Parameterized Partial Differential Equations and the Proper Orthogonal Decomposition

David Amsallem

Stanford University

February 04, 2014
Outline

- Parameterized PDEs
- The steady case
- Dimensionality reduction
- Proper orthogonal decomposition
- Projection-based model reduction
- Snapshot selection
- The unsteady case
Parameterized PDE

- **Parametrized partial differential equation (PDE)**
  \[
  \mathcal{L}(\mathcal{W}, x, t; \mu) = 0
  \]

- **Associated boundary conditions**
  \[
  \mathcal{B}(\mathcal{W}, x_{BC}, t; \mu) = 0
  \]

- **Initial condition**
  \[
  \mathcal{W}_0(x) = \mathcal{W}_{IC}(x, \mu)
  \]

- \( \mathcal{W} = \mathcal{W}(x, t) \in \mathbb{R}^q \): state variable
- \( x \in \Omega \subset \mathbb{R}^d, d \leq 3 \): space variable
- \( t \geq 0 \): time variable
- \( \mu \in \mathcal{D} \subset \mathbb{R}^p \): parameter vector
Model parameterized PDE

- Advection-diffusion-reaction equation: \( \mathcal{W} = \mathcal{W}(x, t; \mu) \) solution of

\[
\frac{\partial \mathcal{W}}{\partial t} + U \cdot \nabla \mathcal{W} - \kappa \Delta \mathcal{W} = f_R(\mathcal{W}, t, \mu_R) \text{ for } x \in \Omega
\]

with appropriate boundary and initial conditions

\[
\mathcal{W}(x, t; \mu) = \mathcal{W}_D(x, t; \mu_D) \text{ for } x \in \Gamma_D
\]

\[
\nabla \mathcal{W}(x, t; \mu) \cdot n(x) = 0 \text{ for } x \in \Gamma_N
\]

\[
\mathcal{W}(x, 0; \mu) = \mathcal{W}_0(x; \mu_{IC}) \text{ for } x \in \Omega
\]

- Parameters of interest

\[
\mu = [U_1, \ldots, U_d, \kappa, \mu_R, \mu_D, \mu_{IC}]
\]
Semi-discretized problem

- The PDE is then discretized in space by one of the following methods
  - Finite Differences approximation
  - Finite Element method
  - Finite Volumes method
  - Discontinuous Galerkin method
  - Spectral method....
- This leads to a system of $N_w = q \times N_{\text{space}}$ ordinary differential equations (ODEs)
  \[
  \frac{dw}{dt} = f(w, t; \mu)
  \]
in terms of the discretized state variable
  \[
  w = w(t; \mu) \in \mathbb{R}^{N_w}
  \]
with initial condition $w(0; \mu) = w(\mu)$
- This is the high-dimensional model (HDM)
Parameterized solutions

- Example: two dimensional advection-diffusion equation

\[
\frac{\partial W}{\partial t} + U \cdot \nabla W - \kappa \Delta W = 0 \text{ for } x \in \Omega
\]

with boundary and initial conditions

\[
W(x, t; \mu) = W_D(x, t; \mu_D) \text{ for } x \in \Gamma_D
\]

\[
\nabla W(x, t; \mu) \cdot n(x) = 0 \text{ for } x \in \Gamma_N
\]

\[
W(x, 0; \mu) = W_0(x) \text{ for } x \in \Omega
\]
Parameterized solutions

- Example: two dimensional advection-diffusion equation

\[
\frac{\partial \mathcal{W}}{\partial t} + \mathbf{U} \cdot \nabla \mathcal{W} - \kappa \Delta \mathcal{W} = 0 \text{ for } x \in \Omega
\]

with boundary and initial conditions

\[
\mathcal{W}(x, t; \mu) = \mathcal{W}_D(x, t; \mu_D) \text{ for } x \in \Gamma_D
\]

\[
\nabla \mathcal{W}(x, t; \mu) \cdot \mathbf{n}(x) = 0 \text{ for } x \in \Gamma_N
\]

\[
\mathcal{W}(x, 0; \mu) = \mathcal{W}_0(x) \text{ for } x \in \Omega
\]

- 4 parameters of interest

\[
\mu = [\mathcal{U}_1, \mathcal{U}_2, \kappa, \mu_D] \in \mathbb{R}^4
\]

