Lecture 5
Dynamical Systems via Group Actions
Linear Transformations and Flows on Manifolds

Moody T. Chu

North Carolina State University

June ***, 2009 @ MCM, CAS

Outline

Outline

Outline

Outline

Outline

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 in Geometric Methods in the Theory of Ordinary Differential Equations, 1988

- What is the simplest form referred to here?
- What kind of continuous change can be employed?

Realization via Transformation

- In linear algebra, transformations are often used to reduce a matrix into its simplest form so as to agilely think and retrieve inherent information.
  - The simplest form usually refers to structures such as diagonal, triangular, or Hessenberg matrices.
  - There are other types of simplest forms such as matrix factorizations.
- The simplest form usually cannot be reached immediately. The intermediate steps form a dynamical system.

Toda Lattice

(Symes'82, Deift el al'83)

$$\begin{aligned} \frac{dX}{dt} &= [X, \Pi_0(X)] \\ X(0) &= X_0. \end{aligned}$$

- Lie algebra decomposition:

$$X = \underbrace{X^- - {X^-}^\top}_{\Pi_0(X)} + \underbrace{X^o + X^+ + {X^-}^\top}_{\Pi_1(X)}.$$

- Sampled at integer times, $\{X(k)\}$ gives the same sequence as does the QR algorithm applied to the matrix $A_0 = exp(X_0)$.

Double Bracket Flow
(Brocket'88, Chu&Driessel'90)

$$\begin{aligned} \frac{dX}{dt} &= [X,[X,N]] \\ X(0) &= \Lambda. \end{aligned}$$

- With $X := Q^T \Lambda Q$, this is a projected gradient flow for

$$\begin{aligned} \text{Minimize } F(Q) &:= \frac{1}{2}||Q^T \Lambda Q - N||^2 \\ \text{Subject to } Q^T Q &= I. \end{aligned}$$

  - Flow $X(t)$ moves in a descent direction to reduce $||X - N||^2$.
- The optimal solution X can be fully characterized by the spectral decomposition of N and is unique (Wiedlandt-Hoffman Theorem).

Equivalence

- (Bloch'90) Suppose X is tridiagonal. Take

$$N = \mathrm{diag}\{n, \ldots, 2, 1\},$$

then

$$[X, N] = \Pi_0(X).$$

- A gradient flow hence becomes a Hamiltonian flow.

Matrix Groups and Actions

Lots of transformations used in numerical linear algebra are the results of group actions.

- What groups can be used?
- What actions can be taken?
- What results can be expected?

Matrix Groups

- A subset of nonsingular matrices (over any field) closed under matrix multiplication and inversion is called a *matrix group*.
  - Matrix groups are central in many parts of mathematics and applications.

Examples of Matrix Groups

Group	Subgroup	Notation	Characteristics
General linear		$\mathcal{G}l(n)$	$\{A \in \mathbb{R}^{n \times n} \mid \det(A) \neq 0\}$
	Special linear	$\mathcal{S}l(n)$	$\{A \in \mathcal{G}l(n) \mid \det(A) = 1\}$
Upper triangular		$\mathcal{U}(n)$	$\{A \in \mathcal{G}l(n) \mid A \text{ is upper triangular}\}$
	Unipotent	$\mathcal{U}nip(n)$	$\{A \in \mathcal{U}(n) \mid a_{ii} = 1 \text{ for all } i\}$
Orthogonal		$\mathcal{O}(n)$	$\{Q \in \mathcal{G}l(n) \mid Q^\top Q = I\}$
Generalized orthogonal		$\mathcal{O}_S(n)$	$\{Q \in \mathcal{G}l(n) \mid Q^\top SQ = S\}$; S is a fixed matrix
	Symplectic	$\mathcal{S}p(2n)$	$\mathcal{O}_J(2n)$; $J := \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$
	Lorentz	$\mathcal{L}or(n, k)$	$\mathcal{O}_L(n+k)$; $L := \text{diag}\{\underbrace{1, \ldots, 1}_{n}, \underbrace{-1, \ldots -1}_{k}\}$
Affine		$\mathcal{A}ff(n)$	$\left\{ \begin{bmatrix} A & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \mid A \in \mathcal{G}l(n), \mathbf{t} \in \mathbb{R}^n \right\}$
	Translation	$\mathcal{T}rans(n)$	$\left\{ \begin{bmatrix} I & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \mid \mathbf{t} \in \mathbb{R}^n \right\}$
	Isometry	$\mathcal{I}som(n)$	$\left\{ \begin{bmatrix} Q & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \mid Q \in \mathcal{O}(n), \mathbf{t} \in \mathbb{R}^n \right\}$
Center of G		$Z(G)$	$\{z \in G \mid zg = gz, \text{ for every } g \in G\}$, G is a given group
Product of G_1 and G_2		$G_1 \times G_2$	$(g_1, g_2) * (h_1, h_2) := (g_1 h_1, g_2 h_2)$

