Evolution of Lax Dynamics and Its Applications

Moody T. Chu

North Carolina State University

presented at

LMS Durham Symposium
July 13-23, 2000

Outline

- Introduction:
  - ◇ Motivation
  - ◇ Basic Form
- General Theory:
  - ◇ QR-type Framework
  - ◇ Gradient-type Framework
- Application:
  - ◇ Isospectral flows
  - ◇ Projected gradient flows
  - ◇ Generalized flows
- Generalization:
  - ◇ Inverse stochastic eigenvalue problem
- Computation:
- Conclusion:

The Eigenvalue Problem

- The mathematical problem:
  - ◇ A symmetric matrix A_0 is given.
  - ◇ Solve the equation
    $$A_0 x = \lambda x$$
    for a nonzero vector x and a scalar λ.
- An iterative method :
  - ◇ The QR decomposition:
    $$A = QR$$
    where Q is orthogonal and R is upper triangular.
  - ◇ The QR algorithm (Francis'61):
    $$\begin{aligned} A_k &= Q_k R_k \\ A_{k+1} &= R_k Q_k. \end{aligned}$$
  - ◇ The sequence $\{A_k\}$ converges to a diagonal matrix.
  - ◇ Every matrix A_k has the same eigenvalues of A_0.

- A continuous method:
  - ◇ Lie algebra decomposition:

$$X = X^o + X^+ + X^-$$

where X^o is the diagonal, X^+ the strictly upper triangular, and X^- the strictly lower triangular part of X.

  - ◇ The Toda lattice (Symes'82, Deift el al'83):

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

  - ◇ Sampled at integer times, $\{X(k)\}$ gives the same sequence as does the QR algorithm applied to the matrix $A_0 = exp(X_0)$.
- Evolution from X_0 to the limit point of Toda flow, which is a diagoal matrix, maintains isospectrum.
  - ◇ What motivates the construction of the Toda lattice?
  - ◇ Why is convergence guaranteed?

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:

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

  - ◇ Equality Constrained Optimization:
    - ▷ Augmented Lagrangian methods.
    - ▷ Sequential quadratic programming methods.
  - ◇ None of these techniques is easy.

- A continuous approach:
  - ◇ The projection of the gradient of F can easily be calculated.
  - ◇ Projected gradient flow (Chu&Driessel'90):

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

    - ▷ $X := Q^T \Lambda Q$.
    - ▷ 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.
- Evolution from a starting point to the limit point, which solves the least squares problem, is built on the basis of systematically reducing the difference between the current position and the target position.

Basic Form

- Lax dynamics:

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

- Parameter dynamics:

$$\begin{aligned} \frac{dg(t)}{dt} &= g(t)k(t) \\ g(0) &= I. \end{aligned}$$

- Isospectral relationship:

$$X(t) = g(t)^{-1} X_0 g(t).$$

- Some choices of $k(t)$:

$$\begin{aligned} k(t) &= X(t)^{-} - X(t)^{-T} \\ k(t) &= [X(t), N] \\ k(t) &= k(X(t)), \text{ where } k \text{ is } \ldots \end{aligned}$$

Notation

$$
\begin{aligned}
Gl(n) &:= \{n \times n \text{ real nonsingular matrices}\} \\
gl(n) &:= \{n \times n \text{ real matrices}\} \\
X_0 &:= \text{A given matrix in } gl(n) \\
M(X_0) &:= \{g^{-1} X_0 g \mid g \in Gl(n)\} \\
[A, B] &:= AB - BA \text{ (Lie bracket)} \\
T &:= \text{Subspace of } gl(n) \\
P_T &:= \text{Projection mapping from } gl(n) \text{ to } T
\end{aligned}
$$

QR-type Framework

- Subspace splitting of $gl(n)$:

$$gl(n) = T_1 + T_2.$$

  - $\diamond$ T_1 and T_2 are subspaces of $gl(n)$.
  - $\diamond$ This is a subspace decomposition only, not necessarily a subalgebra decomposition of $gl(n)$.
  - $\diamond$ Given T_1, one may choose $T_2 = gl(n) - T_1$. This is not necessarily a direct sum decomposition.

