AA216/CME345: MODEL REDUCTION

Projection-based Model Order Reduction

Charbel Farhat
Stanford University
cfarhat@stanford.edu
Outline

1. Solution Approximation
2. Orthogonal and Oblique Projections
3. Galerkin and Petrov-Galerkin Projections
4. Equivalent High-Dimensional Model
5. Error Analysis
6. Preservation of Model Stability
Solution Approximation

High-Dimensional Model

- Ordinary Differential Equation (ODE)

\[
\frac{dw}{dt}(t) = f(w(t), t) \tag{1}
\]

- \( w \in \mathbb{R}^N \): State variable
- initial condition: \( w(0) = w_0 \)

- Output equation

\[
y(t) = g(w(t), t) \tag{2}
\]
Solution Approximation

High-Dimensional Model

- Ordinary Differential Equation (ODE)

\[ \frac{d\mathbf{w}}{dt}(t) = f(\mathbf{w}(t), t) \]  \hspace{1cm} (1)

- \( \mathbf{w} \in \mathbb{R}^N \): State variable
- initial condition: \( \mathbf{w}(0) = \mathbf{w}_0 \)

- Output equation

\[ y(t) = g(\mathbf{w}(t), t) \]  \hspace{1cm} (2)

- Note the absence of a parameter dependence for now
In many cases, the trajectories of the solutions computed using High-Dimensional Models (HDMs) are contained in low-dimensional subspaces.
In many cases, the trajectories of the solutions computed using High-Dimensional Models (HDMs) are contained in low-dimensional subspaces.

Let $S$ denote such a subspace and let $k_S = \dim(S)$.
Solution Approximation

Low Dimensionality of Trajectories

- In many cases, the trajectories of the solutions computed using High-Dimensional Models (HDMs) are contained in low-dimensional subspaces.

- Let $S$ denote such a subspace and let $k_S = \text{dim}(S)$.

- The state variable – or simply, the state – can be written exactly as a linear combination of vectors spanning $S$

$$w(t) = q_1(t)v_1 + \cdots + q_{k_S}(t)v_{k_S}$$

- $V_S = [v_1, \ldots, v_{k_S}] \in \mathbb{R}^{N \times k_S}$ is a time-invariant basis for $S$.

- $(q_1(t), \ldots, q_{k_S}(t))$ are the generalized coordinates for $w(t)$ in $S$.

- $q(t) = [q_1(t), \ldots, q_{k_S}(t)]^T \in \mathbb{R}^{k_S}$ is the reduced-order state vector.
Solution Approximation

Low Dimensionality of Trajectories

- In many cases, the trajectories of the solutions computed using High-Dimensional Models (HDMs) are contained in low-dimensional subspaces
- Let $S$ denote such a subspace and let $k_S = \dim(S)$
- The state variable – or simply, the state – can be written exactly as a linear combination of vectors spanning $S$

$$w(t) = q_1(t)v_1 + \cdots + q_{k_S}(t)v_{k_S}$$

- $V_S = [v_1, \cdots, v_{k_S}] \in \mathbb{R}^{N \times k_S}$ is a time-invariant basis for $S$
- $(q_1(t), \cdots, q_{k_S}(t))$ are the generalized coordinates for $w(t)$ in $S$
- $q(t) = [q_1(t), \cdots, q_{k_S}(t)]^T \in \mathbb{R}^{k_S}$ is the reduced-order state vector
- In matrix form, the above expansion can be written as

$$w(t) = V_S q(t)$$
Often, the exact basis $V_S$ is unknown but can be estimated empirically by a trial basis $V \in \mathbb{R}^{N \times k}$, $k < N$. 

Substituting the above subspace approximation in Eq. (1) and in Eq. (2) leads to:

$$\frac{d}{dt}(Vq(t)) = f(Vq(t), t) + r(t) \approx g(Vq(t), t)$$

where $r(t)$ is the residual due to the subspace approximation.
Often, the exact basis $V_S$ is unknown but can be estimated empirically by a **trial basis** $V \in \mathbb{R}^{N \times k}$, $k < N$

- $k$ and $k_S$ may be different
Often, the exact basis $V_S$ is unknown but can be estimated empirically by a trial basis $V \in \mathbb{R}^{N \times k}$, $k < N$

- $k$ and $k_S$ may be different

The following ansatz (educated guess, assumption, etc. to be verified later) is considered

$$w(t) \approx Vq(t)$$
Often, the exact basis $V_S$ is unknown but can be estimated empirically by a **trial basis** $V \in \mathbb{R}^{N \times k}$, $k < N$.

- $k$ and $k_S$ may be different.
- The following **ansatz** (educated guess, assumption, etc. to be verified later) is considered:

$$w(t) \approx Vq(t)$$

Substituting the above subspace approximation in Eq. (1) and in Eq. (2) leads to:

$$\frac{d}{dt}(Vq(t)) = f(Vq(t), t) + r(t)$$

$$y(t) \approx g(Vq(t), t)$$

where $r(t)$ is the residual due to the subspace approximation.
The residual \( r(t) \in \mathbb{R}^N \) accounts for the fact that \( Vq(t) \) is not in general an exact solution of problem (1).
The residual \( r(t) \in \mathbb{R}^N \) accounts for the fact that \( Vq(t) \) is not in general an exact solution of problem (1).

Since the basis \( V \) is assumed to be time-invariant

\[
\frac{d}{dt} (Vq(t)) = V \frac{dq(t)}{dt}
\]

and therefore

\[
V \frac{dq(t)}{dt} = f(Vq(t), t) + r(t)
\]

\[
y(t) \approx g(Vq(t), t)
\]
The residual $r(t) \in \mathbb{R}^N$ accounts for the fact that $Vq(t)$ is not in general an exact solution of problem (1).