Lie Groups

- A smooth manifold which is also a group where the multiplication and the inversion are smooth maps is called a *Lie group*.
    - The most remarkable feature of a Lie group is that the structure is the same in the neighborhood of each of its elements.
- (Howe'83) Every (non-discrete) matrix group is in fact a Lie group.
    - Algebra and geometry are intertwined in the study of matrix groups.
- The Hessenberg group, $\mathcal{H}ess(n) := \mathcal{U}nip(n)/\mathcal{Z}_n$, is a Lie group, but not a matrix group.

Group Actions

- A function $\mu : G \times \mathbb{V} \longrightarrow \mathbb{V}$ is said to be a *group action* of G on a set $\mathbb{V}$ if and only if
  - $\mu(gh, \mathbf{x}) = \mu(g, \mu(h, \mathbf{x}))$ for all $g, h \in G$ and $\mathbf{x} \in \mathbb{V}$.
  - $\mu(e, \mathbf{x}) = \mathbf{x}$, if e is the identity element in G.
- Given $\mathbf{x} \in \mathbb{V}$, two important notions associated with a group action μ:
  - The *stabilizer* of $\mathbf{x}$ is

$$Stab_G(\mathbf{x}) := \{g \in G | \mu(g, \mathbf{x}) = \mathbf{x}\}.$$

  - The *orbit* of $\mathbf{x}$ is

$$Orb_G(\mathbf{x}) := \{\mu(g, \mathbf{x}) | g \in G\}.$$

Examples of Group Actions

Set $\mathbb{V}$	Group G	Action $\mu(g, A)$	Application
$\mathbb{R}^{n\times n}$	Any subgroup	$g^{-1}Ag$	conjugation
$\mathbb{R}^{n\times n}$	$\mathcal{O}(n)$	$g^{\top}Ag$	orthogonal similarity
$\underbrace{\mathbb{R}^{n\times n} \times \ldots \times \mathbb{R}^{n\times n}}_{k}$	Any subgroup	$(g^{-1}A_1 g, \ldots, g^{-1}A_k g)$	simultaneous reduction
$\mathbb{S}(n) \times \mathbb{S}_{PD}(n)$	Any subgroup	$(g^{\top}Ag, g^{\top}Bg)$	symm. positive definite pencil reduction
$\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$	$\mathcal{O}(n) \times \mathcal{O}(n)$	$(g_1^{\top}Ag_2, g_1^{\top}Bg_2)$	*QZ* decomposition
$\mathbb{R}^{m\times n}$	$\mathcal{O}(m) \times \mathcal{O}(n)$	$g_1^{\top}Ag_2$	singular value decomp.
$\mathbb{R}^{m\times n} \times \mathbb{R}^{p\times n}$	$\mathcal{O}(m) \times \mathcal{O}(p) \times \mathcal{G}l(n)$	$(g_1^{\top}Ag_3, g_2^{\top}Bg_3)$	generalized singular value decomp.

Some Exotic Group Actions (yet to be studied!)