- Examples:
  - $\diamond$ Toda flow:
    - $\triangleright$ T_1 = Subspace of skew symmetric matrices,

$$k(X) := (X^-) - (X^-)^T.$$

  - $\diamond$ General flow:
    - $\triangleright$ T_1 = Arbitrary linear subspace,

$$k(X) := \text{Projection of } X \text{ onto subspace } T_1.$$

    - $\triangleright$ Time-1 mapping of the solution still enjoys a QR-type algorithm.

Dynamical Systems

- Lax dynamics:

$$\begin{aligned} \frac{dX(t)}{dt} &:= [X(t), P_1(X(t))] \\ X(0) &:= X_0. \end{aligned}$$

  ◇ $P_1 :=$ Projection onto T_1.

- Parameter dynamics:

$$\begin{aligned} \frac{dg_1(t)}{dt} &:= g_1(t)P_1(X(t)) \\ g_1(0) &:= I. \end{aligned}$$

  and

$$\begin{aligned} \frac{dg_2(t)}{dt} &:= P_2(X(t))g_2(t) \\ g_2(0) &:= I. \end{aligned}$$

  ◇ $P_2 :=$ Projection onto T_2.

Similarity Property

$$X(t) = g_1(t)^{-1} X_0 g_1(t) = g_2(t) X_0 g_2(t)^{-1}.$$

- Define $Z(t) = g_1(t) X(t) g_1(t)^{-1}$.
- Check

$$\begin{aligned} \frac{dZ}{dt} &= \frac{dg_1}{dt} X g_1^{-1} + g_1 \frac{dX}{dt} g_1^{-1} + g_1 X \frac{dg_1^{-1}}{dt} \\ &= (g_1 P_1(X)) X g_1^{-1} \\ &\quad + g_1 (X P_1(X) - P_1(X) X) g_1^{-1} \\ &\quad + g_1 X (-P_1(X) g_1^{-1}) \\ &= 0. \end{aligned}$$

- Thus $Z(t) = Z(0) = X(0) = X_0$.

Decomposition Property

$$exp(tX_0) = g_1(t)g_2(t).$$

- Trivially $exp(X_0 t)$ satisfies the IVP

$$\frac{dY}{dt} = X_0 Y, Y(0) = I.$$

- Define $Z(t) = g_1(t)g_2(t)$.
- Then $Z(0) = I$ and

$$\begin{aligned}\frac{dZ}{dt} &= \frac{dg_1}{dt}g_2 + g_1\frac{dg_2}{dt} \\ &= (g_1 P_1(X))g_2 + g_1(P_2(X)g_2) \\ &= g_1 X g_2 \\ &= X_0 Z \quad \text{(by Similarity Property).}\end{aligned}$$

- By the uniqueness theorem in the theory of ordinary differential equations, $Z(t) = exp(X_0 t)$.

Reverse Property

$$exp(tX(t)) = g_2(t)g_1(t).$$

- By Decomposition Property,

$$\begin{aligned} g_2(t)g_1(t) &= g_1(t)^{-1}exp(X_0t)g_1(t) \\ &= exp(g_1(t)^{-1}X_0g_1(t)t) \\ &= exp(X(t)t). \end{aligned}$$

Abstraction

- QR-type Decomposition:
  - ◇ Lie algebra decomposition of $gl(n) \Longleftrightarrow$ Lie group decomposition of $Gl(n)$ in the neighborhood of I.
  - ◇ Arbitrary subspace decomposition $gl(n) \Longleftrightarrow$ Product of two nonsingular matrices in the neighborhood of I, i.e.,
$$exp(X_0 t) = g_1(t)g_2(t).$$
  - ◇ The product $g_1(t)g_2(t)$ will be called the *abstract g_1g_2 decomposition* of $exp(X_0t)$.
- QR-type Algorithm:
  - ◇ By setting $t = 1$, we have
$$\begin{aligned} exp(X(0)) &= g_1(1)g_2(1) \\ exp(X(1)) &= g_2(1)g_1(1). \end{aligned}$$
  - ◇ The dynamical system for $X(t)$ is autonomous $\Longrightarrow$ The above phenomenon will occur at every feasible integer time.
  - ◇ 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*.

