Contraction Theory for Control, Computation and Dynamical Systems

Francesco Bullo

Center for Control,
Dynamical Systems & Computation
University of California at Santa Barbara
https://fbullo.github.io

Peking University Engineering and Science Webinar
College of Engineering, Peking University, May 26, 2023

Acknowledgments

Veronica Centorrino
Scuola Sup Meridionale

Alexander Davydov
UC Santa Barbara

Anand Gokhale
UC Santa Barbara

Saber Jafarpour
GeorgiaTech

Anton Proskurnikov
Politecnico Torino

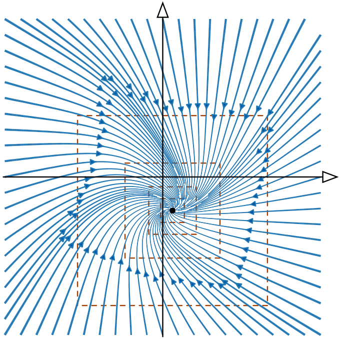
Giovanni Russo
Univ Salerno

John W. Simpson-Porco
University of Toronto

AFOSR ARO DTRA/ERDC

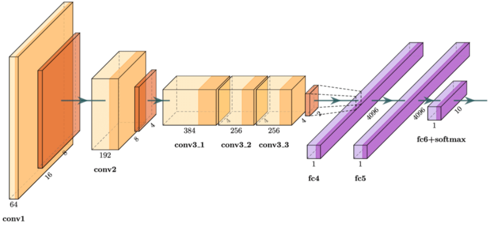
Frederick Leve @AFOSR FA9550-22-1-0059
Edward Palazzolo @ARO W911NF-22-1-0233
Paul Tandy @DTRA W912HZ-22-2-0010
Donald Wagner @AFOSR FA9550-21-1-0203

Outline

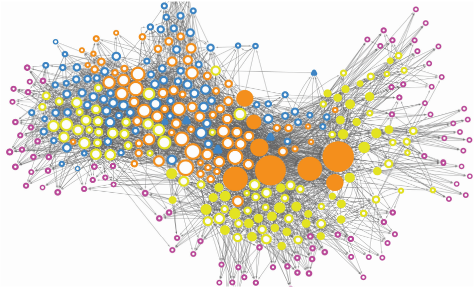


contractivity = robust computationally-friendly stability

fixed point theory + Lyapunov stability theory + geometry of metric spaces



artificial neural network AlexNet ’12

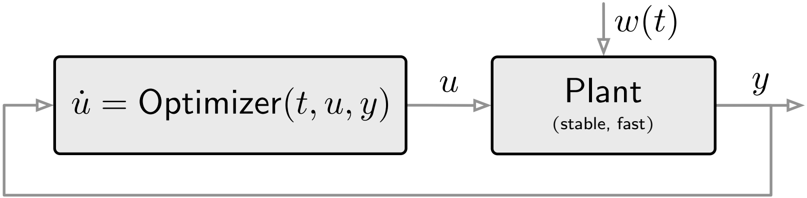


C. elegans connectome ’17

recurrent neural networks

well-posedness, stability, computation and input/output robustness

Chapter 3: Time-varying optimization algorithms



optimization via dynamical systems

online time-varying optimization, optimization-based feedback control, . . .

Outline

Contraction theory: historical notes

- **Origins**

  S. Banach. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fundamenta Mathematicae*, 3(1):133–181, 1922.

- **Dynamics**:

  G. Dahlquist. *Stability and error bounds in the numerical integration of ordinary differential equations*. PhD thesis, (Reprinted in Trans. Royal Inst. of Technology, No. 130, Stockholm, Sweden, 1959), 1958

  S. M. Lozinskii. Error estimate for numerical integration of ordinary differential equations. I. *Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika*, 5:52–90, 1958. URL **http://mi.mathnet.ru/eng/ivm2980**. (in Russian)

- **Computation**:

  C. A. Desoer and H. Haneda. The measure of a matrix as a tool to analyze computer algorithms for circuit analysis. *IEEE Transactions on Circuit Theory*, 19(5):480–486, 1972.

- **Systems and control**:

  W. Lohmiller and J.-J. E. Slotine. On contraction analysis for non-linear systems. *Automatica*, 34(6): 683–696, 1998.

Linear algebra: induced norms

Vector norm	Induced matrix norm	Induced matrix log norm
$\|x\|_1 = \sum_{i=1}^n \|x_i\|$	$\|A\|_1 = \max_{j\in\{1,\dots,n\}} \sum_{i=1}^n \|a_{ij}\|$	$\mu_1(A) = \max_{j\in\{1,\dots,n\}} \left(a_{jj} + \sum_{i=1,i\neq j}^n \|a_{ij}\|\right)$ = max column "absolute sum" of A
$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$	$\|A\|_2 = \sqrt{\lambda_{\max}(A^\top A)}$	$\mu_2(A) = \lambda_{\max}\left(\frac{A + A^\top}{2}\right)$
$\|x\|_\infty = \max_{i\in\{1,\dots,n\}} \|x_i\|$	$\|A\|_\infty = \max_{i\in\{1,\dots,n\}} \sum_{j=1}^n \|a_{ij}\|$	$\mu_\infty(A) = \max_{i\in\{1,\dots,n\}} \left(a_{ii} + \sum_{j=1,j\neq i}^n \|a_{ij}\|\right)$ = max row "absolute sum" of A