- 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 and numerical stability.
  - What could be said if actions of the isometry group are used?
    - Being isometric, stability is guaranteed.
    - The inverse of an isometry matrix is easy.

$$\begin{bmatrix} Q & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}^{-1} = \begin{bmatrix} Q^\top & -Q^\top \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}.$$

    - The isometry group is larger than the orthogonal group.

Actions with Shift or Scaling

- What could be said if actions of the orthogonal group plus shift are used?

$$\mu((Q, s), A) := Q^\top AQ + sI, \quad Q \in \mathcal{O}(n), s \in \mathbb{R}_+.$$

- What could be said if action of the orthogonal group with scaling are used?

$$\mu((Q, s), A) := sQ^\top AQ, \quad Q \in \mathcal{O}(n), s \in \mathbb{R}_\times,$$

or

$$\mu((Q, \mathbf{s}, \mathbf{t}), A) := \mathrm{diag}\{\mathbf{s}\}Q^\top AQ\mathrm{diag}\{\mathbf{t}\}, \quad Q \in \mathcal{O}, \mathbf{s}, \mathbf{t} \in \mathbb{R}^n_\times.$$

Using the Group Actions

- Given a group G and its action μ on a set $\mathbb{V}$, the associated orbit $Orb_G(\mathbf{x})$ characterizes the rule by which $\mathbf{x}$ is to be changed in $\mathbb{V}$.
  - Merely an orbit is often too "wild" to be readily traced for finding the "simplest form" of $\mathbf{x}$.
- Depending on the applications, a path or differential equation needs to be built on the orbit to connect $\mathbf{x}$ to its simplest form.

Embedded Environment

- A differential equation on the orbit $Orb_G(\mathbf{x})$.
  - Lax dynamics on $X(t)$.
  - Usually difficult to maintain the innate properties.
- A differential equation on the group G.
  - Parameter dynamics on $g_1(t)$ or $g_2(t)$.
  - Can take advantage of the group structure.

Following the Orbits

- 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.

Tangent Space in General

- Given a matrix group $G \leq \mathcal{G}l(n)$, the *tangent space* to G at $A \in G$ can be defined as

$$\mathcal{T}_A G := \{\gamma'(0) | \gamma \text{ is a differentiable curve in } G \text{ with } \gamma(0) = A\}.$$

- The tangent space $\mathfrak{g} = \mathcal{T}_I G$ at the identity I is critical.
  - $\mathfrak{g}$ is a Lie subalgebra in $\mathbb{R}^{n \times n}$, i.e.,

$$\text{If } \alpha'(0), \beta'(0) \in \mathfrak{g}, \text{ then } [\alpha'(0), \beta'(0)] \in \mathfrak{g}$$

  - The tangent space of a matrix group has the same structure everywhere, i.e.,

$$\mathcal{T}_A G = A\mathfrak{g}.$$

  - $\mathcal{T}_I G$ can be characterized as the *logarithm* of G, i.e.,

$$\mathfrak{g} = \{M \in \mathbb{R}^{n \times n} | \exp(tM) \in G, \text{ for all } t \in \mathbb{R}\}.$$

Examples of Tangent Spaces

Group G	Algebra $\mathfrak{g}$	Characteristics
$\mathcal{G}l(n)$	$gl(n)$	$\mathbb{R}^{n \times n}$
$\mathcal{S}l(n)$	$sl(n)$	$\{M \in gl(n) \mid \mathrm{trace}(M) = 0\}$
$\mathcal{A}ff(n)$	$aff(n)$	$\{\begin{bmatrix} M & \mathbf{t} \\ \mathbf{0} & 0 \end{bmatrix} \mid M \in gl(n), \mathbf{t} \in \mathbb{R}^n\}$
$\mathcal{O}(n)$	$o(n)$	$\{K \in gl(n) \mid \text{K is skew-symmetric}\}$
$\mathcal{I}som(n)$	$isom(n)$	$\{\begin{bmatrix} K & \mathbf{t} \\ \mathbf{0} & 0 \end{bmatrix} \mid K \in o(n), \mathbf{t} \in \mathbb{R}^n\}$
$G_1 \times G_2$	$\mathcal{T}_{(e_1, e_2)} G_1 \times G_2$	$\mathfrak{g}_1 \times \mathfrak{g}_2$