Since the basis $V$ is assumed to be time-invariant

$$\frac{d}{dt} (Vq(t)) = V\frac{dq}{dt}(t)$$

and therefore

$$V\frac{dq}{dt}(t) = f(Vq(t), t) + r(t)$$

$$y(t) \approx g(Vq(t), t)$$

Set of $N$ differential equations in terms of $k$ unknowns

$$q_1(t), \cdots, q_k(t)$$
Solution Approximation

Low Dimensionality of Trajectories

- The residual $r(t) \in \mathbb{R}^N$ accounts for the fact that $Vq(t)$ is not in general an exact solution of problem (1).
- Since the basis $V$ is assumed to be time-invariant

$$\frac{d}{dt} (Vq(t)) = V\frac{dq}{dt}(t)$$

and therefore

$$V\frac{dq}{dt}(t) = f(Vq(t), t) + r(t)$$

$$y(t) \approx g(Vq(t), t)$$

- Set of $N$ differential equations in terms of $k$ unknowns

$$q_1(t), \cdots, q_k(t)$$

- Over-determined system ($k < N$)
Let $w$ and $y$ be two vectors in $\mathbb{R}^N$. 

- Orthogonal and Oblique Projections
- Orthogonality

- Let $w$ and $y$ be two vectors in $\mathbb{R}^N$. 

- Orthogonal and Oblique Projections
- Orthogonality
Let $w$ and $y$ be two vectors in $\mathbb{R}^N$

- $w$ and $y$ are **orthogonal** to each other with respect to the canonical inner product in $\mathbb{R}^N$ if and only if
  
  \[ w^T y = 0 \]
- Orthogonal and Oblique Projections
- Orthogonality

Let $\mathbf{w}$ and $\mathbf{y}$ be two vectors in $\mathbb{R}^N$

- $\mathbf{w}$ and $\mathbf{y}$ are **orthogonal** to each other with respect to the canonical inner product in $\mathbb{R}^N$ if and only if

  $$\mathbf{w}^T \mathbf{y} = 0$$

- $\mathbf{w}$ and $\mathbf{y}$ are **orthonormal** to each other with respect to the canonical inner product in $\mathbb{R}^N$ if and only if $\mathbf{w}^T \mathbf{y} = 0$, and

  $$\mathbf{w}^T \mathbf{w} = 1, \text{ and } \mathbf{y}^T \mathbf{y} = 1$$
Let $w$ and $y$ be two vectors in $\mathbb{R}^N$.

- $w$ and $y$ are **orthogonal** to each other with respect to the canonical inner product in $\mathbb{R}^N$ if and only if $w^T y = 0$.

- $w$ and $y$ are **orthonormal** to each other with respect to the canonical inner product in $\mathbb{R}^N$ if and only if $w^T y = 0$, and $w^T w = 1$, and $y^T y = 1$.

Let $V$ be a matrix in $\mathbb{R}^{N \times k}$. 


Let \( w \) and \( y \) be two vectors in \( \mathbb{R}^N \).

- \( w \) and \( y \) are **orthogonal** to each other with respect to the canonical inner product in \( \mathbb{R}^N \) if and only if

\[
    w^T y = 0
\]

- \( w \) and \( y \) are **orthonormal** to each other with respect to the canonical inner product in \( \mathbb{R}^N \) if and only if \( w^T y = 0 \), and

\[
    w^T w = 1, \quad \text{and} \quad y^T y = 1
\]

Let \( V \) be a matrix in \( \mathbb{R}^{N\times k} \).

- \( V \) is an **orthogonal** (orthonormal) matrix if and only if

\[
    V^T V = I_k
\]
Definition

A matrix $\Pi \in \mathbb{R}^{N \times N}$ is a projection matrix (or projective matrix, idempotent matrix) if

$$\Pi^2 = \Pi$$

Some direct consequences

- $\text{range}(\Pi)$ is invariant under the action of $\Pi$. 
Definition

A matrix \( \Pi \in \mathbb{R}^{N \times N} \) is a projection matrix (or projective matrix, idempotent matrix) if

\[
\Pi^2 = \Pi
\]

- Some direct consequences
  - \( \text{range}(\Pi) \) is invariant under the action of \( \Pi \)
  - 0 and 1 are the only possible eigenvalues of \( \Pi \)
**Definition**

A matrix $\Pi \in \mathbb{R}^{N \times N}$ is a projection matrix (or projective matrix, idempotent matrix) if

$$\Pi^2 = \Pi$$

- Some direct consequences
  - $\text{range}(\Pi)$ is invariant under the action of $\Pi$
  - 0 and 1 are the only possible eigenvalues of $\Pi$
  - $\Pi$ is diagonalizable (follows from the previous consequence)
Orthogonal and Oblique Projections

Projections

Definition
A matrix $\Pi \in \mathbb{R}^{N \times N}$ is a projection matrix (or projective matrix, idempotent matrix) if
\[ \Pi^2 = \Pi \]

Some direct consequences
- range($\Pi$) is invariant under the action of $\Pi$
- 0 and 1 are the only possible eigenvalues of $\Pi$
- $\Pi$ is diagonalizable (follows from the previous consequence)
- let $k$ be the rank of $\Pi$: then, there exists a basis $X$ such that
\[ \Pi = X \begin{bmatrix} I_k & 0_{N-k} \end{bmatrix} X^{-1} \]
(follows from the two previous consequences)
Consider

\[ \Pi = X \begin{bmatrix} I_k & 0_{N-k} \end{bmatrix} X^{-1} \]
Consider

\[ \Pi = X \begin{bmatrix} I_k & 0_{N-k} \end{bmatrix} X^{-1} \]

- decompose \( X \) as

\[ X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}, \text{ where } X_1 \in \mathbb{R}^{N \times k} \text{ and } X_2 \in \mathbb{R}^{N \times (N-k)} \]

then, \( \forall w \in \mathbb{R}^N \)

