Aposteriori error estimates for Reduced Order Models and adaptive DEIM strategies

Răzvan Ştefănescu and Adrian Sandu

Computational Science Laboratory (CSL)
Department of Computer Science
Virginia Tech

Outline

Motivation

Reduced Order Modeling

POD/DEIM justification and methodology

Aposteriori error estimates

Numerical Results

Motivation

- A posteriori error estimators employ the discrete solution itself to derive estimates of the actual solution errors.
- Derive reduced order solution error estimates - Dihlmann and Haasdonk (2014). Evaluate the error in some QoI computed via reduced order models - Carlberg (2014) - basis splitting.
- The mechanism makes use of adjoint models and allows us to disentangle the QoI error contribution of each discrete space point at every time step.

Motivation

- Tune the accuracy of ROMs by controling some of their features: DEIM points (nonlinear terms) - Chaturantabut and Sorensen (2010); DEIM indexes (Jacobians) - Wirtz and Sorensen (2014), Tonn (2011), Stefanescu and Sandu (2014) and POD basis Carlberg (2014).

- Dual-weighted residuals to guide the selection of DEIM points for approximation of ROM nonlinear terms - Peherstorfer and Willcox (2015) - online optimal rank-one DEIM basis update with respect to the Frobenius norm;

- When using the adjoint approach in combination with ROMs, the reduced space has to be designed so that the adjoint solutions can be approximated well in this space (online estimation).

Reduced Order Modeling

- The desired simulation is well approximated in the input collection Lumley(1967).
- Data analysis is conducted to extract basis functions, from experimental data or detailed simulations of high-dimensional systems.
- Galerkin and Petrov-Galerkin projections that yield low dimensional dynamical models.
- Galerkin POD models - Aubry et al (1998): *Its nonlinear reduced terms still have to be evaluated on the original state space making the simulat ion of the reduced-order system too expensive*.

POD/DEIM justification and methodology

- Model order reduction : Reduce the computational complexity/time of large scale dynamical systems.

- Goal : Construct reduced-order model for different types of discretization method (finite difference (FD), finite element (FEM), finite volume (FV)) of unsteady and/or parametrized nonlinear PDEs. E.g., PDE:

$$\frac{\partial y}{\partial t}(x,t) = \mathrm{L}(y(x,t)) + \mathrm{F}(y(x,t)),\ t \in [0, T]$$

where L is a linear function and F a nonlinear one.

POD/DEIM justification and methodology

- The corresponding FD scheme is a n dimensional ordinary differential system

$$\frac{d}{dt}\mathbf{y}(t) = A\mathbf{y}(t) + \mathbf{F}(\mathbf{y}(t)),\ A \in \mathbb{R}^{n\times n},$$

where $\mathbf{y}(t) = [y_1(t), y_2(t), .., y_n(t)] \in \mathbb{R}^n$. $\mathbf{F}$ is a nonlinear function evaluated at $\mathbf{y}(t)$, i.e. $\mathbf{F} = [\mathrm{F}(y_1(t)), .., \mathrm{F}(y_n(t))]^T$, $\mathrm{F} : I \subset \mathbb{R} \rightarrow \mathbb{R}$.

- A common model order reduction method involves the Galerkin projection with basis $U \in \mathbb{R}^{n\times k}$ obtained from Proper Orthogonal Decomposition (POD), for $k \ll n$, i.e. $\mathbf{y} \approx U\tilde{\mathbf{y}}(\mathbf{t})$, $\tilde{\mathbf{y}}(\mathbf{t}) \in \mathbb{R}^k$. Applying an inner product to the ODE discrete system we get

$$\frac{d}{dt}\tilde{\mathbf{y}}(\mathbf{t}) = \underbrace{U^TAU}_{k\times k}\tilde{\mathbf{y}}(\mathbf{t}) + \underbrace{U^T\mathbf{F}(U\tilde{\mathbf{y}}(\mathbf{t}))}_{\tilde{N}(\tilde{\mathbf{y}})} \quad (1)$$