An Illustration of Projection

- The tangent space of $\mathcal{O}(n)$ at any orthogonal matrix Q is

$$\mathcal{T}_Q\mathcal{O}(n) = Q\mathbb{K}(n)$$

where

$$\mathbb{K}(n) = \{\text{All skew-symmetric matrices}\}.$$

- The normal space of $\mathcal{O}(n)$ at any orthogonal matrix Q is

$$\mathcal{N}_Q\mathcal{O}(n) = Q\mathbb{S}(n).$$

- The space $\mathbb{R}^{n\times n}$ is split as

$$\mathbb{R}^{n\times n} = Q\mathbb{S}(n) \oplus Q\mathbb{K}(n).$$

- A unique orthogonal splitting of $X \in \mathbb{R}^{n \times n}$:

$$X = Q(Q^T X) = Q\left\{\frac{1}{2}(Q^T X - X^T Q)\right\} + Q\{\frac{1}{2}(Q^T X + X^T Q)\right\}.$$

- The projection of X onto the tangent space $\mathcal{T}_Q\mathcal{O}(n)$ is given by

$$\mathrm{Proj}_{\mathcal{T}_Q\mathcal{O}(n)} X = Q\left\{\frac{1}{2}(Q^T X - X^T Q)\right\}.$$

Canonical Forms

- A canonical form refers to a “specific structure" by which a certain conclusion can be drawn or a certain goal can be achieved.
- The superlative adjective “simplest" is a relative term which should be interpreted broadly.
  - A matrix with a specified pattern of zeros, such as a diagonal, tridiagonal, or triangular matrix.
  - A matrix with a specified construct, such Toeplitz, Hamiltonian, stochastic, or other linear varieties.
  - A matrix with a specified algebraic constraint, such as low rank or nonnegativity.

Examples of Canonical Forms

Canonical form	Also know as	Action
Bidiagonal J	Quasi-Jordan Decomp., $A \in \mathbb{R}^{n\times n}$	$P^{-1}AP = J$, $P \in \mathcal{G}l(n)$
Diagonal Σ	Sing. Value Decomp., $A \in \mathbb{R}^{m\times n}$	$U^{\top}AV = \Sigma$, $(U, V) \in \mathcal{O}(m) \times \mathcal{O}(n)$
Diagonal pair (Σ_1, Σ_2)	Gen. Sing. Value Decomp., $(A, B) \in \mathbb{R}^{m\times n} \times \mathbb{R}^{p\times n}$	$(U^{\top}AX, V^{\top}BX) = (\Sigma_1, \Sigma_2)$, $(U, V, X) \in \mathcal{O}(m) \times \mathcal{O}(p) \times \mathcal{G}l(n)$
Upper quasi-triangular H	Real Schur Decomp., $A \in \mathbb{R}^{n\times n}$	$Q^{\top}AQ = H$, $Q \in \mathcal{O}(n)$
Upper quasi-triangular H Upper triangular U	Gen. Real Schur Decomp., $A, B \in \mathbf{R}^{n\times n}$	$(Q^{\top}AZ, Q^{\top}BZ) = (H, U)$, $Q, Z \in \mathcal{O}(n)$
Symmetric Toeplitz T	Toeplitz Inv. Eigenv. Prob., $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{R}$ is given	$Q^{\top}\mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}Q = T$, $Q \in \mathcal{O}(n)$
Nonnegative $N \geq 0$	Nonneg. inv. Eigenv. Prob., $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$ is given	$P^{-1}\mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}P = N$, $P \in \mathcal{G}l(n)$
Linear variety X with fixed entries at fixed locations	Matrix Completion Prob., $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$ is given $X_{i_\nu, j_\nu} = a_\nu, \nu = 1, \ldots, \ell$	$P^{-1}\{\lambda_1, \ldots, \lambda_n\}P = X$, $P \in \mathcal{G}l(n)$
Nonlinear variety with fixed singular values and eigenvalues	Test Matrix Construction, $\Lambda = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$ and $\Sigma = \mathrm{diag}\{\sigma_1, \ldots \sigma_n\}$ are given	$P^{-1}\Lambda P = U^{\top}\Sigma V$ $P \in \mathcal{G}l(n)$, $U, V \in \mathcal{O}(n)$
Maximal fidelity	Structured Low Rank Approx. $A \in \mathbb{R}^{m\times n}$	$\left(\mathrm{diag}\left(USS^{\top}U^{\top}\right)\right)^{-1/2}USV^{\top}$, $(U, S, V) \in \mathcal{O}(m) \times \mathbb{R}^{k}_{\times} \times \mathcal{O}(n)$