- \(\mathbf{w} \in \mathbb{R}^{N_w}\) with \(N_w = 2,701\)
Parameterized solutions

- (0,0,0.0002,950)
- (10,0,950)
- (10,10,0.0004,800)
- (10,2,0.00025,950)
Steady parameterized PDE

- Steady parameterized HDM

\[ f(w; \mu) = 0 \]

- Linear case

\[ A(\mu)w = b(\mu) \]

- Example: steady advection-diffusion equation
Dimensionality reduction

- Consider the manifold of solutions

\[ \mathcal{M} = \{ w(\mu) \text{ s.t. } \mu \in \mathcal{D} \} \subset \mathbb{R}^{N_w} \]

- Often \( \dim(\mathcal{M}) \ll N_w \)

- Therefore, \( \mathcal{M} \) could be described in terms of a much smaller set of variables, rather than \( \{e_1, \cdots, e_{N_w}\} \)

- Hence dimensionality reduction
Dimensionality reduction

- First idea: use solutions of the equation to describe $M$
- Consider pre-computed solution $\{w(\mu_1), \ldots, w(\mu_m)\}$ where $\{\mu_1, \ldots, \mu_m\} \subset D$
- Let $\mu \in D$. Then approximate $w(\mu)$ as
  \[
  w(\mu) \approx \alpha_1(\mu)w(\mu_1) + \cdots + \alpha_m(\mu)w(\mu_m)
  \]
  where $\{\alpha_1(\mu), \ldots, \alpha_m(\mu)\}$ are coefficients to be determined
Reduced-order basis

- There may be redundancies in the solutions \( \{ w(\mu_1), \cdots, w(\mu_m) \} \).
- Better approach: remove the redundancies by considering an equivalent independent set \( \{ v_1, \cdots, v_k \} \) with \( k \leq m \) such that

\[
\text{span} \{ v_1, \cdots, v_k \} = \text{span} \{ w(\mu_1), \cdots, w(\mu_m) \}
\]

- \( V = [v_1, \cdots, v_k] \in \mathbb{R}^{N_w \times k} \) is a reduced-order basis with \( k \ll N_w \).
Basis construction

- Lagrange basis

  \[ \text{span} \{v_1, \ldots, v_k\} = \text{span} \{w(\mu_1), \ldots, w(\mu_m)\} \]

- Hermite basis

  \[ \text{span} \{v_1, \ldots, v_k\} = \text{span} \left\{ w(\mu_1), \frac{\partial w}{\partial \mu_1}(\mu_1), \ldots, \frac{\partial w}{\partial \mu_p}(\mu_1), w(\mu_2), \ldots \right\} \]
Data compression

- It is possible to remove more information from the snapshots \( \{w(\mu_1), \cdots, w(\mu_m)\} \)
- Consider the snapshot matrix \( W = [w(\mu_1), \cdots, w(\mu_m)] \)
- Can we quantify the main information contained in \( W \) and discard the rest (noise)?
- This amount to data compression
- Here orthogonal projection will be used to compress the data
Orthogonal projection

- Let $V \in \mathbb{R}^{N \times k}$ be an orthogonal matrix ($V^T V = I_k$) which columns span $S$, a subspace of dimension $k$.
- Let $x \in \mathbb{R}^{Nw}$. The orthogonal projection of $x$ onto the subspace $S$ is
  \[ VV^T x \]

- Projection matrix
  \[ \Pi_{V, V} = V(V^T V)^{-1}V^T = VV^T \]
- special case 1: if $x$ belongs to $S$
  \[ \Pi_{V, V} x = VV^T x = x \]
- special case 2: if $x$ is orthogonal to $S$
  \[ \Pi_{V, V} x = VV^T x = 0 \]
\[ S_1 = \text{span}(\mathbf{V}) \]
Proper Orthogonal Decomposition

- POD seeks the subspace $S$ of a given dimension $k$ minimizing the projection error of the snapshots \( \{w(\mu_1), \cdots, w(\mu_m)\} \)
- Mathematical formulation $S = \text{range}(V)$ where

\[
V = \arg\min_Y \sum_{i=1}^{m} ||w(\mu_i) - \Pi_Y Yw(\mu_i)||^2_2
\]

s.t. $Y^TY = I_k$
POD by eigenvalue decomposition

- Minimization problem

\[ V = \arg\min_{Y^T Y = I_k} \sum_{i=1}^{m} \| w(\mu_i) - \Pi_{Y_i} Y w(\mu_i) \|_2^2 \]

