Normal Mode Analysis

Jun Wan and Willy Wriggers
School of Health Information Sciences &
Institute of Molecular Medicine
University of Texas – Houston
Contents

• Introduction --- protein dynamics (low-frequency vibration and variables reduction)
• Normal mode analysis (NMA)
• Application and limitation
• Conclusion and future work
Protein Dynamics is Hierarchical

Vibration of bonds: $10^{-15}$ s

Protein folding/unfolding: $10^{-6}$ s, $10^{-3}$ s, s and even longer

Large-scale functional motions
Collective Coordinates and Dimensionality Reduction

\[ z = \frac{\sqrt{2}}{2} x + \frac{\sqrt{2}}{2} y \]
Find a Model

- To investigate low-frequency movement (vibration)
- To reduce the number of degrees of freedom
Normal Mode Analysis

- Theory of vibration
- Harmonic potential
- Close to the potential minimum
- Orthogonal normal modes
- Conformational fluctuation = a superposition of normal modes.
Harmonic Approximation

Approximation:

Potential energy $\Rightarrow$ harmonic

$$E \approx E_0 + \sum_i \frac{\partial E}{\partial x_i} \Delta x_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 E}{\partial x_i \partial x_j} \Delta x_i \Delta x_j$$
Harmonic Oscillator

Newton / Hooke

\[ m \frac{d^2 \Delta x}{dt^2} = -k \Delta x - k \Delta x \]

\[ m \frac{d^2 \Delta x}{dt^2} + 2k \Delta x = 0 \]

\[ \Delta x = A \cos(\omega t + \delta) \]

\[ \omega^2 = 2k / m \]
Coupled oscillators

\[
\begin{align*}
    m \frac{d^2 \Delta x_1}{dt^2} &= -k \Delta x_1 + k(\Delta x_2 - \Delta x_1) \\
    m \frac{d^2 \Delta x_2}{dt^2} &= -k(\Delta x_2 - \Delta x_1) - k\Delta x_2
\end{align*}
\]

Matrix Diagonalization

\[
\begin{pmatrix}
    \Delta x_1 \\
    \Delta x_2
\end{pmatrix} = \frac{1}{\sqrt{2}}
\begin{pmatrix}
    1 & 1 \\
    1 & -1
\end{pmatrix}
\begin{pmatrix}
    \Delta u_1 \\
    \Delta u_2
\end{pmatrix}
\]

\[
\begin{align*}
    m \frac{d^2 (\Delta u_1)}{dt^2} + k \Delta u_1 &= 0 \\
    m \frac{d^2 (\Delta u_2)}{dt^2} + 3k \Delta u_2 &= 0
\end{align*}
\]

\[
\begin{align*}
    \Delta u_1 &= A_1 \cos(\omega_1 t + \delta_1) \\
    \Delta u_2 &= A_2 \cos(\omega_2 t + \delta_2)
\end{align*}
\]

\[
\omega_1^2 = k / m \\
\omega_2^2 = 3k / m
\]
Eigenvalue and Eigenvector Problem

\[ m \frac{d^2}{dt^2} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} + \begin{pmatrix} 2k & -k \\ -k & 2k \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} = 0 \]

\[ \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \Delta u_1 \\ \Delta u_2 \end{pmatrix} \]

\[ \mathbf{U} \]

\[ \mathbf{F} \]

\[ \mathbf{U} \]

\[ \mathbf{U} \]

\[ \mathbf{U}^t \mathbf{F} \mathbf{U} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \]

\[ \mathbf{U}^t \mathbf{U} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \]

Matrix Diagonalization

U is chosen so that it satisfies the following conditions.
Conformational Fluctuation

...is given by a superposition of normal modes:

\[
\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} = \frac{A_1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cos(\omega_1 t + \delta_1) + \frac{A_2}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \cos(\omega_2 t + \delta_2)
\]

Lower frequency mode

\[
\omega_1 = \sqrt{\frac{k}{m}}
\]

Higher frequency mode

\[
\omega_2 = \sqrt{\frac{3k}{m}}
\]
Two-Atomic Molecule

Spring constant
\[ k \]

\[ \leftrightarrow \]

Frequency
\[ \nu \propto \sqrt{k} \]
Three-Atomic Molecule

\[ v_1 + v_2 + v_3 \]
Multi-Atom Molecule

- Low frequencies
  - Large collective motions
- High frequencies
  - Localized motions

© Florence Tama
Success Story
Success Story

EM
Flexible Docking
NMA
X-ray

Darst et al, PNAS, 2002, 99:4296

70% overlap between the direction of the observed displacements with the direction of mode 1
Success Story