Objective Functions

- The orbit of a selected group action only defines the rule by which a transformation is to take place.
- Properly formulated objective functions helps to control the construction of a bridge between the current point and the desired canonical form on a given orbit.
  - The bridge often assumes the form of a differential equation on the manifold.
  - The vector field of the differential equation must distributed over the tangent space of the manifold.
  - Corresponding to each differential equation on the orbit of a group action is a differential equation on the group, and vice versa.
- How to choose appropriate objective functions?

Flows on $Orb_{\mathcal{O}(n)}(X)$ under Conjugation

- Toda lattice arises from a special mass-spring system (Symes'82, Deift el al'83),

$$\begin{aligned} \frac{dX}{dt} &= [X, \Pi_0(X)], \quad \Pi_0(X) = X^- - X^{-\top}, \\ X(0) &= \text{tridiagonal and symmetric.} \end{aligned}$$

  - No specific objective function is used, but physics laws govern the definition of the vector field.

- Generalization to general matrices is totally by brutal force and blindness (*and by the then young and desperate researchers*) (Chu'84, Watkins'84).

$$\frac{dX}{dt} = [X, \Pi_0(G(X))], \quad G(z) \text{ is analytic over spectrum of } X(0).$$

  - But nicely explains the pseudo-convergence and convergence behavior of the classical QR algorithm for general and normal matrices, respectively.
  - Sorting of eigenvalues at the limit point is observed, but not quite clearly understood.

- Double bracket flow (Brockett'88),

$$\frac{dX}{dt} = [X, [X, N]], \qquad N = \text{fixed and symmetric.}$$

  - This is the projected gradient flow of the objective function

$$\begin{aligned} \text{Minimize } F(Q) &:= \frac{1}{2}||Q^T \Lambda Q - N||^2, \\ \text{Subject to } Q^T Q &= I. \end{aligned}$$

    - Sorting is necessary in the first order optimality condition (Wielandt&Hoffman'53).

- Take a special $N = \mathrm{diag}\{n, n-1, \ldots, 2, 1\}$,
  - X is tridiagonal and symmetric $\Longrightarrow$ Double bracket flow $\equiv$ Toda lattice (Bloch'90).
    - Bingo! The classical Toda lattice does have an objective function in mind.
  - X is a general symmetric matrix $\Longrightarrow$ Double bracket = A specially scaled Toda lattice.

- Scaled Toda lattice (Chu'95),

$$\frac{dX}{dt} = [X, K \circ X], \quad K = \text{fixed and skew-symmetric.}$$

  - Flexible in componentwise scaling.
  - Enjoy very general convergence behavior.
  - But still no explicit objective function in sight.

