Equivalence of Contraction and Koopman Approaches, and Its Application in Nonlinear Systems Identification

Bowen Yi

Polytechnique Montréal

In collaboration with Ian R. Manchester and Fletcher Fan (Sydney)

GERAD, June 3rd, 2024

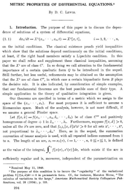
Facts

- Most systems in engineering exhibit nonlinear dynamics
- Analysis of linear systems is much easier (drawbacks)
- Contraction and Koopman are two methods for analysing NL systems exactly and globally by way of linear systems theory

> Useful for other communities!

- Neither is particularly in their basic ideas:
  - Koopman, B.O., **1931**. Hamiltonian systems and transformation in Hilbert space, Proc. Natl. Acad. Sci., USA 17:315–318.
  - Lewis, D.C., **1949**. Metric properties of differential equations. Amer. J. of Math., 71(2): 294–312.
- But both are enjoying a resurgence in popularity in recent years

Bernard Koopman (1900-1981)



METRIC PROPERTIES OF DIFFERENTIAL EQUATIONS.*

By D. C. Lewis.

1. Introduction. The purpose of this paper is to discuss the dependence of solutions of a system of differential equations,

$$dx_i/dt = X^i(x_1, \cdots, x_n, t) = X^i[x, t], \qquad i = 1, 2, \cdots, n, \tag{1.1}$$

on the initial conditions. The classical existence proofs yield inequalities which show that the solutions depend continuously on the initial conditions, at least, if the right hand members satisfy a Lipschitz condition. In this paper we shall refine and supplement these classical inequalities, assuming that the X^i are of class C'. In so doing we call attention to the fundamental importance of a certain quadratic form Q to be introduced in the sequel. Still further, but less useful, refinements may be obtained on the assumption that the X^i are of class C'', in which case a certain biquadratic form B plays the decisive role. It is also indicated, by consideration of simple examples, that our fundamental theorems are the best possible ones of their type. A simple application to the theory of qualitative integration is given.

Our theorems are specified in terms of a metric which we assign to the space of the $(x_1, \cdots, x_n)$. For most purposes it is sufficient to assume a Riemannian space. Much of the analysis, however, is not more difficult, if we assume a general Finsler space.

Let $f[x, \dot{x}] = f(x_1, \cdots, x_n, \dot{x}_1, \cdots, \dot{x}_n)$ be of class C''' and positively homogeneous of degree $+1$ in $\dot{x}_1, \cdots, \dot{x}_n$. Furthermore, suppose $f[x, \dot{x}] > 0$, unless all the $\dot{x}$'s are zero, and that $f_{\dot{x}_i\dot{x}_j}[x, \dot{x}]\lambda^i\lambda^j > 0$ for all sets $\lambda^1, \cdots, \lambda^n$, not proportional to $\dot{x}_1, \cdots, \dot{x}_n$.[1] Here, as in the sequel, the summation convention of tensor analysis is used, with all repeated indices summed from 1 to n. The length of an arc, $x_i = x_i(\tau)$, $i = 1, \cdots, n$, $0 \leq \tau \leq 1$, is defined as the value of the integral, $\int_0^1 f[x(\tau), x'(\tau)]d\tau$, which exists if the arc is sufficiently regular and is, moreover, independent of the parametrization on

* Received May 11, 1948.

[1] The purpose of this condition is to insure the "regularity" of the variational problem $\delta \int f[x, \dot{x}]d\tau = 0$ in parametric form. Cf., for instance, Marston Morse, "The calculus of variations in the large," *American Mathematical Society Colloquium Publications*, vol. 18 (1934), p. 121.

294

D.C. Lewis' article (1949)

Outline

Preliminary: Koopman Operator and Contraction Analysis

Linear Dynamical Systems: Eigendecomposition

Consider an LTI (linear time-invariant) system:

$$\dot{x} = Ax. \tag{1}$$

- If A is diagonalisable, then $A = V\Lambda V^{-1}$ with

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \tag{2}$$

- Solution: $x(t) = e^{At}x_0 = Ve^{\Lambda t}V^{-1}x_0$
- Left eigenvectors v_k are columns of V
- Right eigenvectors r_k are rows of V^{-1}
- Modal decomposition:

$$x(t) = \sum_k (r_k x_0) e^{\lambda_k t} v_k \tag{3}$$

Koopman Operator

Consider a nonlinear system

$$\dot{x} = f(x), \quad x \in \Omega \subset \mathbb{R}^n \tag{4}$$

- Denote its solution at time t (if it exists) as $X(x_0, t)$
- Now consider an arbitrary scalar function of state $\phi(x)$, i.e. an observable, e.g. $x^2, e^x, \sin(x), \ldots$
- [Composition operator] the evolution of ϕ along system solutions:

$$U^t : \phi \;\mapsto\; U^t[\phi] := \phi \circ X(\cdot, t) \tag{5}$$

- Semi-group:
  - $U^{t+s} = U^t \circ U^s = U^s \circ U^t$ for every $s, t \geq 0$
  - $U^0 = \mathrm{Id}$