- \( \Pi w \in \text{range}(X_1) = \text{range}(\Pi) = S_1 \)
Consider

$$\Pi = X \begin{bmatrix} I_k & 0_{N-k} \end{bmatrix} X^{-1}$$

decompose $X$ as

$$X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}, \text{ where } X_1 \in \mathbb{R}^{N \times k} \text{ and } X_2 \in \mathbb{R}^{N \times (N-k)}$$

then, $\forall w \in \mathbb{R}^N$

- $\Pi w \in \text{range}(X_1) = \text{range}(\Pi) = S_1$
- $w - \Pi w \in \text{range}(X_2) = \text{Ker}(\Pi) = S_2$
Consider

\[ \Pi = X \begin{bmatrix} I_k & 0_{N-k} \end{bmatrix} X^{-1} \]

decompose \( X \) as

\[ X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}, \] where \( X_1 \in \mathbb{R}^{N \times k} \) and \( X_2 \in \mathbb{R}^{N \times (N-k)} \)

then, \( \forall w \in \mathbb{R}^N \)

- \( \Pi w \in \text{range}(X_1) = \text{range}(\Pi) = S_1 \)
- \( w - \Pi w \in \text{range}(X_2) = \text{Ker}(\Pi) = S_2 \)

\( \Pi \) defines the projection onto \( S_1 \) parallel to \( S_2 \)

\[ \mathbb{R}^{N} = S_1 \oplus S_2 \]
Consider the case where $S_2 = S_1^\perp$. 

The orthogonal projection of $w$ onto the subspace $S_1$ is $V^T w$. The equivalent projection matrix is $\Pi_V$, $V^T = VV^T$. 

**Special case #1:** If $w$ belongs to $S_1$—that is, $w = Vq$, where $q \in \mathbb{R}^k$, then $\Pi_V w = VV^T Vq = Vq = w$. 

**Special case #2:** If $w$ is orthogonal to $S_1$—that is, $V^T w = 0$—then $\Pi_V w = VV^T w = 0$. 

Consider the case where $S_2 = S_1^\perp$.

Let $V \in \mathbb{R}^{N \times k}$ be an orthogonal matrix whose columns span $S_1$, and let $w \in \mathbb{R}^N$: The orthogonal projection of $w$ onto the subspace $S_1$ is

$$V V^T w$$
Consider the case where \( S_2 = S_1^\perp \).

Let \( V \in \mathbb{R}^{N \times k} \) be an orthogonal matrix whose columns span \( S_1 \), and let \( w \in \mathbb{R}^N \): The orthogonal projection of \( w \) onto the subspace \( S_1 \) is

\[
VV^Tw
\]

the equivalent projection matrix is

\[
\Pi_{V,V} = VV^T
\]
Consider the case where $S_2 = S_1^\perp$

Let $V \in \mathbb{R}^{N \times k}$ be an *orthogonal* matrix whose columns span $S_1$, and let $w \in \mathbb{R}^N$: The orthogonal projection of $w$ onto the subspace $S_1$ is

$$VV^Tw$$

the equivalent projection matrix is

$$\Pi_{V,V} = VV^T$$

special case #1: If $w$ belongs to $S_1$ – that is, $w = Vq$, where $q \in \mathbb{R}^k$

$$\Pi_{V,V}w = VV^Tw = VV^TVq = Vq = w$$
Consider the case where $S_2 = S_1^\perp$.

Let $V \in \mathbb{R}^{N \times k}$ be an orthogonal matrix whose columns span $S_1$, and let $w \in \mathbb{R}^N$: The orthogonal projection of $w$ onto the subspace $S_1$ is

$$VV^Tw$$

the equivalent projection matrix is

$$\Pi_{V,V} = VV^T$$

special case #1: If $w$ belongs to $S_1$ — that is, $w = Vq$, where $q \in \mathbb{R}^k$

$$\Pi_{V,V}w = VV^Tw = VV^TVq = Vq = w$$

special case #2: If $w$ is orthogonal to $S_1$ — that is, $V^Tw = 0$

$$\Pi_{V,V}w = VV^Tw = 0$$
Orthogonal Projections

\[ \Pi_{V, V}w = VV^T w \]

\[ S_1 = \text{range}(V) \]
Example: Helix in 3D \((N = 3)\)

**Orthogonal and Oblique Projections**

**Orthogonal Projections**

- Let \(w(t) \in \mathbb{R}^3\) define a curve parameterized by \(t \in [0, 6\pi]\) as follows:

\[
\mathbf{w}(t) = \begin{bmatrix}
    w_1(t) \\
    w_2(t) \\
    w_3(t)
\end{bmatrix} = \begin{bmatrix}
    \cos(t) \\
    \sin(t) \\
    t
\end{bmatrix}
\]
Orthogonal and Oblique Projections

Orthogonal Projections

Orthogonal projection onto $\text{range}(V) = \text{span}(e_1, e_2)$

$$\Pi_{V, V}w(t) = \begin{bmatrix} \cos(t) \\ \sin(t) \\ 0 \end{bmatrix}$$
Orthogonal projection onto

\[ \Pi_{V, V} w(t) = \begin{bmatrix} 0 \\ \sin(t) \\ t \end{bmatrix} \]
Orthogonal and Oblique Projections

Orthogonal Projections

Orthogonal projection onto

\[ \text{range}(V) = \text{span}(e_1, e_3) \]

\[
\Pi_{V, V} w(t) = \begin{bmatrix} \cos(t) \\ 0 \\ t \end{bmatrix}
\]
The following is the general case, where $S_2$ may be distinct from $S_1$. 

