Theory and Applications of Contracting Dynamical Systems

Francesco Bullo

Center for Control,
Dynamical Systems & Computation
University of California at Santa Barbara
http://motion.me.ucsb.edu

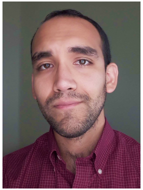
PhD program in "Modeling and Engineering Risk and Complexity"
Scuola Superiore Meridionale, Napoli, November 2022

Acknowledgments

Saber Jafarpour
GeorgiaTech

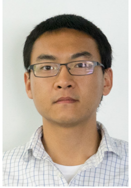
Alex Davydov
UC Santa Barbara

Anton Proskurnikov
Politecnico Torino



Pedro Cisneros-Velarde
University of Illinois

Kevin Smith
UC Santa Barbara



Xiaoming Duan
Shanghai Jiao Tong

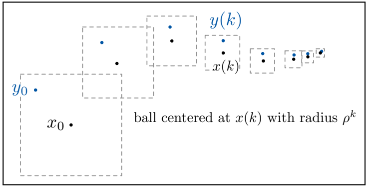
Veronica Centorrino
Scuola Sup Meridionale

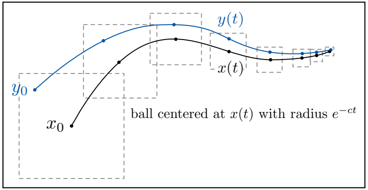
Giovanni Russo
Univ Salerno

Outline and Aims

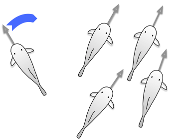
Dynamical Network Systems via Contraction Theory

1. structure and function of dynamical network systems
2. contractivity of dynamical systems
3. perspectives into artificial & biological neural networks

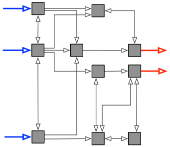




Structure and function for dynamical network systems

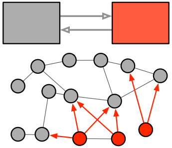


averaging

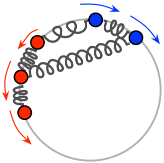


compartmental flows

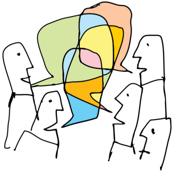
mutualism



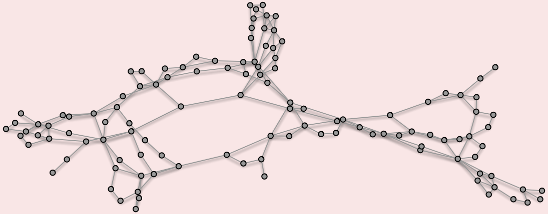
virus spread



coupled oscillators



social power

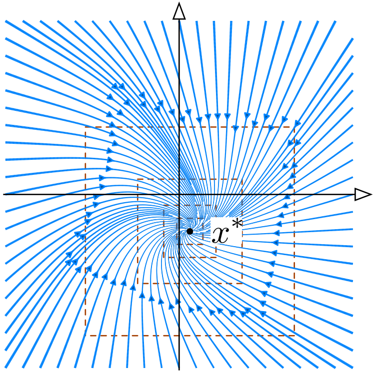


network structure $\Longleftrightarrow$ **function $=$ dynamic behavior**

Structure and function for dynamical network systems

function = dynamic behavior
highly-ordered transient and asymptotic behavior:

1. unique globally exponential stable equilibrium
   & two natural Lyapunov functions
2. robustness properties
   - bounded input, bounded output (iss)
   - 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



contracting dynamical systems

Contraction theory: historical notes

- **Origins**

- **Application in dynamics and control**:

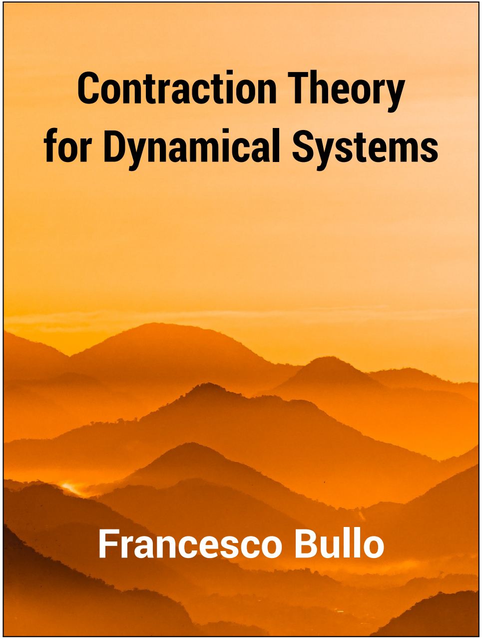
- **Reviews**:

The Banach Contraction Theorem is also referred to as the *Picard-Banach-Caccioppoli*, because of the earlier work by Picard (1890) on the "method of successive approximations" and the later independent work by Renato Caccioppoli (1930).

Figure: Renato Caccioppoli (Napoli, 20 gennaio 1904 – Napoli, 8 maggio 1959) was an Italian mathematician

1921-1932 student and researcher @ Napoli
1931-1934 professor @ Padova
1934-1959 professor @ Napoli

1. **Lotka-Volterra population dynamics** (??):
   ℓ_1-weakly contracting (after a rescaling change of coordinates)
2. **Matrosov-Bellman interconnected stable systems** (??):
   strongly contracting wrt composite norm
3. **Kuramoto coupled oscillators** (?):
   strongly semicontracting wrt (ℓ_2, Π_n) norm, in neighb'd of each phase-cohesive equilibrium
4. **Yorke multigroup SIS epidemic model** (?):
   equilibrium contracting wrt weighted ℓ_1/ℓ_∞ norms (at disease-free and endemic eq.)
5. **Hopfield and cellular neural networks** (?):
   ℓ_1/ℓ_∞-strongly contracting
6. **Daganzo cell transmission model for traffic networks** (?):
   ℓ_1-weakly contracting, when the dynamics is monotone
7. **Chua's diffusively-coupled dynamical systems** (?):
   strongly semi-contracting wrt $(2, p)$ tensor norm on $\mathbb{R}^n \otimes \mathbb{R}^k$
8. ...



Contraction Theory for Dynamical Systems, Francesco Bullo, KDP, 1.0 edition, 2022, ISBN 979-8836646806

1. Content:
(i) Banach contraction theorem and fixed point theory,
(ii) induced norms and induced log norms of matrices
(iii) strongly contracting dynamics over normed spaces,
(iv) weakly-contracting dynamics and monotone dynamics,
(v) semicontracting and partially contracting systems,

(vi) examples: Hopfield neural networks, systems in Lure' form, interconnected systems, gradient and primal dual flows of convex functions, Lotka-Volterra population dynamics, Daganzo traffic models, averaging flows, and diffusively-coupled synchronizing systems.

2. "Continuous improvement is better than delayed perfection"
Mark Twain

– Self-Published and Print-on-Demand at:
https://www.amazon.com/dp/B0B4K1BTF4
– PDF Freely available at
http://motion.me.ucsb.edu/book-ctds

Outline

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

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

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

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

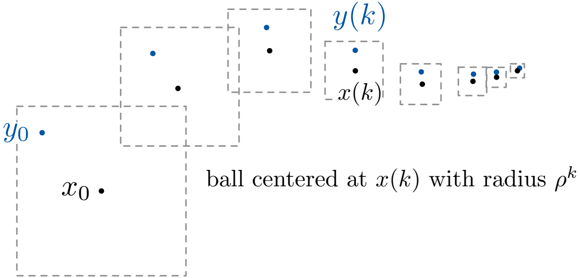
$$\|x\|_{2,P} = (x^\top P x)^{1/2} \qquad \mathsf{Lip}_{2,P}(\mathsf{F}_A) = \|A\|_{2,P} \leq \ell \qquad \iff \qquad A^\top P A \preceq \ell^2 P$$

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

Banach contraction theorem for discrete-time dynamics:

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

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



From induced norms to induced log norms

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

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

Basic properties:

subadditivity:	$\mu(A+B) \leq \mu(A) + \mu(B)$	
scaling:	$\mu(bA) = b\mu(A),$	$\forall b \geq 0$
convexity:	$\mu(\theta A + (1-\theta)B) \leq \theta\mu(A) + (1-\theta)\mu(B),$	$\forall \theta \in [0,1]$

spectral radius $\leq$ induced norm

spectral abscissa $\leq$ induced log norm

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

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

One-sided Lipschitz constant

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

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

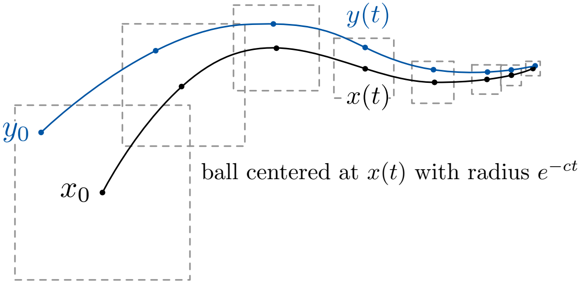
For **affine map** $\mathsf{F}_A(x) = Ax + a$

$$\begin{aligned}\mathsf{osLip}_{2,P}(\mathsf{F}_A) = \mu_{2,P}(A) \leq \ell \quad &\iff \quad A^\top P + AP \preceq 2\ell P \\ \mathsf{osLip}_{\infty,\eta}(\mathsf{F}_A) = \mu_{\infty,\eta}(A) \leq \ell \quad &\iff \quad a_{ii} + \sum_{j \neq i} |a_{ij}|\eta_i/\eta_j \leq \ell\end{aligned}$$

Banach contraction theorem for continuous-time dynamics:

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

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



From inner products to weak pairings

A **weak pairing** is $[\![\cdot,\cdot]\!] : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfying

1. $[\![x_1 + x_2, y]\!] \leq [\![x_1, y]\!] + [\![x_2, y]\!]$ and $x \mapsto [\![x, y]\!]$ is continuous,
2. $[\![bx, y]\!] = [\![x, by]\!] = b\,[\![x, y]\!]$ for $b \geq 0$ and $[\![-x, -y]\!] = [\![x, y]\!]$,
3. $[\![x, x]\!] > 0$, for all $x \neq \mathbb{0}_n$,
4. $|\,[\![x, y]\!]\,| \leq [\![x, x]\!]^{1/2}\,[\![y, y]\!]^{1/2}$,

Given norm $\|\cdot\|$, compatibility: $[\![x, x]\!] = \|x\|^2$ for all x

Key properties

Curve norm derivative formula: $\frac{1}{2}D^+\|x(t)\|^2 = [\![\dot{x}(t), x(t)]\!]$

Sup of non-Euclidean numerical range: $\mu(A) = \sup_{\|x\|=1} [\![Ax, x]\!]$

Example weak pairings

Norms	From inner products to sign and max pairings	From LMIs to log norms
$\|x\|_{2,P^{1/2}}^2 = x^\top P x$	$[\![x,y]\!]_{2,P^{1/2}} = x^\top P y$	$\mu_{2,P^{1/2}}(A) = \min\{b \mid A^\top P + PA \preceq 2bP\}$
$\|x\|_1 = \sum_i \|x_i\|$	$[\![x,y]\!]_1 = \|y\|_1 \operatorname{sign}(y)^\top x$	$\mu_1(A) = \max_j \Big(a_{jj} + \sum_{i \neq j} \|a_{ij}\|\Big)$
$\|x\|_\infty = \max_i \|x_i\|$	$[\![x,y]\!]_\infty = \max_{i \in I_\infty(y)} y_i x_i$	$\mu_\infty(A) = \max_i \Big(a_{ii} + \sum_{j \neq i} \|a_{ij}\|\Big)$

where $I_\infty(x) = \{i \in \{1,\dots,n\}$ such that $|x_i| = \|x\|_\infty\}$

Log Norm bound	Demidovich condition	One-sided Lipschitz condition
$\mu_{2,P}(D\mathsf{F}(x)) \leq b$	$PD\mathsf{F}(x) + D\mathsf{F}(x)^\top P \preceq 2bP$	$(x-y)^\top P\big(\mathsf{F}(x) - \mathsf{F}(y)\big) \leq b\|x-y\|_{P^{1/2}}^2$
$\mu_1(D\mathsf{F}(x)) \leq b$	$\mathrm{sign}(v)^\top D\mathsf{F}(x)v \leq b\|v\|_1$	$\mathrm{sign}(x-y)^\top(\mathsf{F}(x) - \mathsf{F}(y)) \leq b\|x-y\|_1$
$\mu_\infty(D\mathsf{F}(x)) \leq b$	$\max\limits_{i\in I_\infty(v)} v_i\,(D\mathsf{F}(x)v)_i \leq b\|v\|_\infty^2$	$\max\limits_{i\in I_\infty(x-y)}(x_i-y_i)(\mathsf{F}_i(x) - \mathsf{F}_i(y)) \leq b\|x-y\|_\infty^2$

Equivalent contractivity conditions

Background on one-sided Lipschitz continuity

contraction conditions without Jacobians have been studied under many different names:

1. **uniformly decreasing maps** in:
2. no-name in: (Chapter 1, page 5)
3. **one-sided Lipschitz maps** in: (Section 1.10, Exercise 6)
4. **maps with negative nonlinear measure** in:
5. **dissipative Lipschitz maps** in:
6. **maps with negative lub log Lipschitz constant** in:
7. **QUAD maps** in:
8. **incremental quadratically stable maps** in:

Outline

Fot time and input-dependent vector F,

$$\dot{x} = \mathsf{F}(t, x, u(t)), \qquad x(0) = x_0 \in \mathcal{X}, \qquad u(t) \in \mathcal{U} \tag{1}$$

Given norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{U}}$, assume constants $c, \ell > 0$ s.t.

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

Then

1. any soltns: $x(t)$ with input u_x and $y(t)$ with input u_y

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

2. F is **incrementally ISS**, that is, for all x_0, y_0

$$\|x(t)-y(t)\|_{\mathcal{X}} \;\le\; e^{-ct}\|x_0-y_0\|_{\mathcal{X}} \;+\; \frac{\ell(1-e^{-ct})}{c}\sup_{\tau\in[0,t]}\|u_x(\tau)-u_y(\tau)\|_{\mathcal{U}}$$

3. F has **incremental $\mathcal{L}^q_{\mathcal{X},\mathcal{U}}$ gain equal to ℓ/c, for $q\in[1,\infty]$**,

$$\|x(\cdot)-y(\cdot)\|_{\mathcal{X},q} \;\le\; \frac{\ell}{c}\,\|u_x(\cdot)-u_y(\cdot)\|_{\mathcal{U},q} \qquad \text{(for } x_0=y_0\text{)}$$

Given norm $\|\cdot\|_{\mathcal{X}}$ on $\mathbb{R}^n$ (or $\|\cdot\|_{\mathcal{U}}$ on $\mathbb{R}^k$),

- $\mathcal{L}^q_{\mathcal{X}}$, $q \in [1, \infty]$, is vector space of continuous signals, $x : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$, with well-defined bounded norm

$$\|x(\cdot)\|_{\mathcal{X},q} = \begin{cases} \left(\int_0^\infty \|x(t)\|^q_{\mathcal{X}} dt \right)^{1/q} & \text{if } q \in [1, \infty[\\ \sup_{t \geq 0} \|x(t)\|_{\mathcal{X}} & \text{if } q = \infty \end{cases} \tag{2}$$

- Input-state system has **$\mathcal{L}^q_{\mathcal{X},\mathcal{U}}$-induced gain** upper bounded by $\gamma > 0$ if, for all $u \in \mathcal{L}^q_{\mathcal{U}}$, the state x from zero initial state satisfies

$$\|x(\cdot)\|_{\mathcal{X},q} \leq \gamma \, \|u(\cdot)\|_{\mathcal{U},q} \tag{3}$$

From nominal to uncertain systems

Given a norm $\|\cdot\|$, consider

$$\dot{x} = \mathsf{F}(t,x) + \mathsf{G}(t,x) \tag{4}$$

Assume:

- $\mathrm{osLip}_x(\mathsf{F}) \leq -c < 0$
- $\mathrm{osLip}_x(\mathsf{G}) \leq d$

Then

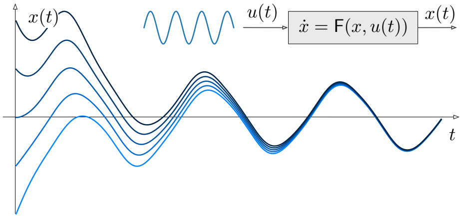
1. **(contractivity under perturbations)** if $d < c$,
   then $\mathsf{F}+\mathsf{G}$ is strongly contracting with rate $c-d$,
2. **(equilibria under perturbations)** if additionally F and G are time-invariant, then the unique equilibrium points x^* of F and x^{**} of $\mathsf{F}+\mathsf{G}$ satisfy

$$\|x^* - x^{**}\| \leq \frac{\|\mathsf{G}(x^*)\|}{c-d} \tag{5}$$

From time-invariant to periodic systems

For time-varying vector field F and norm $\|\cdot\|$

1. $\mathrm{osLip}_x(\mathsf{F}) \leq -c < 0$
2. F is T-periodic



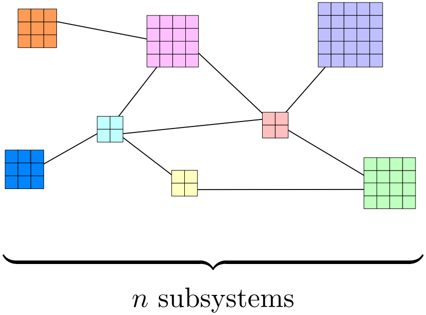
Then

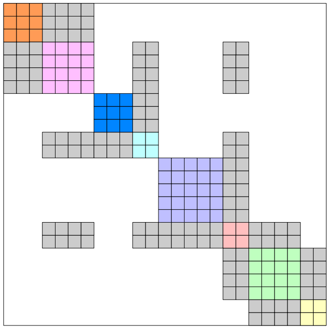
1. there exists a unique periodic solution $x^* : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ with period T
2. for every initial condition x_0,

$$\|x(t, x_0) - x^*(t)\| \leq \mathrm{e}^{-ct}\|x_0 - x^*(0)\| \tag{6}$$

Outline

Composite norms





1. n *local norms* $\|\cdot\|_i$ on $\mathbb{R}^{N_i}$
2. an *aggregating norm* $\|\cdot\|_{\text{agg}}$ on $\mathbb{R}^n$
3. *composite norm*

Networks of contracting systems

Interconnected subsystems: $x_i \in \mathbb{R}^{N_i}$ and $x_{-i} \in \mathbb{R}^{N-N_i}$:

$$\dot{x}_i = \mathsf{F}_i(x_i, x_{-i}), \qquad \text{for } i \in \{1, \dots, n\}$$

Network contraction theorem

- **osLip wrt x_i**: $\mathrm{osLip}_{x_i}(\mathsf{F}_i) \leq -c_i$, uniformly in x_{-i}
- **Lip wrt to x_j**: $\mathrm{Lip}_{x_j}(\mathsf{F}_i) \leq \ell_{ij}$, uniformly in x_{-j}
- the Lipschitz constants matrix $\begin{bmatrix} -c_1 & \dots & \ell_{1n} \\ \vdots & & \vdots \\ \ell_{n1} & \dots & -c_n \end{bmatrix}$ is **Hurwitz**

$\Longrightarrow$ the **interconnected system** is infinitesimally contracting

The network science of Metzler Hurwitz matrices

$$\begin{bmatrix} -c_1 & \dots & \ell_{1n} \\ \vdots & & \vdots \\ \ell_{n1} & \dots & -c_n \end{bmatrix} \text{ is } \textbf{Metzler} \text{ (so that Perron-Frobenius Theorem applies)}$$

Hurwitzness depends upon both topology and edge weights

- directed acyclic interconnections of contracting systems are strongly contracting
- For $n = 2$, Hurwitz if and only if **small gain condition**

$$\text{cycle gain} := \frac{\ell_{12}}{c_1}\frac{\ell_{21}}{c_2} < 1$$

- For $n \geq 3$, Hurwitz if and only if **network small-gain theorem for Metzler matrices**

Hurwitz Metzler Theorem

1. M is Hurwitz,
2. there exists $\eta \in \mathbb{R}^n_{>0}$ such that $\eta^\top M < \mathbb{0}_n^\top$ or, equivalently, $\mu_{1,[\eta]}(M) < 0$,
3. there exists $\xi \in \mathbb{R}^n_{>0}$ such that $M\xi < \mathbb{0}_n$ or, equivalently, $\mu_{\infty,[\xi]^{-1}}(M) < 0$, and
4. there exists a diagonal $P = P^\top \succ 0$ satisfying $M^\top P + PM \prec 0$ or, equivalently, $\mu_{2,P^{1/2}}(M) < 0$.

Input: a Metzler matrix $M \in \mathbb{R}^{n\times n}$
Output: polynomials $\{\gamma_{\mathcal{C}_2}, \dots, \gamma_{\mathcal{C}_n}\}$ in entries of M

1: $\mathcal{C} :=$ set of simple cycles of digraph associated to M
2: $\gamma_\phi :=$ gain of cycle $\phi \in \mathcal{C}$
3: **for** i from 2 to n
4: $\quad \mathcal{C}_i :=$ cycles in $\mathcal{C}$ passing through only nodes $1, \dots, i$
5: $\quad \gamma_{\mathcal{C}_i} := \sum_{\phi\in\mathcal{C}_i} \gamma_\phi - \sum_{\substack{\phi,\psi\in\mathcal{C}_i \\ \phi\perp\psi}} \gamma_\phi\gamma_\psi + \sum_{\substack{\phi,\psi,\rho\in\mathcal{C}_i \\ \phi\perp\psi,\, \phi\perp\rho,\, \psi\perp\rho}} \gamma_\phi\gamma_\psi\gamma_\rho - \cdots$

Network small-gain theorem for Metzler matrices

$$\text{Metzler } M \text{ is Hurwitz} \iff \gamma_{\mathcal{C}_2} < 1, \cdots, \gamma_{\mathcal{C}_n} < 1$$

- not unique: distinct/equivalent conditions after renumbering, redundancy
- computational efficiency: after precomputation of simple cycles

$$M = \begin{bmatrix} -c_1 & 0 & 0 & \ell_{14} \\ 0 & -c_2 & \ell_{23} & \ell_{24} \\ 0 & \ell_{32} & -c_3 & \ell_{34} \\ \ell_{41} & \ell_{42} & \ell_{43} & -c_4 \end{bmatrix}$$

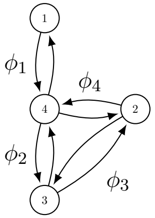


Figure: associated digraph and simple cycles

- $\gamma_{\phi_1} = \frac{\ell_{14}\ell_{41}}{c_1 c_4}$, $\gamma_{\phi_2} = \frac{\ell_{34}\ell_{43}}{c_3 c_4}$, $\gamma_{\phi_3} = \frac{\ell_{23}\ell_{32}}{c_2 c_3}$, and $\gamma_{\phi_4} = \frac{\ell_{24}\ell_{42}}{c_2 c_4}$
- $\mathcal{C}_2 = \emptyset$
- $\mathcal{C}_3 = \{\phi_3\}$: $\gamma_{\mathcal{C}_3} = \gamma_{\phi_3} < 1$ (redundant)
- $\mathcal{C}_4 = \{\phi_1, \ldots, \phi_4\}$: $\gamma_{\mathcal{C}_4} = \sum_{i=1}^{4} \gamma_{\phi_i} - \gamma_{\phi_1}\gamma_{\phi_3} < 1$

$$\begin{bmatrix} -c_1 & 0 & 0 & 0 & \ell_{15} & \ell_{16} \\ 0 & -c_2 & 0 & \ell_{24} & \ell_{25} & 0 \\ 0 & 0 & -c_3 & \ell_{34} & 0 & \ell_{36} \\ 0 & \ell_{42} & \ell_{43} & -c_4 & 0 & 0 \\ \ell_{51} & \ell_{52} & 0 & 0 & -c_5 & 0 \\ \ell_{61} & 0 & \ell_{63} & 0 & 0 & -c_6 \end{bmatrix}$$