Koopman Operator (cont'd)

- Linearity: for two observables ϕ, ψ and $a, b \in \mathbb{R}$

$$U^t[a\phi + b\psi] = aU^t[\phi] + bU^t[\psi]$$

- Infinite-dimensional: acting on $C^0(\Omega)$
- Strongly continuous semigroup of linear contracting operators in $(C^0(\Omega), \|\cdot\|_\infty)$
- Think of this a linear system

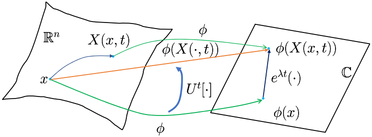
$$\dot{z} = Az \qquad z = \begin{bmatrix} \phi_1(x) \\ \vdots \\ \phi_k(x) \\ \vdots \end{bmatrix}$$

  but operating on an infinite-dimensional space of observables. [What is A?]
- Infinitesimal generator of U^t

$$\frac{\partial U}{\partial t} := \lim_{t \to 0^+} \frac{U^t[\phi] - \phi}{t} \qquad \text{(spectrum?)}$$

Koopman Eigenfunctions

- Infinite-dimensional linear operator can still have eigenvalues and eigenfunctions (generalising eigenvectors)
- Sometimes a finite-dimensional subspace (fixed basis) is enough



For a non-zero observable ϕ_λ, the Koopman eigenvalue is defined as the constant $\lambda \in \mathbb{C}$ s.t.

$$U^t[\phi_\lambda] = e^{\lambda t}\phi_\lambda \tag{6}$$

if it exits, we call ϕ_λ a Koopman eigenfunction.

Equivalent to

$$\frac{\partial \phi_\lambda}{\partial x}(x) f(x) = \lambda \phi_\lambda(x).$$

Examples

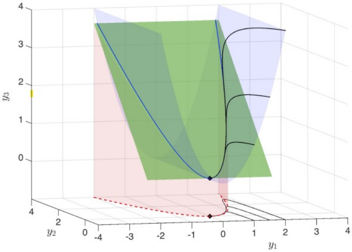
- The energy of a Hamiltonian system is an eigenfunction.
- Consider the nonlinear system (Brunton et al., PLOS 2016)

$$\dot{x} = \begin{bmatrix} \mu x_1 \\ \lambda(x_2 - x_1^2) \end{bmatrix}$$

The Koopman eigenfunctions

$$z := \phi(x) = \begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \end{bmatrix} \implies \dot{z} = \begin{bmatrix} \mu & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & 2\mu \end{bmatrix} z$$

with a high-dimensional dynamics.



Koopman Criteria for Global Asymptotic Stability[1]

Consider the system

$$\dot{x} = f(x)$$

with a hyperbolic equilibrium $x_\star$ with $\nabla f(x_\star) \in \mathbb{C}_{<0}$. If $\exists$ Koopman eigenfunctions $\phi(x) := \text{col}(\phi_1, \ldots, \phi_n)$ s.t.

- (Distinct eigenvalues) Koopman eigenvalues λ_i are different $\phi_i \in C^1$ with $\nabla \phi_i(x_\star) \neq 0$;
- (Stability) $\lambda_i \in \mathbb{C}_{<0}$, and they are eigenvalues of $\nabla f(x_\star)$.

Then, the equilibrium $x_\star$ is globally asymptotically stable (GAS).

[1] Mauroy and Mezić. Global stability analysis using the eigenfunctions of the Koopman operator, IEEE TAC 2016.

Summary of Koopman Approach

- In principle, the infinite dimensional operator acting on the space of observables (functions of the state) is linear.
- Sometimes, finite-dimensional invariant bases can be constructed, from which stability analysis is possible.
- A common *data-driven* approach is to collect samples of $x_t, \dot{x}_t$ and try to find a nonlinear mapping $\phi : x \mapsto z$ with z finite dimensional s.t.

$$\dot{z} \approx Az.$$

  This is called dynamic mode decomposition (DMD) and is closely related to linear system identification.
- From this, various conclusions can be drawn (approximately) regarding stability, domains of attraction, ...