• One mode can represent 70-90% of functionally relevant motion.
• For many observed movements, the first 12 normal modes contain the relevant degrees of freedom
NMA using Molecular Mechanics

Full atomic representation and MM interactions require:

- energy minimization
- diagonalization of the 2\textsuperscript{nd} derivative of the potential energy (3N x 3N Hessian matrix)
Computational Challenges

NMA requires:

- minimization
- diagonalization of the Hessian matrix

Problems for large systems:

- expensive, cumbersome (MM)
- memory requirements
Memory-Efficient Diagonalization

DIMB => Diagonalization in mixed basis
(Perahia & Mouawad, 1995, J. Comp. Chem. 19, 241)

Group theory => Use symmetrical properties of viruses

RTB => Rotation Translation Blocks method gives approximate low-frequency NM

- block = 1 or several residues
- rotation + translation of block => new basis
- expression of Hessian in this new basis
- diagonalization of a matrix $6n_B * 6n_B$

© Florence Tama
Reducing the Number of Variables

**Cartesian coordinate space**
3N-6 variables are necessary
N : number of atoms

**Torsion angle space**
Bond angles and bond lengths are fixed, and only torsion angles are allowed to vary.
Number of variables: ~1/10
Elastic Network Model


Simplified force-field: no MM, already minimized

\[
E(r_a, r_b) = \frac{C}{2} \left( |r_{a,b}| - |r_{a,b}^0| \right)^2
\]

\[
E_p = \sum_{a, b} E(r_a, r_b)
\]

Possibility to reduce level of detail (up to 1 point for 40 residue)
Vector Quantization

Encode data (in $\mathbb{R}^{d=3}$) using a finite set $\{w_j\}$ ($j=1,\ldots,k$) of codebook vectors.

Delaunay triangulation divides $\mathbb{R}^3$ into $k$ Voronoi polyhedra ("receptive fields"): 

![Diagram of Delaunay triangulation]

Encoding Distortion Error:

$$E = \sum_{i (\text{atoms, voxels})} \left\| v_i - w_{j(i)} \right\|^2 m_i$$


Choice of Cut-off

1 codebook vector ≈ 1 residue
⇒ 10-12 Å cut-off OK

Reducing number of codebook vectors
⇒ too sparse connectivity

Inspect the pair-distance distribution of codebook vectors and increase cutoff beyond first peak.

© Florence Tama
Level of Detail not Important

Projection onto atomic normal modes ≈ 1 for the first few modes

Low frequency NM are similar to atomic NM

Models can reproduce functional rearrangements even at 30Å resolution

© Florence Tama
Application to EM Data

RNA Polymerase, S. Darst et al.

Deposition of Density Map
Application to EM Data

RNA Polymerase, S. Darst et al.
Application to EM Data

RNA Polymerase, S. Darst et al.
Application to EM Data

RNA Polymerase, S. Darst et al.

Deposition of Density Map → Vector Quantization → Choice of Cut-off → NMA
Application to EM Data

RNA Polymerase, S. Darst et al.
Examples

Ribosome

RNA Polymerase
What are the Limitations of NMA (I)?

- We do not know \textit{a priori} which is the relevant mode, but the first 12 low-frequency modes are probable candidates.

- The amplitude of the motion is unknown.

- NMA requires additional standards for parameterization, i.e. a screening against complementary experimental data to select the relevant modes and amplitude.

- Expert user input / evaluation required

- Not based on first principles of physics (like MD).
Solution 1: Annotation of Modes

(Essam Metwally: emotion.biomachina.org)
What are the Limitations of NMA (II)?

- Normal modes may break the symmetry of structures due to forced orthogonalization:

Global representation $\rightarrow$ Local description of dynamic feature
Local Feature Analysis
(Zhiyong Zhang)

Global modes replaced by 3N LFA highly correlated output functions

Sparsification: n “seed atoms” + their neighboring correlated regions

T4 Lysozyme:

\[ n=4 \]

\[ C_{a} -109 \]
\[ C_{a} -51 \]
\[ C_{a} -162 \]
\[ C_{a} -1 \]
Conclusion

- Normal mode analysis is an alternative method to study dynamics of molecules.
- Normal mode analysis does not require trajectory, working with single structure.
- Conformational fluctuation is given by a superposition of normal modes.
- We are using normal mode analysis to refine small-angle X-ray scattering profiles.
Acknowledgement

Dr. Zhiyong Zhang
Dr. Pablo Chacon
Dr. Florence Tama
Dr. Atsushi Matsumoto at Wright-Rieman Laboratory, Rutgers University