Relation to Classical Algorithms

	Case 1	Case 2	Case 3
T_1	$o(n)$	$l(n)$	$l(n) + d(n)/2$
T_2	$r(n)+d(n)$	$r(n)+d(n)$	$r(n)+d(n)/2$
$k(t)=P_1(X(t))$	$X^- - X^{-T}$	X^-	$X^- + X^0/2$
$P_2(X(t))$	$X^+ + X^0 + X^{-T}$	$X^+ + X^0$	$X^+ + X^0/2$
$g_1(t)$	$\mathrm{Q}(t) \in O(n)$	$\mathrm{L}(t) \in L(n)$	$\mathrm{G}(t) \in L(n)$
$g_2(t)$	$\mathrm{R}(t) \in R(n)$	$\mathrm{U}(t) \in R(n)$	$\mathrm{H}(t) \in R(n)$
Algorithm	QR	LU	Cholesky

$$
\begin{aligned}
o(n) &:= \{\text{Skew-symmetric matrices in } gl(n)\} \\
O(n) &:= \{\text{Orthogonal matrices in } Gl(n)\} \\
r(n) &:= \{\text{Strictly upper triangular matrices in } gl(n)\} \\
R(n) &:= \{\text{Upper triangular matrices in } Gl(n)\} \\
l(n) &:= \{\text{Strictly lower triangular matrices in } gl(n)\} \\
L(n) &:= \{\text{Lower triangular matrices in } Gl(n)\} \\
d(n) &:= \{\text{Diagonal matrices in } Gl(n)\}
\end{aligned}
$$

$$
\begin{aligned}
X^+ &:= \text{The strictly upper triangular matrix of } X \\
X^o &:= \text{The diagonal matrix of } X \\
X^- &:= \text{The strictly lower triangular matrix of } X
\end{aligned}
$$

Nonclassical Examples

- Assume:

$$
\begin{aligned}
X_0 &:= \text{symmetric} \\
\Delta &:= \text{Active index subset} \\
\hat{X}(t) &:= \text{Portion of } X(t) \text{ conforming to } \Delta \\
P_1(X(t)) &:= \hat{X}(t) - (\hat{X}(t))^T \\
P_2(X(t)) &:= X(t) - P_1(X(t))
\end{aligned}
$$

Then:

$$\text{For all } (i,j) \in \Delta,\ x_{ij}(t) \longrightarrow 0 \text{ as } t \longrightarrow \infty.$$