Contraction and Incremental Stability

Nonlinear system: $\dot{x} = f(x)$

- Incremental exponential stability: all trajectories converge to each other exponentially, i.e.,

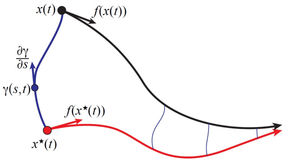
  $$|X(x_a, t) - X(x_b, t)| \leq k_0 |x_a - x_b| e^{-\rho t}$$

  $\forall (x_a, x_b)$, for some $k_0, \rho > 0$.

- Analysis using set stability (Angeli, TAC'02)
- Differential framework: Contraction Analysis

  $$\delta \dot{x} = \frac{\partial f}{\partial x}(x) \delta x$$

  δx is infinitesimal displace between any two trajectories. Not $x - x_\star$ in first-order linearisation!



Linear Systems: Quadratic Lyapunov Functions

Theorem (Lyapunov, 1892) Consider a system

$$\dot{x} = Ax.$$

This system is stable $\iff$ $\exists$ a matrix $P \succ 0$ s.t.

$$A^\top P + PA \prec 0.$$

- Meaning: the function $V(x) = x^\top P x$ is positive for $x \neq 0$, and decreases along flows of system:

$$\begin{aligned}\dot{V} &= \dot{x}^\top P x + x^\top P \dot{x} \\ &= (Ax)^\top P x + x^\top P (Ax) \\ &= x^\top (A^\top P + PA) x \\ &< 0 \quad \forall x \in \mathbb{R}^n \backslash \{0\}\end{aligned}$$

Contraction Analysis

- **Main idea:** study time-varying linearisation along trajectories

$$\delta \dot{x} = F(x)\delta x, \quad F(x) = \frac{\partial f}{\partial x}(x) \tag{7}$$

- **Key result:** (Lohmiller and Slotine, Automatica'98)

  Exp. stability of (7) $\iff$ contraction (IES) of $\dot{x} = f(x)$

- **Contraction metric:** a matrix function $M(x) > 0$ s.t. $\forall \delta x,\ x$

$$\frac{d}{dt}(\delta x^\top M \delta x) = \delta x^\top (\dot{M} + F(x)^\top M(x) + M(x)F(x))\delta x < 0. \tag{8}$$

  True $\iff$ the blue matrix is negative definite

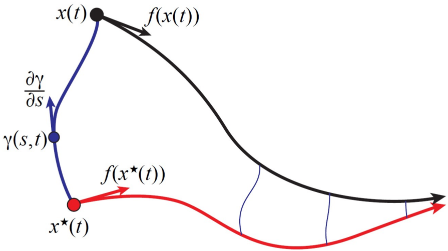
- Convex in $M(x)$ (but infinite-dimensional)
- $M(x)$: a state-dependent family of quadratic Lyapunov function of δx.

Global Convergence

- If a system is contracting, then any pair of solutions x and $x^\star$ converge:

$$|x(t) - x^\star(t)| \to 0. \tag{9}$$

- Why? Intuition: think of a chain connecting pairs of states. If every link in the chain gets shorts, then the states must eventually converge. (Forni and Sepulchre, TAC'13)



Applications of Contraction

- Nonlinear control via convex optimisation
- Observer design and analysis
- Synchronisation of nonlinear oscillators (e.g. power systems)
- Convergence of optimisation algorithms
- Learning (identifying) stable dynamical systems
- Robust machine learning

Koopman vs Contraction

- Koopman: study a nonlinear system by mapping it to a single infinite dimensional LTI (linear time-invariant) system

- Contraction: study a nonlinear system by way of an infinite family of finite dimensional LTV (linear time-varying) systems.

Question: Similar. Any connections?

Slight Extension of Koopman Criteria for GAS

Proposition

Consider the system $\dot{x} = f(x)$ with a hyperbolic equilibrium $x_\star$. If $\exists$ a C^2 mapping $\phi(x) := [\phi_1(x), \dots, \phi_N(x)]^\top$ with $N \geq n$ s.t.

C1 (*immersion*) For a finite N, $\Phi(x) := \frac{\partial \phi}{\partial x}(x)$ is full column rank.

C2 (*stability*) the existence of a Hurwitz matrix A verifying the PDE

$$\frac{\partial \phi}{\partial x}(x) f(x) = A\phi(x). \tag{10}$$

Then, the equilibrium $x_\star$ is GAS.

