CME 345: MODEL REDUCTION

Moment Matching

David Amsallem & Charbel Farhat
Stanford University
cfarhat@stanford.edu

These slides are based on the recommended textbook: A.C. Antoulas, "Approximation of Large-Scale Dynamical Systems," Advances in Design and Control, SIAM, ISBN-0-89871-529-6

Outline

$$\begin{aligned}\frac{d}{dt}\mathbf{x}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)\\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)\\ \mathbf{x}(0) &= \mathbf{x}_0\end{aligned}$$

- $\mathbf{x} \in \mathbb{R}^n$: vector state variables
- $\mathbf{u} \in \mathbb{R}^p$: vector of input variables, typically $p \ll n$
- $\mathbf{y} \in \mathbb{R}^q$: vector output variables, typically $q \ll n$

- Goal: obtain a **Reduced-Order Model** (ROM)

$$\begin{aligned} \frac{d}{dt}\mathbf{x}_r(t) &= \mathbf{A}_r\mathbf{x}_r(t) + \mathbf{B}_r\mathbf{u}(t) \\ \mathbf{y}_r(t) &= \mathbf{C}_r\mathbf{x}_r(t) + \mathbf{D}_r\mathbf{u}(t) \end{aligned}$$

  - $\mathbf{x}_r \in \mathbb{R}^k$: **vector of reduced state variables**
- ROM resulting from **Petrov-Galerkin** projection:

$$\begin{aligned} \mathbf{A}_r &= (\mathbf{W}^T\mathbf{V})^{-1}\mathbf{W}^T\mathbf{A}\mathbf{V} \in \mathbb{R}^{k\times k} \\ \mathbf{B}_r &= (\mathbf{W}^T\mathbf{V})^{-1}\mathbf{W}^T\mathbf{B} \in \mathbb{R}^{k\times p} \\ \mathbf{C}_r &= \mathbf{C}\mathbf{V} \in \mathbb{R}^{q\times k} \\ \mathbf{D}_r &= \mathbf{D} \in \mathbb{R}^{q\times p} \end{aligned}$$

- Let $\mathbf{h}$ denote a general matrix valued function of time

$$\mathbf{h}: \; t \in \mathbb{R} \longmapsto \mathbb{R}^{q \times p}$$

Example: impulse response of an LTI system

$$\mathbf{h}(t) = \mathbf{C}e^{\mathbf{A}t}\mathbf{B} + \mathbf{D}$$

- $\mathbf{H}(s)$ will denote its Laplace transform, that is

$$\mathbf{H}(s) = \int_0^\infty \mathbf{h}(t)e^{-st}dt$$

Example: impulse response of an LTI system

$$\mathbf{H}(s) = \mathbf{C}(s\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}$$

- Let $m \in \{0, \cdots, \}$.
  The m-th **moment** of $\mathbf{h} : \ t \in \mathbb{R} \longmapsto \mathbb{R}^{q \times p}$ at $s_0 \in \mathbb{C}$ is

$$\eta_m(s_0) = \int_0^{\infty} t^m e^{-s_0 t} \mathbf{h}(t) dt$$

- In terms of its transfer function

$$\eta_m(s_0) = (-1)^m \left. \frac{d^m}{ds^m} \mathbf{H}(s) \right|_{s=s_0}$$

  Example: impulse response of an LTI system

$$\begin{array}{rcl} \eta_0(s_0) & = & \mathbf{C}(s_0 \mathbf{I}_n - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D} \\ \eta_m(s_0) & = & m! \ \mathbf{C}(s_0 \mathbf{I}_n - \mathbf{A})^{-(m+1)} \mathbf{B}, \ \forall m \geq 1 \end{array}$$

- Development of $\mathbf{H}(s)$ in Taylor series

$$
\begin{aligned}
\mathbf{H}(s) &= \mathbf{H}(s_0) + \frac{d}{ds}\mathbf{H}(s)\Big|_{s=s_0} \frac{(s-s_0)}{1!} + \cdots \\
&\quad + \frac{d^m}{ds^m}\mathbf{H}(s)\Big|_{s=s_0} \frac{(s-s_0)^m}{m!} + \cdots \\
&= \eta_0(s_0) - \eta_1(s_0)\frac{(s-s_0)}{1!} + \cdots + (-1)^m \eta_m(s_0)\frac{(s-s_0)^m}{m!} + \cdots \\
&= \eta_0(s_0) + \eta_1(s_0)\frac{(s_0-s)}{1!} + \cdots + \eta_m(s_0)\frac{(s_0-s)^m}{m!} + \cdots
\end{aligned}
$$

