Some Dynamical Systems on Manifolds

Group Theory, Linear Transformations, and Flows

Moody T. Chu

North Carolina State University

June 30, 2004

Workshop on Lie Group Methods and Control Theory

Presented at

North Carolina State University
Outline

Motivation

Realization Process

ACaseStudy

Basic Form

Similarity Property

Decomposition Property

Reversal Property

Matrix Groups and Group Actions

Tangent Space and Projection

Canonical Forms

Objective Functions and Dynamical Systems

Examples

Least Squares

New Thoughts

Conclusion

Objective Functions and Dynamical Systems

Examples

Least Squares

Matrix Groups and Group Actions

Tangent Space and Projection

Canonical Forms

Objective Functions and Dynamical Systems

Examples

Least Squares

New Thoughts

Conclusion
Motivation

What is the simplest form to which a family of matrices depending smoothly on the parameters can be reduced by a change of coordinates depending smoothly on the parameters?

V. I. Arnold

What is the simplest form referred to here?

What kind of continuous change can be employed?
Realization Process

Realization, in a sense, means any deducible procedure that we use to rationalize and solve problems.

In mathematics, realization process often appears in the form of an iterative procedure or a differential equation. The steps taken for the realization, i.e., the changes, could be discrete or continuous.

The simplest form refers to the agility to think and draw conclusions.
Motivation

Continuous Realization

Two abstract problems:

\(\square\) One is a make-up and is easy.
\(\square\) The other is the real problem and is difficult.

\(\square\) A bridge: A continuous path connecting the two problems.
\(\square\) A path that is easy to follow.
\(\square\) A method for moving along the bridge.
\(\square\) A method that is readily available.

A numerical method:
Build the Bridge

Motivation

\[ \text{Specified guidance is available.} \]
\[ \text{Such as the projected gradient method.} \]
\[ \text{The bridge is constructed by monitoring the values of certain specified functions.} \]

\[ \text{Specified guidance is available.} \]

\[ \text{Such as the homotopy method.} \]
\[ \text{A bridge is built in a straightforward way.} \]
\[ \text{The path is guaranteed to work.} \]

\[ \text{Specified guidance is available.} \]

\[ \text{Such as the projected gradient method.} \]
\[ \text{A bridge is built similarly by accident.} \]
\[ \text{The path is guaranteed to work.} \]

\[ \text{Specified guidance is available.} \]

\[ \text{Such as the isospectral flows.} \]
\[ \text{Usually deeper mathematical theory is involved.} \]
\[ \text{A bridge is built seemingly by accident.} \]

\[ \text{Specified guidance is available.} \]

\[ \text{Such as the homotopy method.} \]
\[ \text{A bridge is built in a straightforward way.} \]
\[ \text{The path is guaranteed to work.} \]

\[ \text{Specified guidance is available.} \]

\[ \text{Such as the projected gradient method.} \]
\[ \text{A bridge is built seemingly by accident.} \]
\[ \text{The path is guaranteed to work.} \]

\[ \text{Specified guidance is available.} \]
Motivation

Characteristics of a Bridge

- A bridge, if it exists, usually is characterized by an ordinary differential equation.
- The discretization of a bridge, or a numerical method in travelling along a bridge, usually produces an iterative scheme.
Two Examples

Constrained Least Squares Approximation

Eigenvalue Computation

Two Examples
**The Eigenvalue Problem**

- Motivation

The mathematical problem:
- An asymmetric matrix $A$ is given.
- Solve the equation $A\mathbf{x} = \lambda \mathbf{x}$ for a non-zero vector $\mathbf{x}$ and a scalar $\lambda$.

An iterative method:

For a non-zero vector $\mathbf{x}$ and a scalar $\lambda$, solve the equation $A\mathbf{x} = \lambda \mathbf{x}$.

**The Eigenvalue Problem**

**The $QR$ algorithm (Francis' 61):**
- The $QR$ decomposition:
  - An iterative method:
  - For a non-zero vector $\mathbf{x}$ and a scalar $\lambda$.

The sequence $fA_kg$ converges to a diagonal matrix.

- Every matrix $A$ has the same eigenvalues of $A^0$, i.e., $\{\lambda\}$ converges to a diagonal matrix.

- The $QR$ decomposition:
  - Where $Q$ is orthogonal and $R$ is upper triangular.

- The sequence $fA_kg$ converges to a diagonal matrix.

- Every matrix $A$ has the same eigenvalues of $A^0$, i.e., $\{\lambda\}$ converges to a diagonal matrix.

- The $QR$ decomposition:
  - An iterative method:
  - For a non-zero vector $\mathbf{x}$ and a scalar $\lambda$.

