Model reduction of large-scale systems

Lecture I: Overview

Thanos Antoulas

Rice University and Jacobs University

email: aca@rice.edu
URL: www.ece.rice.edu/˜aca

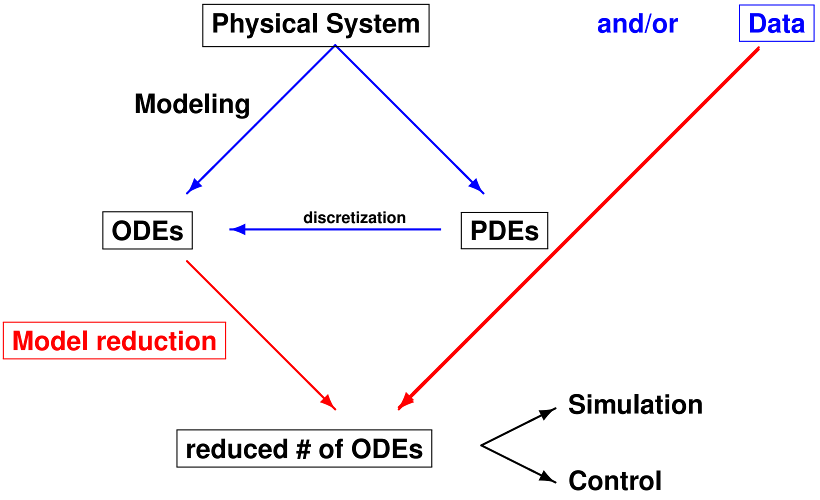
International School, Monopoli, 7-12 September 2008

Outline

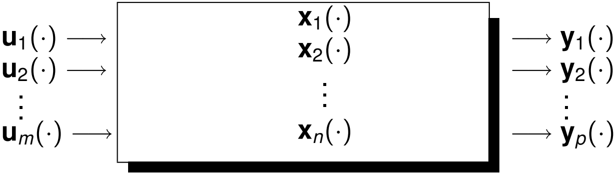
1. **Introduction and model reduction problem**
   - The big picture
   - Problem formulation
   - Projections
2. **Motivating examples**
3. **Overview of approximation methods**
   - SVD
   - POD
   - Balanced truncation
   - Krylov methods
   - Moment matching
4. **Summary – Challenges – References**

Outline

The big picture



Dynamical systems



We consider explicit state equations

$$\Sigma : \quad \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)),\ \ \mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t))$$

with **state** $\mathbf{x}(\cdot)$ of dimension $n \gg m, p$.

Problem statement

Given: dynamical system

$$\Sigma = (\mathbf{f}, \mathbf{h}) \text{ with: } \mathbf{u}(t) \in \mathbb{R}^m,\ \mathbf{x}(t) \in \mathbb{R}^n,\ \mathbf{y}(t) \in \mathbb{R}^p.$$

Problem: Approximate Σ with:

$$\hat{\Sigma} = (\hat{\mathbf{f}}, \hat{\mathbf{h}}) \text{ with : } \mathbf{u}(t) \in \mathbb{R}^m,\ \hat{\mathbf{x}}(t) \in \mathbb{R}^k,\ \hat{\mathbf{y}}(t) \in \mathbb{R}^p,\ k \ll n :$$

(1) Approximation error small - global error bound

(2) Preservation of stability/passivity

(3) Procedure must be computationally efficient

Approximation by projection

Unifying feature of approximation methods: **projections**.

Let $\mathbf{V},\ \mathbf{W} \in \mathbb{R}^{n\times k}$, such that $\mathbf{W}^*\mathbf{V} = \mathbf{I}_k \Rightarrow \Pi = \mathbf{V}\mathbf{W}^*$ is a projection.
Define $\hat{\mathbf{x}} = \mathbf{W}^*\mathbf{x}$. Then

$$\hat{\Sigma} : \begin{cases} \frac{d}{dt}\hat{\mathbf{x}}(t) & = & \mathbf{W}^*\mathbf{f}(\mathbf{V}\hat{\mathbf{x}}(t),\ \mathbf{u}(t)) \\ \mathbf{y}(t) & = & \mathbf{h}(\mathbf{V}\hat{\mathbf{x}}(t),\ \mathbf{u}(t)) \end{cases}$$

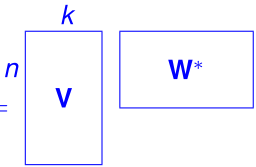
Thus $\hat{\Sigma}$ is "good" approximation of Σ, if $\mathbf{x} - \Pi\mathbf{x}$ is "small".