- The **Markov parameters** $\eta_m(\infty)$ of the system defined by $\mathbf{h}$ are the coefficient in the Laurent expansion of the transfer function at infinity:

$$\mathbf{H}(s) \approx_{s\to\infty} \sum_{i=0}^{\infty} s^{-i}\eta_m(\infty)$$

Example: impulse response of an LTI system

$$\begin{aligned} \eta_0(\infty) &= \mathbf{D} \\ \eta_m(\infty) &= \mathbf{C}\mathbf{A}^{m-1}\mathbf{B},\ \forall m \geq 1 \end{aligned}$$

- *Goal*: Let $s_0 \in \mathbb{C}$ and a FOM $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ defined by its transfer function $\mathbf{H}(s)$.
- Compute a ROM $(\mathbf{A}_r, \mathbf{B}_r, \mathbf{C}_r, \mathbf{D}_r)$ such that the first l moments $\eta_{r,j}(s_0)$ of its transfer function $\mathbf{H}_r$ at s_0 match their counterparts from the FOM.
- In other words

$$\eta_{r,j}(s_0) = \eta_j(s_0) \Leftrightarrow \mathbf{H}^{(j)}(s_0) = \mathbf{H}_r^{(j)}(s_0), \ \forall j = 0, \cdots, l-1$$

  - a direct matching of the moment is a numerically unstable procedure
  - an iterative procedure based on Krylov subspaces addresses that issue

Theorem

Let $\mathbf{V}$ *be a right reduced-order basis such that*

$$span(\mathbf{V}) = span\left\{\mathbf{b}, \mathbf{A}\mathbf{b}, \cdots, \mathbf{A}^{k-1}\mathbf{b}\right\}$$

and $\mathbf{W}$ *be a left reduced-order basis satisfying*

$$\mathbf{W}^T\mathbf{V} = \mathbf{I}$$

Then, the ROM obtained by Petrov-Galerkin projection of the HFM $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ *using* $\mathbf{W}$ *and* $\mathbf{V}$ *satisfies*

$$\eta_{r,j}(\infty) = \eta_j(\infty) \Leftrightarrow \mathbf{H}_r^{(j)}(\infty) = \mathbf{H}^{(j)}(\infty), \ \forall j = 0, \cdots, k-1$$

- span $\left\{\mathbf{b}, \mathbf{A}\mathbf{b}, \cdots, \mathbf{A}^{k-1}\mathbf{b}\right\}$ is the **Krylov subspace** $\mathcal{K}_k(\mathbf{A}, \mathbf{b})$.

To prove the theorem, the following lemma is used

Lemma

The moments of the transfer function of a ROM only depend on the left and right subspaces considered and not the corresponding reduced-order bases.

Proof.

From the lemma, we can choose without loss of generality

$$\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_k] = [\mathbf{b}, \mathbf{A}\mathbf{b}, \cdots, \mathbf{A}^{k-1}\mathbf{b}]$$

Note that $\mathbf{W}^T\mathbf{V} = \mathbf{I}_k$ and $\mathbf{A}\mathbf{V}\mathbf{W}^T\mathbf{v}_i = \mathbf{A}\mathbf{V}\mathbf{e}_i = \mathbf{A}\mathbf{v}_i = \mathbf{v}_{i+1} = \mathbf{A}^i\mathbf{b}$. Then

$$\begin{aligned}
\eta_{r,0}(\infty) &= \mathbf{D} = \eta_0(\infty) \\
\eta_{r,1}(\infty) &= \mathbf{c}_r\mathbf{b}_r = \mathbf{c}\mathbf{V}\mathbf{W}^T\mathbf{b} = \mathbf{c}\mathbf{V}\mathbf{W}^T\mathbf{v}_1 = \mathbf{c}\mathbf{V}\mathbf{e}_1 = \mathbf{c}\mathbf{b} = \eta_1(\infty) \\
\eta_{r,j}(\infty) &= \mathbf{c}_r\mathbf{A}_r^j\mathbf{b}_r = \mathbf{c}\mathbf{V}\mathbf{W}^T(\mathbf{A}\mathbf{V}\mathbf{W}^T)^j\mathbf{b} = \mathbf{c}\mathbf{V}\mathbf{W}^T(\mathbf{A}\mathbf{V}\mathbf{W}^T)^j\mathbf{v}_1 \\
&= \mathbf{c}\mathbf{V}\mathbf{W}^T\mathbf{v}_{j+1} = \mathbf{c}\mathbf{V}\mathbf{e}_{j+1} = \mathbf{c}\mathbf{v}_{j+1} = \mathbf{c}\mathbf{A}^j\mathbf{b} = \eta_j(\infty) \quad \blacksquare
\end{aligned}$$