Discrete-time dynamics and Lipschitz constants

$$x_{k+1} = \mathsf{F}(x_k) \qquad \text{on } \mathbb{R}^n \text{ with norm } \|\cdot\| \text{ and induced norm } \|\cdot\|$$

Lipschitz constant

$$\begin{aligned}\mathsf{Lip}(\mathsf{F}) &= \inf\{\ell > 0 \ \text{such that} \ \|\mathsf{F}(x) - \mathsf{F}(y)\| \leq \ell\|x - y\| \quad \text{for all } x, y\} \\ &= \sup_x \|\mathsf{J}_{\mathsf{F}}(x)\|\end{aligned}$$

For **scalar map** f, $\mathsf{Lip}(f) = \sup_x |f'(x)|$

For **affine map** $\mathsf{F}_A(x) = Ax + a$

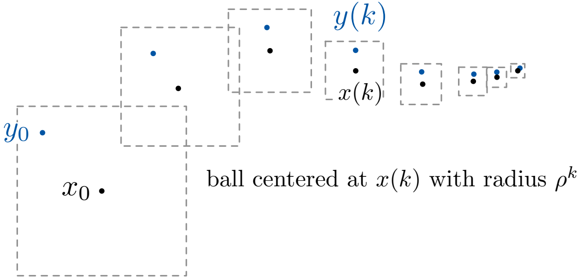
$$\|x\|_{2,P} = (x^\top P x)^{1/2} \qquad \mathsf{Lip}_{2,P}(\mathsf{F}_A) = \|A\|_{2,P} \leq \ell \qquad \iff \qquad A^\top P A \preceq \ell^2 P$$

$$\|x\|_{\infty,\eta} = \max_i |x_i|/\eta_i \qquad \mathsf{Lip}_{\infty,\eta}(\mathsf{F}_A) = \|A\|_{\infty,\eta} \leq \ell \qquad \iff \qquad \eta^\top |A| \leq \ell \eta^\top$$

Banach contraction theorem for discrete-time dynamics:

If $\rho := \mathrm{Lip}(F) < 1$, then

1. F is **contracting** $=$ distance between trajectories decreases exp fast (ρ^k)
2. F has a unique, glob exp stable equilibrium x^*



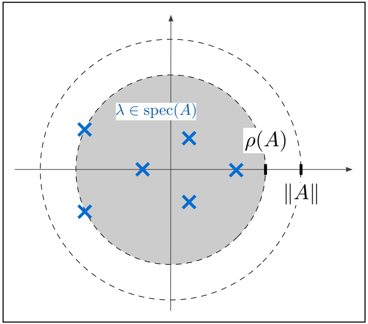
From discrete to continuous time

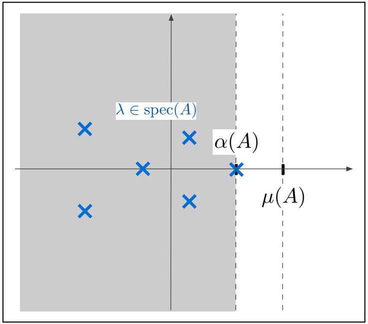
The **induced log norm** of $A \in \mathbb{R}^{n\times n}$ wrt to $\|\cdot\|$:

$$\mu(A) := \lim_{h\to 0^+} \frac{\|I_n + hA\| - 1}{h}$$

subadditivity: $\mu(A+B) \leq \mu(A) + \mu(B)$

scaling: $\mu(bA) = b\mu(A), \quad \forall b \geq 0$





Example induced log norms

Vector norm	Induced matrix norm	Induced matrix log norm
$\|x\|_1 = \sum_{i=1}^n \|x_i\|$	$\|A\|_1 = \max_{j\in\{1,\dots,n\}} \sum_{i=1}^n \|a_{ij}\|$	$\mu_1(A) = \max_{j\in\{1,\dots,n\}} \Big(a_{jj} + \sum_{i=1,i\neq j}^n \|a_{ij}\|\Big)$ $=$ max column "absolute sum" of A
$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$	$\|A\|_2 = \sqrt{\lambda_{\max}(A^\top A)}$	$\mu_2(A) = \lambda_{\max}\Big(\frac{A+A^\top}{2}\Big)$
$\|x\|_\infty = \max_{i\in\{1,\dots,n\}} \|x_i\|$	$\|A\|_\infty = \max_{i\in\{1,\dots,n\}} \sum_{j=1}^n \|a_{ij}\|$	$\mu_\infty(A) = \max_{i\in\{1,\dots,n\}} \Big(a_{ii} + \sum_{j=1,j\neq i}^n \|a_{ij}\|\Big)$ $=$ max row "absolute sum" of A

Continuous-time dynamics and one-sided Lipschitz constants

$$\dot{x} = \mathsf{F}(x) \qquad \text{on } \mathbb{R}^n \text{ with norm } \|\cdot\| \text{ and induced log norm } \mu(\cdot)$$