Special case: linear dynamical systems

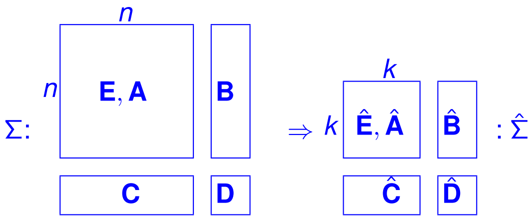
$$\Sigma:\ \mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t),\ \ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$$

$$\Sigma = \left(\begin{array}{c|c} \mathbf{E}, \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$$

Problem: Approximate Σ by **projection**: $\Pi = \mathbf{V}\mathbf{W}^* =$



$$\hat{\Sigma} = \left(\begin{array}{c|c} \hat{\mathbf{E}}, \hat{\mathbf{A}} & \hat{\mathbf{B}} \\ \hline \hat{\mathbf{C}} & \hat{\mathbf{D}} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{W}^*\mathbf{E}\mathbf{V}, \mathbf{W}^*\mathbf{A}\mathbf{V} & \mathbf{W}^*\mathbf{B} \\ \hline \mathbf{C}\mathbf{V} & \mathbf{D} \end{array}\right),\ \ k \ll n$$



Norms:

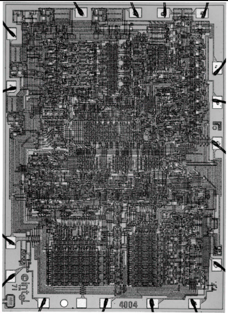
- $\mathcal{H}_\infty$-norm: worst output error $\|\mathbf{y}(t) - \hat{\mathbf{y}}(t)\|$ for $\|\mathbf{u}(t)\| = 1$.
- $\mathcal{H}_2$-norm: $\|\mathbf{h}(t) - \hat{\mathbf{h}}(t)\|$

Outline

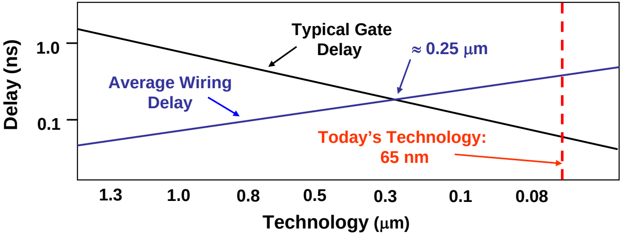
Motivating Examples: Simulation/Control

1. **Passive devices**	• VLSI circuits • Thermal issues • Power delivery networks
2. **Data assimilation**	• North sea forecast • Air quality forecast
3. **Molecular systems**	• MD simulations • Heat capacity
4. **CVD reactor**	• Bifurcations
5. **Mechanical systems**:	•Windscreen vibrations • Buildings
6. **Optimal cooling**	• Steel profile
7. **MEMS**: **Micro Electro-** **-Mechanical Systems**	• Elf sensor
8. **Nano-Electronics**	• Plasmonics

Passive devices: VLSI circuits

		
1960's: IC	1971: Intel 4004	2001: Intel Pentium IV
	10μ details 2300 components 64KHz speed	0.18μ details 42M components 2GHz speed 2km interconnect 7 layers