Theorem

Let $s_0 \in \mathbb{C}$ and $\mathbf{V}$ be a right reduced-order basis such that

$$span(\mathbf{V}) = span\left\{(s_0\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{b}, \cdots, (s_0\mathbf{I}_n - \mathbf{A})^{-k}\mathbf{b}\right\}$$

and $\mathbf{W}$ be a left reduced-order basis satisfying

$$\mathbf{W}^T\mathbf{V} = \mathbf{I}$$

Then, the ROM obtained by Petrov-Galerkin projection of the HFM $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ using $\mathbf{W}$ and $\mathbf{V}$ satisfies

$$\eta_{r,j}(s_0) = \eta_j(s_0) \Leftrightarrow \mathbf{H}_r^{(j)}(s_0) = \mathbf{H}^{(j)}(s_0), \ \forall j = 0, \cdots, k-1$$

- $\left\{(s_0\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{b}, \cdots, (s_0\mathbf{I}_n - \mathbf{A})^{-k}\mathbf{b}\right\}$ is the **Krylov subspace** $\mathcal{K}_k((s_0\mathbf{I}_n - \mathbf{A})^{-1}, (s_0\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{b})$.
- This is a more expensive procedure as each Krylov iteration requires the solution of a large scale linear system

Theorem

Let $s_i \in \mathbb{C}, \; i = 1, \cdots, k$ and $\mathbf{V}$ be a right reduced-order basis such that

$$span(\mathbf{V}) = span\left\{(s_1\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{b}, \cdots, (s_k\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{b}\right\}$$

and $\mathbf{W}$ be a left reduced-order basis satisfying

$$\mathbf{W}^T\mathbf{V} = \mathbf{I}$$

Then, the ROM obtained by Petrov-Galerkin projection of the HFM $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ using $\mathbf{W}$ and $\mathbf{V}$ satisfies

$$\eta_{r,0}(s_i) = \eta_0(s_i) \Leftrightarrow \mathbf{H}_r(s_i) = \mathbf{H}(s_i), \; \forall i = 1, \cdots, k$$

Theorem

Let $s_i \in \mathbb{C}$, $i = 1, \cdots, 2k$, $\mathbf{V}$ be a right reduced-order basis such that

$$span(\mathbf{V}) = span\left\{(s_1\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{b}, \cdots, (s_k\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{b}\right\}$$

and $\mathbf{W}$ be a left reduced-order basis such that

$$span(\mathbf{W}) = span\left\{(s_{k+1}\mathbf{I}_n - \mathbf{A}^*)^{-1}\mathbf{c}^*, \cdots, (s_{2k}\mathbf{I}_n - \mathbf{A}^*)^{-1}\mathbf{c}^*\right\}$$

and $\mathbf{W}^T\mathbf{V}$ is nonsingular.

Then, the ROM obtained by Petrov-Galerkin projection of the HFM $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ using $\mathbf{W}$ and $\mathbf{V}$ satisfies

$$\eta_{r,0}(s_i) = \eta_0(s_i) \Leftrightarrow \mathbf{H}_r(s_i) = \mathbf{H}(s_i),\ \forall i = 1, \cdots, 2k$$

Definition

Let $\mathbf{A} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^n$. The k-th Krylov subspace defined by $\mathbf{A}$ and $\mathbf{b}$ is

$$\mathcal{K}_k(\mathbf{A}, \mathbf{b}) = \text{span}\{\mathbf{b}, \mathbf{Ab}, \cdots, \mathbf{A}^{k-1}\mathbf{b}\}$$

Remark: Constructing $\mathcal{K}_k(\mathbf{A}, \mathbf{b})$ only requires the ability to compute the action of the matrix $\mathbf{A}$ onto vectors. This allows "black-box" types of approaches where the matrix $\mathbf{A}$ is not explicitly formed.

