t).$$

Linear Time-invariant DAEs

System Theoretic Aspects of DAEs

- In frequency domain (after a *Laplace* transform) the transfer function is given as

$$G(s) = C(sE - A)^{-1}B + D$$

.

- Depending on B and C, the transfer function is likely to be improper.
- Since standard model reduction approaches return a *proper* reduced system $\hat{G}$, the error will grow for large frequencies.

Linear Time-invariant DAEs

System Theoretic Aspects of DAEs

- In frequency domain (after a *Laplace* transform) the transfer function is given as

$$G(s) = C(sE - A)^{-1}B + D$$

.

- Depending on B and C, the transfer function is likely to be improper.
- Since standard model reduction approaches return a *proper* reduced system $\hat{G}$, the error will grow for large frequencies.

For an improper it holds that $\|G(s)\| \to \infty \quad \text{as } s \to \infty$.

Linear Time-invariant DAEs

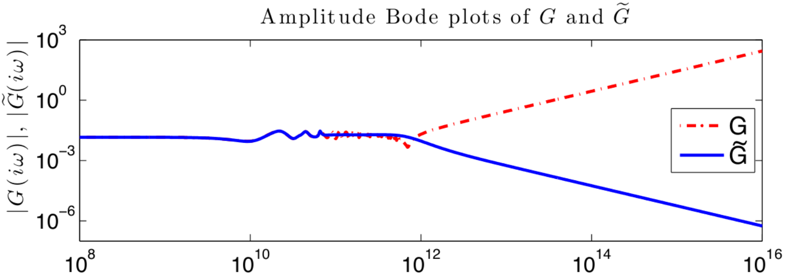
System Theoretic Aspects of DAEs

- In frequency domain (after a *Laplace* transform) the transfer function is given as $G(s) = C(sE - A)^{-1}B + D$.
- Depending on B and C, the transfer function is likely to be improper.
- Since standard model reduction approaches return a *proper* reduced system $\hat{G}$, the error will grow for large frequencies.

Linear Time-invariant DAEs

System Theoretic Aspects of DAEs

- In frequency domain (after a *Laplace* transform) the transfer function is given as $G(s) = C(sE - A)^{-1}B + D$.
- Depending on B and C, the transfer function is likely to be improper.
- Since standard model reduction approaches return a *proper* reduced system $\hat{G}$, the error will grow for large frequencies.



Plot taken from [Gugercin, Stykel, Wyatt '13].

Linear Time-invariant DAEs

System Theoretic Aspects of DAEs

The general problem is:

- the transfer function can have an improper part (frequency domain)
- the system differentiates the input (time domain)

The general approach is:

1. Project the DAE onto the part that is an ODE, i.e. a standard state space system
2. Keep the remainder, i.e. the algebraic or improper part, as it is

This means: no model reduction on the algebraic part!

Linear Time-invariant DAEs

Balanced Truncation for Navier-Stokes Systems

We consider linearized Navier-Stokes equations:

$$\begin{aligned} M\dot{v}(t) &= A_1 v(t) + J^T p(t) + B_1 u(t), \\ Jv(t) &= B_2 u(t), \\ y(t) &= C_1 v(t) + C_2 p(t). \end{aligned}$$

- $v(t) \in \mathbb{R}^n$: state (velocity)
- $p(t) \in \mathbb{R}^p$: state (pressure)
- $u(t) \in \mathbb{R}^m$: input or control
- $y(t) \in \mathbb{R}^q$: the output or measurements

- $M \in \mathbb{R}^{n\times n}$: mass matrix (symmetric
- $A_1 \in \mathbb{R}^{n\times n}$: the system matrix
- $J \in \mathbb{R}^{p\times n}$ is another system matrix (full rank)
- $B_1 \in \mathbb{R}^{n\times m}$, $B_2 \in \mathbb{R}^{p\times m}$: input matrices
- $C_1 \in \mathbb{R}^{q\times n}$, $C_2 \in \mathbb{R}^{q\times p}$: output matrices

Note that this is an $(E; A, B, C, D)$ with

$$E := \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix}, \quad A := \begin{bmatrix} A_1 & -J \\ J^T & 0 \end{bmatrix}, \quad B := \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad \text{and} \quad C := \begin{bmatrix} C_1 & C_2 \end{bmatrix}.$$