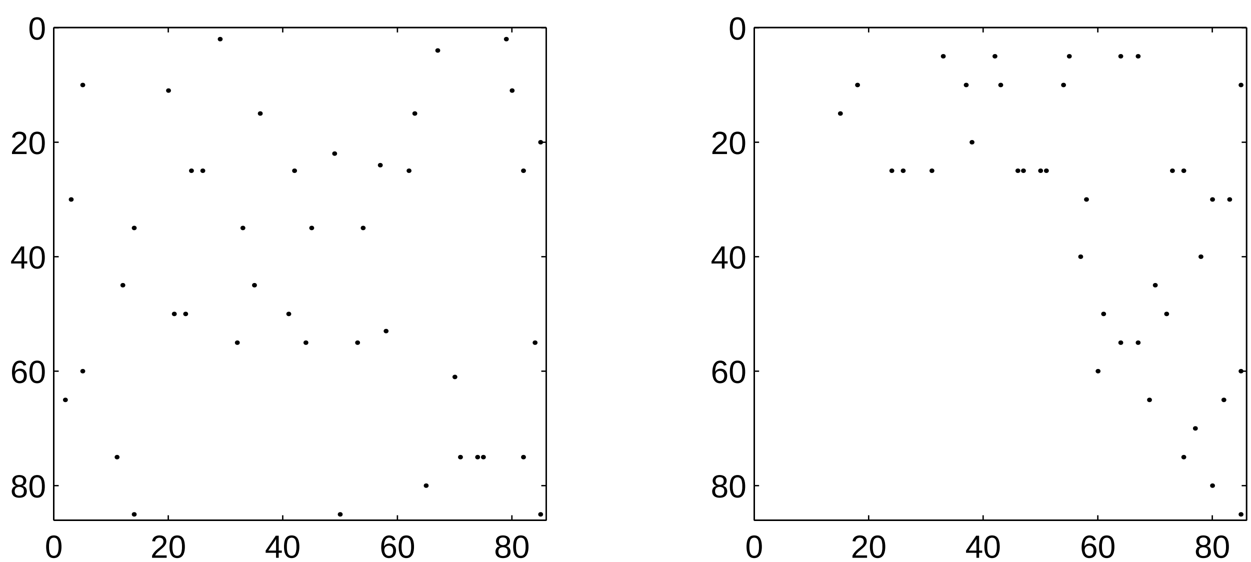
◇ The above result suggests a way to produce (or knock out) any prescribed pattern that is symmetric to the diagonal of a symmetric matrix.

- Assume

$$
\begin{aligned}
X_0 &:= \text{general (distinct eigenvalues)} \\
\Delta &\subset \{(i,j) | 1 \le j < i \le n\} \\
&:= \text{a } \textit{rectangular} \text{ index subset} \\
\hat{X}(t) &:= \text{Portion of } X(t) \text{ conforming to } \Delta \\
P_1(X(t)) &:= \hat{X}(t) - (\hat{X}(t))^T \\
P_2(X(t)) &:= X(t) - P_1(X(t))
\end{aligned}
$$

Then

For all $(i,j) \in \Delta$, $x_{ij}(t) \longrightarrow 0$ as $t \longrightarrow \infty$.



- Assume

$$
\begin{aligned}
X_0 &:= \text{Hamiltonian } \in gl(2n) \\
&:= \begin{bmatrix} A_0, & N_0 \\ K_0, & -A_0^T \end{bmatrix} \\
K, N &:= \text{symmetric } \in gl(n) \\
P_1(X(t)) &:= \begin{bmatrix} 0, & -K(t) \\ K(t), & 0 \end{bmatrix}
\end{aligned}
$$

Then

$a)$ $[X, P_1(X)]$ is Hamiltonian
$b)$ g_1 is both orthogonal and sympletic
$c)$ $X(t)$ remains Hamiltonian
$d)$ $K(t) \longrightarrow 0$ as $t \longrightarrow \infty$.

◇ The Hamiltonian eigenvalue problem for X_0 practically becomes the eigenvalue problem for

$$\lim_{t\to\infty} A(t).$$

◇ No explicit iterative scheme is known for the Hamiltonian eigenvalue problem due to the lack of knowledge of the structure of $g_2(t)$ in the abstract decomposition of $exp(X_0 t)$.

Gradient-type Framework

- Least squares approximations for various types of real and symmetric matrices subject to spectral constraints share a common structure.
- The projected gradient can be formulated explicitly.
- A descent flow can be followed numerically.
- The procedure can be extended to general matrices subject to singular value constraints.

Spectrally Constrained Problem

- Notation:

$$\begin{aligned} S(n) &:= \{\text{All real symmetric matrices}\} \\ O(n) &:= \{\text{All real orthogonal matrices}\} \\ \|X\| &:= \text{ Frobenius matrix norm of } X \\ \Lambda &:= \text{ A given matrix in } S(n) \\ M(\Lambda) &:= \{Q^T \Lambda Q | Q \in O(n)\} \\ V &:= \text{ A single matrix or a subspace in } S(n) \\ P(X) &:= \text{ The projection of } X \text{ into } V \end{aligned}$$

- General problem:

$$\begin{aligned} \text{Minimize } F(X) &:= \frac{1}{2}\|X - P(X)\|^2 \\ \text{Subject to } X &\in M(\Lambda) \end{aligned}$$

- Special cases:
  - ◇ Problem A: Given a symmetric matrix, find its least squares approximation with prescribed spectrum.
  - ◇ Problem B: Construct a symmetric Toeplitz matrix that has a prescribed set of eigenvalues.
  - ◇ Problem C: Find the spectrum of a given a symmetric matrix.