Let $V \in \mathbb{R}^{N \times k}$ and $W \in \mathbb{R}^{N \times k}$ be two full-column rank matrices whose columns span respectively $S_1$ and $S_2^\perp$. The projection of $w \in \mathbb{R}^N$ onto the subspace $S_1$ parallel to $S_2^\perp$ is given by

$$V(W^T V)^{-1} W^T w$$

The equivalent projection matrix is

$$\Pi_{V, W} = V(W^T V)^{-1} W^T V.$$ 

**special case #1:** If $w$ belongs to $S_1$, then $w = Vq$, where $q \in \mathbb{R}^k$, and

$$\Pi_{V, W} w = V(W^T V)^{-1} W^T Vq = Vq.$$ 

**special case #2:** If $w$ is orthogonal to $S_2^\perp$, that is, $W^T w = 0$, then

$$\Pi_{V, W} w = V(W^T V)^{-1} W^T w = 0.$$
The following is the **general case**, where $S_2$ may be distinct from $S_1^\perp$.

- Let $V \in \mathbb{R}^{N \times k}$ and $W \in \mathbb{R}^{N \times k}$ be two **full-column rank** matrices whose columns span respectively $S_1$ and $S_2^\perp$.
Orthogonal and Oblique Projections

Oblique Projections

- The following is the **general case**, where $S_2$ may be distinct from $S_1^\perp$.
- Let $V \in \mathbb{R}^{N \times k}$ and $W \in \mathbb{R}^{N \times k}$ be two **full-column rank** matrices whose columns span respectively $S_1$ and $S_2^\perp$.
- The projection of $w \in \mathbb{R}^N$ onto the subspace $S_1$ parallel to $S_2$ is

$$V(W^TV)^{-1}W^Tw$$
Orthogonal and Oblique Projections

Oblique Projections

- The following is the general case, where $S_2$ may be distinct from $S_1^\perp$.

- Let $V \in \mathbb{R}^{N \times k}$ and $W \in \mathbb{R}^{N \times k}$ be two full-column rank matrices whose columns span respectively $S_1$ and $S_2^\perp$.

- The projection of $w \in \mathbb{R}^N$ onto the subspace $S_1$ parallel to $S_2$ is
  \[ V(W^T V)^{-1} W^T w \]

- The equivalent projection matrix is
  \[ \Pi_{V,W} = V(W^T V)^{-1} W^T \]
Orthogonal and Oblique Projections

Oblique Projections

- The following is the general case, where $S_2$ may be distinct from $S_1^\perp$.
- Let $V \in \mathbb{R}^{N \times k}$ and $W \in \mathbb{R}^{N \times k}$ be two full-column rank matrices whose columns span respectively $S_1$ and $S_2^\perp$.
- The projection of $w \in \mathbb{R}^N$ onto the subspace $S_1$ parallel to $S_2$ is
  \[ V(W^T V)^{-1}W^T w \]

- the equivalent projection matrix is
  \[ \Pi_{V,W} = V(W^T V)^{-1}W^T \]

- special case #1: If $w$ belongs to $S_1$, then $w = Vq$, where $q \in \mathbb{R}^k$, and
  \[ \Pi_{V,W}w = V(W^T V)^{-1}W^T V q = Vq \]
The following is the **general case**, where $S_2$ may be distinct from $S_1^\perp$.

Let $V \in \mathbb{R}^{N \times k}$ and $W \in \mathbb{R}^{N \times k}$ be two **full-column rank** matrices whose columns span respectively $S_1$ and $S_2^\perp$.

The projection of $w \in \mathbb{R}^N$ onto the subspace $S_1$ parallel to $S_2$ is

$$V(W^TV)^{-1}W^Tw$$

the equivalent projection matrix is

$$\Pi_{V,W} = V(W^TV)^{-1}W^T$$

**special case #1**: If $w$ belongs to $S_1$, then $w = Vq$, where $q \in \mathbb{R}^k$, and

$$\Pi_{V,W}w = V(W^TV)^{-1}W^TVq = Vq$$

**special case #2**: If $w$ is orthogonal to $S_2^\perp$ – that is, $W^Tw = 0$, then

$$\Pi_{V,W}w = V(W^TV)^{-1}W^Tw = 0$$
Orthogonal and Oblique Projections

Oblique Projections

\[ \Pi_{V,w}w = V(W^T V)^{-1}W^T w \]
Example: Helix in 3D

bases

\[ V = [e_1, e_2], \quad W = [e_1 + e_3, e_2 + e_3] \]
Example: Helix in 3D

- bases

\[ V = [e_1, e_2], \quad W = [e_1 + e_3, e_2 + e_3] \]

- projection matrix

\[ \Pi_{V,W} = V(W^T V)^{-1} W^T = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \]
Example: Helix in 3D

- bases
  \[ V = [e_1, \ e_2], \ W = [e_1 + e_3, \ e_2 + e_3] \]

- projection matrix
  \[
  \Pi_{V,W} = V(W^TV)^{-1}W^T = \begin{bmatrix}
  1 & 0 & 1 \\
  0 & 1 & 1 \\
  0 & 0 & 0
\end{bmatrix}
  \]

- projected helix equation
  \[
  \Pi_{V,W}w(t) = \begin{bmatrix}
  1 & 0 & 1 \\
  0 & 1 & 1 \\
  0 & 0 & 0
\end{bmatrix}
  \begin{bmatrix}
  \cos(t) \\
  \sin(t) \\
  t
\end{bmatrix}
  = \begin{bmatrix}
  \cos(t) + t \\
  \sin(t) + t \\
  0
\end{bmatrix}
  \]
Orthogonal and Oblique Projections

Oblique Projections
Start from a HDM for the problem of interest

\[
\begin{align*}
\frac{dw}{dt}(t) &= f(w(t), t) \\
y(t) &= g(w(t), t) \\
w(0) &= w_0
\end{align*}
\]

- \(w \in \mathbb{R}^N\): Vector of state variables
- \(y \in \mathbb{R}^q\): Vector of output variables (typically \(q \ll N\))
- \(f(\cdot, \cdot) \in \mathbb{R}^N\): Completes the specification of the HDM-based problem
The goal is to construct a **Projection-based Reduced-Order Model** (PROM)

\[
\frac{dq}{dt}(t) = f_r(q(t), t)
\]

\[
y(t) \approx g_r(q(t), t)
\]

where

- \( q \in \mathbb{R}^k \): **Vector of reduced-order state variables**, \( k \ll N \)
- \( y \in \mathbb{R}^q \): **Vector of output variables**
- \( f_r(\cdot, \cdot) \in \mathbb{R}^k \): Completes the description of the PROM