- No need to require A diagnolisable
- **C1** is used to pull back to the x coordinate
- Extension to infinite N

Equivalence Between Two Approaches

Our Findings (in Plain Language)

Two approaches are equivalent in the nonlinear stability analysis for $\dot{x} = f(x)$:

- Satisfying the Koopman stability criteria $\Longrightarrow$ Contraction (with a compact form of contraction metrics)
- Contracting systems $\Longrightarrow$ Analytical form of finite-dimensional Koopman mappings

Equivalence Between Koopman and Contraction

Theorem

Consider the nonlinear system

$$\dot{x} = f(x). \tag{11}$$

- Assume $\exists$ Koopman mapping ϕ satisfying **C1** and **C2**. If $\Phi^\top \Phi$ ($\Phi := \nabla \phi^\top$) is uniformly bounded, then the system is contracting, a *contracting metric* $M = \Phi^\top P \Phi$ with P the solution of $A^\top P + PA = -I$.
- Conversely, if the system is contracting with the metric $M(x)$ in $\mathtt{cl}(\mathcal{X})$, then $\exists$ a C^1 Koopman mapping ϕ satisfying **C1-C2**.

C1 (*immersion*) For a finite N, $\Phi(x) := \frac{\partial \phi}{\partial x}(x)$ is full column rank.

C2 (*stability*) the existence of an exponentially stable A verifying the PDE

$$\frac{\partial \phi}{\partial x}(x) f(x) = A\phi(x). \tag{12}$$

Sketch of Proof ($\Longrightarrow$)

- Koopman conditions $\Longrightarrow$ $\frac{\partial \phi}{\partial x}(x) f(x) = A\phi(x)$.
- Calculate the partial derivative w.r.t. x, yielding

$$\dot{\Phi}(x) + \Phi(x)F(x) = A\Phi(x), \quad F(x) = \frac{\partial f}{\partial x}(x) \tag{13}$$

- Combining the Lyapunov equation

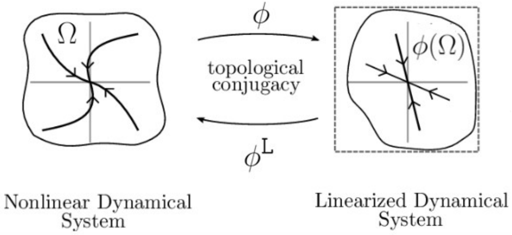
$$\Phi^\top (A^\top P + PA)\Phi = -\Phi^\top \Phi < 0. \tag{14}$$

- From the above,

$$\Phi^\top (A^\top P + PA)\Phi = \dot{M} + F^\top M + MF = -\Phi^\top \Phi$$
$$\leq \frac{1}{\lambda_{\mathtt{max}}\{P\}} M$$

$\left[\text{ Koopman conditions } \textbf{C1-C2} \right] \quad \Longrightarrow \quad$ contration.

- Topological conjugacy



- The proof boils down to the application of contraction of the LTI system $\dot{z} = Az$.
- Though incremental stability is intrinsic (coordinate-free) of a system, $x \mapsto z$ is an immersion rather than diffeomorphism.

Sketch of Proof ($\Longleftarrow$)

- The converse claim $\iff$ find a solution to the PDE

$$\frac{\partial \phi}{\partial x}(x) f(x) = A\phi(x) \tag{15}$$

  s.t. **C1** and **C2**.

- Parameterise ϕ as $\phi(x) = x + T(x)$

$$(15) \iff \frac{\partial T}{\partial x}(x) f(x) = AT(x) + H(x) \tag{16}$$

  with the Hurwitz matrix $A = F(x_\star)$ and $H(x) := -f(x) + F_\star x$.

- The PDE (17) is exactly the one in Kazantzis-Kravaris-Luenberger (KKL) observer. [What is KKL?]

Kazantzis-Kravaris-Luenberger (KKL) Observer

- State observer design for the system

$$\dot{x} = f(x), \quad y = h(x).$$

- Search for a coordinate change $x \mapsto z := T(x)$, the dynamics in which is linear with a nonlinear output injection:

$$\dot{z} = Az + Bh(x)$$

with A Hurwitz and (A, B) controllable.
- Implement the observer (converging by itself)

$$\dot{\xi} = A\xi + By, \quad \hat{x} = \phi^{\mathtt{L}}(\xi)$$

- Solve the partial differential equation

$$(15) \iff \frac{\partial T}{\partial x}(x) f(x) = AT(x) + Bh(x) \tag{17}$$

with the Hurwitz matrix $A = F(x_\star)$ and $H(x) := -f(x) + F_\star x$.
- Always solvable [Andrieu/Praly, SIAM JCO'2006]

$$T(x) = \int_0^{+\infty} \exp(F_\star s) Bh(\breve{X}(x, -s)) ds \tag{18}$$

Sketch of Proof ($\Longleftarrow$) (cont'd)