Reformulation

- Idea:
  1. $X \in M(\Lambda)$ satisfies the spectral constraint.
  2. $P(X) \in V$ has the desirable structure in V.
  3. Minimize the undesirable part $\|X - P(X)\|$.
- Working with the parameter Q is easier:

$$\begin{aligned} \text{Minimize } F(Q) \; &:= \; \frac{1}{2}\langle Q^T\Lambda Q - P(Q^T\Lambda Q), \\ & \qquad Q^T\Lambda Q - P(Q^T\Lambda Q)\rangle \\ \text{Subject to } Q^TQ \; &= \; I \end{aligned}$$

  ◇ $\langle A, B\rangle = \text{trace}(AB^T)$ is the Frobenius inner product.

Feasible Set $O(n)$ & Gradient of F

- The set $O(n)$ is a regular surface.
- The tangent space of $O(n)$ at any orthogonal matrix Q is given by

$$T_Q O(n) = QK(n)$$

where

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

- The normal space of $O(n)$ at any orthogonal matrix Q is given by

$$N_Q O(n) = QS(n).$$

- The Fréchet Derivative of F at a general matrix A acting on B:

$$F'(A)B = 2\langle \Lambda A(A^T \Lambda A - P(A^T \Lambda A)), B\rangle.$$

- The gradient of F at a general matrix A:

$$\nabla F(A) = 2\Lambda A(A^T \Lambda A - P(A^T \Lambda A)).$$

The Projected Gradient

- A splitting of $R^{n\times n}$:

$$\begin{aligned} R^{n\times n} &= T_Q O(n) + N_Q O(n) \\ &= QK(n) + QS(n). \end{aligned}$$

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

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

- The projection of $\nabla F(Q)$ into the tangent space:

$$\begin{aligned} g(Q) &= Q\left\{\frac{1}{2}(Q^T \nabla F(Q) - \nabla F(Q)^T Q)\right\} \\ &= Q[P(Q^T \Lambda Q), Q^T \Lambda Q]. \end{aligned}$$

An Isospectral Descent Flow

- A descent flow on the manifold $O(n)$:

$$\frac{dQ}{dt} = Q[Q^T \Lambda Q, P(Q^T \Lambda Q)].$$

- A descent flow on the manifold $M(\Lambda)$:

$$\begin{aligned} \frac{dX}{dt} &= \frac{dQ^T}{dt}\Lambda Q + Q^T \Lambda \frac{dQ}{dt} \\ &= [X, \underbrace{[X, P(X)]}_{k(X)}]. \end{aligned}$$

- The entire concept can be obtained by utilizing the Riemannian geometry on the Lie group $O(n)$.

The Second Order Derivative

- Extend the projected gradient g to the function
$$G(Z) := Z[P(Z^T \Lambda Z), Z^T \Lambda Z]$$
for general matrix Z.
- The Fréchet derivative of G:
$$\begin{aligned} G'(Z)H &= H[P(Z^T \Lambda Z), Z^T \Lambda Z] \\ &+ Z[P(Z^T \Lambda Z), Z^T \Lambda H + H^T \Lambda Z] \\ &+ Z[P'(Z^T \Lambda Z)(Z^T \Lambda H + H^T \Lambda Z), Z^T \Lambda Z]. \end{aligned}$$
- The projected Hessian at a critical point $X = Q^T \Lambda Q$ for the tangent vector QK with $K \in K(n)$:
$$\langle G'(Q)QK, QK \rangle = \\ \langle [P(X), K] - P'(X)[X, K], [X, K] \rangle.$$

Example: Problem A

- Given $N \Longrightarrow P(X) = N$.
- The descent flow:

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

- Assume
  - ◇ The given eigenvalues are $\lambda_1 > \ldots > \lambda_n$.
  - ◇ The eigenvalues of N are $\mu_1 > \ldots > \mu_n$.
- At a critical point Q, the first order condition:

$$[Q^T \Lambda Q, N] = 0$$

  $\Longrightarrow QNQ^T$ must be a diagonal matrix whose elements must be a permutation of $\mu_1, \ldots, \mu_n$.