POD/DEIM justification and methodology

- The efficiency of POD - Galerkin technique is limited to the linear or bilinear terms. The projected nonlinear term still depends on the dimension of the original system

$$\tilde{N}(\mathbf{y}) = \underbrace{U^T}_{k \times n} \underbrace{\mathbf{F}(U\mathbf{y}(\mathbf{t}))}_{n \times 1}.$$

- To mitigate this inefficiency Chaturantabut and Sorensen (2010) introduces "Discrete Empirical Interpolation Method (DEIM) " for nonlinear approximation. For $m \ll n$

$$\tilde{N}(\mathbf{y}) \approx \underbrace{U^T V (P^T V)^{-1}}_{precomputed\ \ k \times m} \underbrace{\mathbf{F}(P^T U\mathbf{y}(\mathbf{t}))}_{m \times 1}.$$

Aposteriori error estimates

- We consider a time-evolving system governed by the discrete-time full order model

$$\mathbf{x}_{k+1} = M_{k,k+1}(\mathbf{x}_k), \quad k = 0, \ldots, N-1\,, \quad \mathbf{x}_0 = \mathbf{x}_0(\theta)\,, \tag{2}$$

  where $\mathbf{x}_k \in \mathbb{R}^n$ is the state vector at time t_k, $M_{k,k+1}$ is the solution operator that advances the state vector from time t_k to t_{k+1}. At each time t_k the model state $\mathbf{x}_k$ approximates the truth, i.e., the state of the physical system $\mathbf{x}(t_k)$.
- We are also interested in the reduced order model whose projection reads

$$\widehat{\mathbf{x}}_{k+1} = \widehat{M}_{k,k+1}(\widehat{\mathbf{x}}_k), \quad k = 0, 1, \ldots, N-1\,. \tag{3}$$

Aposteriori error estimates

- A quantity of interest defined on the solution has the general form

$$\mathcal{E}\left(\mathbf{x}\right) = \sum_{k=0}^{N} r_k\left(\mathbf{x}_k, \theta\right). \tag{4}$$

- The error in the Qoi due to ROM is

$$\mathcal{E}\left(\mathbf{x}\right) - \mathcal{E}\left(\widehat{\mathbf{x}}\right) \approx \mathcal{E}_{\mathbf{x}} \cdot \left(\mathbf{x} - \widehat{\mathbf{x}}\right) = \sum_{k=0}^{N} \frac{\partial \mathcal{E}}{\partial \mathbf{x}_k}\left(\widehat{\mathbf{x}}_k - \mathbf{x}_k\right).$$

Aposteriori error estimates

- We consider the following Jacobian of the model solution operator with respect to the state

$$\mathcal{M}_{k,k+1} := \left(\mathcal{M}_{k,k+1}(\mathbf{x}_k)\right)_{\mathbf{x}_k}.$$

The adjoint model reads

$$\begin{aligned}\lambda_N &= \left(r_N\left(\mathbf{x_N},\theta\right)\right)^{\mathrm{T}}_{\mathbf{x_N}}, \\ \lambda_k &= \mathcal{M}^{\mathrm{T}}_{k,k+1}\,\lambda_{k+1} + \left(r_k\left(\mathbf{x}_k,\theta\right)\right)^{\mathrm{T}}_{\mathbf{x}_k}, \quad k = N-1,\ldots,0.\end{aligned}$$

Aposteriori error estimates

- The adjoint variables represent sensitivities of the QoI with respect to changes in the state

$$\lambda_k = \left(\frac{\partial \mathcal{E}}{\partial \mathbf{x}_k}\right)^T$$

and the error in the QoI due to ROM can be written as

$$\Delta\mathcal{E} = \mathcal{E}\left(\mathbf{x}\right) - \mathcal{E}\left(\widehat{\mathbf{x}}\right) \approx \sum_{k=0}^{N} \lambda_k^T \cdot \left(\mathbf{x}_k - \widehat{\mathbf{x}}_k\right) \approx \sum_{k=0}^{N} \lambda_k^T \cdot \Delta\mathbf{x}_k.$$