- Equivalent maximization problem

\[ V = \arg\max_{Y^T Y = I_k} \sum_{i=1}^{m} \| Y Y^T w(\mu_i) \|_2^2 \]
\[ = \arg\max_{Y^T Y = I_k} \| Y^T W \|_F^2 \]
\[ = \arg\max_{Y^T Y = I_k} \text{trace} (Y^T W W^T Y) \]

- Solution: \( V \) is the matrix of eigenvectors \( \{ \phi_1, \cdots, \phi_k \} \) associated with the \( k \) largest eigenvalues of \( K = W W^T \)
The method of snapshots

- POD: $V$ is the matrix of eigenvectors $\{\phi_1, \cdots, \phi_k\}$ associated with the $k$ largest eigenvalues of $K = WW^T$
- $K \in \mathbb{R}^{N_w \times N_w}$ is a large, dense matrix
- Its rank is at most $m \ll N_w$
- In 1987, Sirovich developed the method of snapshots by noticing that $R = W^T W \in \mathbb{R}^{m \times m}$ has the same non-zero eigenvalues $\{\lambda_i\}_{i=1}^r$ as $K$
- $r = \text{rank}(R) \leq m \leq N_w$ and $R\psi_i = \lambda_i \psi_i, \ i = 1, \cdots, r$
- Exercise: relationship between $\{\phi_1, \cdots, \phi_r\}$ and $\{\psi_1, \cdots, \psi_r\}$?
The method of snapshots

- Step 1: compute the eigenpairs \( \{\lambda_i, \psi_i\}_{i=1}^{r} \) associated with \( R \)

\[
R \psi_i = \lambda_i \psi_i, \quad i = 1, \cdots, r
\]

- Step 2: compute \( \phi_i = \frac{1}{\sqrt{\lambda_i}} W \psi_i, \quad i = 1, \cdots, r \)

- In matrix form \( \Phi = W \Psi \Lambda^{-\frac{1}{2}} \)

- POD reduced basis of dimension \( k \leq r \):

\[
V = [\phi_1, \cdots, \phi_k]
\]
 POD by singular value decomposition

- The POD basis $V$ can also be computed by singular value decomposition (SVD)
- SVD of $W$

\[ W = U_r \Sigma_r Z_r^T \]

- $U_r = [u_1, \cdots, u_r] \in \mathbb{R}^{N_w \times r}$: left singular vectors ($U_r^T U_r = I_r$)
- $\Sigma_r = \text{diag}(\sigma_1, \cdots, \sigma_r) \in \mathbb{R}^{r \times r}$: singular values
- $Z_r = [z_1, \cdots, z_r] \in \mathbb{R}^{m \times r}$: right singular vectors ($Z_r^T Z_r = I_r$)
- POD reduced basis of dimension $k \leq r$

\[ V = [u_1, \cdots, u_k] \]
POD basis size selection

- The POD basis \( \mathbf{V} \) can also be computed by singular value decomposition (SVD)
- SVD of \( \mathbf{W} \):
  \[
  \mathbf{W} = \mathbf{U}_r \Sigma_r \mathbf{Z}_r^T
  \]
- POD reduced basis of dimension \( k \leq r \)
  \[
  \mathbf{V}_k = [\mathbf{u}_1, \cdots, \mathbf{u}_k]
  \]
- Relative projection error
  \[
  e(k) = \frac{\| \mathbf{W} - \mathbf{V}_k \mathbf{V}_k^T \mathbf{W} \|_F}{\| \mathbf{W} \|_F} = \sqrt{\frac{\sum_{i=k+1}^{r} \sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2}}
  \]
- Typically \( k \) is chosen so that \( e(k) < 0.1 \)
Projection-based model reduction

- High-dimensional model (HDM)
  \[ A(\mu)w = b(\mu) \]

- Reduced-order modeling assumption using a reduced basis \( V \)
  \[ w(\mu) \approx Vq(\mu) \]

  - \( q(\mu) \): reduced (generalized) coordinates

- Inserting in the HDM equation
  \[ AVq \approx b \]

- \( N_w \) equations in terms of \( k \) unknowns \( q \)
- Associated residual
  \[ r(q) = A(\mu)Vq - b(\mu) \]
Galerkin projection