- The projected Hessian:

$$\begin{aligned} \langle G'(Q)QK, QK \rangle &= \langle [N, K], [X, K] \rangle \\ &= \langle E\hat{K} - \hat{K}E, \Lambda\hat{K} - \hat{K}\Lambda \rangle \\ &= 2\sum_{i<j} (\lambda_i - \lambda_j)(\mu_i - \mu_j)\hat{k}_{ij}^2. \end{aligned}$$

- If a matrix Q is optimal, then:
  - ◇ Columns of $Q^T = [q_1, \ldots, q_n]$ must be the normalized eigenvectors of N corresponding in the order to $\mu_1, \ldots, \mu_n$.
  - ◇ The solution to Problem A is unique.
  - ◇ The solution is given by

$$X = \lambda_1 q_1 {q_1}^T + \ldots + \lambda_n q_n {q_n}^T.$$

- We have reproved the Wielandt-Hoffman theorem.
- The dynamics in Problem A enjoys a special sorting property.
  - ◇ Can be applied to data matching problem and a variety of combinatorial optimizations, including simplex method and the interior point methods for the LP problem.

Example: Problem B

- Given $T = \{$All symmetric Toeplitz matrices$\} \Longrightarrow$

$$P(X) = \sum_{i=1}^{n} \langle X, E_i \rangle E_i.$$

  ◇ $E_1, \ldots, E_n$ is a natural basis of T.

- The descent flow:

$$\begin{aligned} \frac{dX}{dt} &= [X, \underbrace{[X, P(X)]}_{k(X)}] \\ X(0) &= \text{Any thing on } M(\Lambda) \text{ but diagonal matrices.} \end{aligned}$$

- The Lax dynamics offers a globally convergent method for solving the inverse Toeplitz eigenvalue problem.

- Better yet dynamics (Toeplitz annihilator):

$$k_{ij} := \begin{cases} x_{i+1,j} - x_{i,j-1}, & \text{if } 1 \le i < j \le n \\ 0, & \text{if } 1 \le i = j \le n \\ x_{i-1,j} - x_{i,j+1}, & \text{if } 1 \le j < i \le n \end{cases}$$

Example: Problem C

- Take $V = \{\text{All diagonal matrices}\}$ and $\Lambda = X_0 =$ the matrix whose eigenvalues are to be found.
- The objective of Problem C is the same as that of the Jacobi method, i.e., to minimize the off-diagonal elements.
- The descent flow:

$$\begin{aligned} \frac{dX}{dt} &= [X, \underbrace{[X, \mathrm{diag}(X)]}_{k(X)}] \\ X(0) &= X_0. \end{aligned}$$

- Let X be a critical point. Then
  - ◇ If X is a diagonal matrix, then X is a global minimizer.
  - ◇ If X is not a diagonal matrix but $\mathrm{diag}(X)$ is a scalar matrix, then X is a global maximizer.
  - ◇ If X is not a diagonal matrix and $\mathrm{diag}(X)$ is not a scalar matrix, then X is a saddle point.