Aposteriori error estimates

- Assume for simplicity that

$$\mathcal{E}(\mathbf{x}) = \sum_{k=0}^{N} \mathbf{x}_k.$$

- Let $\mathbf{x}_i^0$, $\mathbf{x}_i^1$, $\mathbf{x}_i^2$, .. $\mathbf{x}_i^{Nt-1}$ be the solution obtained via the full model using as initial conditions the solution of reduced order model at time t_0, $t_1, .., t_{Nt-1}$ projected onto the full space.
- Estimations for Δx:

$$\Delta \mathbf{x}_0 = \mathbf{x}_0 - U\widehat{\mathbf{x}}_0, \quad \Delta \mathbf{x}_1 = \mathbf{x}_1^0 - U\widehat{\mathbf{x}}_1, ..., \quad \Delta \mathbf{x}_{Nt} = \mathbf{x}_{Nt}^{Nt-1} - U\widehat{\mathbf{x}}_{Nt}$$

Aposteriori error estimates

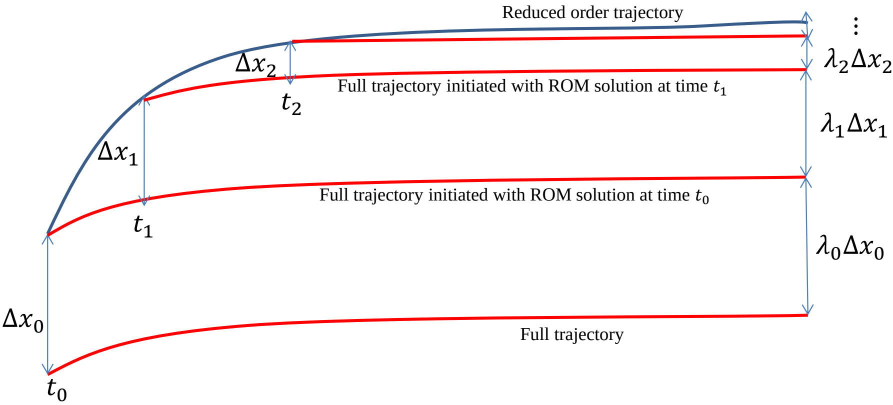


Figure : Geometrical Interpretation of aposteriori error estimates

Aposteriori error estimates

- Using the properties of the adjoint solutions we have the following first order approximations:

$$\begin{aligned}
\lambda_0' \cdot \Delta \mathbf{x}_0 &\approx \mathbf{x}_0 - U\widehat{\mathbf{x}}_0 + \sum_{i=1}^{Nt} (\mathbf{x}_i - \mathbf{x}_i^0) \\
\lambda_1' \cdot \Delta \mathbf{x}_1 &\approx \mathbf{x}_1^0 - U\widehat{\mathbf{x}}_1 + \sum_{i=2}^{Nt} (\mathbf{x}_i^0 - \mathbf{x}_i^1) \\
\cdots\cdots &\cdots \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\lambda_{Nt}' \cdot \Delta \mathbf{x}_{Nt} &\approx \mathbf{x}_{Nt}^{Nt-1} - U\widehat{\mathbf{x}}_{Nt}
\end{aligned}$$

- By adding all the above equations we obtain

$$\mathcal{E}(\mathbf{x}) - \mathcal{E}(\widehat{\mathbf{x}}) = \sum_{i=1}^{Nt} (\mathbf{x}_i - U\widehat{\mathbf{x}}_i) \approx \sum_{i=0}^{Nt} \lambda_i' \cdot \Delta \mathbf{x}_i$$

Aposteriori error estimates

- Efficient estimates of $\lambda_{k+1}^T \Delta \mathbf{x}_{k+1}$
- Explicit and Implicit Euler

$$- \lambda_{k+1}^T r(U\widehat{\mathbf{x}}_k)$$
$$h(\lambda_k^T - (r_{k+1}(\mathbf{x}_{k+1}, \theta))_{\mathbf{x}_{k+1}}^{\mathrm{T}})r(U\widehat{\mathbf{x}}_{k+1})$$

- One can approximate the high-fidelity adjoint model solution using a reduced order version for high qualitative reduced manifold.