Linear Time-invariant DAEs

Balanced Truncation for Navier-Stokes Systems

We consider linearized Navier-Stokes equations:

$$\begin{aligned} M\dot{v}(t) &= A_1 v(t) + J^T p(t) + B_1 u(t), \\ Jv(t) &= B_2 u(t), \\ y(t) &= C_1 v(t) + C_2 p(t). \end{aligned}$$

- $v(t) \in \mathbb{R}^n$: state (velocity)
- $p(t) \in \mathbb{R}^p$: state (pressure)
- $u(t) \in \mathbb{R}^m$: input or control
- $y(t) \in \mathbb{R}^q$: the output or measurements

- $M \in \mathbb{R}^{n\times n}$: mass matrix (symmetric
- $A_1 \in \mathbb{R}^{n\times n}$: the system matrix
- $J \in \mathbb{R}^{p\times n}$ is another system matrix (full rank)
- $B_1 \in \mathbb{R}^{n\times m}$, $B_2 \in \mathbb{R}^{p\times m}$: input matrices
- $C_1 \in \mathbb{R}^{q\times n}$, $C_2 \in \mathbb{R}^{q\times p}$: output matrices

Note that this is an $(E; A, B, C, D)$ with

$$E := \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix}, \quad A := \begin{bmatrix} A_1 & -J \\ J^T & 0 \end{bmatrix}, \quad B := \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad \text{and} \quad C := \begin{bmatrix} C_1 & C_2 \end{bmatrix}.$$

Linear Time-invariant DAEs

Decoupling Differential and Algebraic Parts

$$\begin{aligned} M\dot{v}(t) &= A_1 v(t) + J^T p(t) + B_1 u(t), \\ Jv(t) &= B_2 u(t), \\ y(t) &= C_1 v(t) + C_2 p(t). \end{aligned}$$

Consider the projector

$$P := I - M^{-1}J^{\mathsf{T}}(JM^{-1}J^{\mathsf{T}})^{-1}J$$

and see that with $v = Pv + (I - P)v =: v_d + v_a$ the system writes as

$$\begin{aligned} M\dot{v}_d(t) &= P^{\mathsf{T}}A_1 v_d(t) + P^{\mathsf{T}}A_1 v_a(t) + P^{\mathsf{T}}B_1 u(t), \\ v_a(t) &= -M^{-1}J^{\mathsf{T}}(JM^{-1}J^{\mathsf{T}})^{-1}JB_2 u(t), \\ p(t) &= -(JM^{-1}J^{\mathsf{T}})^{-1}\big[JM^{-1}[A(v_a(t) + v_d(t)) + B_1 u(t)] - B_2\dot{u}(t)\big], \\ y(t) &= C_1 v_d(t) + C_1 v_a(t) + C_2 p(t). \end{aligned}$$

Linear Time-invariant DAEs

Decoupling Differential and Algebraic Parts

Since v_a and p depend linearly on v_d, u, and $\dot{u}$ is an $(E; A, B, C, D)$ system with the state v_d and

$$\begin{aligned} E &:= M, \\ A &:= P^\mathsf{T} A, \\ B &:= P^\mathsf{T}[B_1 - AM^{-1}J^\mathsf{T}(JM^{-1}J^\mathsf{T})^{-1}JB_2], \\ C &:= C_1 - C_2(JM^{-1}J^\mathsf{T})^{-1}JM^{-1}A, \\ D &:= D_1 + D_2, \end{aligned}$$

with

$$D_1 := -C_1M^{-1}J^\mathsf{T}(JM^{-1}J^\mathsf{T})^{-1}JB_2 + C_2(JM^{-1}J^\mathsf{T})^{-1}JM^{-1}AM^{-1}J^\mathsf{T}(JM^{-1}J^\mathsf{T})^{-1}JB_1$$

$$D_2 := -C_2(JM^{-1}J^\mathsf{T})^{-1}B_2\frac{\mathrm{d}}{\mathrm{d}t}.$$

Linear Time-invariant DAEs
Decoupling Differential and Algebraic Parts

$$D_1 = -C_1 M^{-1} J^\mathsf{T} (J M^{-1} J^\mathsf{T})^{-1} J B_2 + C_2 (J M^{-1} J^\mathsf{T})^{-1} J M^{-1} A M^{-1} J^\mathsf{T} (J M^{-1} J^\mathsf{T})^{-1} J B_1$$
$$D_2 = -C_2 (J M^{-1} J^\mathsf{T})^{-1} B_2 \tfrac{\mathrm{d}}{\mathrm{d}t}.$$