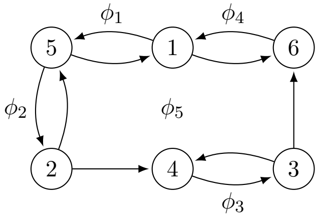
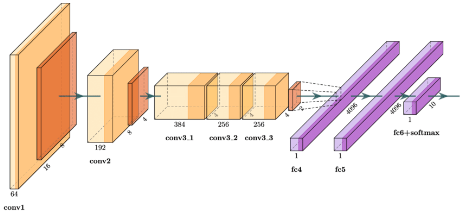


Figure: associated digraph and simple cycles

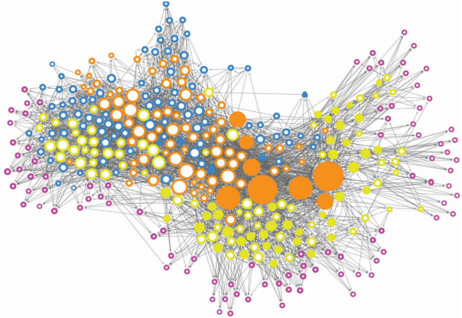
- $\mathcal{C}_2$, $\mathcal{C}_3$ empty
- $\mathcal{C}_4 = \{\phi_3\}$: $\gamma_3 < 1$ (redundant)
- $\mathcal{C}_5 = \{\phi_1, \phi_2, \phi_3\}$: $\gamma_{\mathcal{C}_5} = \gamma_1 + \gamma_2 + \gamma_3 - \gamma_1\gamma_3 - \gamma_2\gamma_3 < 1$
- $\mathcal{C}_6 = \{\phi_1, \ldots, \phi_5\}$: $\gamma_{\mathcal{C}_6} = \sum_{i=1}^{5} \gamma_i - \gamma_1\gamma_3 - \gamma_2\gamma_3 - \gamma_3\gamma_4 - \gamma_2\gamma_4 + \gamma_2\gamma_3\gamma_4 < 1$

Outline

Artificial and biological neural networks



artificial neural network AlexNet ’12



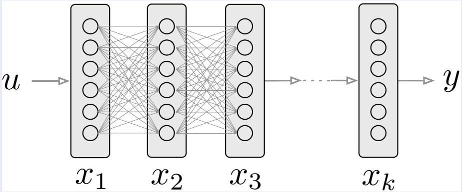
C. elegans connectome ’17

Aim: understand the dynamics and functionality of neural networks, so that

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

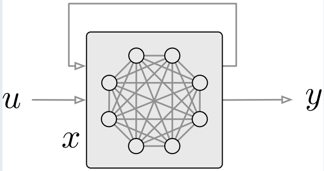
Artificial and biological neural networks – mathematization

Feedforward NN



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

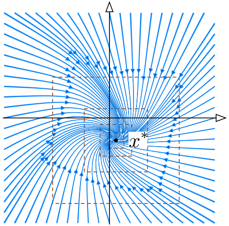
Implicit/Recurrent NN



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

Aim:

- well-posedness of the static model
- dynamic input/output models
- highly-ordered transient+asymptotic dynamic behavior
- biologically-plausible optimization



From continuous-time contracting dynamics to discrete-time computation

$$x = \mathsf{G}(x)$$

Banach contraction theorem

If $\mathrm{Lip}(\mathsf{G}) < 1$ that is $\|\mathsf{G}(u) - \mathsf{G}(v)\| \leq \mathrm{Lip}(\mathsf{G})\|u - v\|$,

then *Picard iteration* $x_{k+1} = \mathsf{G}(x_k)$ is a Banach contraction

For $\mathrm{Lip}(\mathsf{G}) \geq 1$, define the *average iteration*

$$x_{k+1} = (1 - \alpha)x_k + \alpha \mathsf{G}(x_k)$$

Infinitesimal Contraction Theorem

1. there exists $0 < \alpha < 1$ such that the average iteration is a Banach contraction
2. the map G satisfies $\mathrm{osLip}(\mathsf{G}) < 1$
3. the dynamics $\dot{x} = -x + \mathsf{G}(x)$ is infinitesimally contracting

Average iteration on the inner product space $(\mathbb{R}^n, \ell_2)$

Given $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}^n$

$$x^* \in \mathrm{zero}(\mathsf{F}) \qquad \Longleftrightarrow \qquad x^* \in \mathrm{fixed}(G), \text{ where } \mathsf{G} = \mathsf{Id} + \mathsf{F}$$

consider **forward step = Euler integration** for F $\quad$ = average iteration for G:

$$x_{k+1} = (\mathsf{Id} + \alpha \mathsf{F}) x_k = x_k + \alpha \mathsf{F}(x_k) \qquad = (1-\alpha)\,\mathsf{Id} + \alpha \mathsf{G}$$

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

1. the map $\mathsf{Id} + \alpha \mathsf{F}$ is a contraction map with respect to $\|\cdot\|_{2,P^{1/2}}$ for

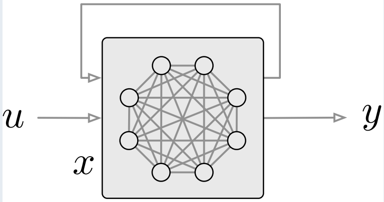
$$0 < \alpha < \frac{2}{c\kappa^2}$$

2. the optimal step size minimizing and minimum contraction factor:

$$\alpha^*_{\mathsf{E}} = \frac{1}{c\kappa^2}$$

$$\ell^*_{\mathsf{E}} = 1 - \frac{1}{2\kappa^2} + \mathcal{O}\Big(\frac{1}{\kappa^4}\Big)$$

ℓ_∞-contracting recurrent neural networks



$$x = \Phi(Wx + Bu) \qquad (\textit{fixed point})$$

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

$$x_{k+1} = (1-\alpha)x_k + \alpha\Phi(Wx_k + Bu) \qquad (\textit{average iter.n})$$

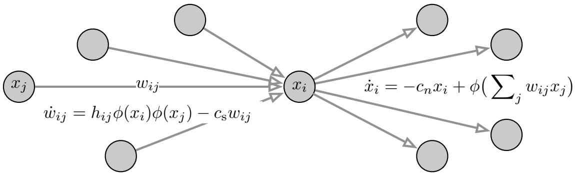
Maximizing a convex function over polytope:

$$\mathsf{osLip}_\infty\big(-x + \Phi(Wx + Bu)\big) = \sup_{x,u} \mu_\infty(-I_n + D\Phi \cdot W) = -1 + \mu_\infty(W)_+$$

If $\mu_\infty(W) < 1$ $\left(\text{i.e., } a_{ii} + \sum_j |a_{ij}| < 1 \text{ for all } i\right)$

- dynamics is contracting with rate $1 - \mu_\infty(W)_+$
- average iteration is Banach with factor $1 - \dfrac{1 - \mu_\infty(W)_+}{1 - \min_i(a_{ii})_-}$ at $\alpha = \dfrac{1}{1 - \min_i(a_{ii})_-}$

Coupled neural-synaptic networks with Hebbian learning



coupled neural-synaptic dynamics

$$x_i = -c_{\mathsf{n}} x_i + \Phi\Big(\sum_{j=1}^{n} w_{ij} x_j + u_i\Big),$$

$$\dot{w}_{ij} = h_{ij}\Phi(x_i)\Phi(x_j) - c_{\mathsf{s}} w_{ij} + U_{ij}$$

network small gain condition:

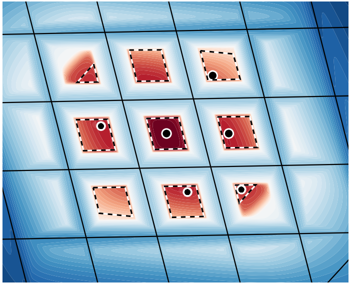
$c_{\mathsf{n}} c_{\mathsf{s}} >$ interconnection strength