$x\lambda = x^0\lambda$
Motivation

A continuous method:

\[ X = X_0 + X_+ + X_- \]

where \( X_0 \) is the diagonal, \( X_+ \) is the strictly upper triangular, and \( X_- \) is the strictly lower triangular part of \( X \).

\( X + \_X + \_X = X \)

The Toda lattice is based on the physics.

This is a Hamiltonian system.

Evolution starts from \( X_0 \) and converges to the limit point of Toda flow, which is a diagonal matrix, maintains the spectrum.

Sampled at integer times, \( \{ (y(t))X \} \) gives the same sequence as does the algorithm applied to the matrix \( X_0 = exp(X) \).

Deft (Symes'82, Deift et al.'83):

Define \( \Pi X \) to be the strictly upper triangular part of \( X \) and \( \Pi X + \_X + \_X = X \).

The construction of the Toda lattice is based on the physics.

Sampled at integer times, \( \{ (y(t))X \} \) gives the same sequence as does the algorithm applied to the matrix \( X_0 = exp(X) \).

The convergence is guaranteed by 'nature'.
Motivation

Least Squares Matrix Approximation

- The mathematical problem:
  - A symmetric matrix $N$ and a set of real values $\lambda_1, \ldots, \lambda_n$ are given.
  - Find a least squares approximation of $N$ that has the prescribed eigenvalues.
- A standard formulation:
  
  \[
  \begin{align*}
  \text{Minimize } & \quad F(Q) := \frac{1}{2} |Q^T \Lambda Q - N|^2 \\
  \text{Subject to } & \quad Q^T Q = I.
  \end{align*}
  \]

- Equality Constrained Optimization:
  - Augmented Lagrangian methods.
  - Sequential quadratic programming methods.
- None of these techniques is easy.
- The constraint carries lots of redundancies.
A continuous approach:

The projection of the gradient of $F$ can easily be calculated.

Projected gradient flow (Brockett '88, Chu & Driessel '90):

\[ \frac{dX}{dt} = X; [X; N] \]

\[ X(0) = \delta \]

\[ X : = QT \delta \]

\[ X(t) \text{ moves in a descent direction to reduce } \|N - X\|_2^2 \]

\[ \frac{dV}{dt} = X \frac{\partial V}{\partial X} \]

The optimal solution $X$ can be fully characterized by the spectral decomposition of $N$ and is unique.

Evolution starts from an initial value and converges to the limit point, which solves the least squares problem.

This is a descent flow. The flow is built on the basis of systematically reducing the difference between the current position and the target position.

A continuous approach.
Motivation

Equivalence

Suppose $X$ is tridiagonal. Take

\[ \{ i, 2, \ldots, n \} \text{ diag } = N \]

A gradient flow hence becomes a Hamiltonian flow:

\[ (X)^0 H = [N, X] \]

then

(Bloch 90) Suppose $X$ is tridiagonal. Take
\( \dot{X} = (X)^{\gamma} + (X)^{\gamma} \odot \)

and

Parameter dynamics:

\[
\begin{align*}
I &= (0)^{\beta} \\
\frac{d}{dt}((iX)^{\gamma}) &= \frac{ip}{(i)^{\beta}p} \\

0X &= (0)X \\
((iX)^{\gamma})X &= \frac{ip}{(i)XP}
\end{align*}
\]
\[ \begin{align*} \mathcal{Z}(t) &= \mathcal{Z}(0) = X(0) = X_0. \\
\mathcal{Z}(t) &= \left(1_{16}(X)^{16} - X^{16}\right) + \left(1_{16}(X)(X)^{16} - (X)^{16}X^{16}\right) + \left(1_{16}(X)(X)^{16}(X)^{16}\right) = 0. \\
\text{Check:} \quad \mathcal{X}(t) &= 1_{16}(i)^{16}(i)(X)^{16} = (i)X. \\
\text{Define:} \quad 1_{-1}(i)^{16}(i)^{16}X_{-1}(i)^{16} &= (i)X. \\
\end{align*} \]
\[(i^{0}X)d\varepsilon x = (i)Z\]

By the uniqueness theorem in the theory of ordinary differential equations, \[\exp(tX) = g_{1}(t)g_{2}(t)\]

Trivially \(\exp(tX)\) satisfies the IVP

\[\frac{dY}{dt} = X; Y(0) = I\]

Define \[Z(t) = g_{1}(t)g_{2}(t)\]

Then \(Z(0) = I\)

and

\[I = (0)X = \frac{\partial p}{\partial p}\]

\[\frac{dZ}{dt} = (g_{1}) (g_{2})^{-1} g_{2} + g_{1} (g_{2})^{-1} g_{2} = \frac{\partial p}{\partial p}\]

Trivially \(\exp(tX)\) satisfies the IVP

\[(i^{0}X)d\varepsilon x = (i)Z\]