Passive devices: VLSI circuits



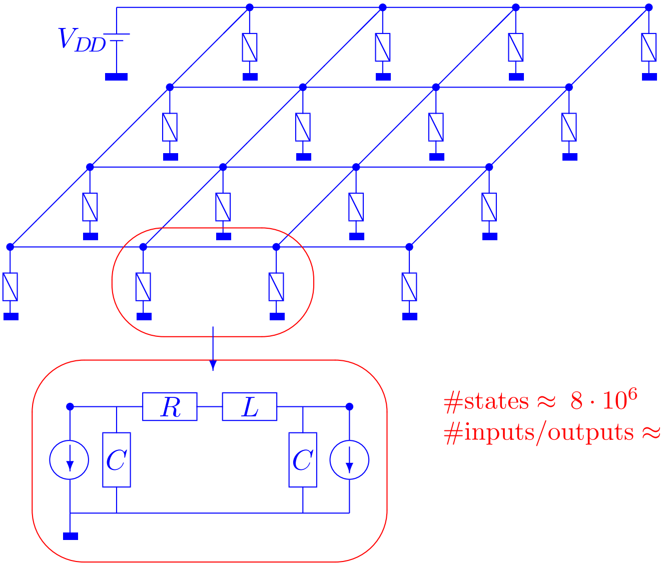
65nm technology: gate delay $<$ interconnect delay!

Conclusion: Simulations are required to verify that internal electromagnetic fields do not significantly delay or distort circuit signals. Therefore interconnections must be modeled.

⇒ Electromagnetic modeling of packages and interconnects ⇒ resulting models very complex: using PEEC methods (discretization of Maxwell's equations): $n \approx 10^5 \cdots 10^6$ ⇒ SPICE: inadequate

- Source: van der Meijs (Delft)

Power delivery network for VLSI chips



#states $\approx$ $8 \cdot 10^6$
#inputs/outputs $\approx 1 \cdot 10^6$

Mechanical systems: cars

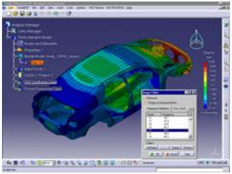
Car windscreen simulation subject to acceleration load.

Problem: compute noise at points away from the window.
PDE: describes deformation of a structure of a specific material; FE discretization: 7564 nodes (3 layers of 60 by 30 elements). Material: glass with Young modulus $7 \cdot 10^{10}\,\mathrm{N/m^2}$; density $2490\,\mathrm{kg/m^3}$; Poisson ratio 0.23 $\Rightarrow$ coefficients of FE model determined experimentally.
The discretized problem has dimension: 22,692.

Notice: this problem yields 2nd order equations:

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{f}(t).$$

- Source: Meerbergen (Free Field Technologies)

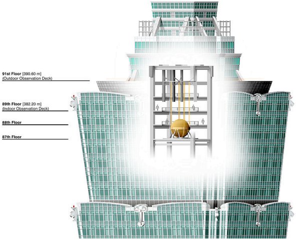


Mechanical Systems: Buildings

Earthquake prevention

Taipei 101: 508m

Damper between 87-91 floors



730 ton damper

Building	Height	Control mechanism	Damping frequency Damping mass
CN Tower, Toronto	533 m	Passive tuned mass damper	
Hancock building, Boston	244 m	Two passive tuned dampers	0.14Hz, 2x300t
Sydney tower	305 m	Passive tuned pendulum	0.1,0.5z, 220t
Rokko Island P&G, Kobe	117 m	Passive tuned pendulum	0.33-0.62Hz, 270t
Yokohama Landmark tower	296 m	Active tuned mass dampers (2)	0.185Hz, 340t
Shinjuku Park Tower	296 m	Active tuned mass dampers (3)	330t
TYG Building, Atsugi	159 m	Tuned liquid dampers (720)	0.53Hz, 18.2t

Outline

Large-scale systems

What is the problem with very large systems?