The discussion of the initial condition is deferred to later
Galerkin and Petrov-Galerkin Projections

Requirements

- A Projection-based Model Order Reduction (PMOR) method should
  - be computationally tractable
A Projection-based Model Order Reduction (PMOR) method should
  - be computationally tractable
  - be applicable to a large class of dynamical systems
A Projection-based Model Order Reduction (PMOR) method should:

- be computationally tractable
- be applicable to a large class of dynamical systems
- minimize a certain measure of the error between the solution computed using the HDM and that computed using the PROM (error criterion)
A Projection-based Model Order Reduction (PMOR) method should
- be computationally tractable
- be applicable to a large class of dynamical systems
- minimize a certain measure of the error between the solution computed using the HDM and that computed using the PROM (error criterion)
- preserve as many properties of the HDM as possible
Recall the residual $r(t) \in \mathbb{R}^{N \times k}$ introduced by approximating $w(t)$ as $Vq(t)$

$$V \frac{dVq}{dt}(t) = f(Vq(t), t) + r(t) \iff r(t) = V \frac{dVq}{dt}(t) - f(Vq(t), t)$$
Recall the residual $r(t) \in \mathbb{R}^{N \times k}$ introduced by approximating $w(t)$ as $Vq(t)$

$$V \frac{dq}{dt}(t) = f(Vq(t), t) + r(t) \iff r(t) = V \frac{dq}{dt}(t) - f(Vq(t), t)$$

Constrain this residual to be orthogonal to a subspace $\mathcal{W}$ defined by a test basis $W \in \mathbb{R}^{N \times k}$ – that is, compute $q(t)$ such that

$$W^T r(t) = 0$$
Recall the residual $r(t) \in \mathbb{R}^{N \times k}$ introduced by approximating $w(t)$ as $Vq(t)$

$$V \frac{dq}{dt}(t) = f(Vq(t), t) + r(t) \iff r(t) = V \frac{dq}{dt}(t) - f(Vq(t), t)$$

Constrain this residual to be orthogonal to a subspace $\mathcal{W}$ defined by a test basis $W \in \mathbb{R}^{N \times k}$ – that is, compute $q(t)$ such that

$$W^T r(t) = 0$$

This leads to the descriptive form of the governing equations of the Petrov-Galerkin PROM

$$W^T V \frac{dq}{dt}(t) = W^T f(Vq(t), t)$$
Assume that $W^T V$ is non-singular: In this case, the PROM can be re-written in the *non-descriptive form*

\[
\begin{align*}
\frac{dq}{dt}(t) &= (W^T V)^{-1}W^T f(Vq(t), t) \\
y(t) &\approx g(Vq(t), t)
\end{align*}
\]
Assume that $W^T V$ is non-singular: In this case, the PROM can be re-written in the *non-descriptive form*

$$\frac{dq(t)}{dt} = (W^T V)^{-1} W^T f(Vq(t), t)$$

$$y(t) \approx g(Vq(t), t)$$

After the above reduced-order equations have been solved, the subspace approximation of the high-dimensional state vector can be reconstructed, if needed, as follows

$$w(t) \approx Vq(t)$$
If \( W = V \), the projection method is called a **Galerkin** projection and the resulting PROM is called a **Galerkin** PROM.
If \( W = V \), the projection method is called a **Galerkin** projection and the resulting PROM is called a **Galerkin** PROM.

If in addition \( V \) is orthogonal, the reduced-order equations become

\[
\frac{dq}{dt}(t) = V^T f(Vq(t), t) \\
y(t) \approx g(Vq(t), t)
\]
Special case: **Linear Time-Invariant (LTI)** systems

\[
\begin{align*}
    f(w(t), t) &= Aw(t) + Bu(t) \\
    g(w(t), t) &= Cw(t) + Du(t)
\end{align*}
\]

\(u \in \mathbb{R}^p\): Vector of input variables
Special case: **Linear Time-Invariant (LTI) systems**

\[
\begin{align*}
f(w(t), t) &= Aw(t) + Bu(t) \\
g(w(t), t) &= Cw(t) + Du(t)
\end{align*}
\]

- **u ∈ ℝ^p**: Vector of input variables
- corresponding Petrov-Galerkin PROM

\[
\begin{align*}
\frac{dq}{dt}(t) &= (W^T V)^{-1} W^T (AVq(t) + Bu(t)) \\
y(t) &= CVq(t) + Du(t)
\end{align*}
\]
Galerkin and Petrov-Galerkin Projections

Linear Time-Invariant Systems

Special case: **Linear Time-Invariant (LTI) systems**

\[
\begin{align*}
    f(w(t), t) &= Aw(t) + Bu(t) \\
    g(w(t), t) &= Cw(t) + Du(t)
\end{align*}
\]

- \( u \in \mathbb{R}^p \): Vector of input variables
- corresponding Petrov-Galerkin PROM

\[
\begin{align*}
\frac{dq}{dt}(t) &= (W^T V)^{-1} W^T (AVq(t) + Bu(t)) \\
y(t) &= CVq(t) + Du(t)
\end{align*}
\]

- reduced-order LTI operators

\[
\begin{align*}
A_r &= (W^T V)^{-1} W^T AV \in \mathbb{R}^{k \times k}, \ k \ll N \\
B_r &= (W^T V)^{-1} W^T B \in \mathbb{R}^{k \times p} \\
C_r &= CV \in \mathbb{R}^{q \times k} \\
D_r &= D \in \mathbb{R}^{q \times p}
\end{align*}
\]
Galerkin and Petrov-Galerkin Projections