- In our case, a feasible solution

$$\begin{aligned} T(x) &= \int_0^{+\infty} \exp(F_\star s) H(\breve{X}(x,-s)) ds \\ \phi^0(x) &= x + T(x) \end{aligned} \tag{19}$$

- Locally injective but not global
- Redesigning

$$\phi(x) = e^{-At_x} \phi^0(X(x,t_x))$$

  with a large $t_x > 0$ can guarantee **C1** (immersion). [Different from the way to impose immersion in KKL.]

Example

Consider again the system

$$\dot{x} = \begin{bmatrix} -x_1 \\ -x_2 + x_1^2 \end{bmatrix} \tag{20}$$

Selecting the metric $M(x) = \text{diag}(1 + 4x_1^2, 1)$, we may verify contraction

$$\dot{M}(x) + M(x)F(x) + F(x)^\top M(x) = \begin{bmatrix} -2 - 16x_1^2 & 2x_1 \\ 2x_1 & -2 \end{bmatrix} \prec 0.$$

The flow is given by

$$X(x,t) = \begin{bmatrix} e^{-t}x_1 \\ e^{-t}x_1^2 + e^{-t}x_2 - e^{-2t}x_1^2 \end{bmatrix}.$$

The Koopman mapping is calculated as

$$\phi(x) = x + \int_0^{+\infty} \exp(F_\star s) H(\breve{X}(x, -s)) ds = \begin{bmatrix} x_1 \\ -2x_1 + x_1^2 + x_2 \end{bmatrix}.$$

It is easy to verify

$$\frac{d}{dt}\phi(x) = \begin{bmatrix} -x_1 \\ 2x_1 - x_1^2 - x_2 \end{bmatrix} = F(x_\star)\phi(x).$$

Applications: Learning Contraction Metrics and Stable Koopman Embeddings

Applications

Is the proposed equivalence useful?

Expressivity: All the nonlinear contracting systems can be learned in the Koopman framework.

- Learning contraction metrics from data
- System identification of nonlinear stable systems

Algorithms

Learning Contraction Metrics From Data

Learning contraction metrics has recently been explored for robust motion planning and control:

- D. Sun, S. Jha, C. Fan. Learning certified control using contraction metric, *Proc. Conf. Robot Learning*, pp. 1519–1539, 2020.
- H. Tsukamoto, S.-J. Chung, J.-J. E. Slotine, Neural stochastic contraction metrics for learning-based control and estimation, *IEEE Control Syst. Lett.*, vol. 5, 2020.
- G. Chou, N. Ozay, D. Berenson, Model error propagation via learned contraction metrics for safe feedback motion planning of unknown systems, *ArXiv*, 2021.

Koopman operator may simplify this problem to linear syst identification.

Problem. (*Data-driven contraction metrics learning*) For a given contracting system, assume that only a set of state trajectory data $\{\tilde{x}_k, \dot{\tilde{x}}_k\}_{k=0}^{T}$. Our task is to estimate the contraction metric $M(x)$ using the information of data only.

- Intuitive idea: Contraction $\implies \exists\phi$, ϕ^{L} and Hurwitz matrix A, and the contraction metric is

$$M(x) = \nabla\phi(x)P(\nabla\phi(x))^\top \quad \text{s.t. } A^\top P + PA = -I.$$

- Approach [similar to DMD]:
  - Select basis functions $w(x) \in \mathbb{R}^N$, with $N \gg n$,
  - Parameterise the Koopman mapping as

$$\phi(x) = \theta^\top w(x), \quad \theta \in \mathbb{R}^{N\times n}. \tag{21}$$

    Now the PDE (12) becomes

$$\theta^\top W(x)\dot{x} = A\theta^\top w(x), \quad W(x) := \frac{\partial w}{\partial x}(x) \tag{22}$$

  - Solve the optimization problem

$$\begin{aligned} \min_{\theta, A} \quad & \sum_{k=1}^{n_k} \left|\theta^\top Y_1(k) - A\theta^\top Y_2(k)\right|^2 \\ \text{s.t.} \quad & \text{rank } \theta = n \end{aligned} \tag{23}$$

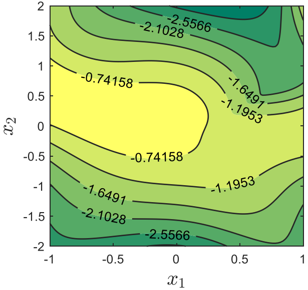
    with avaialbe $Y_1(k) := W(\tilde{x}(k))\dot{\tilde{x}}(k)$ and $Y_2(k) := w(\tilde{x}(k))$.

