Model Reduction for Variable-Fidelity Optimization Frameworks

Karen Willcox
Aerospace Computational Design Laboratory
Department of Aeronautics and Astronautics
Massachusetts Institute of Technology

Workshop on Model Order Reduction, Coupled Problems and Optimization
Sept. 19-22, 2005
Leiden University
Model Reduction for Large-Scale Systems

- Many model reduction methods for large-scale systems
  - Krylov-based, proper orthogonal decomposition (POD), balanced truncation, modal analysis, Fourier model reduction, *et al.*

- Methodology “mature” for linear time-invariant systems with few inputs/few outputs
  - Open challenges: nonlinear, parametrically varying systems
  - Model reduction for optimization versus just simulation
Model Reduction in Design and Optimization: Key Challenges

- Optimization-ready reduced-order models
  - What is a “good” reduced model for design?
  - Metrics beyond the transfer function?

- High dimensionality of parametric input space
  - Reduction methods rely on small number of inputs (curse of dimensionality)

- Fidelity management
  - Can the results of the reduced model be trusted?
  - When is it time to update the model?
Outline

• Projection framework for model reduction
• An optimization framework for model reduction
  – Joint work with O. Ghattas (UTexas), B. van Bloemen Waanders & B. Bader (Sandia National Laboratories)
• Variable-fidelity frameworks in design optimization
• Design variable mapping
  – PhD research of T. Robinson (MIT), joint work with M. Eldred (Sandia National Laboratories) and R. Haimes (MIT)
Dynamical Systems

\[ \dot{x} = Ax + Bu \]
\[ y = Cx \]

\[ \dot{x} = f(x, u) \]
\[ y = g(x) \]

\( x(t) \in \mathbb{R}^n \): state vector
\( u(t) \in \mathbb{R}^p \): input vector
\( y(t) \in \mathbb{R}^q \): output vector
Reduced-Order Projection

\[ x(t) = \sum_{i=1}^{m} V_i \alpha_i(t) \]

\[ x(t) = V x_r(t) \]

\[
\begin{bmatrix}
  x \\
  x_r
\end{bmatrix}
= 
\begin{bmatrix}
  V
\end{bmatrix}
\begin{bmatrix}
  x_r
\end{bmatrix}
\]

\[ W^T V = I \]
Reduced-Order Dynamical Systems

\[ n \times 1 \left\{ \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} V \\ \end{bmatrix} \begin{bmatrix} x_r \end{bmatrix} \right\} m \times 1 \Rightarrow W^T V = I \]

\[
\begin{align*}
\dot{x} &= f(x, u) \\
y &= g(x) \\
\dot{x}_r &= W^T f(V x_r, u) \\
y_r &= g(V x_r) \\
\dot{x} &= A x + B u \\
y &= C x \\
\dot{x}_r &= A_r x_r + B_r u_r \\
y_r &= C_r x_r \\
A_r &= W^T A V, \quad B_r = W^T B, \quad C_r = C V
\end{align*}
\]
Reduced-Order Basis

- Determine the projection \( x = V x_r \)
  where \( V \) contains \( m \) basis vectors

\[
V = \begin{bmatrix} v_1 & v_2 & \cdots & v_m \end{bmatrix}
\]

so that \( m \ll n \) and system dynamics are captured accurately: \( y_r \approx y \)

- Most, but not all, reduction techniques use projection framework (all reduced-order models can be represented in projection framework)
Proper Orthogonal Decomposition

- Consider $M$ snapshots $x_1, x_2, \ldots, x_M \in \mathcal{R}^n$ (instantaneous state solutions)

- Construct kernel $K \in \mathcal{R}^{n \times n}$

  \[
  K = \sum_{j=1}^{M} x_j x_j^T
  \]

- $V = [v_1 \ v_2 \ v_3 \cdots]$ are eigenvectors of $K$ with $\lambda_1 > \lambda_2 > \lambda_3 > \cdots$

- If $V_m$ contains the first $m$ eigenvectors, then $Q_m = V_m V_m^T$ is the optimal projection in a least squares sense:

  \[
  \min_{Q_m} \sum_{i=1}^{M} \| x_i - Q_m x_i \|_2^2 = \sum_{i=m+1}^{M} \lambda_i
  \]

- Note: optimality and error bound applies to the reconstruction of sampled data, not to the ROM.
Method of Snapshots