Initial Condition

- High-dimensional initial condition

\[ \mathbf{w}(0) = \mathbf{w}_0 \in \mathbb{R}^N \]
Galerkin and Petrov-Galerkin Projections

Initial Condition

- High-dimensional initial condition

\[ w(0) = w_0 \in \mathbb{R}^N \]

- Reduced-order initial condition (Petrov-Galerkin PROM)

\[ q(0) = (W^T V)^{-1} W^T w_0 \in \mathbb{R}^k \]
Galerkin and Petrov-Galerkin Projections

Initial Condition

- High-dimensional initial condition
  \[ \mathbf{w}(0) = \mathbf{w}_0 \in \mathbb{R}^N \]

- Reduced-order initial condition (Petrov-Galerkin PROM)
  \[ \mathbf{q}(0) = (\mathbf{W}^T \mathbf{V})^{-1} \mathbf{W}^T \mathbf{w}_0 \in \mathbb{R}^k \]

  in the high-dimensional state space, this gives
  \[ \mathbf{Vq}(0) = \mathbf{V}(\mathbf{W}^T \mathbf{V})^{-1} \mathbf{W}^T \mathbf{w}_0 \in \mathbb{R}^k \]
Galerkin and Petrov-Galerkin Projections

Initial Condition

- High-dimensional initial condition
  \[ \mathbf{w}(0) = \mathbf{w}_0 \in \mathbb{R}^N \]

- Reduced-order initial condition (Petrov-Galerkin PROM)
  \[ \mathbf{q}(0) = (\mathbf{W}^T \mathbf{V})^{-1} \mathbf{W}^T \mathbf{w}_0 \in \mathbb{R}^k \]

- in the high-dimensional state space, this gives
  \[ \mathbf{Vq}(0) = \mathbf{V}(\mathbf{W}^T \mathbf{V})^{-1} \mathbf{W}^T \mathbf{w}_0 \in \mathbb{R}^k \]

- this is an oblique projection of \( \mathbf{w}_0 \) onto \( \text{range}(\mathbf{V}) \) parallel to \( \text{range}(\mathbf{W}) \)
Galerkin and Petrov-Galerkin Projections

Initial Condition

- High-dimensional initial condition

\[ \mathbf{w}(0) = \mathbf{w}_0 \in \mathbb{R}^N \]

- Reduced-order initial condition (Petrov-Galerkin PROM)

\[ \mathbf{q}(0) = (\mathbf{W}^T \mathbf{V})^{-1} \mathbf{W}^T \mathbf{w}_0 \in \mathbb{R}^k \]

- in the high-dimensional state space, this gives

\[ \mathbf{Vq}(0) = \mathbf{V}(\mathbf{W}^T \mathbf{V})^{-1} \mathbf{W}^T \mathbf{w}_0 \in \mathbb{R}^k \]

- this is an oblique projection of \( \mathbf{w}_0 \) onto \( \text{range}(\mathbf{V}) \) parallel to \( \text{range}(\mathbf{W}) \)

- Error in the subspace approximation of the initial condition

\[ \mathbf{w}(0) - \mathbf{Vq}(0) = (\mathbf{I}_N - \mathbf{V}(\mathbf{W}^T \mathbf{V})^{-1} \mathbf{W}^T)\mathbf{w}_0 \]
Galerkin and Petrov-Galerkin Projections

Initial Condition

- High-dimensional initial condition
  \[ \mathbf{w}(0) = \mathbf{w}_0 \in \mathbb{R}^N \]

- Reduced-order initial condition (Petrov-Galerkin PROM)
  \[ \mathbf{q}(0) = (\mathbf{W}^T \mathbf{V})^{-1} \mathbf{W}^T \mathbf{w}_0 \in \mathbb{R}^k \]

- In the high-dimensional state space, this gives
  \[ \mathbf{Vq}(0) = \mathbf{V}(\mathbf{W}^T \mathbf{V})^{-1} \mathbf{W}^T \mathbf{w}_0 \in \mathbb{R}^k \]

- This is an oblique projection of \( \mathbf{w}_0 \) onto \( \text{range}(\mathbf{V}) \) parallel to \( \text{range}(\mathbf{W}) \)

- Error in the subspace approximation of the initial condition
  \[ \mathbf{w}(0) - \mathbf{Vq}(0) = (\mathbf{I}_N - \mathbf{V}(\mathbf{W}^T \mathbf{V})^{-1} \mathbf{W}^T)\mathbf{w}_0 \]

- Alternative: use an affine approximation \( \mathbf{w}(t) = \mathbf{w}(0) + \mathbf{Vq}(t) \) (see Homework #1)
Question: Which HDM would produce the same solution as that given by the following Petrov-Galerkin PROM? (this notion will prove to be useful for the stability analysis of a PROM)
Question: Which HDM would produce the same solution as that given by the following Petrov-Galerkin PROM? (this notion will prove to be useful for the stability analysis of a PROM)

- recall the reduced-order equations

\[
\frac{dq}{dt}(t) = (W^T V)^{-1} W^T f(Vq(t), t) \\
y(t) = g(Vq(t), t)
\]
Question: Which HDM would produce the same solution as that given by the following Petrov-Galerkin PROM? (this notion will prove to be useful for the stability analysis of a PROM)

- recall the reduced-order equations

\[ \frac{dq}{dt}(t) = (W^T V)^{-1} W^T f(Vq(t), t) \]

\[ y(t) = g(Vq(t), t) \]

- the corresponding reconstructed high-dimensional state solution is

\[ \tilde{w}(t) = Vq(t) \]
Question: Which HDM would produce the same solution as that given by the following Petrov-Galerkin PROM? (this notion will prove to be useful for the stability analysis of a PROM)

- recall the reduced-order equations

\[
\frac{dq}{dt}(t) = (W^T V)^{-1}W^T f(Vq(t), t) \\
y(t) = g(Vq(t), t)
\]