Outline

Conclusions and future research

Contraction theory for dynamical system

1. from discrete-time to continuous-time
2. from single system to networks of systems
3. Metzler Hurwitz, fixed point computation, ...
4. applications to neural networks



Future work

1. open problems
   1. local contractivity in multistable systems
   2. network theory of Metzler Hurwitz matrices
   3. contractivity of Lyapunov-diagonally-stable neural networks
2. applications to networks, control and optimization
3. learning strategies in neuroscience and ML

References

Contraction theory on normed spaces:

Contracting neural networks:

Tutorial, text and extensions:

Advanced Topics

Acknowledgments

Giulia De Pasquale
ETH

Elena Valcher
Universita di Padova

John W. Simpson-Porco
University of Toronto

Outline

Optimization and Fixed Point Theory

For differentiable $V : \mathbb{R}^n \to \mathbb{R}$, equivalent statements:

1. V is **strongly convex** with parameter m
2. $-\operatorname{grad} V$ is m-**strongly infinitesimally contracting**, that is

$$\big(-\operatorname{grad} V(x) + \operatorname{grad} V(y)\big)^\top (x-y) \leq -m\|x-y\|_2^2$$

For map $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}^n$, equivalent statements:

1. F is a **monotone operator**[a] (or a **coercive operator**) with parameter m,
2. $-\mathsf{F}$ is m-**strongly contracting**

[a] $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}^n$ is a *monotone operator* if $\langle\!\langle \mathsf{F}(x) - \mathsf{F}(y), x-y \rangle\!\rangle \geq 0$

On fixed point algorithms and Banach contractions

$$x = \mathsf{G}(x)$$

Banach Contraction Theorem

If $\mathsf{Lip}(\mathsf{G}) < 1$ that is $\|\mathsf{G}(u) - \mathsf{G}(v)\| \leq \mathsf{Lip}(\mathsf{G})\|u - v\|$,
then *Picard iteration* $x_{k+1} = \mathsf{G}(x_k)$ is a Banach contraction

For $\mathsf{Lip}(\mathsf{G}) \geq 1$, define the *average iteration*

$$x_{k+1} = (1-\alpha)x_k + \alpha\mathsf{G}(x_k)$$

Infinitesimal Contraction Theorem

1. there exists $0 < \alpha < 1$ such that the average iteration is a Banach contraction
2. the map G satisfies $\mathsf{osLip}(\mathsf{G}) < 1$
3. the dynamics $\dot{x} = -x + \mathsf{G}(x)$ is infinitesimally strongly contracting

Robustness of fixed point algorithms

Robustness based upon Contraction

x_u^* is a fixed point of $x = \mathsf{G}(x,u)$ and $\mathsf{Lip}_x \mathsf{G} < 1$, then

$$\|x_u^* - x_v^*\| \leq \frac{\mathsf{Lip}_u \mathsf{G}}{1 - \mathsf{Lip}_x \mathsf{G}} \|u - v\|$$

Robustness based upon Infinitesimal Contraction

x_u^* is a fixed point of $x = \mathsf{G}(x,u)$
x_v^* is a fixed point of $x = \mathsf{G}(x,v) + \mathsf{D}(x,v)$, and
$\mathsf{osLip}_x(\mathsf{G}+\mathsf{D}) < 1$, then

$$\|x_u^* - x_v^*\| \leq \frac{1}{1 - \mathsf{osLip}_x(\mathsf{G}+\mathsf{D})} \Big(\mathsf{Lip}_u(\mathsf{G}+\mathsf{D}) \|u - v\| + \|\mathsf{D}(x_u^*, u)\| \Big)$$

Average iteration on the inner product space $(\mathbb{R}^n, \ell_2)$

Given $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}^n$

$$x^* \in \text{zero}(\mathsf{F}) \qquad \iff \qquad x^* \in \text{fixed}(G), \text{ where } \mathsf{G} = \mathsf{Id} + \mathsf{F}$$

consider **forward step = Euler integration** for F = average iteration for G:

$$x_{k+1} = (\mathsf{Id} + \alpha \mathsf{F}) x_k = x_k + \alpha \mathsf{F}(x_k) \qquad = (1-\alpha)\,\mathsf{Id} + \alpha \mathsf{G}$$

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

1. the map $\mathsf{Id} + \alpha \mathsf{F}$ is a contraction map with respect to $\|\cdot\|_{2,P^{1/2}}$ for

$$0 < \alpha < \frac{2}{c\kappa^2}$$

2. the optimal step size minimizing and minimum contraction factor:

$$\alpha_{\mathsf{E}}^* = \frac{1}{c\kappa^2}$$
$$\ell_{\mathsf{E}}^* = 1 - \frac{1}{2\kappa^2} + \mathcal{O}\Big(\frac{1}{\kappa^4}\Big)$$

Average iteration on the normed vector spaces $(\mathbb{R}^n, \ell_1/\ell_\infty)$

Consider a norm $\|\cdot\|$ with compatible weak pairing $[\![\cdot,\cdot]\!]$
Recall **forward step method** $\quad x_{k+1} = (\mathsf{Id} + \alpha \mathsf{F}) x_k = x_k + \alpha \mathsf{F}(x_k)$

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

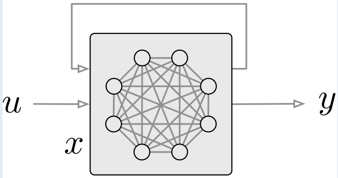
1. the map $\mathsf{Id} + \alpha \mathsf{F}$ is a contraction map with respect to $\|\cdot\|$ for

$$0 < \alpha < \frac{1}{c\kappa(1+\kappa)}$$

2. the optimal step size minimizing and minimum contraction factor:

$$\alpha^*_{\mathsf{nE}} = \frac{1}{c}\Big(\frac{1}{2\kappa^2} - \frac{3}{8\kappa^3} + \mathcal{O}\Big(\frac{1}{\kappa^4}\Big)\Big)$$
$$\ell^*_{\mathsf{nE}} = 1 - \frac{1}{4\kappa^2} + \frac{1}{8\kappa^3} + \mathcal{O}\Big(\frac{1}{\kappa^4}\Big)$$

Application: ℓ_∞-contracting neural networks



$$
\begin{aligned}
x &= \Phi(Ax + Bu + b) && (\textit{INN fixed point})\\
\dot{x} &= -x + \Phi(Ax + Bu + b) && (\textit{Recurrent NN})\\
x_{k+1} &= (1-\alpha)x_k + \alpha\Phi(Ax_k + Bu + b) && (\textit{Average iter.n})
\end{aligned}
$$

If

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

- dynamics is contracting with rate $1 - \mu_\infty(A)_+$
- average iteration is Banach with factor $1 - \dfrac{1-\mu_\infty(A)_+}{1-\min_i(a_{ii})_-}$ at $\alpha = \dfrac{1}{1-\min_i(a_{ii})_-}$
- input-output Lipschitz constant $\mathrm{Lip}_{u\to y} = \dfrac{\|B\|_\infty \|C\|_\infty}{1-\mu_\infty(A)_+}$

Background on Infinitesimal Contraction Theorem

1. there exists $0 < \alpha < 1$ such that the average iteration is a Banach contraction
2. the map G satisfies $\mathrm{osLip}(\mathsf{G}) < 1$
3. the dynamics $\dot{x} = \mathsf{F}(x) := -x + \mathsf{G}(x)$ is infinitesimally contracting

- the equivalence (2) $\iff$ (3) is just a transcription:
  - $\mathsf{F} = -\mathrm{Id} + \mathsf{G}$ contracting with rate $c \iff \mathrm{osLip}(\mathsf{F}) < -c \iff \mathrm{osLip}(\mathsf{G}) < 1 - c$, for $c > 0$
  - in (ℓ_2, P), $\mathrm{osLip}(\mathsf{F}) < -c$ is usual Krasovskii: $PJ(x) + J(x)^\top P \preceq -2cP$ for all x and $J = D\mathsf{F}$
- (2) $\implies$ (1): known in monotone operator theory (page 15 “forward step method” in[1])
  - vector field F is contracting with rate $c \iff -\mathsf{F}$ is strongly monotone with parameter c
- Theorem 1 in[2] proves the equivalence (1) $\iff$ (2) for any norm, i.e., the implication (2) $\implies$ (1) for any norm (with proper osLip definitions) and the converse direction (1) $\implies$ (2) for ℓ_2, P. Theorem 3 in[2] proves partly the “Robustness based upon infinitesimal contraction”.