Isospectral Flows

- QR flow for normal matrices (Chu'84).
- Generalized Toda flow (Chu'84, Watkins'84),
$$\frac{dX}{dt} = [X, \Pi_0(G(X))]$$
where $G(z)$ is analytic over spectrum of $X(0)$.
- QZ flow (Chu'86).
- Continuous Rayleigh quotient flow (Chu'86).
- SVD flow (Chu'86),
$$\frac{dY}{dt} = YN - MY$$
where
$$\begin{aligned} M(t) &:= \Pi_0\left(Y(t)Y(t)^T\right) \\ N(t) &:= \Pi_0\left(Y(t)^T Y(t)\right). \end{aligned}$$
- Abstract QR-type flow (Chu'88).
- Scaled Toda-like flow (Chu'95),
$$\frac{dX}{dt} = [X, A \circ X].$$

Projected Gradient Flows

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

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

- Inverse eigenvalue problem for non-negative matrices (Chu&Driessel'91).
- Inverse singular value problem (Chu'92).

Generalized Flows

- Matrix differential equations (Chu'92).
- Schur-Horn theorem (Chu'95),

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

- Least squares inverse eigenvalue problem (Chu&Chen'96).
- Inverse generalized eigenvalue problem (Chu&Guo'98).
- Inverse stochastic eigenvalue problem (Chu&Guo'98).
- Adaptive Optics with Deformable Mirror Control,

$$\max_{U \in O(n)} \sum_{i=1}^{n} \max_{1 \le j \le m} \left\{ \left(U^T M_j U \right)_{ii} \right\}.$$

  ◇ M_j = Adaptive optics performance characterization and U = Basis of control modes.

  ◇ Parameter dynamics:

$$\begin{aligned} \frac{dU}{dt} &= U \left(U^T \mathcal{K}(MU) - (\mathcal{K}(MU))^T U \right) \\ U(0) &= \text{any orthogonal matrix} \end{aligned}$$

  where

$$\begin{aligned} \mathcal{K}(MU) &:= [M_{k_1} u_1, \ldots, M_{k_n} u_n] \\ k_i &:= \arg\max_j \left\{ \left(U^T M_j U \right)_{ii} \right\}. \end{aligned}$$

Inverse Stochastic Eigenvalue Problem

- Construct a stochastic matrix with prescribed spectrum — A hard problem (Karpelevic'51, Minc'88).
  - ◇ No strings of symmetry.
- Reformulation:

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

  - ◇ J = Real matrix carrying spectral information.
  - ◇ $\circ$ = Hadamard product.
- Steepest descent flow:

$$\begin{aligned} \frac{dP}{dt} &= [(PJP^{-1})^T, \alpha(P,R)]P^{-T} \\ \frac{dR}{dt} &= 2\alpha(P,R) \circ R. \end{aligned}$$

  - ◇ $\alpha(P,R) := PJP^{-1} - R \circ R$.

- ASVD flow for P (Bunse-Gerstner et al'91, Wright'92):

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

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

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

  - ◇ Z, W are skew-symmetric matrices obtainable from Q and S.

Numerical Computation

- Special features in Lax dynamics or parameter dynamics:
  - ◇ Only asymptotically stable equilibria are needed for the original problem.
  - ◇ An explicit Lyapunov function is available.
  - ◇ Orbits are required to stay on a prescribed manifold.
- Challenge to the current numerical ODE techniques:
  - ◇ The size of the differential system can easily be large.
  - ◇ Need an ODE solver that can effectively approximate the asymptotically stable equilibrium point.
  - ◇ Need an ODE solver that can trace trajectories on a manifold constraint (DAE).
- Lots of ongoing research:
  - ◇ Numerical Hamiltonian methods (Sanz-Serna'94).
  - ◇ Projected unitary schemes (Dieci et al'94).
  - ◇ Modified GL RK methods (Calvo et al'95).
  - ◇ Adaptive neural networks method (Dehaene'95).
  - ◇ This conference — many experts, many approaches.

Conclusion

- Area of applications is broad.
- Sheds critical insights into the understanding of the dynamics of discrete methods.
- Unifies different discrete methods as special cases of its discretization and often gives rise to the design of new numerical algorithms.
- May be used as benchmark problems for testing new ODE techniques.
- New ODE techniques may further benefit the numerical computation.
- Enable to tackle existence problems that are seemingly impossible to be solved by conventional discrete methods.
- Usually offers a global method for solving the underlying problem.