Note that

- The transfer function is given as $G = C(sE - A)^{-1}B + D_1 + sD_2$
- if B_2 or C_2 is zero, then D_2 is zero,
  - no $\dot{u}$ in the output
  - no obviously improper part sD_2 in G
- if B_2 is zero, then D_1, $D_2 = 0$
  - we obtain a standard $(E; A, B, C, D)$ system
  - no improper parts in G

Linear Time-invariant DAEs

Decoupling Differential and Algebraic Parts

$$D_1 = -C_1 M^{-1} J^\mathsf{T} (JM^{-1}J^\mathsf{T})^{-1} JB_2 + C_2 (JM^{-1}J^\mathsf{T})^{-1} JM^{-1} AM^{-1} J^\mathsf{T} (JM^{-1}J^\mathsf{T})^{-1} JB_1$$
$$D_2 = -C_2 (JM^{-1}J^\mathsf{T})^{-1} B_2 \tfrac{\mathrm{d}}{\mathrm{d}t}.$$

Note that

- The transfer function is given as $G = C(sE - A)^{-1}B + D_1 + sD_2$
- if B_2 or C_2 is zero, then D_2 is zero,
  - no $\dot{u}$ in the output
  - no obviously improper part sD_2 in G
- if B_2 is zero, then D_1, $D_2 = 0$
  - we obtain a standard $(E; A, B, C, D)$ system
  - no improper parts in G

Linear Time-invariant DAEs

Decoupling Differential and Algebraic Parts

$$D_1 = -C_1 M^{-1} J^\mathsf{T} (J M^{-1} J^\mathsf{T})^{-1} J B_2 + C_2 (J M^{-1} J^\mathsf{T})^{-1} J M^{-1} A M^{-1} J^\mathsf{T} (J M^{-1} J^\mathsf{T})^{-1} J B_1$$
$$D_2 = -C_2 (J M^{-1} J^\mathsf{T})^{-1} B_2 \frac{\mathrm{d}}{\mathrm{d}t}.$$

Note that

- The transfer function is given as $G = C(sE - A)^{-1} B + D_1 + s D_2$
- if B_2 or C_2 is zero, then D_2 is zero,
  - no $\dot{u}$ in the output
  - no obviously improper part sD_2 in G
- if B_2 is zero, then $D_1,\ D_2 = 0$
  - we obtain a standard $(E; A, B, C, D)$ system
  - no improper parts in G

Linear Time-invariant DAEs

Decoupling Differential and Algebraic Parts

If B_2 and C_2 are zero, then we have a standard $(A, B, C, -)$ system:

$$M\dot{v}_d = P^\mathsf{T} A v_d + P^\mathsf{T} B_1 u,$$
$$y = C_1 v.$$

If we want to apply Balanced Truncation, we need to cope with the following difficulties:

- The system is not minimal
  - $\rightarrow$ this is automatically *fixed* by BT, if we can find the right solutions of the nonregular Lyapunov equations like

$$MXP^T A + APXM + P^\mathsf{T} BB^\mathsf{T} P = 0.$$

- The system is not stable
  - $\rightarrow$ Combine BT with *LQG*-stabilization [Benner and Heiland, '15]
- Explicit computation of the projector P is not possible for large scale systems
  - $\rightarrow$ use algorithms that do not need P explicitly, cf. [Gugercin, Stykel, Wyatt '13], [Heinkenschloss, Sorensen, Sun '08], and [Benner and Heiland, '15]

Linear Time-invariant DAEs

Decoupling Differential and Algebraic Parts

If B_2 and C_2 are zero, then we have a standard $(A, B, C, -)$ system:

$$M\dot{v}_d = P^\mathsf{T} A v_d + P^\mathsf{T} B_1 u,$$
$$y = C_1 v.$$

If we want to apply Balanced Truncation, we need to cope with the following difficulties:

- The system is not minimal
  - $\rightarrow$ this is automatically *fixed* by BT, if we can find the right solutions of the nonregular Lyapunov equations like

$$MXP^T A + APXM + P^\mathsf{T} BB^\mathsf{T} P = 0.$$

- The system is not stable
  - $\rightarrow$ Combine BT with *LQG*-stabilization [Benner and Heiland, '15]
- Explicit computation of the projector P is not possible for large scale systems
  - $\rightarrow$ use algorithms that do not need P explicitly, cf. [Gugercin, Stykel, Wyatt '13], [Heinkenschloss, Sorensen, Sun '08], and [Benner and Heiland, '15]

Linear Time-invariant DAEs

Decoupling Differential and Algebraic Parts

If B_2 and C_2 are zero, then we have a standard $(A, B, C, -)$ system:

$$M\dot{v}_d = P^\mathsf{T} A v_d + P^\mathsf{T} B_1 u,$$
$$y = C_1 v.$$

If we want to apply Balanced Truncation, we need to cope with the following difficulties:

- The system is not minimal
  - $\rightarrow$ this is automatically *fixed* by BT, if we can find the right solutions of the nonregular Lyapunov equations like
  $$MXP^T A + APXM + P^\mathsf{T} BB^\mathsf{T} P = 0.$$
- The system is not stable
  - $\rightarrow$ Combine BT with *LQG*-stabilization [Benner and Heiland, '15]
- Explicit computation of the projector P is not possible for large scale systems
  - $\rightarrow$ use algorithms that do not need P explicitly, cf. [Gugercin, Stykel, Wyatt '13], [Heinkenschloss, Sorensen, Sun '08], and [Benner and Heiland, '15]

Linear Time-invariant DAEs

Decoupling Differential and Algebraic Parts

If B_2 and C_2 are zero, then we have a standard $(A, B, C, -)$ system:

$$\begin{aligned} M\dot{v}_d &= P^{\mathsf{T}} A v_d + P^{\mathsf{T}} B_1 u, \\ y &= C_1 v. \end{aligned}$$

If we want to apply Balanced Truncation, we need to cope with the following difficulties:

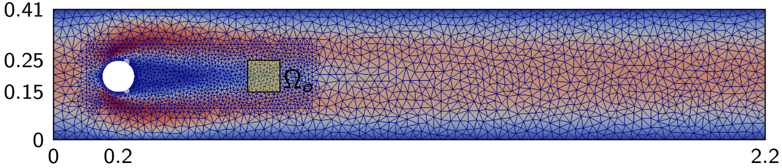
- The system is not minimal
  - $\rightarrow$ this is automatically *fixed* by BT, if we can find the right solutions of the nonregular Lyapunov equations like

$$MXP^T A + APXM + P^{\mathsf{T}} BB^{\mathsf{T}} P = 0.$$

- The system is not stable
  - $\rightarrow$ Combine BT with *LQG*-stabilization [Benner and Heiland, '15]
- Explicit computation of the projector P is not possible for large scale systems
  - $\rightarrow$ use algorithms that do not need P explicitly, cf. [Gugercin, Stykel, Wyatt '13], [Heinkenschloss, Sorensen, Sun '08], and [Benner and Heiland, '15]

Linear Time-invariant DAEs

Numerical Example NSE

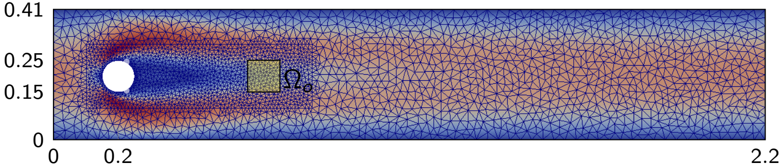


- 2D cylinder wake
- Navier-Stokes Equations
- $Re = 100$
- *Taylor-Hood* finite elements
- 30000 velocity nodes

- Boundary control at 2 outlets
- distributed observation with 6 degrees of freedom
- LQGBT-reduced order observer and controller of state dimension $r = 13$
- Target: stabilization of the steady-state solution

Linear Time-invariant DAEs

Numerical Example NSE



- 2D cylinder wake
- Navier-Stokes Equations
- $Re = 100$
- *Taylor-Hood* finite elements
- 30000 velocity nodes

- Boundary control at 2 outlets
- distributed observation with 6 degrees of freedom
- LQGBT-reduced order observer and controller of state dimension $r = 13$
- Target: stabilization of the steady-state solution