- the corresponding reconstructed high-dimensional state solution is

\[
\tilde{w}(t) = Vq(t)
\]

- pre-multiplying the above reduced-order equations by \( V \) leads to

\[
\frac{d\tilde{w}}{dt}(t) = V(W^T V)^{-1}W^T f(\tilde{w}(t), t) \\
\tilde{y}(t) = g(\tilde{w}(t), t)
\]
Question: Which HDM would produce the same solution as that given by the following Petrov-Galerkin PROM? (this notion will prove to be useful for the stability analysis of a PROM)

- recall the reduced-order equations

\[
\frac{dq}{dt}(t) = (W^T V)^{-1}W^T f(Vq(t), t)
\]
\[
y(t) = g(Vq(t), t)
\]

- the corresponding reconstructed high-dimensional state solution is

\[
\tilde{w}(t) = Vq(t)
\]

- pre-multiplying the above reduced-order equations by \( V \) leads to

\[
\frac{d\tilde{w}}{dt}(t) = V(W^T V)^{-1}W^T f(\tilde{w}(t), t)
\]
\[
\tilde{y}(t) = g(\tilde{w}(t), t)
\]

- the associated initial condition is

\[
\tilde{w}(0) = Vq(0) = V(W^T V)^{-1}W^T w(0)
\]
Recall the projector $\Pi_{V,W}$

$$\Pi_{V,W} = V(W^T V)^{-1} W^T$$

**Definition**

Equivalent HDM

$$\frac{d\tilde{w}}{dt}(t) = \Pi_{V,W} f(\tilde{w}(t), t)$$

$$\tilde{y}(t) = g(\tilde{w}(t), t)$$

with the initial condition

$$\tilde{w}(0) = \Pi_{V,W} w(0)$$

The equivalent dynamical function is

$$\tilde{f}(\cdot, \cdot) = \Pi_{V,W} f(\cdot, \cdot)$$
Consider the Petrov-Galerkin PROM

\[
\frac{dq}{dt}(t) = (W^T V)^{-1} W^T f(Vq(t), t)
\]

\[
y(t) \approx g(Vq(t), t)
\]

\[
q(0) = (W^T V)^{-1} W^T w(0)
\]

Lemma

*Choosing two different bases $V'$ and $W'$ that respectively span the same subspaces $V$ and $W$ results in the same reconstructed solution $w(t)$*

In other words, subspaces are more important than bases ...
Consequences

- given a HDM, a corresponding PROM is uniquely defined by its associated Petrov-Galerkin projector $\Pi_{V,W}$
Consequences

- given a HDM, a corresponding PROM is uniquely defined by its associated Petrov-Galerkin projector $\Pi_{V,W}$
- this projector is itself uniquely defined by the two subspaces

$$W = \text{range}(W) \quad \text{and} \quad V = \text{range}(V)$$
Consequences

- Given a HDM, a corresponding PROM is uniquely defined by its associated Petrov-Galerkin projector $\Pi_{V,W}$.
- This projector is itself uniquely defined by the two subspaces $\mathcal{W} = \text{range}(W)$ and $\mathcal{V} = \text{range}(V)$.

Hence

$$\text{PROM} \iff (\mathcal{W}, \mathcal{V})$$
Consequences

- given a HDM, a corresponding PROM is uniquely defined by its associated Petrov-Galerkin projector $\Pi_{V,W}$
- this projector is itself uniquely defined by the two subspaces

$$W = \text{range}(W) \quad \text{and} \quad V = \text{range}(V)$$

- hence

$$\text{PROM} \Leftrightarrow (W, V)$$

- $W$ and $V$ belong to the Grassmann manifold $G(k, N)$, which is the set of all subspaces of dimension $k$ in $\mathbb{R}^N$
Question: Can we characterize the error of the solution computed using a PROM relative to the solution obtained using the HDM?

\[ \epsilon_{\text{PROM}}(t) = w(t) - \tilde{w}(t) \]

\[ = w(t) - Vq(t) \]

assume here a Galerkin projection and an associated orthogonal basis

\[ V^T V = I_k \]

projector \( \Pi_{V,V} = VV^T \)
Question: Can we characterize the error of the solution computed using a PROM relative to the solution obtained using the HDM?

\[ E_{\text{PROM}}(t) = w(t) - \tilde{w}(t) = w(t) - Vq(t) \]

- assume here a Galerkin projection and an associated orthogonal basis
- \( V^T V = I_k \)
- projector \( \Pi_{V,V} = VV^T \)

- the error vector can be decomposed into two orthogonal components

\[ E_{\text{PROM}}(t) = w(t) - \Pi_{V,V}w(t) + \Pi_{V,V}w(t) - Vq(t) = (I_N - \Pi_{V,V})w(t) + V \left( V^T w(t) - q(t) \right) = E_{V\perp}(t) + E_V(t) \]
Error component orthogonal to $\mathbf{V}$

$$\mathbf{E}_{V \perp}(t) = (I_N - \Pi_{V,V}) \mathbf{w}(t)$$

Interpretation: The exact trajectory does not strictly belong to $\mathcal{V} = \text{range}(\mathbf{V}) \Rightarrow \text{projection error}$
Error Analysis

Orthogonal Components of the Error Vector

- Error component orthogonal to $V$
  \[
  \mathcal{E}_{V \perp}(t) = (I_N - \Pi_{V,V}) w(t)
  \]
  Interpretation: The exact trajectory does not strictly belong to $V = \text{range}(V) \Rightarrow \text{projection error}$

- Error component parallel to $V$
  \[
  \mathcal{E}_V(t) = V \left( V^T w(t) - q(t) \right)
  \]
  Interpretation: An “equivalent” but different dynamical system is solved $\Rightarrow \text{modeling error}$
Error component orthogonal to $V$

$$\mathcal{E}_{V\perp}(t) = (I_N - \Pi_{V,V})w(t)$$

Interpretation: The exact trajectory does not strictly belong to $V = \text{range}(V) \Rightarrow \text{projection error}$