Numerical Example (Learning Contraction Metrics)

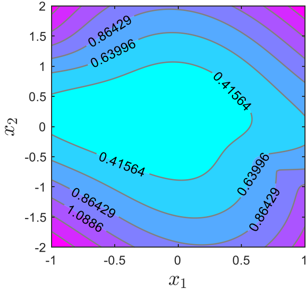
- The data are generated from the system

$$\dot{x} = \begin{bmatrix} -x_1 \\ -x_2 + x_1^2 \end{bmatrix} \tag{24}$$

- Learned metric with polynomial bases



(a) $\lambda_{\max}(\partial_f M + MF + F^\top M)$



(b) $\lambda_{\min}(M)$

- Question: How to select the bases (or dictionary, latent features)?

Learning Stable Koopman Embeddings

Problem. System identification for the discrete-time contracting systems:

$$x_{t+1} = f(x_k) \tag{25}$$

from the dataset $\{x_k\}_{k \in \mathbb{Z}_+}$.

Our targets:

- Propose a new model class for learning stable discrete-time nonlinear systems
- This model class
  1. is guaranteed to be stable (important - physical priori knowledge),
  2. is unconstrained in its parameters,
  3. embeds the system in a linear (Koopman) subspace, and
  4. contains all contracting (stable) systems.

Equivalence for Discrete Dynamical Systems

- Discrete-time Koopman operator:

$$
\begin{aligned}
U:\ L^p(\mathcal{X}) &\rightarrow L^p(\mathcal{X}) \\
&\mapsto U[\phi] := \phi \circ f
\end{aligned}
$$

- Extension to Discrete-time case:

Theorem: The system (25) is contracting iff $\exists$ a mapping ϕ such that

C1 there exists a Schur stable matrix A satisfying

$$\phi(f(x)) = A\phi(x), \quad x \in \mathcal{X} \tag{26}$$

C2 $\Phi(x) := \frac{\partial \phi}{\partial x}$ has full rank and $\Phi(x)^\top \Phi(x)$ is uniformly bounded.

Model Class

We propose the following model class:

Definition (Koopman model)

$$x_k = \phi^{\mathtt{L}}(A^k \phi(x_0)), \tag{27}$$

where A is Schur stable, $\phi : \mathbb{R}^n \to \mathbb{R}^N$ and $\phi^{\mathtt{L}}(\phi(x)) = x$.

- It can be thought of as an LTV system with a nonlinear output:

$$\begin{aligned} z_0 &= \phi(x_0) \\ z_{k+1} &= A z_k \\ x_k &= \phi^{\mathtt{L}}(z_k) \end{aligned}$$

- Algorithms to search for stable A, ϕ and $\phi^{\mathtt{L}}$.

Direct Parameterization of A

An unconstrained parameterisation of A which is guaranteed to be stable:[2]

$$A(L, R) = (M_{11} + M_{22} + R - R^\top)^{-1} M_{21}, \tag{28}$$

where

$$M := \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = LL^\top + \epsilon I. \tag{29}$$

with $\epsilon > 0$.

Proposition: A matrix $A \in \mathbb{R}^{n\times n}$ is Schur stable iff there exists L and R such that $A = A(L, R)$.

Hint:

$$\left[\begin{array}{c|c} E + E^\top - P & F^\top \\ \hline F & P \end{array}\right] \succ \gamma I, \gamma > 0 \iff \text{Schur stability of } E^{-1}F$$

[2]Tobenkin, Manchester and Megretski, IEEE TAC, 2017.

Parameterization of ϕ

- We parametrize the observables as:

$$\phi(x) = Cx + \varphi(x, \theta_{\mathtt{NN}}), \tag{30}$$

  where $C = [I_n, 0_{n\times(N-n)}]^\top$.
- The linear part discourages the model from converging to the trivial solution $\phi(x) = 0$.
- For the nonlinear part $\varphi(x, \theta_{\mathtt{NN}})$, we use a feedforward fully-connected neural network with rectified linear (ReLU) activation functions. [Or any differentiable approximators]
- We also parameterize the left inverse function as another neural network $\phi^{\mathtt{L}}(z, \theta_{\mathtt{L}})$.