$\Downarrow$

- **Storage**
- **Computational speed**
- **Accuracy**
- **System theoretic properties**

Approximation methods: Overview

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

SVD

Nonlinear systems	Linear systems
• POD methods	• Balanced truncation
• Empirical Gramians	• Hankel approximation

Krylov/SVD Methods

Approximation methods

The Singular value decomposition: SVD

$$A = U\Sigma V^* \in \mathbb{R}^{n\times m}$$

- Singular values: $\Sigma = \mathrm{diag}\,(\sigma_1, \cdots, \sigma_n)$, $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$
  $\Rightarrow \sigma_i = \sqrt{\lambda_i(A^*A)}$
- left singular vectors: $U = (u_1 \;\; u_2 \;\; \cdots \;\; u_n),\; UU^* = I_n$
- right singular vectors: $V = (v_1 \;\; v_2 \;\; \cdots \;\; v_m),\; VV^* = I_m$
- Dyadic decomposition:

$$A \;=\; \sigma_1 u_1 v_1^* + \sigma_2 u_2 v_2^* + \cdots + \sigma_n u_n v_n^*$$

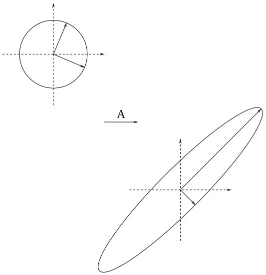
- σ_1: 2-induced norm of A

Reminder: 2-norm

Vectors:

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \Rightarrow \| x \|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$$

Matrices/Operators: induced 2-norm



$$\Rightarrow \| A \|_2 = \max_{\| x \| = 1} \| Ax \|_2 = \sigma_1$$

Optimal approximation in the 2-norm

- Given: $A \in \mathbb{R}^{n \times m}$
- find: $X \in \mathbb{R}^{n \times m}$, $\operatorname{rank} X = k < \operatorname{rank} A$
- Criterion: norm(error) is minimized, where
  error: $E = A - X$, norm: 2-norm

Theorem (*Schmidt-Mirsky, Eckart-Young*)

$$\min_{\operatorname{rank} X \leq k} \| A - X \|_2 = \sigma_{k+1}(A)$$

Minimizer (non-unique): truncation of dyadic decomposition of A:

$$X_{\#} = \sigma_1 u_1 v_1^* + \sigma_2 u_2 v_2^* + \cdots \sigma_k u_k v_k^*$$

Remarks.

(a) Importance of Schmidt-Mirsky: establishes a relationship between the rank k of the approximant, and the $(k+1)^{st}$ largest singular value of A.

(b) Other minimizers:

$$X(\eta_1, \cdots, \eta_k) := \sum_{i=1}^{k} (\sigma_i - \eta_i) u_i v_i^*$$

where $0 \leq \eta_i \leq \sigma_{k+1}$.

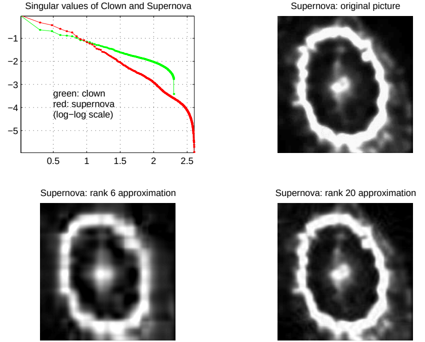
(c) The problem of minimizing the 2-induced norm of $A - X$ over all matrices X of rank at most k, is *non-convex*.

(d) Problem can also be solved in the Frobenius norm. ∎

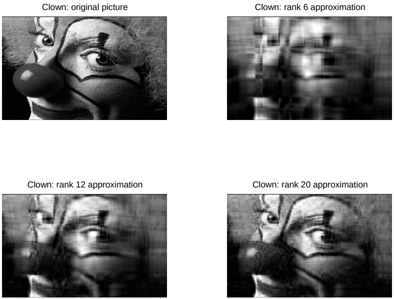
SVD Approximation methods

A prototype approximation problem – the **SVD**
(Singular Value Decomposition): $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^*$.

Supernova



Clown



Singular values provide trade-off between accuracy and complexity

POD: Proper Orthogonal Decomposition

Consider: $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)),\ \mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t))$.