Error component parallel to $V$

$$\mathcal{E}_V(t) = V(V^T w(t) - q(t))$$

Interpretation: An “equivalent” but different dynamical system is solved $\Rightarrow \text{modeling error}$

Note that $\mathcal{E}_{V\perp}(t)$ can be computed without executing the PROM and therefore can provide an a priori error estimate
Error Analysis

Orthogonal Components of the Error Vector

\( (I_N - \Pi_{V,V})w(t) \)
Error Analysis

Orthogonal Components of the Error Vector

- Again, consider the case of an orthogonal Galerkin projection
- One can derive an ODE governing the behavior of the error component lying in $\mathcal{V}$ in terms of that lying in $\mathcal{V}^\perp$

\[
\frac{d\mathcal{E}_\mathcal{V}}{dt}(t) = \Pi_{\mathcal{V},\mathcal{V}} (f(w(t), t) - f(w(t) - \mathcal{E}_\mathcal{V}(t) - \mathcal{E}_{\mathcal{V}^\perp}(t), t))
\]

\[
\mathcal{E}_\mathcal{V}(0) = 0
\]

- In the case of an autonomous linear system

\[
\frac{dw}{dt}(t) = Aw(t)
\]

the error ODE has the simple form

\[
\frac{d\mathcal{E}_\mathcal{V}}{dt}(t) = \Pi_{\mathcal{V},\mathcal{V}} (A\mathcal{E}_\mathcal{V}(t)) + \Pi_{\mathcal{V},\mathcal{V}} (A\mathcal{E}_{\mathcal{V}^\perp}(t))
\]

where $\mathcal{E}_{\mathcal{V}^\perp}(t)$ acts as a forcing term
Then, one can then derive the following error bound

**Theorem**

\[
\| \mathcal{E}_{\text{PROM}}(t) \| \leq (\| F(T, V^T AV) \|_2 \| V^T AV^\perp \|_2 + 1) \| \mathcal{E}_{V^\perp}(t) \|
\]

where \( \| \cdot \| \) denotes the \( L_2([0, T], \mathbb{R}^N) \) function norm,
\[
\| f \|_2 = \sqrt{\int_0^T \| f(\tau) \|_2^2 d\tau}, \text{ and } F(T, M) \text{ denotes the linear operator defined by}
\]

\[
F(T, M) : \mathcal{L}_2([0, T], \mathbb{R}^N) \rightarrow \mathcal{L}_2([0, T], \mathbb{R}^N)
\]

\[
u \mapsto t \mapsto \left( \int_0^t e^{M(t-\tau)} u(\tau) d\tau \right)
\]

Error bounds for the nonlinear case can be found in *A New Look at Proper Orthogonal Decomposition*, Rathiman and Petzold, SIAM Journal of Numerical Analysis, Vol. 41, No. 5, 2003
If $A$ is \textbf{symmetric} and the projection is an \textbf{orthogonal Galerkin} projection, the stability of the HDM is preserved during the reduction process (Hint: Consider the equivalent HDM and analyze the sign of $\frac{d}{dt}(\tilde{w}^T \tilde{w})$)

\[ A = \begin{bmatrix} 1 & -3 \\ 0 & .6 & -2 \end{bmatrix} \]

the eigenvalues of $A$ are $\{-0.1127, -0.8873\}$ (stable model)

Consider next the reduced-order basis $V = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and therefore the Galerkin PROM is not asymptotically stable.
Preservation of Model Stability

- If $A$ is symmetric and the projection is an orthogonal Galerkin projection, the stability of the HDM is preserved during the reduction process (Hint: Consider the equivalent HDM and analyze the sign of $\frac{d}{dt}(\tilde{w}^T\tilde{w})$).

- However, if $A$ is not symmetric, the stability of the HDM is not preserved: For example, consider a linear HDM characterized by the following unsymmetric matrix

\[
A = \begin{bmatrix}
1 & -3.5 \\
0.6 & -2
\end{bmatrix}
\]

- Consider next the reduced-order basis $V$

\[
V = \begin{bmatrix}
1 \\
0
\end{bmatrix}
\]
Preservation of Model Stability

- If \( A \) is **symmetric** and the projection is an **orthogonal Galerkin** projection, the stability of the HDM is preserved during the reduction process (Hint: Consider the equivalent HDM and analyze the sign of \( \frac{d}{dt}(\tilde{w}^T \tilde{w}) \))

- However, if \( A \) is not symmetric, the stability of the HDM is not preserved: For example, consider a linear HDM characterized by the following unsymmetric matrix

\[
A = \begin{bmatrix}
1 & -3.5 \\
0.6 & -2
\end{bmatrix}
\]

- the eigenvalues of \( A \) are \( \{-0.1127, -0.8873\} \) (stable model)
- consider next the reduced-order basis \( V \)

\[
V = \begin{bmatrix}
1 \\
0
\end{bmatrix}
\]
Preservation of Model Stability

- If \( A \) is symmetric and the projection is an orthogonal Galerkin projection, the stability of the HDM is preserved during the reduction process (Hint: Consider the equivalent HDM and analyze the sign of \( \frac{d}{dt}(\tilde{w}^T\tilde{w}) \)).

- However, if \( A \) is not symmetric, the stability of the HDM is not preserved: For example, consider a linear HDM characterized by the following unsymmetric matrix

\[
A = \begin{bmatrix}
1 & -3.5 \\
0.6 & -2
\end{bmatrix}
\]

- The eigenvalues of \( A \) are \( \{-0.1127, -0.8873\} \) (stable model).
- Consider next the reduced-order basis \( V \)

\[
V = \begin{bmatrix}
1 \\
0
\end{bmatrix}
\]

- \( A_r = [1] \) and therefore the Galerkin PROM is not asymptotically stable.