One-sided Lipschitz constant

$$\begin{aligned}\mathrm{osLip}(\mathsf{F}) &= \inf\{b \in \mathbb{R} \text{ such that } \langle\!\langle \mathsf{F}(x) - \mathsf{F}(y), x - y \rangle\!\rangle \leq b\|x-y\|^2 \quad \text{for all } x, y\} \\ &= \sup_x \mu(\mathsf{J}_{\mathsf{F}}(x))\end{aligned}$$

For **scalar map** f, $\quad \mathrm{osLip}(f) = \sup_x f'(x)$

For **affine map** $\mathsf{F}_A(x) = Ax + a$

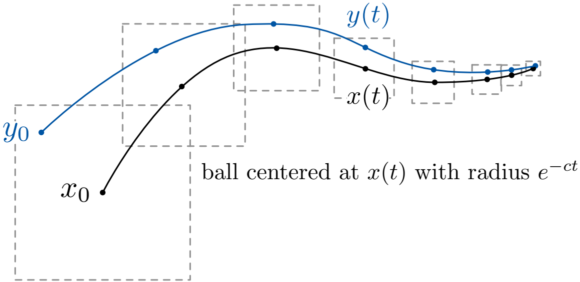
$$\mathrm{osLip}_{2,P}(\mathsf{F}_A) = \mu_{2,P}(A) \leq \ell \qquad \Longleftrightarrow \qquad A^\top P + AP \preceq 2\ell P$$

$$\mathrm{osLip}_{\infty,\eta}(\mathsf{F}_A) = \mu_{\infty,\eta}(A) \leq \ell \qquad \Longleftrightarrow \qquad a_{ii} + \sum_{j \neq i} |a_{ij}|\eta_i/\eta_j \leq \ell$$

Banach contraction theorem for continuous-time dynamics:

If $-c := \mathsf{osLip}(\mathsf{F}) < 0$, then

1. F is **infinitesimally contracting** = distance between trajectories decreases exp fast (e^{-ct})
2. F has a unique, glob exp stable equilibrium x^*



Detour: From inner products to weak pairings

$$\tfrac{1}{2}\frac{d}{dt}\|x(t)\|_2^2 \;=\; \dot{x}^\top x \;=\; \langle\!\langle \dot{x}, x \rangle\!\rangle$$

$$\Longrightarrow \qquad \tfrac{1}{2}D^+\|x(t)\|^2 \;=:\; [\![\dot{x}, x]\!]$$

- D^+ is upper-right Dini derivative
- **weak pairing** $[\![\cdot,\cdot]\!] : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ exists for each norm, i.e.,

$$[\![y, x]\!]_1 \;:=\; \|x\|_1 \operatorname{sign}(x)^\top y \qquad \text{(**sign pairing**)}$$

$$[\![y, x]\!]_\infty \;:=\; \max_{i \in \mathcal{A}_\infty(x)} x_i y_i \qquad \text{for } \mathcal{A}_\infty(x) = \{i \mid |x_i| = \|x\|_\infty\} \qquad \text{(**max pairing**)}$$

theory of weak pairings: computational properties and applications to monotone operators

Log norm bounds	Demidovich conditions	One-sided Lipschitz conditions
$\mu_{2,P}(\mathsf{J}_\mathsf{F}(x)) \leq -c$	$P\mathsf{J}_\mathsf{F}(x) + \mathsf{J}_\mathsf{F}(x)^\top P \preceq -2cP$	$(x-y)^\top P\big(\mathsf{F}(x) - \mathsf{F}(y)\big) \leq -c\\|x-y\\|_{P^{1/2}}^2$
$\mu_1(\mathsf{J}_\mathsf{F}(x)) \leq -c$	$\operatorname{sign}(v)^\top \mathsf{J}_\mathsf{F}(x)v \leq -c\\|v\\|_1$	$\operatorname{sign}(x-y)^\top(\mathsf{F}(x) - \mathsf{F}(y)) \leq -c\\|x-y\\|_1$
$\mu_\infty(\mathsf{J}_\mathsf{F}(x)) \leq -c$	$\max_{i\in\mathcal{A}_\infty(v)} v_i\,(\mathsf{J}_\mathsf{F}(x)v)_i \leq -c\\|v\\|_\infty^2$	$\max_{i\in\mathcal{A}_\infty(x-y)}(x_i-y_i)(\mathsf{F}_i(x)-\mathsf{F}_i(y)) \leq -c\\|x-y\\|_\infty^2$

Each row = three equivalent statements. To be understood for all $x, y \in \mathbb{R}^n$ and all $v \in \mathbb{R}^n$.

Example #1: Gradient flow for strongly convex function

Given strongly convex $f : \mathbb{R}^n \to \mathbb{R}$ with parameter μ, **gradient dynamics**

$$\dot{x} = f_{\mathsf{G}}(x) := -\nabla f(x)$$

f_{G} is infinitesimally contracting wrt $\|\cdot\|_2$ with rate μ

If f is twice-differentiable, then $\mathrm{Hess}\, f(x) \succeq \mu I_n$ for all x

$$\begin{aligned} & \mathsf{J}_{(-\nabla f)}(x) = -\,\mathrm{Hess}\, f(x) \preceq -\mu I_n \\ \iff \qquad & I_n\, \mathsf{J}_{(-\nabla f)}(x) + \mathsf{J}_{(-\nabla f)}(x)^\top I_n \preceq -2\mu I_n \end{aligned}$$