Snapshots of the state:

$$\mathcal{X} = [\mathbf{x}(t_1)\ \ \mathbf{x}(t_2)\ \ \cdots\ \ \mathbf{x}(t_N)] \in \mathbb{R}^{n\times N}$$

SVD: $\mathcal{X} = \mathbf{U}\Sigma\mathbf{V}^* \approx \mathbf{U}_k\Sigma_k\mathbf{V}_k^*,\ \ k \ll n$. Approximate the state:

$$\hat{\mathbf{x}}(t) = \mathbf{U}_k^*\mathbf{x}(t) \ \ \Rightarrow \ \ \mathbf{x}(t) \approx \mathbf{U}_k\hat{\mathbf{x}}(t),\ \ \hat{\mathbf{x}}(t) \in \mathbb{R}^k$$

Project state and output equations. Reduced order system:

$$\boxed{\dot{\hat{\mathbf{x}}}(t) = \mathbf{U}_k^*\mathbf{f}(\mathbf{U}_k\hat{\mathbf{x}}(t), \mathbf{u}(t)),\ \ \mathbf{y}(t) = \mathbf{h}(\mathbf{U}_k\hat{\mathbf{x}}(t), \mathbf{u}(t))}$$

$\Rightarrow \hat{\mathbf{x}}(t)$ evolves in a **low-dimensional** space.

Issues with POD:

(a) Choice of snapshots, (b) singular values not I/O invariants.

SVD methods: balanced truncation

Trade-off between accuracy and complexity for linear dynamical systems is provided by the **Hankel Singular Values**. Define the **gramians** as solutions of the Lyapunov equations

$$\left.\begin{array}{l} \mathbf{A}\mathbf{P} + \mathbf{P}\mathbf{A}^* + \mathbf{B}\mathbf{B}^* = \mathbf{0}, \quad \mathbf{P} > \mathbf{0} \\ \mathbf{A}^*\mathbf{Q} + \mathbf{Q}\mathbf{A} + \mathbf{C}^*\mathbf{C} = \mathbf{0}, \quad \mathbf{Q} > \mathbf{0} \end{array}\right\} \Rightarrow \boxed{\sigma_i = \sqrt{\lambda_i(\mathbf{PQ})}}$$

σ_i: **Hankel singular values** of the system. There exists balanced basis where $\mathbf{P} = \mathbf{Q} = \mathbf{S} = \mathrm{diag}\,(\sigma_1, \cdots, \sigma_n)$. In this basis partition:

$$\mathbf{A} = \left(\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}\right),\ \mathbf{B} = \left(\begin{array}{c} \mathbf{B}_1 \\ \hline \mathbf{B}_2 \end{array}\right),\ \mathbf{C} = (\mathbf{C}_1 \mid \mathbf{C}_2),\ \mathbf{S} = \left(\begin{array}{c|c} \Sigma_1 & 0 \\ \hline 0 & \Sigma_2 \end{array}\right).$$

The reduced system is obtained by balanced truncation

$$\left(\begin{array}{c|c} \mathbf{A}_{11} & B_1 \\ \hline C_1 & \end{array}\right),$$ where Σ_2 contains the small Hankel singular values.

Properties of balanced reduction

1. **Stability is preserved**
2. **Global error bound**:

   $$\sigma_{k+1} \leq \| \Sigma - \hat{\Sigma} \|_\infty \leq 2(\sigma_{k+1} + \cdots + \sigma_n)$$

Drawbacks

1. **Dense** computations, matrix factorizations and inversions $\Rightarrow$ may be ill-conditioned
2. Need **whole** transformed system in order to truncate $\Rightarrow$ number of operations $\mathcal{O}(n^3)$
3. **Bottleneck**: **solution of two Lyapunov equations**

Approximation methods: Krylov methods

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

SVD

Nonlinear systems	Linear systems
• POD methods	• Balanced truncation
• Empirical Gramians	• Hankel approximation

Krylov/SVD Methods

The basic Krylov iteration

Given $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{b} \in \mathbb{R}^n$, let $\mathbf{v}_1 = \frac{\mathbf{b}}{\|\mathbf{b}\|}$. At the k^{th} step:

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

where

$\mathbf{e}_k \in \mathbb{R}^k$: canonical unit vector
$\mathbf{V}_k = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_k] \in \mathbb{R}^{k\times k}$, $\mathbf{V}_v^*\mathbf{V}_k = \mathbf{I}_k$
$\mathbf{H}_k = \mathbf{V}_k^*\mathbf{A}\mathbf{V}_k \in \mathbb{R}^{k\times k}$

$\Rightarrow \quad \mathbf{v}_{k+1} = \frac{\mathbf{f}_k}{\|\mathbf{f}_k\|} \in \mathbb{R}^n$

Computational complexity for k steps: $\mathcal{O}(n^2k)$; storage $\mathcal{O}(nk)$.