Decomposition Property

Basic Form
\[ (1(i)X)^{dx} = \]
\[ (1(i)\delta_{X_1} - (i)\delta)^{dx} = \]
\[ (i)\delta_{X}^{dx} - (i) \delta = (i)\delta(i) \delta \]

By Decomposition Property.

\[ (i)\delta(i) \delta = ((i)X) i^{dx} \]

Reversal Property
Abstraction

QR-type Decomposition:

- Lie algebra decomposition of $g(\mathfrak{gl}(n))$ in the neighborhood of $I$.
- Arbitrary subspace decomposition $g(\mathfrak{gl}(n)) \leftrightarrow$ Factorization of a one-parameter semigroup in the neighborhood of $I$ as the product of two nonsingular matrices, i.e.,

$$\exp(X_0 t) = g_1(t)g_2(t).$$

QR-type Algorithm:

- The product $g_1(t)g_2(t)$ will be called the abstract $g_1g_2$ decomposition of $\exp(X_0 t)$.
- By setting $t = 1$, we have

$$\exp(X(0)) = g_1(1)g_2(1).$$

Corresponding to the abstract $g_1g_2$ decomposition, the above iterative process for all feasible integers will be called the abstract $g_1g_2$ algorithm.

The dynamical system for $X(t)$ is autonomous $\rightarrow$ The above phenomenon will occur at every feasible integer time.
A subset of nonsingular matrices (over any field) which are closed under matrix multiplication and inversion is called a matrix group.
<table>
<thead>
<tr>
<th>Group</th>
<th>Notation</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>$GL(n)$</td>
<td>$f$</td>
</tr>
<tr>
<td>Special Linear</td>
<td>$SL(n)$</td>
<td>$j\det(A) = 1$</td>
</tr>
<tr>
<td>Upper Triangular</td>
<td>$U(n)$</td>
<td>$A$ is upper triangular</td>
</tr>
<tr>
<td>Unipotent</td>
<td>$U(n)$</td>
<td>$a_{ii} = 1$ for all $i$</td>
</tr>
<tr>
<td>Orthogonal</td>
<td>$O(n)$</td>
<td>$jQ = Q &gt; 0$</td>
</tr>
<tr>
<td>Generalized Orthogonal</td>
<td>$OS(n)$</td>
<td>$jQ &gt; SQ$</td>
</tr>
<tr>
<td>Symplectic</td>
<td>$Sp(2n)$</td>
<td>$J$</td>
</tr>
<tr>
<td>Lorentz</td>
<td>$LO(n;k)$</td>
<td>$L := \text{diag}(1;\ldots;1,-1;\ldots;-1)$</td>
</tr>
<tr>
<td>An affine</td>
<td>$Aff(n)$</td>
<td>$\begin{pmatrix} A &amp; t \ 0 &amp; 1 \end{pmatrix}$</td>
</tr>
<tr>
<td>Translation</td>
<td>$Trans(n)$</td>
<td>$I_t 0 1$</td>
</tr>
<tr>
<td>Isometry</td>
<td>$Isom(n)$</td>
<td>$Q t 0 1$</td>
</tr>
<tr>
<td>Center of $G$</td>
<td>$Z(G)$</td>
<td>$z^g = gz$ for every $g \in G$</td>
</tr>
<tr>
<td>Product of $G_1$ and $G_2$</td>
<td>$G_1 \times G_2$</td>
<td>$g_1; g_2$</td>
</tr>
<tr>
<td>Quotient $G/N$</td>
<td>$G/N$</td>
<td>$N/\langle G \rangle$</td>
</tr>
<tr>
<td>Hessenberg</td>
<td>$\mathcal{H}$</td>
<td>$\langle G \rangle$</td>
</tr>
</tbody>
</table>

**Characteristics**

- $\langle G \rangle$ is a fixed normal subgroup of $G$ and $G_1$ and $G_2$ are given groups.
- For every $g \in G$, $\langle g \rangle = \langle g \rangle$.