Linear Time-invariant DAEs
LQGBT Reduction - Bode Plot

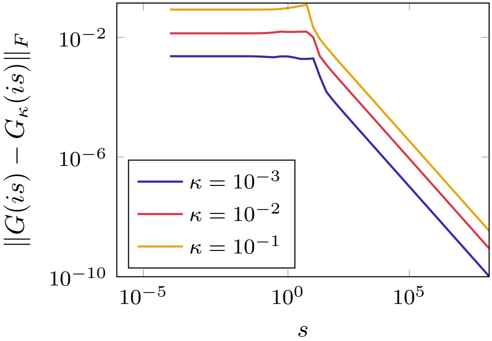


Figure : The error in the frequency response for varying thresholds κ measured in the Frobenius norm with i denoting the imaginary unit and the transfer functions in frequency domain as defined, e.g., in [4].

This plot was taken from [Benner and Heiland, '15]. The varying thresholds κ correspond to reduced models of order $r = 4, 13, 22$.

Linear Time-invariant DAEs

Cylinder Wake Stabilization

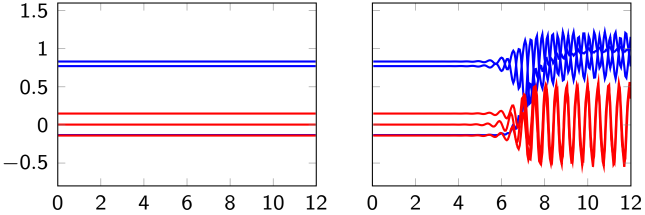
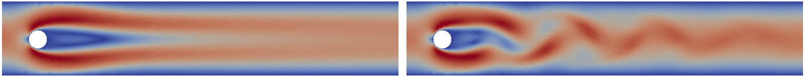


Figure : Measured signal y versus time $t \in [0, 12]$ of the perturbed closed loop system with a reduced controller of dimension $r = 13$ (left), compared to the response of the uncontrolled system (right). Blue corresponds to the x-component of the velocity and red to y-component. Below, a snapshot of the magnitude of the velocity solutions at $t = 12$.



Linear Time-invariant DAEs

Conclusion

- Linear Time Invariant DAEs typically have improper transfer functions
- One can decouple a DAE to extract the differential/proper part of the system
- The differential part is a standard $(A, B, C, -)$ and can be reduced with standard methods
- The algebraic part must not be reduced
- The efficient implementation requires further effort
- For Navier-Stokes equations there are examples of efficient application of BT related methods

Linear Time-invariant DAEs

Conclusion

- Linear Time Invariant DAEs typically have improper transfer functions
- One can decouple a DAE to extract the differential/proper part of the system
- The differential part is a standard $(A, B, C, -)$ and can be reduced with standard methods
- The algebraic part must not be reduced
- The efficient implementation requires further effort
- For Navier-Stokes equations there are examples of efficient application of BT related methods

Linear Time-invariant DAEs

Conclusion

- Linear Time Invariant DAEs typically have improper transfer functions
- One can decouple a DAE to extract the differential/proper part of the system
- The differential part is a standard $(A, B, C, -)$ and can be reduced with standard methods
- The algebraic part must not be reduced
- The efficient implementation requires further effort
- For Navier-Stokes equations there are examples of efficient application of BT related methods

Linear Time-invariant DAEs

Conclusion

- Linear Time Invariant DAEs typically have improper transfer functions
- One can decouple a DAE to extract the differential/proper part of the system
- The differential part is a standard $(A, B, C, -)$ and can be reduced with standard methods
- The algebraic part must not be reduced
- The efficient implementation requires further effort
- For Navier-Stokes equations there are examples of efficient application of BT related methods

Linear Time-invariant DAEs

Conclusion

- Linear Time Invariant DAEs typically have improper transfer functions
- One can decouple a DAE to extract the differential/proper part of the system
- The differential part is a standard $(A, B, C, -)$ and can be reduced with standard methods
- The algebraic part must not be reduced
- The efficient implementation requires further effort
- For Navier-Stokes equations there are examples of efficient application of BT related methods

Linear Time-invariant DAEs
Conclusion

- Linear Time Invariant DAEs typically have improper transfer functions
- One can decouple a DAE to extract the differential/proper part of the system
- The differential part is a standard $(A, B, C, -)$ and can be reduced with standard methods
- The algebraic part must not be reduced
- The efficient implementation requires further effort
- For Navier-Stokes equations there are examples of efficient application of BT related methods

Literature on DAE-NSE I

P. Kunkel and V. Mehrmann.
Differential-Algebraic Equations. Analysis and Numerical Solution.
European Mathematical Society Publishing House, Zürich, Switzerland, 2006.