DEIM: Algorithm for Interpolation Indices

INPUT: $\{v_l\}_{l=1}^{m} \subset \mathbb{R}^n$ (linearly independent):

OUTPUT: $\vec{\rho} = [\rho_1, .., \rho_m] \in \mathbb{R}^m$

1. $[|\psi| \;\; \rho_1] = max|v_1|, \psi \in \mathbb{R}$ and ρ_1 is the component position of the largest absolute value of v_1, with the smallest index taken in case of a tie.
2. $V = [v_1],\; P = [e_{\rho_1}],\; \vec{\rho} = [\rho_1].$
3. For $l = 2, .., m$ do
   - a Solve $(P^T V)c = P^T v_l \text{ for } c$
   - b $r = u_l - Vc$
   - c $[|\psi| \;\; \rho_l] = max\{|r|\}$
   - d $U \leftarrow [V \;\; v_l],\; P \leftarrow [P \;\; e_{\rho_l}],\; \vec{\rho} \leftarrow \begin{bmatrix} \vec{\rho} \\ \rho_l \end{bmatrix}$
4. end for.

DEIM adaptive1: Algorithm for Interpolation Indices

INPUT: $\{v_l\}_{l=1}^m \subset \mathbb{R}^n$ (linearly independent), $\lambda_l . * \Delta \mathbf{x}_l, l = 2, .., N_t$

OUTPUT: $\vec{\rho} = [\rho_1, .., \rho_m] \in \mathbb{R}^m$

1. $[|\psi| \;\; \rho_1] = max|\lambda_l . * \Delta \mathbf{x}_l|, \psi \in \mathbb{R}$ and ρ_1 is the component position of the largest absolute value of $\lambda_l . * \Delta \mathbf{x}_l$.
2. $V = [v_1], \; P = [e_{\rho_1}], \; \vec{\rho} = [\rho_1]$.
3. For $l = 2, .., m$ do
   a Solve $(P^T V)c = P^T(\lambda_l . * \Delta \mathbf{x}_l)$ for c
   b $r = (\lambda_l . * \Delta \mathbf{x}_l) - Vc$
   c $[|\psi| \;\; \rho_l] = max\{|r|\}$
   d $U \leftarrow [V \;\; v_l], \; P \leftarrow [P \;\; e_{\rho_l}], \; \vec{\rho} \leftarrow \begin{bmatrix} \vec{\rho} \\ \rho_l \end{bmatrix}$
4. end for.

DEIM adaptive2: Algorithm for Interpolation Indices

INPUT: $\{v_l\}_{l=1}^{m} \subset \mathbb{R}^n$ (linearly independent), $\{w_l\}_{l=1}^{m} \subset \mathbb{R}^n$ (linearly independent)$\lambda_l . * \Delta \mathbf{x}_l, \text{ all } l = 2, .., N_t$

OUTPUT: $\vec{\rho} = [\rho_1, .., \rho_m] \in \mathbb{R}^m$

1. $[|\psi| \;\; \rho_1] = max|w_1|, \psi \in \mathbb{R}$.
2. $V = [v_1], \; P = [e_{\rho_1}], \; \vec{\rho} = [\rho_1].$
3. For $l = 2, .., m$ do
   - a Solve $(P^T V)c = P^T w_l \text{ for } c$
   - b $r = w_l - Vc$
   - c $[|\psi| \;\; \rho_l] = max\{|r|\}$
   - d $U \leftarrow [V \;\; v_l], \; P \leftarrow [P \;\; e_{\rho_l}], \; \vec{\rho} \leftarrow \begin{bmatrix} \vec{\rho} \\ \rho_l \end{bmatrix}$
4. end for.