<table>
<thead>
<tr>
<th>Group</th>
<th>Notation</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>$GL(n)$</td>
<td>$f$</td>
</tr>
<tr>
<td>Special Linear</td>
<td>$SL(n)$</td>
<td>$j\det(A) = 1$</td>
</tr>
<tr>
<td>Upper Triangular</td>
<td>$U(n)$</td>
<td>$A$ is upper triangular</td>
</tr>
<tr>
<td>Unipotent</td>
<td>$U(n)$</td>
<td>$a_{ii} = 1$ for all $i$</td>
</tr>
<tr>
<td>Orthogonal</td>
<td>$O(n)$</td>
<td>$jQ = Q &gt; 0$</td>
</tr>
<tr>
<td>Generalized Orthogonal</td>
<td>$OS(n)$</td>
<td>$jQ &gt; SQ$</td>
</tr>
<tr>
<td>Symplectic</td>
<td>$Sp(2n)$</td>
<td>$J$</td>
</tr>
<tr>
<td>Lorentz</td>
<td>$LO(n;k)$</td>
<td>$L := \text{diag}(1;\ldots;1,-1;\ldots;-1)$</td>
</tr>
<tr>
<td>An affine</td>
<td>$Aff(n)$</td>
<td>$\begin{pmatrix} A &amp; t \ 0 &amp; 1 \end{pmatrix}$</td>
</tr>
<tr>
<td>Translation</td>
<td>$Trans(n)$</td>
<td>$I_t 0 1$</td>
</tr>
<tr>
<td>Isometry</td>
<td>$Isom(n)$</td>
<td>$Q t 0 1$</td>
</tr>
<tr>
<td>Center of $G$</td>
<td>$Z(G)$</td>
<td>$z^g = gz$ for every $g \in G$</td>
</tr>
<tr>
<td>Product of $G_1$ and $G_2$</td>
<td>$G_1 \times G_2$</td>
<td>$g_1; g_2$</td>
</tr>
<tr>
<td>Quotient $G/N$</td>
<td>$G/N$</td>
<td>$N/\langle G \rangle$</td>
</tr>
<tr>
<td>Hessenberg</td>
<td>$\mathcal{H}$</td>
<td>$\langle G \rangle$</td>
</tr>
</tbody>
</table>
A function $\varphi : G \to \text{End}(V)$ is said to be a group action of $G$ on a set $V$ if and only if:

1. \( \varphi(e) = \text{id}_V \) for the identity element $e$ in $G$. 
2. \( \varphi(gh)(x) = \varphi(g)(\varphi(h)(x)) \) for all $g, h \in G$ and $x \in V$.

Given $x \in V$, two important notions associated with a group action $\varphi$ are:

1. **The orbit** of $x$ is $\text{Orb}_\varphi(x) := \{x = (x, b) \mid x \in G\}$.
2. **The stabilizer** of $x$ is $\text{Stab}_\varphi(x) := \{g \mid \varphi(g)(x) = x\}$.

A function $\varphi : G \to \text{End}(V)$ is said to be a group action of $G$ on a set $V$ if and only if:

- $\varphi(e) = \text{id}_V$ for the identity element $e$ in $G$. 
- $\varphi(gh)(x) = \varphi(g)(\varphi(h)(x))$ for all $g, h \in G$ and $x \in V$. 
- $\text{Orb}_\varphi(x) := \{x = (x, b) \mid x \in G\}$ for all $x \in V$. 
- $\text{Stab}_\varphi(x) := \{g \mid \varphi(g)(x) = x\}$ for all $x \in V$. 

**Group Actions**
<table>
<thead>
<tr>
<th>Matrix Groups</th>
<th>Functions</th>
<th>Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Matrix Groups</strong></td>
<td><strong>Functions</strong></td>
<td><strong>Groups</strong></td>
</tr>
<tr>
<td><strong>Set</strong></td>
<td><strong>Action</strong></td>
<td><strong>Group</strong></td>
</tr>
<tr>
<td><strong>V</strong></td>
<td><strong>(V, G)</strong></td>
<td><strong>G</strong></td>
</tr>
</tbody>
</table>

**Applications**

- **Orthogonal Simplicity**
  - $bV \perp \rho$
  - $\mathcal{O}$

- **Conjugation**
  - $bV_{1-\rho}$
  - $\mathcal{G}$

- **Application**
  - $(V, \rho)$

**Groups**

- **Simultaneous Reduction**
  - $\mathcal{G}_1 \times \cdots \times \mathcal{G}_m$

- **Simultaneous Value Decompostion**
  - $\mathcal{G}_1 \times \cdots \times \mathcal{G}_m$

- **Simultaneous Value Decompostion ZD**
  - $\mathcal{G}_1 \times \cdots \times \mathcal{G}_m$

- **Generalized Simultaneous Value Decompostion**
  - $\mathcal{G}_1 \times \cdots \times \mathcal{G}_m$

- **Symmetric Positive Definite Simultaneous Reduction**
  - $\mathcal{G}_1 \times \cdots \times \mathcal{G}_m$

- **Orthogonal Simplicity**
  - $\mathcal{O}_1 \times \cdots \times \mathcal{O}_m$

- **Conjugation**
  - $\mathcal{G}_1 \times \cdots \times \mathcal{G}_m$
In numerical analysis, it is customary to use actions of the orthogonal group to perform the change of coordinates for the sake of cost efficiency.