Example #2: Primal-dual gradient dynamics

strongly convex function f s.t. $0 \prec \mu_{\text{min}} I_n \preceq \text{Hess}\, f \preceq \mu_{\text{max}} I_n$

constraint matrix A s.t. $0 \prec a_{\text{min}} I_m \preceq AA^\top \preceq a_{\text{max}} I_m$

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & Ax = b \end{aligned}$$

primal-dual gradient dynamics:

$$\begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = f_{\text{PDG}}(x, \lambda) := \begin{bmatrix} -\nabla f(x) - A^\top \lambda \\ Ax - b \end{bmatrix}$$

f_{PDG} is infinitesimally contracting wrt weighted $\|\cdot\|_{2,P^{1/2}}$ with rate c

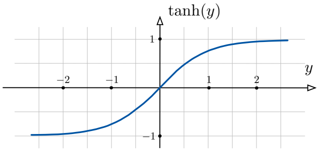
$$P = \begin{bmatrix} I_n & \alpha A^\top \\ \alpha A & I_m \end{bmatrix}, \quad \alpha = \frac{1}{3} \min\left\{ \frac{1}{\mu_{\text{max}}}, \frac{\mu_{\text{min}}}{a_{\text{max}}} \right\}, \qquad \text{and} \qquad c = \frac{5}{18} \min\left\{ \frac{a_{\text{min}}}{\mu_{\text{max}}}, \frac{a_{\text{min}}}{a_{\text{max}}} \mu_{\text{min}} \right\}$$

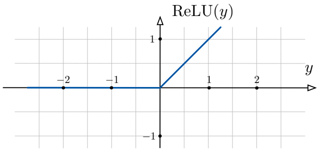
For each $\mu_{\text{min}} I_n \preceq Q \preceq \mu_{\text{max}} I_n$, $\quad \begin{bmatrix} -Q & -A^\top \\ A & 0 \end{bmatrix}^\top P + P \begin{bmatrix} -Q & -A^\top \\ A & 0 \end{bmatrix} \preceq -2cP$

Example #3: Firing-rate recurrent neural network

$$\dot{x} = f_{\mathsf{FR}}(x) := -x + \Phi(Ax + Bu)$$

sigmoid, hyperbolic tangent
ReLU $= \max\{x, 0\} = (x)_+$
$0 \leq \Phi_i'(y) \leq 1$





f_{FR} is infinitesimally contracting wrt $\|\cdot\|_\infty$ with rate $1 - \mu_\infty(A)_+$ if

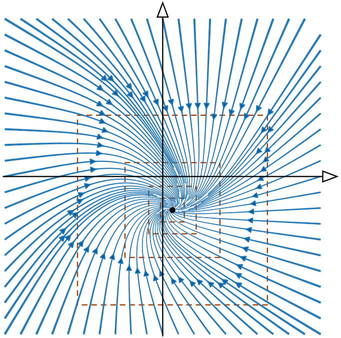
$$\mu_\infty(A) < 1 \qquad \text{(i.e., } a_{ii} + \textstyle\sum_j |a_{ij}| < 1 \text{ for all } i\text{)}$$

$$\begin{aligned}
\mathsf{osLip}_\infty(f_{\mathsf{FR}}) &= \sup_{x,u} \mu_\infty\big(-I_n + (\mathsf{J}_\Phi(Ax+Bu))A\big) = -1 + \sup_{x,u} \mu_\infty\big(\mathsf{J}_\Phi(Ax+Bu)A\big) \\
&= -1 + \max_{d \in [0,1]^n} \mu_\infty(\mathrm{diag}(d)A) \qquad \text{(max convex polytope, } 2^n \text{ vertices)} \\
&= -1 + \max\big\{\mu_\infty(0), \mu_\infty(A)\big\} = -1 + \mu_\infty(A)_+
\end{aligned}$$

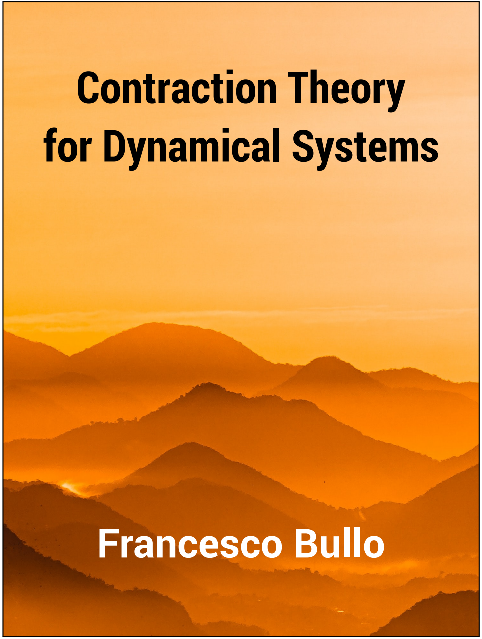
contractivity = robust computationally-friendly stability
fixed point theory + Lyapunov stability theory + geometry of metric spaces

highly-ordered transient and asymptotic behavior:

1. unique globally exponential stable equilibrium
   & two natural Lyapunov functions
2. robustness properties
   - bounded input, bounded output (iss)
   - finite input-state gain
   - robustness margin wrt unmodeled dynamics
   - robustness margin wrt delayed dynamics
3. periodic input, periodic output
4. modularity and interconnection properties
5. accurate numerical integration and equilibrium point computation



search for contraction properties
design engineering systems to be contracting



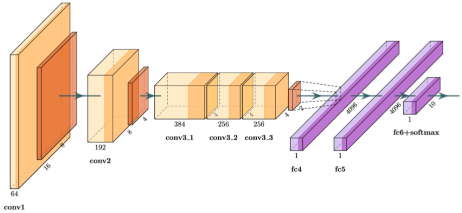
Contraction Theory for Dynamical Systems, Francesco Bullo, KDP, 1.1 edition, 2023, ISBN 979-8836646806

1. Textbook with exercises and answers. Format: textbook, slides, and paperbook
2. Content:
   Fixed point theory
   Theory of contracting dynamics on vector spaces
   Applications to nonlinear and interconnected systems
3. Self-Published and Print-on-Demand at:
   https://www.amazon.com/dp/B0B4K1BTF4
4. PDF Freely available at
   https://fbullo.github.io/ctds
5. 10h minicourse on youtube:
   https://youtu.be/RvR47ZbqJjc
6. Future version to include: systems on Riemannian manifolds, homogeneous spaces, and solid cones

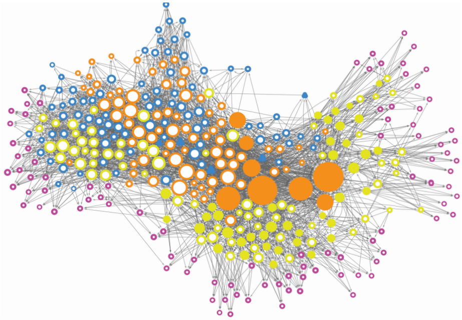
   "Continuous improvement is better than delayed perfection"
   Mark Twain

Outline

While most ML architectures are feedforward,
biological neural networks are recurrent and resemble implicit ML architectures



artificial neural network AlexNet '12



C. elegans connectome '17

Aim: understand the dynamics of neural networks, so that

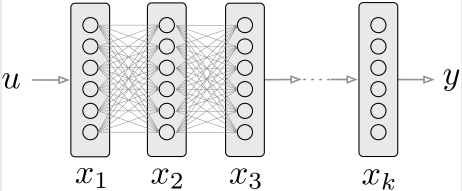
- **reproducible behavior, i.e., equilibrium response as function of stimula**
- robust behavior in face of uncertain stimuli and dynamics
- learning models, efficient computational tools, periodic behaviors ...

A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 2012

G. Yan, P. E. Vértes, E. K. Towlson, Y. L. Chew, D. S. Walker, W. R. Schafer, and A.-L. Barabási. Network control principles predict neuron function in the Caenorhabditis elegans connectome. *Nature*, 550(7677):519–523, 2017.

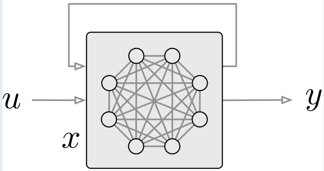
Fixed point computation

Feedforward NN



$$x_{i+1} = \Phi(A_i x_i + b_i), \quad x_0 = u,$$
$$y = C x_k + d$$

Implicit/Recurrent NN

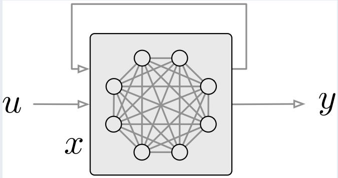


$$x = \Phi(Ax + Bu + b),$$
$$y = Cx + d$$

Fixed point strategies in data science = simplifying and unifying framework to model, analyze, and solve advanced convex optimization methods, Nash equilibria, monotone inclusions, etc.

P. L. Combettes and J.-C. Pesquet. Fixed point strategies in data science. *IEEE Transactions on Signal Processing*, 2021.

Application: ℓ_∞-contracting neural networks



$$\dot{x} = -x + \Phi(Ax + Bu + b) \qquad \text{(\textit{recurrent NN})}$$

$$x = \Phi(Ax + Bu + b) \qquad \text{(\textit{implicit NN})}$$

$$x_{k+1} = (1-\alpha)x_k + \alpha\Phi(Ax_k + Bu + b) \qquad \text{(\textit{forward Euler})}$$

If

$$\mu_\infty(A) < 1 \qquad \Big(\text{i.e., } a_{ii} + \sum_j |a_{ij}| < 1 \text{ for all } i\Big)$$

- recurrent NN is infinitesimally contracting with rate $1 - \mu_\infty(A)_+$
- implicit NN is well posed
- forward Euler is contracting with factor $1 - \dfrac{1-\mu_\infty(A)_+}{1-\min_i(a_{ii})_-}$ at $\alpha = \dfrac{1}{1-\min_i(a_{ii})_-}$

Forward Euler theorem

Forward Euler theorem for contracting dynamics

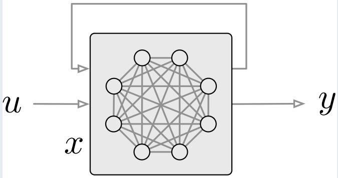
Given arbitrary norm $\|\cdot\|$, equivalent statements