1

2

Outline

Euclidean vs. non-Euclidean contractions

Most foundational results in systems theory are based on ℓ_2 linear-quadratic theory;
their ℓ_1/ℓ_∞ analogs are yet to be worked out.

NonEuclidean contractions: biological transcriptional systems (**?**), Hopfield neural networks (**??**), chemical reaction networks (**?**), traffic networks (**???**), multi-vehicle systems (**?**), and coupled oscillators (**??**)

Advantages of non-Euclidean approaches

1. *especially well suited for certain class of systems*
2. *computational advantages*: non-Euclidean log-norm constraints lead to LPs, whereas ℓ_2 constraints leads to LMIs. Parametrization of log-norm constrained matrices is polytopic.
3. *guaranteed robustness to structural perturbations*: ℓ_∞ contractivity ensures:
   1. absolute contractivity = with respect to a class of activation functions
   2. total contractivity = remove any node and all its incident connections
   3. connective contractivity = remove any set of edges
4. *adversarial input-output analysis*
   ℓ_∞ better suited for the analysis of adversarial examples than ℓ_2: in high dimensions, large inner product between two vectors is possible even when one vector has small ℓ_∞ norm
5. *fully asynchronous distributed model*: ℓ_∞ contractions

Outline

Semicontraction Theory, Dual Seminorms and Ergodicity

Semicontraction: history and setup

For *row-stochastic* A, consider averaging and dynamical flow systems:

$$x(k+1) = Ax(k) \qquad \text{(averaging flow, consensus algorithm)}$$
$$\pi(k+1) = A^\top \pi(k) \qquad \text{(dynamical flow system, Markov chain)}$$

Similarly, let L be a Laplacian matrix and consider the continuous-time counterparts:

$$\dot{x}(t) = -Lx(t) \qquad \text{(Laplacian flow)}$$
$$\dot{\pi}(t) = -L^\top \pi(t) \qquad \text{(continuous-time Markov chain, routing dynamics)}$$

For row-stochastic A, define the *Markov-Dobrushin ergodic coefficient*:

$$\tau_1(A) := \max_{\|z\|_1=1,\, z^\top \mathbb{1}_n=0} \|A^\top z\|_1$$

Simple calculations and remarkable similarity

For $\pi(k+1) = A^\top \pi(k)$, Markov showed any two solutions $\pi(k), \sigma(k)$ satisfy

$$d_{\mathrm{TV}}\big(\pi(k) - \sigma(k)\big) \leq \ \tau_1(A)^k \ d_{\mathrm{TV}}\big(\pi(0) - \sigma(0)\big) \tag{9}$$

$$d_{\mathrm{TV}}(\pi, \sigma) = \tfrac{1}{2} \sum\nolimits_i |\pi_i - \sigma_i|$$

(**total variation distance**)

For $x(k+1) = Ax(k)$, it is known in the consensus literature that

$$\|x(k)\|_{\mathrm{dist},\infty} \leq \ \tau_1(A)^k \ \|x(0)\|_{\mathrm{dist},\infty} \tag{10}$$

$$\|x\|_{\mathrm{dist},\infty} = \tfrac{1}{2}\Big(\max_i\{x_i\} - \min_j\{x_j\}\Big)$$

(**disagreement seminorm**)

Open questions

1. Why is the same ergodic coefficient τ_1 relevant for the contraction properties of both dynamical flows and averaging? Is it the tightest such bound?
2. What is the relationship between d_{TV} and $|||\cdot|||_{\mathrm{dist},\infty}$? How does one generalize bounds (**??**) and (**??**) to τ_p ergodic coefficients defined wrt ℓ_p norms (instead of ℓ_1 in (**??**))?
3. What are canonical Lyapunov functions for both systems, whose discrete-time variation along the flow is described by $\tau_p(A)$?
4. How does one define ergodic coefficients for continuous-time systems?
5. Is there a *contraction theoretic framework* that applies to time-varying and nonlinear systems with generalized invariance or conservation properties?

Seminorms

A function $|||\cdot||| : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is a *seminorm* on $\mathbb{R}^n$ if, for all $x, y \in \mathbb{R}^n$ and $a \in \mathbb{R}$:

$$\begin{aligned} &\text{(homogeneity):} \quad |||ax||| = |a|\,|||x|||, \text{ and} \\ &\text{(subadditivity):} \quad |||x+y||| \leq |||x||| + |||y|||. \end{aligned}$$

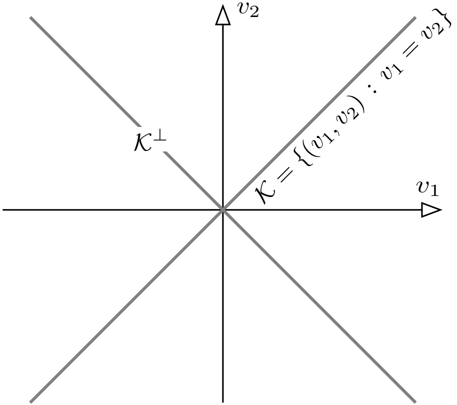
The *kernel* of $|||\cdot|||$ is the vector subspace $\mathcal{K} = \ker(|||\cdot|||) = \{x \in \mathbb{R}^n : |||x||| = 0\}$

1. *dual seminorm* is the function $|||\cdot|||_\star : V^\star \to \mathbb{R}$ defined by

$$|||y|||_\star \triangleq \max_{\substack{|||x||| \leq 1 \\ x \perp \mathcal{K}}} \langle y, x \rangle$$

2. *induced matrix seminorm on $\mathbb{R}^{n \times n}$* $|||\cdot||| : n \times n \to \mathbb{R}_{\geq 0}$ is

$$|||A||| \triangleq \max_{\substack{|||x||| \leq 1 \\ x \perp \mathcal{K}}} |||Ax|||$$



- On $\mathbb{R}^2$, the function $(v_1, v_2) \mapsto \sqrt{(v_1 - v_2)^2} = |v_1 - v_2|$ is seminorm
- $\mathcal{K} = \{(v_1, v_2)$ such that $v_1 = v_2\} = \mathrm{span}\{(1,1)^\top\}$ and $\mathcal{K}^\perp = \mathrm{span}\{(1,-1)^\top\}$.
- The orthogonal projection matrices onto $\mathcal{K}$ and $\mathcal{K}^\perp$ are

$$\Pi_{\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \Pi_{\perp} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Projection and distance seminorms

Given any subspace $\mathcal{K}$, let

$$\Pi_\perp \in \mathbb{R}^{n\times n} \;:=\; \text{orthogonal projection onto } \mathcal{K}^\perp$$

For each $p \in [1,\infty]$, define the *projection seminorm*

$$|||x|||_{\mathrm{proj},p} \triangleq \|\Pi_\perp x\|_p$$

and the *distance seminorm*

$$|||x|||_{\mathrm{dist},p} \triangleq \mathrm{dist}_p(x,\mathcal{K}) = \min_{u\in\mathcal{K}} \|x-u\|_p.$$

Consensus seminorm = a seminorm with kernel $\mathcal{K} = \operatorname{span}\{\mathbb{1}_n\}$

1. define $x_{\text{avg}} = \frac{1}{n}\mathbb{1}_n^\top x$:

$$|||x|||_{\text{proj},1} = \sum_{i=1}^{n} |x_i - x_{\text{avg}}|, \qquad |||x|||_{\text{proj},\infty} = \max_i |x_i - x_{\text{avg}}|$$

$$|||x|||_{\text{proj},2} = \Big(\frac{1}{n}\sum\nolimits_{i,j}(x_i - x_j)^2\Big)^{1/2}$$