Some Exotic Group Actions (Yet to be studied)
Given a group $G$ and its action on a set $\Delta$, the associated orbit $O(x)$ characterizes the rule by which $x$ is to be changed in $\Delta$. A differential equation on the orbit $O(x)$ is equivalent to a differential equation on the group $G$. Depending on the applications, a path/bridge/highway/differential equation needs to be built on the orbit to connect $x$ to its simplest form. Depending on the group $G$, an orbit is often too "wild" to be really needed for finding the "simplest form" of $x$. To stay in either the orbit or the group, the vector field of the dynamical system must be distributed in the tangent space of the corresponding manifold. Most of the tangent spaces for the matrix groups can be calculated explicitly.

If some kind of objective function has been used to control the connecting bridge, its gradient should be projected to the tangent space.

To stay in either the orbit or the group, the vector field of the dynamical system must be distributed in the tangent space of the corresponding manifold.
Given a matrix group $G$, the tangent space $T_G$ at $A = \exp(tM)$ for all $t \in \mathbb{R}$, where $M$ is a Lie algebra in $\mathbb{R}^n \times \mathbb{R}^n$, i.e.,

$$\{ V = (0) \cup \text{a differentiable curve in } G : \text{is a differentiable curve in } G \} = \mathfrak{g}$$

is a Lie algebra in $\mathbb{R}^n \times \mathbb{R}^n$, i.e.,

$$\mathfrak{g}$$

is a Lie algebra at the identity $I$. The tangent space $T_I$ is given by

$$\mathfrak{g} = \{ V = (0) : \exp(tM) \} = \mathfrak{g}$$

for all $t \in \mathbb{R}$, where $M$ is a Lie algebra in $\mathbb{R}^n \times \mathbb{R}^n$. Given a matrix group $G$, the tangent space at $A \in G$ can be defined as

$$\mathfrak{g} = \{ V = (0) : \exp(tM) \}$$

for all $t \in \mathbb{R}$, where $M$ is a Lie algebra in $\mathbb{R}^n \times \mathbb{R}^n$. The tangent space $T_I$ is critical.
<table>
<thead>
<tr>
<th>$\mathfrak{g} \times \mathfrak{h}$</th>
<th>$\mathfrak{g} \times \mathfrak{h} \supset \mathfrak{g}$</th>
<th>$\mathfrak{g} \times \mathfrak{h}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>${ u \vdash 1. { u } \in \mathfrak{g} } \begin{bmatrix} 0 &amp; 0 \ 1 &amp; \mathfrak{g} \end{bmatrix}$</td>
<td>$(u)_{\text{isom}}$</td>
<td>$(u)_{\text{isom}}$</td>
</tr>
<tr>
<td>${ \alpha \vdash \text{is skew-symmetric} } \begin{bmatrix} \mathfrak{m} { u } \in \mathfrak{g} \end{bmatrix}$</td>
<td>$(u)_{\mathfrak{m}}$</td>
<td>$(u)_{\mathfrak{m}}$</td>
</tr>
<tr>
<td>${ \alpha \vdash 1. { u } \in \mathfrak{n} } \begin{bmatrix} 0 &amp; 0 \ 1 &amp; \mathfrak{n} \end{bmatrix}$</td>
<td>$(u)_{\mathfrak{n}}$</td>
<td>$(u)_{\mathfrak{n}}$</td>
</tr>
<tr>
<td>${ 0 = (\mathfrak{n})_{\text{trace}}({ u } \in \mathfrak{n} } }$</td>
<td>$(u)_{\mathfrak{n}}$</td>
<td>$(u)_{\mathfrak{n}}$</td>
</tr>
<tr>
<td>$u \times \mathfrak{h}$</td>
<td>$(u)_{\mathfrak{h}}$</td>
<td>$(u)_{\mathfrak{h}}$</td>
</tr>
</tbody>
</table>

Characteristics | Algebra | Group
An Illustration of Projection

The tangentspace of $O(n)$ at any orthogonal matrix $Q$ is $T_Q O(n) = QK(n)$ where $K(n) = \mathbb{f}$ all skew-symmetric matrices $g$.

Then normalspace of $O(n)$ at any orthogonal matrix $Q$ is $N_Q O(n) = QS(n)$.

The space $\mathbb{R}^n \otimes \mathbb{R}^n$ is split as $\mathbb{R}^n \otimes \mathbb{R}^n = QS(n) \oplus QK(n)$.

A unique orthogonal splitting of $\mathbb{R}^n \otimes \mathbb{R}^n$ is $X = Q\left(Q^T X - X^T Q\right) = \frac{1}{2}Q\left(Q^T X - X^T Q\right)$.

