Principle Component Analysis

and

Model Reduction for Dynamical Systems

D.C. Sorensen

Virginia Tech 12 Nov 2004
Protein Substate Modeling and Identification

PCA Dimension Reduction Using the SVD

• Tod Romo
• George Phillips

Outline

- Brief Intro to Model Reduction for Dynamical Systems
- Reduced Basis Trajectory Time Integration for MD
- The Symmetric SVD: Reduced Dimension MD Simulation
**LTI Systems and Model Reduction**

**Time Domain**

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

\[A \in \mathbb{R}^{n \times n}, \ B \in \mathbb{R}^{n \times m}, \ C \in \mathbb{R}^{p \times n}, \ n >> m, p\]

**Frequency Domain**

\[
sx = Ax + Bu
\]
\[
y = Cx
\]

**Transfer Function**

\[
H(s) \equiv C(sI - A)^{-1}B, \quad y(s) = H(s)u(s)
\]
Model Reduction

Construct a new system \( \{\hat{A}, \hat{B}, \hat{C}\} \) with LOW dimension \( k << n \)

\[
\begin{align*}
\dot{x} &= \hat{A}\hat{x} + \hat{B}u \\
\hat{y} &= \hat{C}\hat{x}
\end{align*}
\]

**Goal:** Preserve system response

\[\hat{y} \text{ should approximate } y\]

**Projection:** \( x(t) = Vx(\hat{t}) \) and \( \dot{\hat{x}} = AV\hat{x} + Bu \)
Approximate $\mathbf{x} \in \mathcal{S}_V = \text{Range}(V)$ $k$-diml. subspace
i.e. Put $\mathbf{x} = V\hat{\mathbf{x}}$, and then force

$$W^T[V\hat{\mathbf{x}} - (AV\hat{\mathbf{x}} + B\mathbf{u})] = 0$$

$$\hat{\mathbf{y}} = CV\hat{\mathbf{x}}$$

If $W^T V = I_k$, then the $k$ dimensional reduced model is

$$\dot{\hat{\mathbf{x}}} = \hat{A}\hat{\mathbf{x}} + \hat{B}\mathbf{u}$$

$$\hat{\mathbf{y}} = \hat{C}\hat{\mathbf{x}}$$

where $\hat{A} = W^TAV$, $\hat{B} = W^TB$, $\hat{C} = CV$. 
Moment Matching $\leftrightarrow$ Krylov Subspace Projection

Padé via Lanczos (PVL)
Freund, Feldmann
Bai

Multipoint Rational Interpolation
Grimme
Gallivan, Grimme, Van Dooren
Gugercin, Antoulas, Beattie
Gramian Based Model Reduction

Proper Orthogonal Decomposition (POD)
Principle Component Analysis (PCA)

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

The gramian

\[ P = \int_{-\infty}^{\infty} x(\tau)x(\tau)^T d\tau \]

Eigenvectors of \( P \)

\[ P = VS^2V^T \]

Orthogonal Basis

\[ x(t) = VS_w(t) \]
PCA or POD Reduced Basis

Low Rank Approximation

\[ \mathbf{x} \approx \mathbf{V}_k \hat{\mathbf{x}}_k(t) \]

Galerkin condition – Global Basis

\[ \dot{\hat{\mathbf{x}}}_k = \mathbf{V}_k^T \mathbf{f}(\mathbf{V}_k \hat{\mathbf{x}}_k(t), \mathbf{u}(t)) \]

Global Approximation Error \((\mathcal{H}_2\text{ bound for LTI})\)

\[ \|\mathbf{x} - \mathbf{V}_k \hat{\mathbf{x}}_k\|_2 \approx \sigma_{k+1} \]

Snapshot Approximation to \(\mathcal{P}\)

\[ \mathcal{P} \approx \frac{1}{m} \sum_{j=1}^{m} \mathbf{x}(t_j)\mathbf{x}(t_j)^T \]
SVD of Snapshot Trajectory (Conformations)

\[ \mathbf{X} = [\mathbf{x}(t_1), \mathbf{x}(t_2), \ldots, \mathbf{x}(t_m)] \]

SVD of \( \mathbf{X} \):

\[
\mathbf{X} = \mathbf{VSW}^T \approx \mathbf{V}_k \mathbf{S}_k \mathbf{W}_k^T
\]

where

\[
\mathbf{V}^T \mathbf{V} = \mathbf{W}^T \mathbf{W} = \mathbf{I}_n \quad \mathbf{S} = \text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n)
\]

with \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \).
SVD Compression

Advantage of SVD Compression

\[ k (m + n) \]

versus

\[ m \times n \]