2. sort the entries of x according to $x_{(1)} \geq x_{(2)} \geq \cdots \geq x_{(n)}$:

$$|||x|||_{\text{dist},1} = \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} x_{(i)} - \sum_{i=\lceil \frac{n}{2} \rceil+1}^{n} x_{(i)}, \qquad |||x|||_{\text{dist},\infty} = \frac{1}{2}\left(x_{(1)} - x_{(n)}\right).$$

$$|||x|||_{\text{dist},2} = \Big(\frac{1}{n}\sum\nolimits_{i,j}(x_i - x_j)^2\Big)^{1/2}$$

Total variation d_{TV} and ℓ_1 projection seminorm: $d_{\text{TV}}(x,y) = \frac{1}{2}|||x-y|||_{\text{proj},1}$ for $x, y \in \Delta_n$

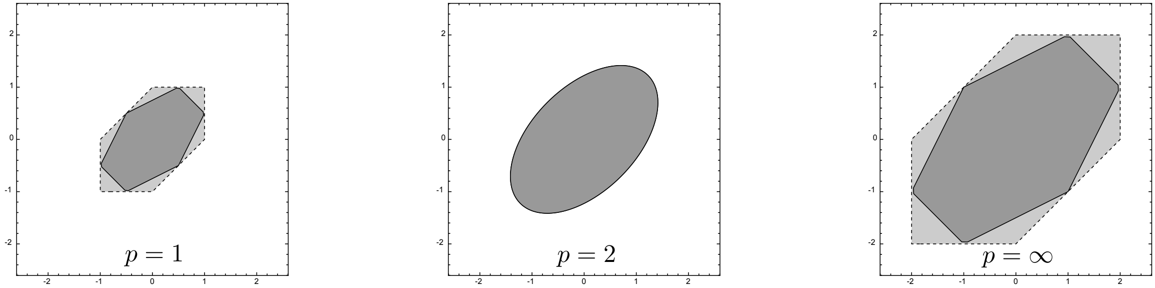


Figure: Two-dimensional sections of three-dimensional unit disks of projection (solid contours) and distance (dashed contours) consensus seminorms. We plot the sections corresponding to $(x_1, x_2, x_3 = 0)$.

1. ℓ_p and ℓ_q norms are dual, for $1/p + 1/q = 1$

$$|||\cdot|||_p = (|||\cdot|||_q)_\star \qquad\qquad |||\cdot|||_q = (|||\cdot|||_p)_\star$$

2. dual norm satisfies (sharp) *Hölder inequality*: $x^\top y \leq \|x\|_p \|y\|_q$
3. dual norm induces duality: $\|A\|_p = \|A^\top\|_q$
4. induced norm is submultiplicative: $\|AB\| \leq \|A\|\|B\|$

Key theorems about dual and induced seminorms

Projection and distance seminorms are dual

$$|||\cdot|||_{\mathrm{dist},p} = (|||\cdot|||_{\mathrm{proj},q})_\star \qquad\qquad |||\cdot|||_{\mathrm{proj},q} = (|||\cdot|||_{\mathrm{dist},p})_\star$$

Properties of dual and induced seminorms

1. dual seminorm satisfies (sharp) *Markov inequality*: $x^\top \Pi_\perp y \le |||x|||_{\mathrm{dist},p} |||y|||_{\mathrm{proj},q}$
2. dual seminorm induces duality: $|||A|||_{\mathrm{dist},p} = |||A^\top|||_{\mathrm{proj},q}$
3. induced seminorm is submultiplicative: $|||AB||| \le |||A|||\,|||B|||$ if $A\mathcal{K} \subseteq \mathcal{K}$ or $B\mathcal{K}^\top \subseteq \mathcal{K}^\top$

Ergodic coefficients are induced seminorms

If $A\mathcal{K} \subseteq \mathcal{K}$, then $$|||A|||_{\mathrm{dist},p} = |||A^\top|||_{\mathrm{proj},q} = \tau_q(\mathcal{K}, A) := \max_{\|z\|_q=1,\ z\perp\mathcal{K}} \|A^\top z\|_q$$

How Markov and Banach's results meet

Given $\mathcal{K} \subset \mathbb{R}^n$ and $p, q \in [1, \infty]$ with $p^{-1} + q^{-1} = 1$, consider $\{A(k)\}_{k \in \mathbb{Z}_{\geq 0}} \subset \mathbb{R}^{n \times n}$ satisfying:

$$A(k)\mathcal{K} \subseteq \mathcal{K} \quad \text{for all } k \in \mathbb{Z}_{\geq 0}, \qquad \text{(invariance)}$$

$$\rho \triangleq \sup_{k \in \mathbb{Z}_{\geq 0}} \tau_p(\mathcal{K}, A(k)) < 1. \qquad \text{(semicontraction)}$$

1. the *averaging system*

$$x(k+1) = A(k)x(k) + b, \quad b \in \mathbb{R}^n,$$

is *strongly semicontracting with rate ρ wrt $|||\cdot|||_{\mathrm{dist},q}$*

$$|||x(k) - y(k)|||_{\mathrm{dist},q} \leq \rho^k |||x(0) - y(0)|||_{\mathrm{dist},q}$$

2. the *dynamical flow system*

$$x(k+1) = A^\top(k)x(k) + b, \quad b \in \mathbb{R}^n,$$

is *strongly semicontracting with rate ρ wrt $|||\cdot|||_{\mathrm{proj},p}$* and, for any $x(0) - y(0) \in \mathcal{K}^\perp$,

$$|||x(k) - y(k)|||_{\mathrm{proj},p} \leq \rho^k |||x(0) - y(0)|||_{\mathrm{proj},p}$$

Continuous-time semicontraction theory

The *induced log seminorm* of $A \in \mathbb{R}^{n\times n}$ is

$$\mu_{|||\cdot|||}(A) \triangleq \lim_{h\to 0^+} \frac{|||I_n + hA||| - 1}{h}$$

Theorem (Dual logarithmic seminorms)

Let $p, q \in [1, \infty]$ *such that* $p^{-1} + q^{-1} = 1$. *For any matrix* $M \in \mathbb{R}^{n\times n}$, *and any kernel* $\mathcal{K}$,

$$\mu_{\mathrm{dist},p}(M) = \mu_{\mathrm{proj},q}(M^\top)$$

Formulas for induced log seminorm of Laplacian matrices

Outline

Indirect optimal control via contraction theory and iISS

Optimal control problem:

$$\begin{cases} \dot{x} = \mathsf{F}(x, u) \\ \mathcal{J}[u] = \int_0^T \text{running cost} + \text{final cost} \end{cases}$$

Pontryagin minimum principle

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

$$\dot{\lambda} = \mathsf{Adjoint}(x, u, \lambda)$$

$$u = \operatorname{argmin}_{\tilde{u}} \mathcal{H}(x, \tilde{u}, \lambda)$$

Method of successive approximations

Input: initial guess $u^{(0)}$, init value x_0

1: **for** $i \in \{1, \dots, N\}$ **do**
2: $\quad x^{(i)} \leftarrow$ forward with $u^{(i-1)}$
3: $\quad \lambda^{(i)} \leftarrow$ backward with $x^{(i)}$ and $u^{(i-1)}$
4: $\quad u^{(i)} \leftarrow \operatorname{argmin}_{\tilde{u}} H(x^{(i)}, \lambda^{(i)}, \tilde{u})$
5: **end for**
6: **return** $u^{(N)}$

Contractivity of adjoint dynamics and MSA

1. $\mathsf{osLip}_{\|\cdot\|^*}(\mathsf{Adjoint}^{\leftarrow}) = \mathsf{osLip}_{\|\cdot\|}(\mathsf{F})$
2. $\mathsf{Lip}(\mathsf{MSA}) \to 0^+$ as $T \to 0^+$ or $\mathsf{osLip}(\mathsf{F}) \to -\infty$
3. MSA is a contraction for (short T or highly contracting F)