The projection of $X$ onto the tangentspace $T_Q O(n)$ is given by $\operatorname{Proj}_{T_Q O(n)} X = \frac{1}{2}Q\left(Q^T X - X^T Q\right)$.
A canonical form refers to a specific structure by which a certain conclusion can be drawn or a certain goal can be achieved. The superlatives adjective "simplest" is a relative term which should be interpreted broadly.

- A matrix with a specified algebraic constraint, such as low rank or nonnegativity.
- A matrix with a specified construct, such as Toeplitz, Hamiltonian, stochastic, or other linear varieties.
- A matrix with a specified pattern of zeros, such as a diagonal, tridiagonal, or triangular matrix.
- A matrix with a specified algebraic constraint, such as low rank or nonnegativity.
<table>
<thead>
<tr>
<th>Canonical Forms</th>
<th>Maximal Likelihood</th>
<th>Structured Low Rank Approx.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(u)\mathcal{O} \times \mathcal{S} \times (u)\mathcal{O} \in (\Lambda',\mathcal{S}';\mathcal{N}),$</td>
<td>$V \in \mathbb{R}^{m \times n}$</td>
<td>$\text{Test Matrix Construction}$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td>$\text{and eigenvalues with fixed singular values}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Nonnegative variety}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \Lambda \triangleleft \mathcal{N}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Linear variety } X$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td>$\text{Nonnegative values at fixed locations with fixed entries}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Linear variety } X$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Nonnegative}$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td>$\text{Symmetric Toepolitz}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Symmetric Toepolitz}$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Symmetric Toepolitz}$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Real Schur Decomposition}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Real Schur Decomposition}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Real Schur Decomposition}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Real Schur Decomposition}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Real Schur Decomposition}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Real Schur Decomposition}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Real Schur Decomposition}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Real Schur Decomposition}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Real Schur Decomposition}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Real Schur Decomposition}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Real Schur Decomposition}$</td>
</tr>
<tr>
<td>$(u)\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$\Lambda \mathcal{S} \mathcal{N} \triangleleft \mathcal{N}$</td>
<td></td>
<td>$\text{Upper quasi-triangular } \mathcal{H}$</td>
</tr>
<tr>
<td>$d \in \mathcal{V}_{1-d}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mathcal{O} \in \mathcal{A} \bigtriangleup \mathcal{N}$</td>
<td></td>
<td>$\text{Real Schur Decomposition}$</td>
</tr>
</tbody>
</table>
The orbit of a selected group action only defines the rule by which a transformation is to take place.

How to choose appropriate objective functions?

- Corresponding to each differential equation on the orbit of a group action is a differential equation on the group, and vice versa.
- The vector field of the differential equation must be distributed over the tangent space of the manifold.
- The bridge often assumes the form of a differential equation on the manifold.
- How to choose appropriate objective functions?

Property formulated objective functions helps to control the construction of a bridge between the current point and the desired canonical form on a given orbit.
Objective Functions

Some flows on $O_q \cdot O$ under conjugation $(X)^{(u)} O_q \cdot O$
But still no explicit objective function in sight.

- Enjoys very general convergence behavior.
- Flexible in componentwise scaling.

This is the projected gradient flow of the objective function

\[ \text{Minimize} \quad F(Q) := \frac{1}{2} \| Q^T Q - I \|_F^2. \]

Subject to

\[ Q^T Q = I. \]

\[ \frac{\text{Optimization}}{\text{Bingo!}} \]

The classical Toda lattice does have an objective function in mind.

- Double bracket flow \( \text{Toda lattice (Bloch'90)} \).
- \( X \) is a general symmetric matrix \( \iff \) Double bracket flow \( \text{Toda lattice} \).
- \( X \) is tri-diagonal and symmetric \( \iff \) Double bracket flow \( \text{Toda lattice (Bloch'90)} \).
- Take a special \( N \) = diag \( \{1, 3, \ldots, u - 1, u\} \).

"Flexible in componentwise scaling."

"Enjoy very general convergence behavior."

"But still no explicit objective function in sight."

Double bracket flow \( \text{Doublebracketflow (Brockett'88)} \).

\[ [L, X^T X] = \frac{np}{X^p} \]

\[ \text{Flexible in componentwise scaling.} \]

\[ \text{Doublebracketflow (Brockett'88).} \]
The objective in the design of this flow was to maintain the bidiagonal structure of \( Y(t) \) for all \( t \).

Some flows on \( \mathcal{O}_1 \times \mathcal{O}_q \) under equivalence.
Given a continuous matrix group $G$ and a fixed $X \in \mathbb{R}^{n \times n}$, where $\mathbb{R}^{n \times n}$ is a subset of matrices.