Storage

D.C. Sorensen
Image Compression - Feature Detection

original

rank = 10

rank = 30

rank = 50
POD in CFD

Extensive Literature

Karhunen-Loéve, L. Sirovich
Burns, King
Kunisch and Volkwein
Many, many others

Incorporating Observations – Balancing

Lall, Marsden and Glavaski
K. Willcox and J. Peraire
POD for LTI systems

Impulse Response: \[ \mathcal{H}(t) = C(tI - A)^{-1}B, \quad t \geq 0 \]

Input to State Map: \[ x(t) = e^{At}B \]

Controllability Gramian:

\[ \mathcal{P} = \int_0^\infty x(\tau)x(\tau)^T d\tau = \int_0^\infty e^{A\tau}BB^T e^{A^T\tau}d\tau \]

State to Output Map: \[ y(t) = Ce^{At}x(0) \]

Observability Gramian:

\[ Q = \int_0^\infty e^{A^T\tau}C^T Ce^{A\tau}d\tau \]
**Balanced Reduction (Moore 81)**

Lyapunov Equations for system Gramians

\[
A P + PA^T + BB^T = 0 \quad A^T Q + QA + C^T C = 0
\]

With \( P = Q = S \): Want Gramians Diagonal and Equal

States Difficult to Reach are also Difficult to Observe

Reduced Model

\[
A_k = W_k^T A V_k, \quad B_k = W_k^T B, \quad C_k = C_k V_k
\]

\[
\mathcal{P} V_k = W_k S_k \quad Q W_k = V_k S_k
\]

Reduced Model Gramians \( \mathcal{P}_k = S_k \) and \( Q_k = S_k \).
Hankel Norm Error estimate (Glover 84)

Why Balanced Realization?

- Hankel singular values $= \sqrt{\lambda(PQ)}$
- Model reduction $\mathcal{H}_\infty$ error (Glover)
  $$\|y - \hat{y}\|_2 \leq 2 \times \left( \text{sum neglected singular values} \right) \|u\|_2$$
- Extends to MIMO
- Preserves Stability

Key Challenge

- Approximately solve large scale Lyapunov Equations in Low Rank Factored Form
CD Player Impulse Response

\[ n = 120 \quad k = 11 \quad \text{tol} = 5 \times 10^{-3} \]
CD Player Impulse Response

\( k = 17 \), \( tol = 5e-4 \)

[Graph of CD Player Impulse Response]
CD Player Impulse Response

\[ k = 31 , \text{ tol } = 5e-5 \]
CD Player - Hankel Singular Values

Hankel Singular Values

D.C. Sorensen 20
Reduction of Second Order Systems

\[
\begin{align*}
M\ddot{x} + G\dot{x} + Kx &= Bu \\
y(t) &= Cx(t) \\
\rightarrow \\
\hat{M}\ddot{\hat{x}} + \hat{G}\dot{\hat{x}} + \hat{K}\hat{x} &= \hat{B}u \\
\hat{y}(t) &= \hat{C}\hat{x}(t)
\end{align*}
\]

where \( \hat{M} = V^T MV \), etc. with \( V^T V = I \).

**Key Point:** *Preserve Second Order Form*  
– DO NOT convert to First Order Sys.

Keeps Physical Meaning - can be built
Applications

Mechanical Systems  Electrical Systems  MEMS devices

- e.g. Building Model

\[ N = 26394, k = 200 \text{ (ROM)} \]

Frequency Response $\alpha = 0.67, \beta = 0.0032$

Response $|G(j\omega)|$

D.C. Sorensen
Error Bound for Second Order Systems

A. Antoulas, C. Teng

Controllability Gramian - Impulse Response

\[ P := \int_{0}^{\infty} x(t)x(t)^{*} \, dt. \]

Reduce with Dominant Eigenspace \( P \):

\[ PV_1 = V_1 S_1 \]

Bounded \( H_2 \) norm of error system \( \mathcal{E} = \Sigma - \hat{\Sigma} \)

\[ \| \mathcal{E} \|_{H_2}^2 \leq \text{tr} \left\{ S_2 \right\} \]

**Key:** Expression for \( P \) in frequency domain.
PCA Model Reduction for Molecular Dynamics

- Rachel Vincent
- Monte Pettitt
Molecular dynamics (MD) simulation is a computational tool used to study a molecular system as it evolves through time. Newton’s second law of motion governs atomic motion in MD:

\[
M \ddot{\mathbf{r}}(t) = -\nabla \mathcal{V}(\mathbf{r}(t)).
\]

- \( \mathbf{r}(t) = \) vector of atomic coordinates at time \( t \)
  \[
  = [x_1 t \ y_1 t \ z_1 t \ \cdots \ x_{N_a} t \ y_{N_a} t \ z_{N_a} t]^T
  \]
- \( M = \) diagonal matrix of atomic masses
- \( \mathcal{V}(\mathbf{r}(t)) = \) potential energy function
Time Step Barrier $\propto \text{fs}$

Example: DHFR (dihydrofolate reductase), 23,558 atoms

To realize a microsecond simulation with a time step of 2 fs would require about 13 months of simulation time when utilizing 126 processors.