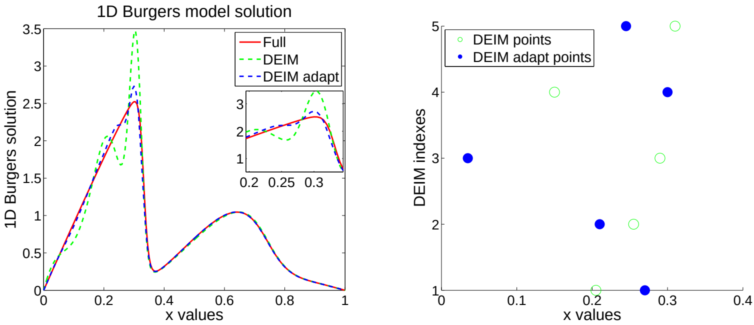
DEIM adaptive3: Algorithm for Interpolation Indices

INPUT: $\{v_l\}_{l=1}^{m} \subset \mathbb{R}^n$ (linearly independent), $\{w_l\}_{l=1}^{m} \subset \mathbb{R}^n$ (linearly independent)$\lambda_l . * \Delta \mathbf{x}_l, \text{ all } l = 2, .., N_t$

OUTPUT: $\vec{\rho} = [\rho_1, .., \rho_m] \in \mathbb{R}^m$

1. $[|\psi| \ \ \rho_1] = max|w_1|, \psi \in \mathbb{R}$.
2. $V = [v_1], \ P = [e_{\rho_1}], \ \vec{\rho} = [\rho_1]$.
3. For $l = 2, .., m$ do
   a Solve $(P^T V)c = P^T v_l \text{ for } c \quad (P^T V)d = P^T w_l \text{ for } d$
   b $r_1 = v_l - Vc, \quad r_2 = w_l - Vd$
   c $[|\psi| \ \ \rho_l] = max\{|\alpha r_1 + (1-\alpha) r_2|\}, \text{ for some } \alpha$
   d $U \leftarrow [V \ \ v_l], \ P \leftarrow [P \ \ e_{\rho_l}], \ \vec{\rho} \leftarrow \begin{bmatrix} \vec{\rho} \\ \rho_l \end{bmatrix}$
4. end for.

Numerical Results



(a) Adaptive 1D-Burgers (b) DEIM Adaptive points

Figure : Adjoint based DEIM adaptivity - version 3

Numerical Results

	m = 2		m=3		m = 5		m = 10	
	DEIM	ADEIM3	DEIM	ADEIM3	DEIM	ADEIM3	DEIM	ADEIM3
Cond. number	1.251	13.34	2.133	14.755	2.48	20.596	4.357	111.18
POD DEIM Sol. Error	6.741	8.6319	6.646	2.321	2.599	0.802	0.097	0.1306
QoI Error	-1.893	1.206	-0.6764	-0.095	0.018	0.0042	0.008	0.002

Table : Adaptive ROM performances using adaptiv DEIM version III

m=5	DEIM	ADEIM1	ADEIM2	ADEIM3
Cond. number	2.48	141.52	83.21	20.596
POD DEIM Sol. Error	2.59	2.44	1.73	0.802
QoI Error	$1.8e^{-2}$	$3.77e^{-5}$	$1.92e^{-4}$	$4.2e^{-3}$

Discussions and Conclusions

- Stabilization issues - condition number of the $(P^T V)^{-1}$; greedy algorithm that relaxes the condition of selecting the location of the largest absolute value of the residuals;
- The error bounds proposed by Chaturantabut and Sorensen (2010) are still valid;
- Comparison with the recent proposed updated optimized rank-one approximation -Peherstorfer and Willcox (2015) - using basis vectors of dual weighted residuals;
- Assuming high-quality ROM basis one can update the DEIM location on-line using reduced order adjoints and only projecting the ROM adjoint solution to the full space.

Discussions and Conclusions

- Extension to ROM optimization and adapt on the fly the DEIM interpolation location using aposteriori error estimates for the sub-optimal solution.
- Incorporate the adaptive mechanism selecting the MDEIM index points for the reduced order sparse approximation of Jacobians (implicit models).
- We already applied the adaptive DEIM strategy for goal-oriented approximations of reduced SWE non-linear terms.