P. Benner and J. Heiland.
LQG-balanced truncation low-order controller for stabilization of laminar flows.
In R. King, editor, *Active Flow and Combustion Control 2014*, volume 127 of *Notes on Numerical Fluid Mechanics and Multidisciplinary Design*, pages 365–379. Springer International Publishing, 2015.

S. Gugercin, T. Stykel, and S. Wyatt.
Model reduction of descriptor systems by interpolatory projection methods.
SIAM J. Sci. Comput., 35(5):B1010–B1033, 2013.

Literature on DAE-NSE II

M. Heinkenschloss, D. C. Sorensen, and K. Sun.
Balanced truncation model reduction for a class of descriptor systems with applications to the Oseen equations.
SIAM J. Sci. Comput., 30(2):1038–1063, 2008.

Further Reading — Model Order Reduction

1. G. Obinata and B.D.O. Anderson.
Model Reduction for Control System Design.
Springer-Verlag, London, UK, 2001.

2. Z. Bai.
Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems.
Appl. Numer. Math, 43(1–2):9–44, 2002.

3. R. Freund.
Model reduction methods based on Krylov subspaces.
Acta Numerica, 12:267–319, 2003.

4. P. Benner, E.S. Quintana-Ortí, and G. Quintana-Ortí.
State-space truncation methods for parallel model reduction of large-scale systems.
Parallel Comput., 29:1701–1722, 2003.

5. P. Benner, V. Mehrmann, and D. Sorensen (editors).
Dimension Reduction of Large-Scale Systems.
Lecture Notes in Computational Science and Engineering, Vol. 45,
Springer-Verlag, Berlin/Heidelberg, Germany, 2005.

6. A.C. Antoulas.
Lectures on the Approximation of Large-Scale Dynamical Systems.
SIAM Publications, Philadelphia, PA, 2005.

7. P. Benner, R. Freund, D. Sorensen, and A. Varga (editors).
Special issue on *Order Reduction of Large-Scale Systems*.
Linear Algebra Appl., June 2006.

8. W.H.A. Schilders, H.A. van der Vorst, and J. Rommes (editors).
Model Order Reduction: Theory, Research Aspects and Applications.
Mathematics in Industry, Vol. 13,
Springer-Verlag, Berlin/Heidelberg, 2008.

9. P. Benner, J. ter Maten, and M. Hinze (editors).
Model Reduction for Circuit Simulation.
Lecture Notes in Electrical Engineering, Vol. 74,
Springer-Verlag, Dordrecht, 2011.

Further Reading — Matrix Equations

1. V. Mehrmann.
The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution.
Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, July 1991.

2. P. Lancaster and L. Rodman.
The Algebraic Riccati Equation.
Oxford University Press, Oxford, 1995.

3. P. Benner.
Computational methods for linear-quadratic optimization
Rendiconti del Circolo Matematico di Palermo, Supplemento, Serie II, 58:21–56, 1999.

4. T. Penzl.
Lyapack Users Guide.
Technical Report SFB393/00-33, Sonderforschungsbereich 393 *Numerische Simulation auf massiv parallelen Rechnern*, TU Chemnitz, 09107 Chemnitz, FRG, 2000.
Available from `http://www.tu-chemnitz.de/sfb393/sfb00pr.html`.

5. H. Abou-Kandil, G. Freiling, V. Ionescu, and G. Jank.
Matrix Riccati Equations in Control and Systems Theory.
Birkhäuser, Basel, Switzerland, 2003.

6. P. Benner.
Solving large-scale control problems.
IEEE Control Systems Magazine, 24(1):44–59, 2004.

7. D. Bini, B. Iannazzo, and B. Meini.
Numerical Solution of Algebraic Riccati Equations.
SIAM, Philadelphia, PA, 2012.

8. P. Benner and J. Saak.
Numerical solution of large and sparse continuous time algebraic matrix Riccati and Lyapunov equations: a state of the art survey.
GAMM-Mitteilungen, 36(1):32-Â–52, 2013.

9. V. Simoncini.
Computational methods for linear matrix equations (survey article).
March 2013.
`http://www.dm.unibo.it/~simoncin/matrixeq.pdf`.