Flows on $Orb_{\mathcal{O}(m)\times\mathcal{O}(n)}(X)$ under Equivalence

- Any flow on the orbit $Orb_{\mathcal{O}(m)\times\mathcal{O}(n)}(X)$ under equivalence must be of the form

$$\frac{dX}{dt} = X(t)h(t) - k(t)X(t), \quad h(t) \in \mathbb{K}(n), \quad k(t) \in \mathbb{K}(m).$$

- *QZ* flow (Chu'86),

$$\frac{dX_1}{dt} = X_1\Pi_0(X_2^{-1}X_1) - \Pi_0(X_1X_2^{-1})X_1,$$
$$\frac{dX_2}{dt} = X_2\Pi_0(X_2^{-1}X_1) - \Pi_0(X_1X_2^{-1})X_2,.$$

- *SVD* flow (Chu'86),

$$\frac{dY}{dt} = Y\Pi_0\left(Y(t)^\top Y(t)\right) - \Pi_0\left(Y(t)Y(t)^\top\right)Y,$$
$$Y(0) = \text{bidiagonal.}$$

  - The "objective" in the design of this flow was to maintain the bidiagonal structure of $Y(t)$ for all t.
  - The flow gives rise to the Toda flows for $Y^\top Y$ and $YY^\top$.
  - Votka-Volterra equation (Nakamura'01).

Projected Gradient Flows

- Given
  - A continuous matrix group $G \subset \mathcal{G}l(n)$.
  - A fixed $X \in \mathbb{V}$ where $\mathbb{V} \subset \mathbb{R}^{n \times n}$ be a subset of matrices.
  - A differentiable map $f : \mathbb{V} \longrightarrow \mathbb{R}^{n \times n}$ with a certain “inherent" properties, e.g., symmetry, isospectrum, low rank, or other algebraic constraints.
  - A group action $\mu : G \times \mathbb{V} \longrightarrow \mathbb{V}$.
  - A projection map P from $\mathbb{R}^{n \times n}$ onto a singleton, a linear subspace, or an affine subspace $\mathbb{P} \subset \mathbb{R}^{n \times n}$ where matrices in $\mathbb{R}$ carry a certain desired structure, e.g., the canonical form.
- Minimize the functional $F : G \longrightarrow \mathbb{R}$

$$F(g) := \frac{1}{2}\|f(\mu(g, X)) - P(\mu(g, X))\|_F^2.$$

Flow Approach

- Compute $\nabla F(g)$.
- Project $\nabla F(g)$ onto $\mathcal{T}_g G$.
- Follow the projected gradient until convergence.

Some Old Examples

- Brockett's double bracket flow (Brockett'88).
- Least squares approximation with spectral constraints (Chu&Driessel'90, Nakamura'92-98).

$$\frac{dX}{dt} = [X, [X, P(X)]].$$

- Simultaneous reduction problem (Chu'91),

$$\begin{aligned} \frac{dX_i}{dt} &= \left[X_i, \sum_{j=1}^{p} \frac{[X_j, P_j^T(X_j)] - [X_j, P_j^T(X_j)]^T}{2} \right] \\ X_i(0) &= A_i \end{aligned}$$

- Nearest normal matrix problem (Chu'91),

$$\begin{aligned} \frac{dW}{dt} &= \left[W, \frac{1}{2}\{[W, diag(W^*)] - [W, diag(W^*)]^*\} \right] \\ W(0) &= A. \end{aligned}$$

- Matrix with prescribed diagonal entries and spectrum (Schur-Horn Theorem) (Chu'95),

$$\dot{X} = [X, [\mathrm{diag}(X) - \mathrm{diag}(a), X]]$$

- Inverse generalized eigenvalue problem for symmetric-definite pencil (Chu&Guo'98).

$$
\begin{aligned}
\dot{X} &= -\left((XW)^T + XW\right), \\
\dot{Y} &= -\left((YW)^T + YW\right), \\
W &:= X(X - P_1(X)) + Y(Y - P_2(Y)).
\end{aligned}
$$

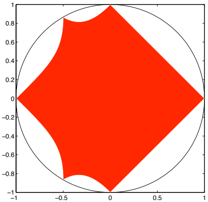
- Various structured inverse eigenvalue problems (Chu&Golub'02).