Optimization Formulation

- Since our model parameters are unconstrained, we can optimize any differentiable objective function using an auto-differentiation toolbox.
- The optimization formulation is:

$$\min_{\theta \in \Theta} \frac{1}{T} \sum_{k=0}^{T} \underbrace{\Big| \underbrace{\phi(\tilde{x}_k)}_{\text{data}} - \underbrace{A(L,R)^k \phi(\tilde{x}_0)}_{\text{prediction}} \Big|_2^2}_{\text{simulation error}} + \alpha \underbrace{\Big| \tilde{x}_t - \phi^{\mathtt{L}}(\phi(\tilde{x}_t)) \Big|_2^2}_{\text{reconstruction loss}}.$$

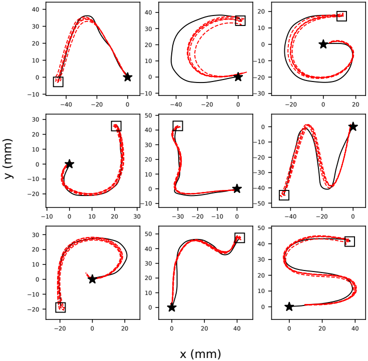
- Empirically, we found that minimizing the simulation error in ϕ-space performed better than directly minimizing the simulation error in x.
- Implementation:
  - ML framework: PyTorch + Adam Optimizer
  - Fast matrix power computation:

$$A^k = (U \Lambda U^{-1})^k = U \Lambda^k U^{-1} \qquad \text{[if diagonalisable]}$$

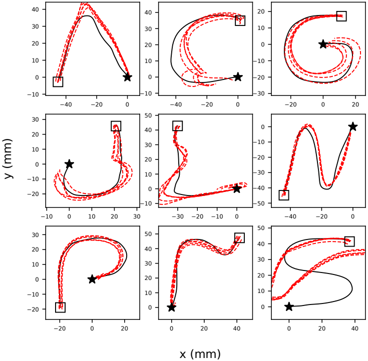
  - 2 hidden layers, 50 nodes each, and an output dimensionality of 20.

Trajectories With and Without Stability Guarantee

LASA handwriting datase (Khansari-Zadeh and A. Billard, TRO'11)



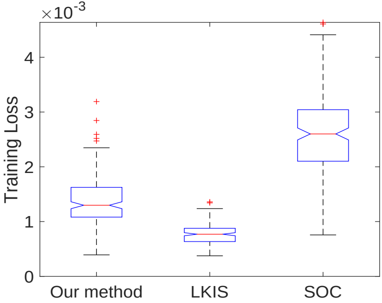
With stability (our method)

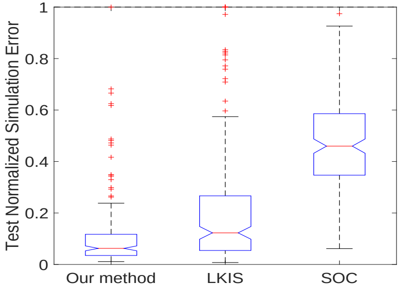


Without stability (LKIS[2])

[2] Takeishi, Kawahara and Yairi, Learning Koopman invariant subspaces for dynamic mode decomposition, *Adv Neural Inf Process Syst*, 2017.

Comparison of Training and Test Errors





SOC: Mamakoukas, Xherija, and Murphey, Memory-Efficient Learning of Stable Linear Dynamical Systems for Prediction and Control, 2020.

Extensions

Nonlinear Time-Varying Systems

Consider the nonlinear time-varying system

$$\dot{x} = f(x, t)$$

with its solution $X(t; x, s)$ from the initial condition $x(s)$.

- Non-autonomous Koopman operator

$$U^{(t,s)}[\phi(x, s)] = \phi(X(t; x, s), t)$$

- Koopman eigenfunction ϕ and eigenvalue λ defined by

$$U^{(t,s)}[\phi] = \exp\left(\int_s^t \lambda(\tau) d\tau\right) \phi$$

- Equivalent to the PDE

$$\frac{\partial \phi}{\partial t} + \langle f(x, t), \nabla_x \phi \rangle = \lambda \phi.$$

- Equivalence between Koopman and contraction still holds.

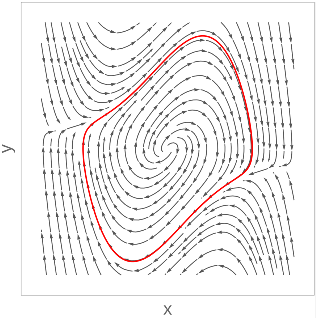
Orbital Stability (Limit Cycle)

Consider the system $\dot{x} = f(x)$.