Outline

Incremental ISS for strongly contracting delay ODEs

$$\dot{x}(t) = f\big(x(t), x(t-s), u(t)\big), 0 \le s \le S, \qquad \|\cdot\|_{\mathcal{X}}, \|\cdot\|_{\mathcal{U}} \tag{11}$$

assume there exist positive constants c, $\ell_{\mathcal{U}}$, $\ell_{\mathcal{X}}$ such that, for all variables,

$$\begin{aligned}
&\textbf{osL } x: && [\![f(x,d,u) - f(y,d,u), x-y]\!]_{\mathcal{X}} \le -c\|x-y\|_{\mathcal{X}}^2 && (12)\\
&\textbf{Lip } x(t-s): && \|f(x,x_1,u) - f(x,x_2,u)\|_{\mathcal{X}} \le \ell_{\mathcal{X}}\|x_1-x_2\|_{\mathcal{X}} && (13)\\
&\textbf{Lip } u: && \|f(x,d,u) - f(x,d,v)\|_{\mathcal{X}} \le \ell_{\mathcal{U}}\|u-v\|_{\mathcal{U}} && (14)
\end{aligned}$$

By the curve norm derivative formula, subadditivity, and Cauchy-Schwarz inequality,

$$\begin{aligned}
\|x(t)-y(t)\|_{\mathcal{X}} D^+\|x(t)-y(t)\|_{\mathcal{X}} &= \big[\!\big[f(x(t), x(t-s), u_x(t)) - f(y(t), y(t-s), u_y(t)), x(t)-y(t) \big]\!\big]_{\mathcal{X}}\\
&\le [\![f(x(t), x(t-s), u_x(t)) - f(y(t), x(t-s), u_x(t)), x(t)-y(t)]\!]_{\mathcal{X}}\\
&\quad + [\![f(y(t), x(t-s), u_x(t)) - f(y(t), y(t-s), u_x(t)), x(t)-y(t)]\!]_{\mathcal{X}}\\
&\quad + \big[\!\big[f(y(t), y(t-s), u_x(t)) - f(y(t), y(t-s), u_y(t)), x(t)-y(t) \big]\!\big]_{\mathcal{X}}\\
&\le -c\|x(t)-y(t)\|_{\mathcal{X}}^2 + \ell_{\mathcal{X}}\|x(t-s)-y(t-s)\|_{\mathcal{U}}\|x(t)-y(t)\|_{\mathcal{X}},\\
&\quad + \ell_{\mathcal{U}}\|u_x(t)-u_y(t)\|_{\mathcal{U}}\|x(t)-y(t)\|_{\mathcal{X}}.
\end{aligned}$$

Thus, with $m(t) = \|x(t)-y(t)\|_{\mathcal{X}}$, delay differential inequality:

$$D^+m(t) \le -cm(t) + \ell_{\mathcal{X}} \sup_{0\le s\le S} m(t-s) + \ell_{\mathcal{U}}\|u_x(t)-u_y(t)\|_{\mathcal{U}}, \tag{15}$$

Halanay inequality is applicable. If $c > \ell_{\mathcal{X}}$, then

$$m(t) \le m_0 \mathrm{e}^{-\rho(t-t_0)} + \ell_{\mathcal{U}} \int_{t_0}^{t} \mathrm{e}^{-\rho(t-\tau)}\|u_x(\tau)-u_y(\tau)\|_{\mathcal{U}} d\tau, \tag{16}$$

where $\rho > 0$ is the unique positive root of $\rho = c - \ell_{\mathcal{X}}\mathrm{e}^{\rho S}$ and $m_0 = \sup_{0\le s\le S} m(t_0 - s)$.

Networks of contracting systems with time delays

Interconnected subsystems $i \in \{1, \ldots, n\}$

$$\dot{x}_i = f_i(x_i, x_{-i}, x_{-i}(t-s), u_i), \qquad 0 \le s \le S, \qquad \|\cdot\|_i, \|\cdot\|_{i,\mathcal{U}} \tag{17}$$

Assume there exist positive constants st

osL x_i : $\llbracket f_i(x_i, \ldots) - f_i(y_i, \ldots), x_i - y_i \rrbracket_i \le -c_i \|x_i - y_i\|_i^2$

Lip x_{-i} : $\|f_i(\ldots, x_{-i}, \ldots) - f_i(\ldots, y_{-i}, \ldots)\|_i \le \sum_{j=1, j\neq i}^n \gamma_{ij} \|x_j - y_j\|_j$

Lip x_{-1}^{-s} : $\|f_i(\ldots, x_{-i}^{-s}, \ldots) - f_i(\ldots, y_{-i}^{-s}, \ldots)\|_i \le \sum_{j=1, j\neq i}^n \widehat{\gamma}_{ij} \|x_j^{-s} - y_j^{-s}\|_j$

Lip u_i : $\|f_i(\ldots, u_i) - f_i(\ldots, v_i)\|_i \le \ell_{i,\mathcal{U}} \|u_i - v_i\|_{i,\mathcal{U}}$

With $m_i(t) = \|x_i(t) - y_i(t)\|_i$, delay differential inequality:

$$D^+ m(t) \le -C m(t) + \Gamma m(t) + \widehat{\Gamma} \sup_{0 \le s \le S} m(t-s) + \ell_{\mathcal{U}} \|u_x(t) - u_y(t)\|_{\mathcal{U}}$$

and, if the Metzler matrix $-C + \Gamma + \widehat{\Gamma}$ is Hurwitz, then (**??**) is incremental ISS

Outline

Contraction theory on Riemannian manifolds

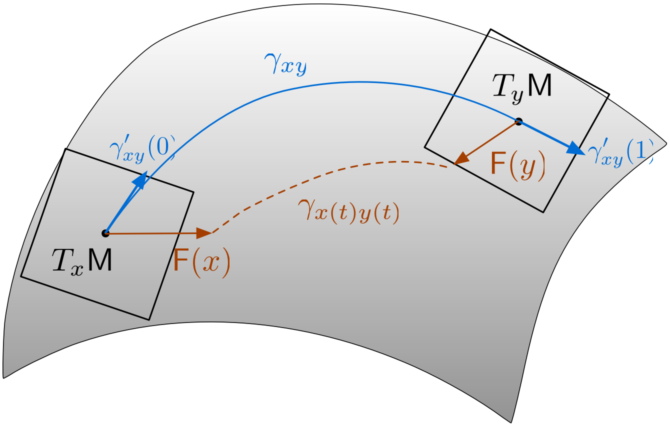
Contraction theory on Riemannian manifolds originates in

A formal coordinate-free analysis (with connection to monotone operators) is given in

In the differential geometry literature, geodesically monotonic vector fields are studied by

Contraction theory on Riemannian manifold $(\mathsf{M}, \mathbb{G})$

F **contracting** if geodesic distances from x to y diminishes along the flow of F



integral test: the inner product between F and the geodesic velocity vector γ'_{xy} at x and y
differential test: condition on covariant differential of F

$$\mathbb{G}(x)D\mathsf{F}x(x) + D\mathsf{F}x(x)^\top\mathbb{G}(x) + \dot{\mathbb{G}}(x) \preceq -2c\mathbb{G}(x)$$

Strong infinitesimal contraction on a Riemannian manifold

Given a time-independent vector field X on a Riemannian manifold $(\mathsf{M}, \mathbb{G})$ and $c > 0$, the following statements are equivalent:

1. for any $x, y \in \mathsf{M}$ and geodesic curve $\gamma_{xy} : [0,1] \to \mathsf{M}$ with $\gamma_{xy}(0) = x$, $\gamma_{xy}(1) = y$,
$$\big\langle\!\big\langle X(y), \gamma'_{xy}(1) \big\rangle\!\big\rangle_{\mathbb{G}} - \big\langle\!\big\langle X(x), \gamma'_{xy}(0) \big\rangle\!\big\rangle_{\mathbb{G}} \leq -c \, \mathrm{d}_{\mathbb{G}}(x,y)^2$$
2. for all $v_x \in T_x\mathsf{M}$
$$\langle\!\langle A_X(x) v_x, v_x \rangle\!\rangle_{\mathbb{G}} \leq -c \|v_x\|^2_{\mathbb{G}},$$
where the *covariant differential* $A_X(x) : T_x\mathsf{M} \to T_x\mathsf{M}$ is defined by $A_X(x)v_x = \nabla_{v_x} X(x)$
3. $D^+ \mathrm{d}_{\mathbb{G}}(x(t), y(t)) \leq -c \, \mathrm{d}_{\mathbb{G}}(x(t), y(t))$, for all solutions $x(\cdot), y(\cdot)$