1. $\dot{x} = \mathsf{F}(x)$ is infinitesimally contracting
2. there exists $\alpha > 0$ such that $x_{k+1} = x_k + \alpha\mathsf{F}(x_k)$ is contracting

Given *contraction rate c* and *Lipschitz constant ℓ*, define *condition number $\kappa = \dfrac{\ell}{c} \geq 1$*

1. $\mathsf{Id} + \alpha\mathsf{F}$ is contracting for
$$0 < \alpha < \frac{1}{c\kappa(1+\kappa)}$$
2. the optimal step size minimizing and minimum contraction factor:
$$\alpha^* = \frac{1}{c}\Big(\frac{1}{2\kappa^2} - \frac{3}{8\kappa^3} + \mathcal{O}\Big(\frac{1}{\kappa^4}\Big)\Big)$$
$$\ell^* = 1 - \frac{1}{4\kappa^2} + \frac{1}{8\kappa^3} + \mathcal{O}\Big(\frac{1}{\kappa^4}\Big)$$

Application: ℓ_∞-contracting neural networks



$$\dot{x} = -x + \Phi(Ax + Bu + b) \qquad \text{(recurrent NN)}$$

$$x = \Phi(Ax + Bu + b) \qquad \text{(implicit NN)}$$

$$x_{k+1} = (1-\alpha)x_k + \alpha\Phi(Ax_k + Bu + b) \qquad \text{(forward Euler)}$$

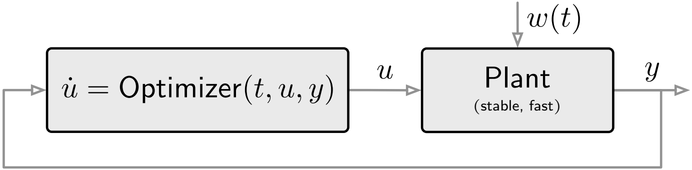
If

$$\mu_\infty(A) < 1 \qquad \Big(\text{i.e., } a_{ii} + \sum_j |a_{ij}| < 1 \text{ for all } i\Big)$$

- recurrent NN is contracting with rate $1 - \mu_\infty(A)_+$
- implicit NN is well posed
- forward Euler is contracting with factor $1 - \dfrac{1-\mu_\infty(A)_+}{1-\min_i(a_{ii})_-}$ at $\alpha^* = \dfrac{1}{1-\min_i(a_{ii})_-}$
- input-state Lipschitz constant $\quad \mathrm{Lip}_{u\to x} = \dfrac{\|B\|_\infty}{1-\mu_\infty(A)_+}$

Outline

Solving optimization problems via dynamical systems



- studies in linear and nonlinear programming (Arrow, Hurwicz, and Uzawa 1958)
- neural networks (Hopfield and Tank 1985) and analog circuits (Kennedy and Chua 1988)
- optimization on manifolds (Brockett 1991)
- ...
- power grids (Bolognani, Carli, Cavraro, Zampieri 2013)
- online and dynamic feedback optimization (Dall'Anese, Dörfler, Simonetto, ...)

Convexity and contractivity

Kachurovskii's Theorem: For differentiable $f : \mathbb{R}^n \to \mathbb{R}$, equivalent statements:

1. f is **strongly convex** with parameter m
2. $-\nabla f$ is **(strongly) infinitesimally contracting** with respect to $\|\cdot\|_2$ with rate m

Also: global minimum of f = globally-exponentially stable equilibrium of $-\nabla f$

R. I. Kachurovskii. Monotone operators and convex functionals. *Uspekhi Matematicheskikh Nauk*, 15(4):213–215, 1960

From convex optimization to contracting dynamics – time-varying

Many convex optimization problems can be solved with contracting dynamics

$$\dot{x} = \mathsf{F}(x, \theta)$$

	Convex Optimization	Contracting Dynamics
Unconstrained	$\min_{x \in \mathbb{R}^n} \quad f(x, \theta)$	$\dot{x} = -\nabla_x f(x, \theta)$
Constrained	$\min_{x \in \mathbb{R}^n} \quad f(x, \theta)$ s.t. $\quad x \in \mathcal{X}(\theta)$	$\dot{x} = -x + \mathrm{Proj}_{\mathcal{X}(\theta)}(x - \gamma \nabla_x f(x, \theta))$
Composite	$\min_{x \in \mathbb{R}^n} \quad f(x, \theta) + g(x, \theta)$	$\dot{x} = -x + \mathrm{prox}_{\gamma g_\theta}(x - \gamma \nabla_x f(x, \theta))$
Equality	$\min_{x \in \mathbb{R}^n} \quad f(x, \theta)$ s.t. $\quad Ax = b(\theta)$	$\dot{x} = -\nabla_x f(x, \theta) - A^\top \lambda,$ $\dot{\lambda} = Ax - b(\theta)$
Inequality	$\min_{x \in \mathbb{R}^n} \quad f(x, \theta)$ s.t. $\quad Ax \leq b(\theta)$	$\dot{x} = -\nabla f(x, \theta) - A^\top \nabla M_{\gamma, b(\theta)}(Ax + \gamma\lambda),$ $\dot{\lambda} = \gamma(-\lambda + \nabla M_{\gamma, b(\theta)}(Ax + \gamma\lambda))$

Tracking equilibrium trajectories

For parameter-dependent vector field $\mathsf{F} : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^n$ and differentiable $\theta : \mathbb{R}_{\geq 0} \to \mathbb{R}^d$

$$\dot{x}(t) = \mathsf{F}(x(t), \theta(t))$$

Assume there exist norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\Theta}$ s.t.