- Residual equation
  \[ r(q) = A(\mu)Vq - b(\mu) \]
- \( N_w \) equations with \( k \) unknowns
- Galerkin projection enforces the orthogonality of \( r(q) \) to range(\( V \)):
  \[ V^T r(q) = 0 \]
- Reduced equations:
  \[ V^T A(\mu)Vq = V^T b(\mu) \]
- \( k \) equations in terms of \( k \) unknowns
  \[ A_k(\mu)q = b_k(\mu) \]
Application to the steady advection diffusion equation

- \( m = 5 \) snapshots \( \{\mu_i\}_{i=1}^{5} \)

\( (0,0,0.0002,950) \)
\( (10,0,0,950) \)
\( (10,10,0.0004,800) \)

\( (10,5,5e-05,950) \)
\( (10,2,0.00025,950) \)

\( (0,0,0.0002,950) \)
\( (10,0,0,950) \)
\( (10,10,0.0004,800) \)

\( (10,5,5e-05,950) \)
\( (10,2,0.00025,950) \)
Application to the steady advection diffusion equation

- \( m = 5 \) snapshots \( \{\mu_i\}_{i=1}^{5} \) \( \Rightarrow \) POD basis of dimension \( k = 5 \)
Application to the steady advection diffusion equation

- Projection error

![Graph showing the projection error as a function of k.](image)
Application to the steady advection diffusion equation

- $\mu_5 = (U_1, U_2, \kappa, \mu_D) = (1, 10, 3 \times 10^{-4}, 850)$
Exercise: Petrov-Galerkin projection

- Other approach to get a unique solution to

\[ A(\mu)Vq \approx b(\mu) \]

- Least-squares approach

\[ q = \arg\min_y \| A(\mu)Vy - b(\mu) \|_2 \]

- Exercise: give the equivalent set of equations satisfied by \( y \)
- Solution:

\[ V^T A(\mu)^T A(\mu) Vq = V^T A(\mu)^T b(\mu) \]
Offline/online decomposition for parametric systems

- Offline phase: computation of $V$ from snapshots $\{w(\mu_1), \ldots, w(\mu_m)\}$
- Online phase: construction and solution of $V^T A(\mu) V q = V^T b(\mu)$
- Issue: constructing $A_k(\mu) = V^T A(\mu) V$ and $b_k(\mu) = V^T b(\mu)$ is expensive
- Exception in the case of affine parameter dependence: $q_A \ll N_w$ and $q_b \ll N_w$

$$A(\mu) = \sum_{i=1}^{q_A} f_A^{(i)}(\mu) A^{(i)}, \quad b(\mu) = \sum_{i=1}^{q_b} f_b^{(i)}(\mu) b^{(i)}$$

- Then

$$A_k(\mu) = \sum_{i=1}^{q_A} f_A^{(i)}(\mu) V^T A^{(i)} V, \quad b(\mu) = \sum_{i=1}^{q_b} f_b^{(i)}(\mu) V^T b^{(i)}$$

- The following small dimensional matrices can be computed offline

$$A_k^{(i)} = V^T A^{(i)} V \in \mathbb{R}^{k \times k}, \quad i = 1, \ldots, q_A$$

$$b_k^{(i)} = V^T b^{(i)} \in \mathbb{R}^k, \quad i = 1, \ldots, q_b$$
Snapshot selection

- For a given number of snapshots $m$, what are the best snapshot parameter locations
  $$\mu_1, \cdots, \mu_m \in \mathcal{D}?$$
- The snapshots $\{w(\mu_1), \cdots, w(\mu_m)\}$ should be optimally placed so that they capture the physics over the parameter space $\mathcal{D}$
- Difficult problem $\Rightarrow$ use a heuristic approach (Greedy algorithm)
Greedy approach

- Start by randomly selecting a parameter value $\mu_1$ and compute $w(\mu_1)$
- For $i = 1, \cdots m$, find the parameter $\mu_i$ which presents the highest error between the ROM solution $V_q(\mu)$ and the HDM solution $w(\mu)$
- This however requires knowing the HDM solution (unknown)
- Instead, for $i = 1, \cdots m$, find the parameter $\mu_i$ for which the residual $r(q(\mu)) = AV_q(\mu) - b(\mu)$ is the highest

$$\mu_i = \arg\min_{\mu \in D} \| A(\mu)V_q(\mu) - b(\mu) \|_2$$

where $V^T A(\mu) V_q(\mu) = V^T b(\mu)$
- The parameter domain $\mu \in D$ can be in practice replaced by a search over a finite set

$$\mu \in \{ \mu^{(1)}, \cdots, \mu^{(N)} \} \subset D$$
Application to the steady advection diffusion equation