Time with respect to simulation using the NAMD program (Not Another Molecular Dynamics program) on an Origin 2000 R10000/250.
PCA Reduced Basis Simulation

\[ MV\ddot{y}(t) = -\nabla \mathcal{V}(Vy(t)) \rightarrow \hat{M}\ddot{y}(t) = -V^T \nabla \mathcal{V}(Vy(t)) \]

1. Initial Basis \( V \): truncated SVD of short traditional MD trajectory using ARPACK.
2. Approximate the reduced basis potential energy with Radial Basis Fit.
3. Update reduced basis positions \( y \) and velocities \( \dot{y} \) in \( k \) dimensions using the approximate potential.
4. Reconstruct 3\( ND \) trajectory \( r = Vy \).
5. Update and truncate reduced basis and perform full space correction as needed.
Remarks

- **Butane (n = 42):**
  80% to 90% of the total motion with 10 - 15 LSVs (24% - 36% DOF)
  Reduced Simulation times order of seconds
  Traditional MD simulation took several minutes.

- **BPTI (n = 2700):**
  80% to 90% of the total motion with 300 - 500 LSVs (11% - 19% DOF)
  Reduced Simulation times order of minutes
  Traditional MD simulation took several hours.
Symmetry Preserving SVD (Mili Shah)

Collaboration with the Physical and Biological Computing Group

- Lydia Kavraki
- Mark Moll
- David Schwarz
- Amarda Shehua
- Allison Heath
Symmetry in HIV-1 protease

Backbone representation of HIV-1 protease (from M. Moll)

bound to an inhibitor (shown in orange)
Uses PCA dimension reduction of Molecular Dynamics Simulations

Symmetry across a plane should be present
Animation: Symmetric SVD Approximation

Comparison of Regular and Symmetric Preserving SVD

click below figures for movies
Finding the Plane of Symmetry

Suppose

\[ X = [x_1, x_2, \ldots, x_n] \quad \text{and} \quad Y = [y_1, y_2, \ldots, y_n] \]

are two sets of points symmetric across a plane

Exact symmetry condition:

\[ Y = (I - 2ww^T)X, \]

where \( w \) is the normal to the (hyper-) Plane of Symmetry

\[ \mathcal{H} = \{ x : w^T x = 0 \} \]

Remark: In Numerical Linear Algebra \((I - 2ww^T)\) is a Householder Transformation or Elementary Reflector
Best Approximate Plane of Symmetry

Symmetry condition with Noise:

\[ Y = (I - 2w_ow_o^T)X + E, \]

**Problem:** Compute a unit vector \( w \) that gives the best
Approximate Plane of Symmetry

\[ \min_w \|Y - (I - 2ww^T)X\|_F, \]

**Solution:**

\[ (XY^T + YX^T)v = v\lambda_{min}, \quad w = v \]

gives the normal \( w \) to the best approximate plane of symmetry
Supressing Outlier Effects

Iteratively determine diagonal \textit{weighting matrix} \( D_w \)

The \( i \)-th diagonal of \( D_w \) is \( 1/\text{discrepancy} \),

\[
\text{discrepancy} = \|y_i - (I - 2ww^T)x_i\|
\]

\textbf{Problem:} Compute a unit vector \( w \) that gives the best 

\textit{Weighted Approximate Plane of Symmetry}

\[
\min_v \| [Y - (I - 2vv^T)X]D_w \|_F,
\]

\textbf{Solution:}

\[
(XD_w^2Y^T + YD_w^2X^T)v = v\lambda_{\text{min}}, \quad w \leftarrow v
\]

gives the normal \( w \) to the best weighted approximate plane of symmetry
Finding Normal to “Best” Plane of Symmetry

click below for movie
The Symmetric SVD Approximation

If \( \mathbf{WX}_2 = \mathbf{X}_1 + \mathbf{E} \) where \( \mathbf{W} = \text{blockdiag}(\mathbf{I} - 2\mathbf{w}\mathbf{w}^T) \)

\[
\min_{\mathbf{W}\hat{\mathbf{X}}_2 = \hat{\mathbf{X}}_1} \left\| \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} - \begin{pmatrix} \hat{\mathbf{X}}_1 \\ \hat{\mathbf{X}}_2 \end{pmatrix} \right\|_F^2 \quad \text{and} \quad \begin{pmatrix} \hat{\mathbf{X}}_1 \\ \hat{\mathbf{X}}_2 \end{pmatrix} = \mathbf{USV}^T
\]