- **contractivity wrt x**: $\mathrm{osLip}_x(\mathsf{F}) \leq -c < 0$, uniformly in u
- **Lipschitz wrt u**: $\mathrm{Lip}_u(\mathsf{F}) \leq \ell$, uniformly in x

Theorem: Incremental ISS any two soltns: $x(t)$ with input u_x and $y(t)$ with input u_y

$$D^+\|x(t) - y(t)\|_{\mathcal{X}} \;\leq\; -c\|x(t) - y(t)\|_{\mathcal{X}} \;+\; \ell\|u_x(t) - u_y(t)\|_{\Theta}$$

Tracking equilibrium trajectories

For parameter-dependent vector field $\mathsf{F} : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^n$ and differentiable $\theta : \mathbb{R}_{\geq 0} \to \mathbb{R}^d$

$$\dot{x}(t) = \mathsf{F}(x(t), \theta(t))$$

Assume there exist norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\Theta}$ s.t.

- **contractivity wrt x:** $\mathrm{osLip}_x(\mathsf{F}) \leq -c < 0$, uniformly in u
- **Lipschitz wrt u:** $\mathrm{Lip}_u(\mathsf{F}) \leq \ell$, uniformly in x

Theorem: Equilibrium tracking for contracting dynamics

1. for each fixed θ, there exists a unique equilbrium $x^\star(\theta)$
2. the equilibrium map $x^\star(\cdot)$ is Lipschitz with constant $\dfrac{\ell}{c}$
3. $\dfrac{d}{dt}\|x(t) - x^\star(\theta(t))\|_{\mathcal{X}} \leq -c\|x(t) - x^\star(\theta(t))\|_{\mathcal{X}} + \dfrac{\ell}{c}\|\dot{\theta}(t)\|_{\Theta}$

Consequences for tracking error

$$\frac{d}{dt}\|x(t)-x^\star(\theta(t))\|_{\mathcal{X}} \;\leq\; -c\|x(t)-x^\star(\theta(t))\|_{\mathcal{X}} \;+\; \frac{\ell}{c}\,\|\dot{\theta}(t)\|_{\Theta}$$

- bounded input, bounded error
  with asymptotic bound:

$$\limsup_{t\to\infty}\|x(t)-x^\star(\theta(t))\|_{\mathcal{X}} \quad\leq\quad \frac{\ell}{c^2}\,\limsup_{t\to\infty}\|\dot{\theta}(t)\|_{\Theta}$$

- bounded energy input, bounded energy error
- vanishing input, vanishing error
- exponentially vanishing input, exponentially vanishing error
- periodic input, periodic error

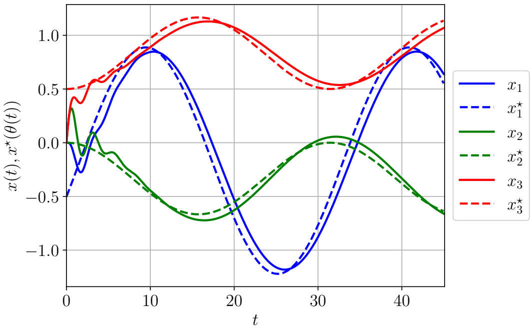
Numerical simulations

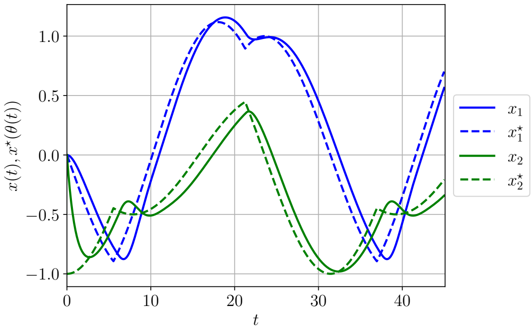
$$\min_{x\in\mathbb{R}^3} \quad \frac{1}{2}\|x-r(t)\|_2^2$$
$$\text{s.t.} \quad x_1+2x_2+x_3=\sin(\omega t),$$

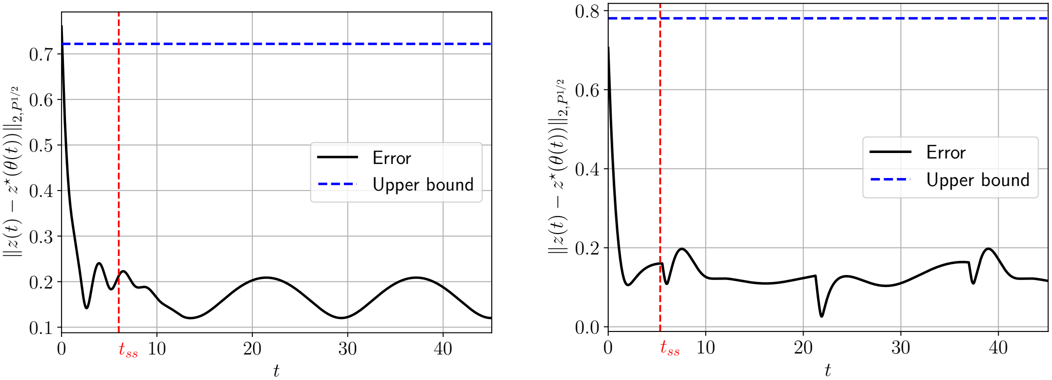
$$\min_{x\in\mathbb{R}^2} \quad \frac{1}{2}\|x+r(t)\|_2^2$$
$$\text{s.t.} \quad -x_1+x_2\le\cos(\omega t),$$