Sirovich: the eigenvectors of the kernel are linear combinations of the snapshots:

$$v_i = \sum_{j=1}^{M} \alpha^i_j x_j$$

The eigenvalue problem becomes:

$$R\alpha^i = \lambda_i \alpha^i \quad \alpha^i = \begin{bmatrix} \alpha^i_1 \\ \alpha^i_2 \\ \vdots \end{bmatrix}$$

where $R$ is the $M \times M$ correlation matrix:

$$R_{ij} = x^T_i x_j$$
Reduction via POD

1. Simulate the high-order system to get $M$ snapshots $x_j, j=1, 2, \ldots, M$ (could be for different parameters)
2. Construct the correlation matrix $R_{ij} = x_i^T x_j$
3. Calculate the eigenvectors $\alpha^i$ and eigenvalues $\lambda_i$ of $R$
4. Construct the basis functions $v_i = \sum_{j=1}^{M} \alpha^i_j x_j$
5. Select the most energetic basis functions using $\lambda_i$
6. Project the governing equations onto the reduced basis

- Reduction via POD offers no guarantees of ROM quality (accuracy/stability).
- POD is a data-driven approach. Basis and ROM do not respect the governing equations.
Outline

• Projection framework for model reduction
  • An optimization framework for model reduction
    – Joint work with O. Ghattas (UTexas), B. van Bloemen Waanders & B. Bader (Sandia National Laboratories)
  • Variable-fidelity frameworks in design optimization
  • Design variable mapping
    – PhD research of T. Robinson (MIT), joint work with M. Eldred (Sandia National Laboratories) and R. Haimes (MIT)
\[ \begin{align*}
\min_{V, x_r} \quad & \frac{1}{2} \sum_{k=1}^{S} \int_{0}^{t_f} (y^k - y^r_k)^T (y^k - y^r_k) \, dt \\
& + \frac{\beta}{2} \sum_{j=1}^{m} (1 - v^T_j v_j)^2
\end{align*} \]

subject to
\[ \begin{align*}
V^T V \dot{x}^k_r &= V^T A^k V x^k_r + V^T B^k u^k, \quad k = 1, \ldots, S \\
V x^k_r(0) &= x^k(0), \quad k = 1, \ldots, S \\
y^k_r &= C^k V x^k_r, \quad k = 1, \ldots, S
\end{align*} \]
Optimization Framework: Goal-Oriented

\[
\min_{V, x_r} G = \frac{1}{2} \sum_{k=1}^{S} \int_{0}^{t_f} (y^k - y_r^k)^T (y^k - y_r^k) \, dt + \frac{\beta}{2} \sum_{j=1}^{m} (1 - V_j^T V_j)^2
\]

subject to

\[
V^T V \dot{x}_r^k = V^T A^k V x_r^k + V^T B^k u^k, \quad k = 1, \ldots, S
\]

\[
V x_r^k(0) = x^k(0), \quad k = 1, \ldots, S
\]

\[
y_r^k = C^k V x_r^k, \quad k = 1, \ldots, S
\]
Optimization Framework

\[
\min_{V(x_r)} \mathcal{G} = \frac{1}{2} \sum_{k=1}^{S} \int_{0}^{t_f} (y^k - y_r^k)^T (y^k - y_r^k) \, dt \\
+ \frac{\beta}{2} \sum_{j=1}^{m} \left(1 - v_j^T v_j\right)^2
\]

subject to

\[
V^T V \dot{x}_r^k = V^T A^k V x_r^k + V^T g \quad k = 1, \ldots, S
\]

\[
V x_r^k(0) = x^k(0), \quad k = 1, \ldots, S
\]

\[
y_r^k = C^k V x_r^k, \quad k = 1, \ldots, S
\]

Regularization term to yield basis vectors of unit length
Optimization Framework: Model-Based

\[
\min_{V, x_r} \mathcal{G} = \frac{1}{2} \sum_{k=1}^{S} \int_{0}^{t_f} (y^k - y^k_r)^T (y^k - y^k_r) \, dt
\]
\[
+ \frac{\beta}{2} \sum_{j=1}^{m} (1 - V^T_j V_j)^2
\]

subject to

\[
V^T V \dot{x}^k_r = V^T A^k V x^k_r + V^T B^k u^k, \quad k = 1, \ldots, S
\]
\[
V x^k_r (0) = x^k (0), \quad k = 1, \ldots, S
\]
\[
y^k_r = C^k V x^k_r, \quad k = 1, \ldots, S
\]

Reduced output predictions from solution of governing equations
PDE-Constrained Optimization

General PDE-constrained optimization:

minimize $\mathcal{J}(x, u)$
subject to $c(x, u) = 0$

Define Lagrangian function ($\lambda$):

$\mathcal{L}(x, u, \lambda) := \mathcal{J}(x, u) + \langle \lambda, c(x, u) \rangle$

Resulting optimality conditions:

$c(x, u) = 0$ \hspace{1cm} \text{state equation}
$c^*_x(x, u) \lambda = -\mathcal{J}_x(x, u)$ \hspace{1cm} \text{adjoint equation}
$c^*_u(x, u) \lambda = -\mathcal{J}_u(x, u)$ \hspace{1cm} \text{optimization equation}
Basis Computation

- Time integrals in objective function are replaced by a summation over a finite number of discrete time instants
  - Method requires a priori computation of a snapshot set over $S$ parameter instances and $T$ time instants
- Assume the basis vectors are linear combinations of snapshots
  \[ v_j = \sum_{i=1}^{ST} \gamma_i^j x_i \quad j = 1, \ldots, m \]
  - Reduces number of unknowns from $mn$ to $mST$
  - Typically $ST \ll n$
- Use POD basis as initial guess for optimizer
Optimization Framework vs. POD

\[
\min_{V, x_r} G = \frac{1}{2} \sum_{k=1}^{S} \int_0^{t_f} (x^k - \hat{x}^k)^T C_k T C_k (x^k - \hat{x}^k) \, dt \\
+ \frac{\beta}{2} \sum_{j=1}^{m} (1 - V_j^T V_j)^2
\]

subject to

\[
V^T V \dot{x}_r^k = V^T A^k V x_r^k + V^T B^k u^k, \quad k = 1, \ldots, S
\]

\[
V x_r^k(0) = x^k(0), \quad k = 1, \ldots, S
\]

\[
\hat{x}^k = V x_r^k
\]

\[
\min_{V} G_{\text{pod}} = \frac{1}{2} \sum_{k=1}^{S} \int_0^{t_f} (x^k - \tilde{x}^k)^T (x^k - \tilde{x}^k) \, dt \\
+ \frac{\beta}{2} \sum_{j=1}^{m} (1 - V_j^T V_j)^2
\]

subject to

\[
\tilde{x}^k(t_j) = V V^T x^k(t_j)
\]

OPT: min error in computed data

POD: min error in projected data
Outline

• Projection framework for model reduction
• An optimization framework for model reduction
  – Joint work with O. Ghattas (UTexas), B. van Bloemen Waanders & B. Bader (Sandia National Laboratories)
• Variable-fidelity frameworks in design optimization
• Design variable mapping
  – PhD research of T. Robinson (MIT), joint work with M. Eldred (Sandia National Laboratories) and R. Haimes (MIT)
Large-Scale Design Optimization

minimize $f(u)$
subject to $c(u) < 0$

- Evaluation of $f(u)$ and $c(u)$ could be computationally expensive
- $c(u)$ might comprise constraints from many different disciplines
- $u$ may have large dimensionality

For complex large-scale multidisciplinary systems, optimization may be computationally intractable with high-fidelity models.
Variable-Fidelity Models

\[
\begin{align*}
\text{minimize} & \quad f(u) & \text{minimize} & \quad f(\hat{u}) \\
\text{subject to} & \quad c(u) < 0 & \text{subject to} & \quad \tilde{c}(\hat{u}) < 0
\end{align*}
\]

- Reduced complexity of \( f(\cdot) \) and \( c(\cdot) \)
  - Simplified physics
  - Model order reduction
  - Other surrogate models (data fit, multigrid, etc.)

- Reduced complexity of \( u \)
  - May need mapping between \( u \) and \( \hat{u} \)
A Hierarchy of Models

CFD: Euler/Navier-Stokes

Linearized Panel Code

Classical Aerodynamics

Low-fidelity EM model: 30,000 DOF

High-fidelity EM model: 800,000 DOF

J.J. Alonso, I. Kroo, Stanford University

J. Castro, A. Giunta Sandia National Labs
Trust Region Model Management

- How to manage the fidelity of the reduced model?
  - Can the results of the reduced model be trusted?
  - When is it time to update the model?

\[
\begin{align*}
\min & \quad f(u) \\
\text{s.t.} & \quad c(u) < 0
\end{align*}
\]

\[
\begin{align*}
\text{minimize} & \quad \tilde{f}^k(u) \\
\text{subject to} & \quad \tilde{c}^k(u) < 0 \\
& \quad \|u - u^k_c\|_{\infty} \leq \Delta^k
\end{align*}
\]

- $\tilde{f}^k(u)$: surrogate model of $f(u)$ at iteration $k$
- $u^k_c$: center point of the trust region at iteration $k$
- $\Delta^k$: size of the trust region at iteration $k$

Alexandrov, Lewis et al., 1997
Corrections

\[
\begin{align*}
\text{minimize} & \quad \tilde{f}^k(u) \\
\text{subject to} & \quad \tilde{c}^k(u) < 0 \\
& \quad \| u - u_c^k \|_\infty \leq \Delta^k
\end{align*}
\]

\[
\tilde{f}^k(u_c^k) = f^k(u_c^k) \quad \nabla \tilde{f}^k(u_c^k) = \nabla f^k(u_c^k)
\]