- Orbit stability: having an attractive periodically non-trivial solution

  $$X(t) = X(t+T), \quad t \geq 0,$$

  with the orbit $\gamma := \{x \in \mathbb{R}^n : x = X(t),\ 0 \leq t \leq T\}$.
- Useful in many engineering problems
  - Bipedal robotics
  - Rotational electromechanical systems (motors)
  - Biology



Orbital Stability (cont'd)

- Koopman operator for orbital stability (Mauroy and Mezic, TAC'16):
  - Similar to the hyperbolic equilibrium case
  - Consider $n-1$ eigenfunctions
  - Zero value on the orbit, i.e. $\phi(x)|_{x\in\gamma} = 0$
- Transverse contraction:
  - Local: a positive definite metric $M(x)$ (Manchester and Slotine, SCL'14)

$$\dot{M} + \frac{\partial f(x)}{\partial x}^\top M + M\frac{\partial f(x)}{\partial x} - \rho(x) f(x) f(x)^\top \prec 0 \qquad (31)$$

  - Global: a positive semidefinite metric $\mathbb{M}(x)$ (Yi et al., TAC'21)

$$\dot{\mathbb{M}}(x) + \mathbb{M}(x)\frac{\partial f(x)}{\partial x} + \frac{\partial f(x)}{\partial x}^\top \mathbb{M}(x) \prec -k\mathbb{M}(x), \quad k > 0. \qquad (32)$$

- Equivalence can also be established for orbital stability.

Example: Limit Cycle in Induction Motor

- The model of induction motors

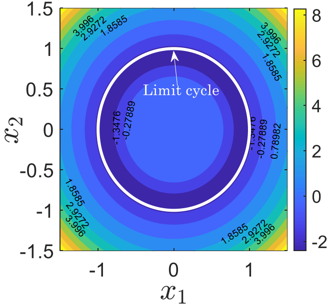
$$\dot{\psi}_r = -R\phi_r + \omega \mathbb{J}\psi_r + Ru$$
$$\dot{\omega} = u^\top \mathbb{J}\psi_r - \tau_L, \quad \mathbb{J} := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \tag{33}$$

- Flux $\psi_r \in \mathbb{R}^2$ and angular speed $\omega \in \mathbb{R}$
- Field-oriented control (FOC)

$$u = \left[\beta_\star I_2 - \frac{k}{\beta_\star}(\omega - \omega_\star)\mathbb{J}\right]\frac{\psi_r}{|\psi_r|}, \quad k > 0.$$

- Koopman eigenfunctions

$$\phi_1(x) = 1 - \frac{1}{|x_p|}, \quad \phi_2(x) = \frac{x_3 - 1}{|x_p|}.$$



Transverse contraction metric

$$M(x) = \left[\frac{\partial \phi}{\partial x}(x)\right] P \left[\frac{\partial \phi}{\partial x}(x)\right]^\top + \Theta(x)^\top \Theta(x)$$

Nonlinear Control Systems

Consider the nonlinear control system

$$\dot{x} = f(x, u).$$

- Control contraction metric: If we can find a uniformly bounded metric $M(x)$ and a function $K(x)$ satisfying

$$\dot{M} + MF + F^\top M + MGK + (GK)^\top M \prec 0, \tag{34}$$

  then we call $M(x)$ a (strong) control contraction metric, in which Jacobians $F(x, u) := \frac{\partial f(x,u)}{\partial x}$ and $G(x, u) := \frac{\partial f(x,u)}{\partial u}$.

- Koopman mapping ϕ satisfying

$$\frac{\partial \phi}{\partial x}(x) f(x, u) = A\phi(x) + Bu, \ \forall u \in \mathbb{R}^m, \tag{35}$$

- Koopman approach $\implies$ the existence of CCM.
- Usually consider the bi-linear case.
- The converse claim?

Summary

- Establish the equivalence between Koopman and contraction approaches for nonlinear stability analysis.
- Consider the autonomous, time-varying, limit cycle and controlled systems.
- Useful for robust learning of dynamical systems
  - model stability guarantee
  - unconstrained in parameters
  - covering all stable (contracting) systems
- Future works:
  - Learning framework for controlled systems
  - The converse result
  - Integrate in control and estimation problems

References:

1. B. Yi and I.R. Manchester, On the equivalence of contraction and Koopman approaches for nonlinear stability and control, IEEE TAC, pp. 1–16, 2024.
2. F. Fan, B. Yi, D. Rye, G. Shi, and I.R. Manchester, Learning stable Koopman embedding for identification and control, ArXiv Preprint, 2024.

THANKS

bowen.yi@polymtl.ca