1. Consider the functional $F : G \rightarrow \mathbb{R}$
   
   $F(g) = \frac{1}{2} \| (X \cdot b) \|_2^2 - \| (X \cdot b) \|_2^2 =: (b)_F$

2. Want to minimize $F$ over $G$.

   Flow approach:
   
   - Compute $\frac{\partial F}{\partial g}$.
   - Project $\frac{\partial F}{\partial g}$ onto $T_g G$.
   - Follow the projected gradient until convergence.

3. Flow approach:

   - Follow the projected gradient until convergence.
   - Project $(b)_F$ onto $L$.
   - Compute $(b)_F$.
Some Old Examples

Brockett's double bracket flow (Brockett'88).

Least squares approximation with spectral constraints (Chu&Driessel'90).

Simultaneous reduction problem (Chu'91).

Nearest normal matrix problem (Chu'91).

Brockett's double bracket flow (Brockett'88).
Objective Functions

Objective Functions

- Various structured inverse eigenvalue problems (Chu & Golub'02).
- Inverse generalized eigenvalue problem for symmetric-definite pencils (Chu, 98).
- Inverse generalized eigenvalue problem for symmetric-definite pencils (Chu, 98).

\[((\Lambda)^2 - \Lambda)A + ((X)^2 H - X)X \quad = \quad M\]
\[
\begin{align*}
(\Lambda M + M\Lambda) & \quad = \quad X \\
(\Lambda X + M\Lambda) & \quad = \quad X
\end{align*}
\]

Matrix with prescribed diagonal entries and spectrum (Schaaf-Horn Theorem) (Chu, 98).

Remember the list of applications that Nicoletta gave on Monday!!!
The idea of group actions, least squares, and the corresponding gradient flows can be generalized to other structures such as ➢

Some advantages of using the isometry group over the orthogonal group:

• Using the product topology to describe separate groups and actions might broaden the applications.
• Using the product topology to describe separate groups and actions might broaden the applications.
• Low rank approximation.
• Semigroups.
• Cone of nonnegative matrices.
• The manifold of orthogonal matrices.
• The manifold of orthogonal matrices.

\{ aI = (\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}) | a \in \mathbb{R} \} = (w)O

\{ bI = (\begin{pmatrix} 0 & d \\ 0 & 0 \end{pmatrix}) | b \in \mathbb{R} \} = (b,d)O

The idea of group actions, least squares, and the corresponding gradient flows can be generalized to other structures such as.
Stochastic Inverse Eigenvalue Problem

Would be done if the non-negative inverse eigenvalue problem is solved — a long standing open question.

Construct a stochastic matrix with prescribed spectrum.

Figure 1: By the Karpelevic theorem.

A hard problem (Karpelevic, 1951; Minc, 1988).
Leastsquares formulation:

\[
\text{Minimize } F(g; R) := \frac{1}{2} \| \mathbf{gg}^\top - \mathbf{I} \|_2^2 \quad \text{subject to } \mathbf{g}^\top \mathbf{1} = \mathbf{1} \]

\( J = \text{realmatrix carrying spectral information.} \)
\( \circ \) Hadamard product.

\[ \mathcal{H} \circ (y,b)^{\circ 2} = \frac{\partial p}{\partial p} \]
\[ \mathcal{H} \circ (y,b)^{\circ 1} \circ (1 - y^2 b^2) = \frac{\partial p}{\partial p} \]

Steepest descent flow:

\[ \mathbf{g}_t = \mathbf{g} - \mathbf{J} \mathcal{H} \circ (y,b)^{\circ 1} \circ (1 - y^2 b^2) \cdot \mathbf{g}_t \]

\[ \mathbf{R}_t = \frac{\partial F}{\partial (y,b)} \]

Subject to

\[ \mathbf{g}_t \in \mathbb{R} \quad \mathbf{R}_t \in \mathbb{R} \]

\[ \left\| \mathcal{H} \circ (y,b) - 1 - y^2 b^2 \right\|_1^2 = (y,b)^{\circ 1} \]

\[ \text{subject to } \mathbf{g}^\top \mathbf{1} = \mathbf{1} \]

Least squares formulation:

\[ \mathbf{H} \circ \mathbf{R} - 1 - y^2 b^2 = : (y,b)^{\circ 1} \]
\[ M \mathbf{x} = \frac{\mathbf{p}}{\lambda \mathbf{p}} \]
\[ Z \mathbf{x} = \frac{\mathbf{p}}{\lambda \mathbf{p}} \]
\[ (\mathcal{D})^\text{new} = \frac{\mathbf{p}}{\lambda \mathbf{p}} \]