- Guarantee convergence to optimum of high-fidelity model
- Additive
  \[
  \tilde{f}^k(u) = \tilde{f}^k(u) + a(u)
  \]
- Multiplicative
  \[
  \tilde{f}^k(u) = \tilde{f}^k(u) b(u)
  \]

surrogate model

low-fidelity (reduced-order) model
Variable Design Descriptions

• How to create a reduced model for a large number of inputs (dimensionality of \( \mathbf{u} \) high)?
• Solution: reduce the complexity of \( \mathbf{u} \) (physically or numerically)

\[
\begin{align*}
\text{minimize} & \quad f^k(\hat{\mathbf{u}}) \\
\text{subject to} & \quad \bar{c}^k(\hat{\mathbf{u}}) < 0 \\
& \quad \| \mathbf{u} - \hat{\mathbf{u}}_c^k \|_\infty \leq \Delta^k
\end{align*}
\]

• Need mapping between \( \mathbf{u} \) and \( \hat{\mathbf{u}} \)
  – Space mapping
  – POD mapping: based on gappy POD theory
Space Mapping

- Developed by Bandler for electronics problems
- Particular form assumed for mapping between high- and low-fidelity design vectors:

\[ \hat{u} = P(u, p) \]

- Parameters, \( p \), determined by solving an optimization problem:

\[ p = \arg \min_p \sum_{i=1}^{k} \left( \| f(u_i) - \hat{f}(P(u_i, p)) \|_2 \right) \]

- Need to choose sample points and form of the mapping
- Trust-region model management methods using space-mapping are not provably convergent
Corrected Space Mapping

- Trust region model management using space mapping can be made rigorous by correcting space-mapped model to at least first order.
- Correcting the model after space mapping distorts global fit.
- Better solution: Perform space mapping and correction simultaneously.
Model Reduction in Design and Optimization: Key Components

\[
\begin{align*}
\text{minimize} & \quad f(u) \\
\text{subject to} & \quad c(u) < 0
\end{align*}
\]

\[
\begin{align*}
\text{minimize} & \quad \tilde{f}^k(\tilde{u}) \\
\text{subject to} & \quad \tilde{c}^k(\tilde{u}) < 0 \\
& \quad \| \tilde{u} - \tilde{u}_c^k \|_\infty \leq \Delta^k
\end{align*}
\]

- An optimization framework to determine a reduced model \( \tilde{f}(\tilde{u}) \)
- A mapping method to connect high- and low-fidelity design descriptions
- A trust region model management framework to ensure convergence of the design optimization
2D Heat Conduction Example

\[
\frac{\partial \bar{u}}{\partial t} - k \nabla^2 \bar{u} = 0 \text{ in } \Omega \\
\bar{u} = \bar{u}_c \text{ on } \Gamma \\
\bar{u} = \bar{u}_0 \text{ in } \Omega \text{ for } t = 0
\]

Objective function value

Number of modes

S=5 parameter instances
T=20 time steps
n=480 states
q=47 outputs
Output Errors vs. State Errors

\[ S = 5 \text{ parameter instances} \]
\[ T = 20 \text{ time steps} \]
\[ n = 480 \text{ states} \]
\[ q = 9 \text{ outputs} \]
Snapshot Reconstruction Errors

Error in snapshot for optimized basis

Error in snapshot for POD basis
Extended Rosenbrock Problem

\[ f(x) = \sum_{k=1}^{N} \left[ 4(x_k + 1 - x_k^2)^2 + (1 - x_k^2)^2 \right] \]

Example:

3D extended Rosenbrock (N=3)
Space Mapping Examples

- Low-fidelity: 2D quadratic
- High-fidelity: 2D Rosenbrock Problem
Space Mapping Examples

- Low-fidelity: 2D Rosenbrock Problem
- High-fidelity: 3D Extended Rosenbrock Problem
Space Mapping Examples

High fidelity: 3D Rosenbrock; Low fidelity: 2D quadratic

- Optimization in high-fidelity space
- Multifidelity using uncorrected space mapping
- Multifidelity using corrected space mapping

Objective function value vs. High-fidelity function evaluations
Airfoil Design Problem

- Inverse design of an airfoil
- High-fidelity model: XFOIL potential flow solver
  - Design variables are magnitudes of 36 bump functions
- Low-fidelity model: Joukowski transform
  - Design variables are two Joukowski parameters
POD mapping example

- Airfoil design problem using POD mapping

![Graph showing comparison between single fidelity BFGS and multifidelity with first order consistency.](image)