- Partial realization requires the construction of $\mathcal{K}_k(\mathbf{A}, \mathbf{b})$, that is the knowledge of the action of $\mathbf{A}$ onto vectors

- Rational interpolation requires the following Krylov space

$$\mathcal{K}_k((s_0\mathbf{I}_n - \mathbf{A})^{-1}, (s_0\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{b}).$$

  Since the knowledge of the action of $(s_0\mathbf{I}_n - \mathbf{A})^{-1} \in \mathbb{R}^{n\times n}$ is needed, two computationally efficient approaches are possible:
  - if n is small enough, an LU factorization of $s_0\mathbf{I}_n - \mathbf{A}$ can be performed and $(s_0\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{v}$ computed by forward and backward substitution for any vector $\mathbf{v} \in \mathbb{R}^n$
  - if n is too large for an LU factorization to be performed, Krylov subspace recycling techniques allowing the reuse of Krylov subspaces for multiple right-hand sides can be used

- $\mathcal{K}_k(\mathbf{A}, \mathbf{b})$ can be efficiently constructed using the Arnoldi factorization method

  Input: $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{b} \in \mathbb{R}^n$
  Output: Orthogonal basis $\mathbf{V}_k \in \mathbb{R}^{n \times k}$ for $\mathcal{K}_k(\mathbf{A}, \mathbf{b})$

- The following recursion is satisfied:

$$\mathbf{A}\mathbf{V}_k = \mathbf{V}_k\mathbf{H}_k + \mathbf{f}_k\mathbf{e}_k^*$$

  with $\mathbf{H}_k = \mathbf{V}_k^*\mathbf{A}\mathbf{V}_k$ an upper Hessenberg matrix, $\mathbf{V}_k^*\mathbf{V}_k = \mathbf{I}_k$ and $\mathbf{V}_k^*\mathbf{f}_k = 0$.

Algorithm:

Input: $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{b} \in \mathbb{R}^n$

Output: Orthogonal basis $\mathbf{V}_k \in \mathbb{R}^{n \times k}$ for $\mathcal{K}_k(\mathbf{A}, \mathbf{b})$

$\mathbf{v}_1 = \mathbf{b}/\|\mathbf{b}\|$;
$\mathbf{w} = \mathbf{A}\mathbf{v}_1$; $\alpha_1 = \mathbf{v}_1^* \mathbf{w}$;
$\mathbf{f}_1 = \mathbf{w} - \alpha_1 \mathbf{v}_1$;
$\mathbf{V}_1 = [\mathbf{v}_1]$; $\mathbf{H} = [\alpha_1]$;
for $j = 1, \cdots, k-1$ **do**
    $\beta_j = \|\mathbf{f}_j\|$; $\mathbf{v}_{j+1} = \mathbf{f}_j / \beta_j$;
    $\mathbf{V}_{j+1} = [\mathbf{V}_j, \ \mathbf{v}_{j+1}]$;
    $\hat{\mathbf{H}}_j = \begin{bmatrix} \mathbf{H}_j \\ \beta_j \mathbf{e}_j^* \end{bmatrix}$
    $\mathbf{w} = \mathbf{A}\mathbf{v}_{j+1}$;
    $\mathbf{h} = \mathbf{V}_{j+1}^* \mathbf{w}$; $\mathbf{f}_{j+1} = \mathbf{w} - \mathbf{V}_{j+1}\mathbf{h}$
    $\mathbf{H}_{j+1} = [\hat{\mathbf{H}}_j, \ \mathbf{h}]$;
end for

- $\mathcal{K}_k(\mathbf{A}, \mathbf{b})$ and $\mathcal{K}_k(\mathbf{A}^*, \mathbf{c}^*)$ can be efficiently simultaneously constructed using the Two-sided Lanczos process

  Input: $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{c}^* \in \mathbb{R}^n$
  Output: Bi-orthogonal bases $\mathbf{V}_k \in \mathbb{R}^{n \times k}$ and $\mathbf{W}_k \in \mathbb{R}^{n \times k}$ for $\mathcal{K}_k(\mathbf{A}, \mathbf{b})$ and $\mathcal{K}_k(\mathbf{A}^*, \mathbf{c}^*)$ respectively satisfying $\mathbf{W}_k^T \mathbf{V}_k = \mathbf{I}_k$

- The following recursions are satisfied:

$$\mathbf{A}\mathbf{V}_k = \mathbf{V}_k \mathbf{T}_k + \mathbf{f}_k \mathbf{e}_k^*,$$

$$\mathbf{A}^*\mathbf{W}_k = \mathbf{W}_k \mathbf{T}_k^* + \mathbf{g}_k \mathbf{e}_k^*,$$

  with $\mathbf{T}_k = \mathbf{W}_k^* \mathbf{A} \mathbf{V}_k$ a tridiagonal matrix, $\mathbf{W}_k^* \mathbf{V}_k = \mathbf{I}_k$, $\mathbf{W}_k^* \mathbf{g}_k = 0$ and $\mathbf{V}^* \mathbf{f}_k = 0$.