Solved by:

\[
\mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{pmatrix}, \quad \mathbf{S} = \sqrt{2}\mathbf{S}_1, \quad \mathbf{V} = \mathbf{V}_1. \quad \text{and} \quad \mathbf{U}_2 = \mathbf{WU}_1,
\]

with

\[
\mathbf{U}_1\mathbf{S}_1\mathbf{V}_1^T = \frac{1}{2} (\mathbf{X}_1 + \mathbf{WX}_2)
\]
Symmetric Major Modes: HIV-1 protease

- Major mode regular SVD is red
- Major mode SYMMETRIC SVD is blue
- 3120 atoms (3*3120=9360 degrees of freedom)
- MD trajectory consisted of 10000 conformations (NAMD)
- SVD and SymSVD used P_ARPACK on a Linux cluster
- dual-processor nodes; 1600MHz AMD Athlon processors, 1GB RAM per node. 1GB/s Ethernet connection. 12 Processors = 6 nodes.
- First 10 standard singular vectors: 88 secs.
- First 10 symmetric singular vectors: 131 secs.
Animation: Symmetric SVD on HIV1 Protease

Red = Unsymmetric

Blue = Symmetric

First SVD mode – Symmetric vs. Unsymmetric
Rotational Symmetry

\[ X_j = WX_{j-1}, \quad j = 1 : k - 1, \quad \text{where} \quad W = I - QGQ^T \]

\( I_p - G \) is a rotation \( \quad X_k = X_0. \)
Finding the Axis of Rotation

\( q \) is an axis of rotation iff \( Q^T q = 0 \)

\[ q^T W = q^T (I - QGQ^T) = q^T \quad \Rightarrow \quad q^T X_0 = q^T X_j \]

Let \( M = (k - 1)X_0 - \sum_{j=1}^{k-1} X_j \)

\[ \min_{\|q\|=1} \| M^T q \| \quad (= 0 \ \text{if exact symmetry holds}) \]

Good for noisy data

(for another condition see Minovic, Ishikawa and Kato)
**Best Rotationally Symmetric Approximation**

If \( W^{k-j}X_j = X_0 + E_j, \ j = 1 : k - 1 \)

\[
\min_{\hat{X}_{j+1}=w\hat{X}_j} \left\| \begin{bmatrix} X_0 \\ \vdots \\ X_{k-1} \end{bmatrix} - \begin{bmatrix} \hat{X}_0 \\ \vdots \\ \hat{X}_{k-1} \end{bmatrix} \right\|_F^2 = \frac{1}{k} \sum_{j=0}^{k-1} \| E_j \|_F^2,
\]

\[
\begin{bmatrix} \hat{X}_0^T & \cdots & \hat{X}_{k-1}^T \end{bmatrix}^T = USV^T \quad \text{with} \quad U = \frac{1}{\sqrt{k}} \begin{bmatrix} U_0^T & \cdots & U_{k-1}^T \end{bmatrix}^T.
\]

\[
U_0S_0V_0^T = \frac{1}{k}(X_0 + W^{k-1}X_1 + W^{k-2}X_2 + \ldots + WX_{k-1}).
\]

\[
S = \sqrt{k}S_0 \quad V = V_0 \quad U_j = W^jU_0, \ j = 0, 1, 2, \ldots, k - 1
\]
Animation: Rotationally Symmetric SVD Approximation

click below figures for movies
Animation: Rotationally Symmetric SVD on HIV1

Red = Unsymmetric  Blue = Symmetric

Second SVD mode – Rotationally Symmetric vs. Unsymmetric
Potential for Symmetric SVD

- Obtain a Symmetric PCA reduced dimension approximate trajectory
- Test Hypothesis of Symmetry in an Unknown Protein
- Locate Symmetric Sub-Structures

Things to Do:
- Improve convergence rate for finding $w$
- Give a complete analysis of convergence
- Give a complete analysis of discrepancy weighting
- Extend to more complex symmetries
- Find New Applications
Contact Info.

e-mail: sorensen@rice.edu

web page: www.caam.rice.edu/~sorensen/

ARPACK: www.caam.rice.edu/software/ARPACK/