- Parameter domain \((U_1, U_2) \in [0, 10] \times [0, 10]\)
Application to the steady advection diffusion equation

initial snapshot

iteration = 1

iteration = 2

iteration = 3

iteration = 4

iteration = 5

iteration = 6

iteration = 7

iteration = 8
Application to the steady advection diffusion equation

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Normalized quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>0.3</td>
</tr>
<tr>
<td>4</td>
<td>0.4</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
</tr>
<tr>
<td>6</td>
<td>0.6</td>
</tr>
<tr>
<td>7</td>
<td>0.7</td>
</tr>
<tr>
<td>8</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Maximum Residual vs Maximum Error

David Amsallem
Parameterized Partial Differential Equations and the Proper Orthogonal D
Application to the steady advection diffusion equation

![Graph showing the normalized quantity against iteration for Maximum Residual and Maximum Error.]

- Maximum Residual
- Maximum Error

David Amsallem
Parameterized Partial Differential Equations and the Proper Orthogonal D
POD in time

- Consider an unsteady linear parametric problem

\[ \frac{d\mathbf{w}}{dt}(t) = A(\mu)\mathbf{w}(t) - b(\mu)\mathbf{u}(t) \]

where \( \mathbf{u}(t) \) is given for a time interval \( t \in [t_0, t_{N_t}] \)

- For a given parameter \( \mu \), POD can also provide an optimal reduced basis associated with the minimization problem

\[
\min_{\mathbf{V}^T\mathbf{V} = I_k} \int_{t_0}^{t_{N_t}} \| \mathbf{w}(t, \mu) - \mathbf{V}\mathbf{V}^T\mathbf{w}(t, \mu) \|^2 dt
\]

- Solution by the method of snapshots: consider the solutions in time \( \mathbf{w}(t_0, \mu), \ldots, \mathbf{w}(t_{N_t}, \mu) \). An approximation of the minimization problem is

\[
\min_{\mathbf{V}^T\mathbf{V} = I_k} \sum_{i=0}^{N_t} \delta_i \| \mathbf{w}(t_i, \mu) - \mathbf{V}\mathbf{V}^T\mathbf{w}(t_i, \mu) \|^2
\]

where \( \delta_0, \ldots, \delta_{N_t} \) are appropriate quadrature weights
**POD in time (continued)**

- Equivalent maximization problem

\[
\max_{\mathbf{v}^T \mathbf{v} = \mathbf{I}_k} \sum_{i=0}^{N_t} \delta_i \| \mathbf{v}^T \mathbf{w}(t_i, \mu) \|^2_2
\]

- Solution by POD by the method of snapshots with the associated snapshot matrix

\[
\mathbf{W} = [\sqrt{\delta_0} \mathbf{w}(t_0, \mu), \ldots, \sqrt{\delta_{N_t}} \mathbf{w}(t_{N_t}, \mu)]
\]
POD: Application to linearized aeroelasticity

- $M_\infty = 0.99$
- $N_w = 2,188,394$, $m = 99$ snapshots, $k = 60$ retained POD vectors

![Graph showing lift vs time for HDM and ROM](image)

- HDM (2,188,394)
- ROM (69)
Global vs. local strategies

- **Global strategy**
  - build a ROB $V$ that captures the behavior of unsteady systems for all $\mu \in D$
  - POD based on snapshots
    \[
    \{w(t_0, \mu_1), \cdots, w(t_{N_t}, \mu_1), w(t_0, \mu_2), \cdots, w(t_{N_t}, \mu_m)\}
    \]
  - not optimal for a given $\mu_i$

- **Local strategy**
  - build a separate ROB $V(\mu)$ for every value of $\mu \in D$
  - database approach: build offline a set of ROBs $\{V(\mu_i)\}_{i=1}^m$ and use them online to build $V(\mu)$ for $\mu \in D$
  - each ROB $V(\mu_i)$ is optimal at $\mu = \mu_i$
  - requires an adaptation approach online (see Lecture 3)
References