Algorithm:
Input: $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{b} \in \mathbb{R}^{n}$, $\mathbf{c}^* \in \mathbb{R}^{n}$
Output: Bi-orthogonal bases $\mathbf{V}_k \in \mathbb{R}^{n\times k}$ and $\mathbf{W}_k \in \mathbb{R}^{n\times k}$ for $\mathcal{K}_k(\mathbf{A}, \mathbf{b})$ and $\mathcal{K}_k(\mathbf{A}^*, \mathbf{c}^*)$ respectively satisfying $\mathbf{W}_k^T \mathbf{V}_k = \mathbf{I}_k$

$\beta_1 = \sqrt{|\mathbf{b}^*\mathbf{c}^*|}$, $\gamma_1 = \text{sign}(\mathbf{b}^*\mathbf{c}^*)\beta_1$
$\mathbf{v}_1 = \mathbf{b}/\beta_1$, $\mathbf{w}_1 = \mathbf{c}^*/\gamma_1$
for $j = 1, \cdots, k-1$ **do**
    $\alpha_j = \mathbf{w}_j^* \mathbf{A} \mathbf{v}_j$;
    $\mathbf{r}_j = \mathbf{A}\mathbf{v}_j - \alpha_j \mathbf{v}_j - \gamma_j \mathbf{v}_{j-1}$;
    $\mathbf{q}_j = \mathbf{A}^*\mathbf{w}_j - \alpha_j \mathbf{w}_j - \beta_j \mathbf{w}_{j-1}$;
    $\beta_{j+1} = \sqrt{|\mathbf{r}_j^* \mathbf{q}_j|}$, $\gamma_{j+1} = \text{sign}(\mathbf{r}_j^* \mathbf{q}_j)\beta_{j+1}$
    $\mathbf{v}_{j+1} = \mathbf{r}_j/\beta_{j+1}$;
    $\mathbf{w}_{j+1} = \mathbf{q}_j/\gamma_{j+1}$;
end for
$\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_k]$, $\mathbf{W} = [\mathbf{w}_1, \cdots, \mathbf{w}_k]$

Definition

The $\mathcal{H}_2$ norm of a continuous dynamical system $S = (\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ is the $\mathcal{L}_2$ norm of its associated impulse response $\mathbf{h}(\cdot)$.
When $\mathbf{A}$ is stable and $\mathbf{D} = 0$, the norm is bounded and

$$\|S\|_{\mathcal{H}_2} = \int_0^\infty \mathrm{Tr}\left(\mathbf{h}^*(t)\mathbf{h}(t)\right) dt$$

- Using Parseval's theorem, one can obtain the expression in the frequency domain using the transfer function $\mathbf{H}(\cdot)$

$$\|S\|_{\mathcal{H}_2} = \frac{1}{2\pi}\int_{-\infty}^\infty \mathrm{Tr}\left(\mathbf{H}^*(-i\omega)\mathbf{H}(i\omega)\right) d\omega$$

- One can also derive the expression of $\|S\|_{\mathcal{H}_2}$ in terms of the reachability and observability gramians.

$$\|S\|_{\mathcal{H}_2} = \sqrt{\mathrm{Tr}\left(\mathbf{B}^*\mathcal{Q}\mathbf{B}\right)} = \sqrt{\mathrm{Tr}\left(\mathbf{C}\mathcal{P}\mathbf{C}^*\right)}$$

- In the Single Input-Single Output case, the transfer function is a rational function. Assuming for simplicity that it has distinct poles $\lambda_i,\ i = 1, \cdots, n$ associated with the residues h_i, one can write it as

$$\mathbf{H}(s) = \sum_{i=1}^{n} \frac{h_i}{s - \lambda_i}$$

- One can then establish the following theorem:

Theorem

Let $\mathbf{H}_r(\cdot)$ be the transfer function associated with the system $\mathcal{S}_r$ resulting from moment matching using the Lanczos procedure of the underlying system $\mathcal{S}$. Denoting by $h_{r,i}$ and $\lambda_{r,i}$ the respective residues and poles of $\mathbf{H}_r(\cdot)$, the following result holds:

$$\|\mathcal{S}-\mathcal{S}_r\|_{\mathcal{H}_2} = \sum_{i=1}^{n} h_i \left(\mathbf{H}(-\lambda_i^*) - \mathbf{H}_r(-\lambda_i^*)\right) + \sum_{i=1}^{k} h_{r,i} \left(\mathbf{H}_r(-\lambda_{r,i}) - \mathbf{H}(-\lambda_{r,i})\right)$$

- This error equality suggests matching the transfer function at the points $-\lambda_i^*$, mirror images of the poles λ_i of the original system