New Thoughts

- The idea of group actions, least squares, and the corresponding gradient flows can be generalized to other structures such as
  - Stiefel manifold $\mathcal{O}(p,q) := \{Q \in \mathbb{R}^{p\times q} | Q^T Q = I_q\}$.
  - The manifold of oblique matrices $\mathcal{OB}(n) := \{Q \in \mathbb{R}^{n\times n} | \text{diag}(Q^\top Q) = I_n\}$.
  - Cone of nonnegative matrices.
  - Semigroups.
  - Low rank approximation.
- Using the product topology to describe separate groups and actions might broaden the applications.
- Any advantages of using the isometry group over the orthogonal group?

Stochastic Inverse Eigenvalue Problem

- Construct a stochastic matrix with prescribed spectrum
  - A hard problem (Karpelevic'51, Minc'88).



  - Would be done if the nonnegative inverse eigenvalue problem is solved – a long standing open question.

Least Squares Formulation

$$\begin{array}{ll} \text{Minimize} & F(g,R) := \frac{1}{2}||gJg^{-1} - R \circ R||^2 \\ \text{Subject to} & g \in Gl(n),\ R \in gl(n). \end{array}$$

- J = Real matrix carrying spectral information.
- $\circ$ = Hadamard product.

Steepest Descent Flow

$$\frac{dg}{dt} = [(gJg^{-1})^T, \alpha(g,R)]g^{-T}$$
$$\frac{dR}{dt} = 2\alpha(g,R) \circ R.$$

- $\alpha(g,R) := gJg^{-1} - R \circ R.$

ASVD Flow for g

(Bunse-Gerstner et al'91, Wright'92)

$$
\begin{aligned}
g(t) &= X(t)S(t)Y(t)^T \\
\dot{g} &= \dot{X}SY^T + X\dot{S}Y^T + XS\dot{Y}^T \\
X^T\dot{g}Y &= \underbrace{X^T\dot{X}}_{Z}S + \dot{S} + S\underbrace{\dot{Y}^TY}_{W}
\end{aligned}
$$

Define $Q := X^T\dot{g}Y$. Then

$$
\begin{aligned}
\frac{dS}{dt} &= \operatorname{diag}(Q). \\
\frac{dX}{dt} &= XZ. \\
\frac{dY}{dt} &= YW.
\end{aligned}
$$

Nonnegative Matrix Factorization

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

$$\min_{0 \leq V \in \mathbb{R}^{m \times k}, 0 \leq H \in \mathbb{R}^{k \times n}} \frac{1}{2} \|A - VH\|_F^2.$$

  - Relatively new techniques for dimension reduction applications.
    - Image processing — no negative pixel values.
    - Data mining — no negative frequencies.
  - No firm theoretical foundation available yet (Tropp'03).

- Relatively easy by flow approach!

$$\min_{E\in\mathbb{R}^{m\times k},F\in\mathbb{R}^{k\times n}} \frac{1}{2}\|A-(E\circ E)(F\circ F)\|_F^2.$$

- Gradient flow:

$$\begin{aligned}\frac{dV}{dt} &= V\circ(A-VH)H^\top),\\ \frac{dH}{dt} &= H\circ(V^\top(A-VH)).\end{aligned}$$

  - Once any entry of either V or H hits 0, it stays zero. This is a natural barrier!
  - The first order optimality condition is clear.

Conclusion

- Many operations used to transform matrices can be considered as matrix group actions.
- The view unifies different transformations under the same framework of tracing orbits associated with corresponding group actions.
  - More sophisticated actions can be composed that might offer the design of new numerical algorithms.
  - As a special case of Lie groups, tangent space structure of a matrix group is the same at every of its element. Computation is easy and cheap.

- It is yet to be determined how a dynamical system should be defined over a group so as to locate the simplest form.
  - The notion of “simplicity" varies according to the applications.
  - Various objective functions should be used to control the dynamical systems.
  - Usually offers a global method for solving the underlying problem.
- Group actions together with properly formulated objective functions can offer a channel to tackle various classical or new and challenging problems, even those where conventional discrete methods seems to be impossible.
- New computational techniques for structured dynamical systems on matrix group will further extend and benefit the scope of this interesting topic.