$$r(t)=(\sin(\omega t),\cos(\omega t),1),\omega=0.2$$

$$r(t)=(\sin(\omega t),\cos(\omega t)),\omega=0.2$$







$\|z(t) - z^\star(\theta(t))\|_{2,P^{1/2}}$
Error
Upper bound
t_{ss}
t

Summary and future work

Summary:

1. from convex optimization to contracting dynamics
2. tracking-bounds for time-varying contracting systems
3. applications to standard convex optimization problems

Ongoing work and open problems:

1. contracting predictor-corrector methods
2. tracking bounds in time-varying norms
3. convex but not strongly convex problems

Outline

Robust and computationally-friendly stability theory

1. contractivity conditions on normed vector spaces
2. application to recurrent and implicit neural networks
3. application to time-varying convex optimization

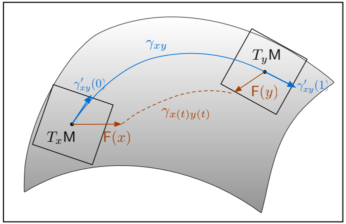
	Lyapunov Theory	Contraction Theory for Dynamical Systems
	F admits global Lyapunov function	F is strongly contracting
existence of equilibrium	assumed	implied + computational methods
Lyapunov function	arbitrary	$\|x - x^*\|$ and $\|\mathsf{F}(x)\|$
inputs	ISS via $\mathcal{KL}$ and $\mathcal{L}$ functions	iISS via explicit formulas

search for contraction properties

design engineering systems to be contracting

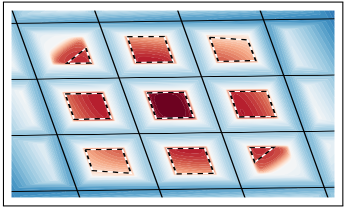
Theoretical frontiers

- higher order contraction
- relationship with monotone operator theory
- metric spaces: seminorms, Hilbert metrices ...



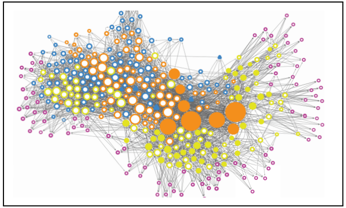
Limitations: not all stable systems are contractive:

- Lyapunov-diagonally-stable networks
- multistable and locally contracting systems
- biochemical networks
- control contraction design



Application to control and learning

1. control: optimization-based control design
2. ML: implicit models and energy-based learning
3. neuroscience: robust dynamical modeling



References (1/2)

Contraction theory on normed spaces:

- A. Davydov, S. Jafarpour, and F. Bullo. Non-Euclidean contraction theory for robust nonlinear stability. *IEEE Transactions on Automatic Control*, 67(12):6667–6681, 2022a.
- S. Jafarpour, A. Davydov, and F. Bullo. Non-Euclidean contraction theory for monotone and positive systems. *IEEE Transactions on Automatic Control*, 2023.

Contractivity in optimal control:

- K. D. Smith and F. Bullo. Contractivity of the method of successive approximations for optimal control. *IEEE Control Systems Letters*, 7:919–924, 2023.

Contracting neural networks and fixed point theory:

- S. Jafarpour, A. Davydov, A. V. Proskurnikov, and F. Bullo. Robust implicit networks via non-Euclidean contractions. In *Advances in Neural Information Processing Systems*, Dec. 2021.
- A. Davydov, A. V. Proskurnikov, and F. Bullo. Non-Euclidean contractivity of recurrent neural networks. In *American Control Conference*, pages 1527–1534, Atlanta, USA, May 2022b.
- F. Bullo, P. Cisneros-Velarde, A. Davydov, and S. Jafarpour. From contraction theory to fixed point algorithms on Riemannian and non-Euclidean spaces. In *IEEE Conf. on Decision and Control*, Dec. 2021.

References (2/2)

Equilibrium tracking for time-varing convex optimization:

- A. Davydov, V. Centorrino, A. Gokhale, G. Russo, and F. Bullo. Contracting dynamics for time-varying convex optimization. *Arxiv*, 2023.
- V. Centorrino, A. Gokhale, A. Davydov, G. Russo, and F. Bullo. Euclidean contractivity of neural networks with symmetric weights. *IEEE Control Systems Letters*, 2023.

Resources on contraction theory for dynamics, control and learning

1. free online book and 10h minicourse
   https://fbullo.github.io/ctds
   https://youtu.be/RvR47ZbqJjc
2. upcoming Workshop on "Contraction Theory for Systems, Control, and Learning" at the 2023 American Control Conference in San Diego, California:
   http://motion.me.ucsb.edu/contraction-workshop-2023