Define \[ Q := \mathbf{X}^T g \mathbf{Y} \]. Then
\[ d\mathbf{s}/dt = \text{diag}(Q) \]
\[ d\mathbf{x}/dt = \mathbf{XZ} \]
\[ d\mathbf{y}/dt = \mathbf{YW} \]

\[ \lambda_{\mathbf{L}} \mathbf{S} + \mathbf{S} + \mathbf{X}_{\mathbf{L}} \mathbf{X} = \lambda_{\mathbf{L}} \mathbf{X} \]
\[ \mathbf{L}_{\mathbf{S}} \mathbf{X} + \mathbf{L}_{\mathbf{S}} \mathbf{X} + \mathbf{L}_{\mathbf{S}} \mathbf{X} = \mathbf{b} \]
\[ \mathbf{L}(i) \lambda(i) s(i) \mathbf{X} = (i)b \]

\[ \text{ASVD now for } \beta \text{ (Bunse-Cerstner et al., 1991, Wigle et al., 1992)} \]
Nonnegative Matrix Factorization

For various applications, given an nonnegative matrix $A \in \mathbb{R}^{m \times n}$ we want to

\[
\min_{V \in \mathbb{R}^{m \times k}, H \in \mathbb{R}^{k \times n}} \|A - VH\|_F^2
\]

\[
\min_{V \in \mathbb{R}^{m \times k}, H \in \mathbb{R}^{k \times n}} \|A - VH\|_F^2
\]

The first order optimality condition is clear.

Once any entry of either $H$ or $A$ hits $0$, it stays zero. This is a natural barrier.

\[
(HA - V) \circ H = \frac{HP}{HP}
\]

\[
(HA - V) \circ A = \frac{AP}{AP}
\]

Gradient flow:

\[
dV/dt = V - (A - VH)H^T
\]

\[
dH/dt = H - (V^T (A - VH))
\]

Relatively easy by flow approach.

No firm theoretical foundation available yet (Tropp'03).

Data mining — no nonnegative frequencies.

Image processing — no nonnegative pixel values.

Relatively new techniques for dimension reduction applications.

Relatively new techniques for dimension reduction applications.

For various applications, given a nonnegative matrix $A \in \mathbb{R}^{m \times n}$ we want to
Assume images are composite objects in many articulations and poses.

Each column $a_j$ of an non-negative matrix $A$ now represents $m$ pixel values of one image.

The columns $v_k$ of $V$ are $k$ basis elements in $\mathbb{R}^m$.

The columns of $H$, belonging to $\mathbb{R}^k$, can be thought of as coefficient sequences representing the $k$ intrinsic "parts" that make up the object being imaged by multiple observers.

Factorization would enable the identification and classification of intrinsic "parts" that make up the object being imaged by multiple observers.

Assume images are compositible objects in many articulations and poses.
A ∈ R^{19200×10}

Representing 10 Gray-scale 120 × 160 Irides

What are the basis parts of these irises?
Basis Irises with $k = 2$
New Thoughts

Wrong? Basis Irises with $k = 4$
Many operations used to transform matrices can offer a channel to tackle various classical or new and challenging problems. Continuous realization methods often enable to tackle existence problems that are seemingly impossible to be solved by conventional discrete methods.

Continuous realization methods offer a global method for solving the underlying problem. However, they usually offer a global method for solving the underlying problem. A variety of practical and useful algorithms should be used to control the dynamical systems. Various objective functions should be used to control the dynamical systems. Various objective functions should be used to control the dynamical systems.

The notion of "simplicity" varies according to the applications. The notion of "simplicity" varies according to the applications. The notion of "simplicity" varies according to the applications.

It is yet to be determined how a dynamical system should be defined over a group so as to locate the simplest form. It is yet to be determined how a dynamical system should be defined over a group so as to locate the simplest form. It is yet to be determined how a dynamical system should be defined over a group so as to locate the simplest form.

Continuous realization methods are easy and cheap. Continuous realization methods are easy and cheap. Continuous realization methods are easy and cheap.

As a special case of Lie groups, the structure of a matrix group is the same at every of its elements. As a special case of Lie groups, the structure of a matrix group is the same at every of its elements. As a special case of Lie groups, the structure of a matrix group is the same at every of its elements.

More sophisticated actions can be composed that might offer the design of new numerical algorithms. More sophisticated actions can be composed that might offer the design of new numerical algorithms. More sophisticated actions can be composed that might offer the design of new numerical algorithms.

The view unifies different transformations under the same framework of tracing orbits associated with corresponding group actions. The view unifies different transformations under the same framework of tracing orbits associated with corresponding group actions. The view unifies different transformations under the same framework of tracing orbits associated with corresponding group actions.
Some basic ideas and examples have been outlined in this talk.

More sophisticated actions can be composed that might offer the design of new numerical algorithms.

The list of applications continues to grow.

New computational techniques for structured dynamical systems on matrix group will further extend and benefit the scope of this interesting topic.

Need ODE techniques specially tailored for gradient flows.

Need ODE techniques suitable for very large-scale dynamical systems.

Help! Help! Help!