The **Lanczos** and the **Arnoldi** algorithms result.

The **Krylov iteration** involves the subspace $\mathcal{R}_k = \left[\mathbf{b}, \mathbf{A}\mathbf{b}, \cdots, \mathbf{A}^{k-1}\mathbf{b}\right]$.

- Arnoldi iteration $\Rightarrow$ arbitrary $\mathbf{A} \Rightarrow \mathbf{H}_k$ upper Hessenberg.
- Symmetric (one-sided) Lanczos iteration $\Rightarrow$ symmetric $\mathbf{A} = \mathbf{A}^*$
  $\Rightarrow \mathbf{H}_k$ tridiagonal and symmetric.
- Two-sided Lanczos iteration with two starting vectors $\mathbf{b}$, $\mathbf{c}$
  $\Rightarrow$ arbitrary $\mathbf{A} \Rightarrow \mathbf{H}_k$ tridiagonal.

Three uses of the Krylov iteration

(1) Iterative solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$: approximate the solution $\mathbf{x}$ iteratively.

(2) Iterative approximation of the eigenvalues of $\mathbf{A}$. In this case $\mathbf{b}$ is not fixed apriori. The eigenvalues of the projected $\mathbf{H}_k$ approximate the dominant eigenvalues of $\mathbf{A}$.

(3) Approximation of linear systems by moment matriching.

$\Rightarrow$ **Item (3) is of interest in the present context.**

Approximation by moment matching

Given $\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t),\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$, expand transfer function around s_0:

$$\mathbf{G}(s) = \eta_0 + \eta_1(s - s_0) + \eta_2(s - s_0)^2 + \eta_3(s - s_0)^3 + \cdots$$

Moments at s_0: η_j.

Find $\hat{\mathbf{E}}\dot{\hat{\mathbf{x}}}(t) = \hat{\mathbf{A}}\hat{\mathbf{x}}(t) + \hat{\mathbf{B}}\mathbf{u}(t),\ \mathbf{y}(t) = \hat{\mathbf{C}}\hat{\mathbf{x}}(t) + \hat{\mathbf{D}}\mathbf{u}(t)$, with

$$\hat{\mathbf{G}}(s) = \hat{\eta}_0 + \hat{\eta}_1(s - s_0) + \hat{\eta}_2(s - s_0)^2 + \hat{\eta}_3(s - s_0)^3 + \cdots$$

such that for appropriate s_0 and ℓ:

$$\boxed{\eta_j = \hat{\eta}_j,\ \ j = 1, 2, \cdots, \ell}$$

Projectors for Krylov and rational Krylov methods

Given:

$$\Sigma = \left(\begin{array}{c|c} \mathbf{E}, \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$$ by **projection**: $\Pi = \mathbf{V}\mathbf{W}^*$, $\Pi^2 = \Pi$ obtain

$$\hat{\Sigma} = \left(\begin{array}{c|c} \hat{\mathbf{E}}, \hat{\mathbf{A}} & \hat{\mathbf{B}} \\ \hline \hat{\mathbf{C}} & \hat{\mathbf{D}} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{W}^*\mathbf{E}\mathbf{V}, \mathbf{W}^*\mathbf{A}\mathbf{V} & \mathbf{W}^*\mathbf{B} \\ \hline \mathbf{C}\mathbf{V} & \mathbf{D} \end{array}\right),$$ where $k < n$.

Krylov (Lanczos, Arnoldi): let $\mathbf{E} = \mathbf{I}$ and

$$\mathbf{V} = \left[\mathbf{B},\ \mathbf{A}\mathbf{B},\ \cdots,\ \mathbf{A}^{k-1}\mathbf{B}\right] \in \mathbb{R}^{n \times k}$$

$$\bar{\mathbf{W}}^* = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \vdots \\ \mathbf{C}\mathbf{A}^{k-1} \end{bmatrix} \in \mathbb{R}^{k \times n}$$

$$\Rightarrow \mathbf{W}^* = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*$$

then the **Markov parameters match**:

$$\mathbf{C}\mathbf{A}^{\mathbf{i}}\mathbf{B} = \hat{\mathbf{C}}\hat{\mathbf{A}}^{\mathbf{i}}\hat{\mathbf{B}}$$

Rational Krylov: let

$$\mathbf{V} = \left[(\lambda_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\ \cdots\ (\lambda_k\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\right] \in \mathbb{R}^{n \times k}$$

$$\bar{\mathbf{W}}^* = \begin{bmatrix} \mathbf{C}(\lambda_{k+1}\mathbf{E} - \mathbf{A})^{-1} \\ \mathbf{C}(\lambda_{k+2}\mathbf{E} - \mathbf{A})^{-1} \\ \vdots \\ \mathbf{C}(\lambda_{2k}\ \mathbf{E} - \mathbf{A})^{-1} \end{bmatrix} \in \mathbb{R}^{k \times n}$$

$$\Rightarrow \mathbf{W}^* = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*$$

then the **moments of $\hat{\mathbf{G}}$ match those of $\mathbf{G}$ at λ_i**:

$$\mathbf{G}(\lambda_\mathbf{i}) = \mathbf{D} + \mathbf{C}(\lambda_\mathbf{i}\mathbf{E} - \mathbf{A})^{-1}\mathbf{B} = \hat{\mathbf{D}} + \hat{\mathbf{C}}(\lambda_\mathbf{i}\hat{\mathbf{E}} - \hat{\mathbf{A}})^{-1}\hat{\mathbf{B}} = \hat{\mathbf{G}}(\lambda_\mathbf{i})$$

Properties of Krylov methods

(a) Number of operations: $\mathcal{O}(kn^2)$ or $\mathcal{O}(k^2n)$ vs. $\mathcal{O}(n^3)$ $\Rightarrow$ efficiency

(b) Only matrix-vector multiplications are required. No matrix factorizations and/or inversions. No need to compute transformed model and then truncate.

(c) **Drawbacks**

- global error bound?
- $\hat{\Sigma}$ may not be stable.

Q: **How to choose the projection points?**

Outline

Approximation methods: Summary

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

Properties

- numerical efficiency
- n $\gg 10^3$
- choice of matching moments

SVD

Nonlinear systems	Linear systems
• POD methods	• Balanced truncation
• Empirical Gramians	• Hankel approximation

Properties

- Stability
- Error bound
- n $\approx 10^3$

Krylov/SVD Methods

(Some) Challenges in complexity reduction

- Model reduction of uncertain systems
- Model reduction of differential-algebraic (DAE) systems
- Domain decomposition methods
- Parallel algorithms for sparse computations in model reduction
- Development/validation of control algorithms based on reduced models
- Model reduction and data assimilation (weather prediction)
- Active control of high-rise buildings
- MEMS and multi-physics problems
- VLSI design
- Molecular Dynamics (MD) simulations
- CAD tools for nanoelectronics

References

- Passivity preserving model reduction
  - Antoulas SCL (2005)
  - Sorensen SCL (2005)
  - Ionutiu, Rommes, Antoulas IEEE TCAD (2008)
- Optimal $\mathcal{H}_2$ model reduction
  - Gugercin, Antoulas, Beattie SIMAX (2008)
- Low-rank solutions of Lyapunov equations
  - Gugercin, Sorensen, Antoulas, Numerical Algorithms (2003)
  - Sorensen (2006)
- Model reduction from data
  - Mayo, Antoulas LAA (2007)
  - Lefteriu, Antoulas, Tech. Report (2008)
- General reference: Antoulas, SIAM (2005)