Computer assisted proofs in differential equations
Lecture 1 - Dec. 10, 2019

Marcio Gameiro

University of São Paulo at São Carlos - ICMC-USP

and

Rutgers University

12th Americas Conference on Differential Equations and Nonlinear Analysis (Americas XII)
December 9-13, 2019, Mathematics Research Center (CIMAT)
Guanajuato, México

Motivation

- Why study differential equations?

Motivation

- Why study differential equations?
  - It is an important branch of mathematics and applied mathematics.

Motivation

- Why study differential equations?
  - It is an important branch of mathematics and applied mathematics.
  - It is widely used in applications in engineering, physics, biology, etc.

Motivation

- Why study differential equations?
  - It is an important branch of mathematics and applied mathematics.
  - It is widely used in applications in engineering, physics, biology, etc.

- Why use computational methods?

Motivation

- Why study differential equations?
  - It is an important branch of mathematics and applied mathematics.
  - It is widely used in applications in engineering, physics, biology, etc.

- Why use computational methods?
  - Many problems are very hard (if not impossible) to solve by hand, but numerical solutions can be (easily) obtained.

Motivation

- Why study differential equations?
  - It is an important branch of mathematics and applied mathematics.
  - It is widely used in applications in engineering, physics, biology, etc.

- Why use computational methods?
  - Many problems are very hard (if not impossible) to solve by hand, but numerical solutions can be (easily) obtained.
    - Lorenz attractor.

Motivation

- Why study differential equations?
  - It is an important branch of mathematics and applied mathematics.
  - It is widely used in applications in engineering, physics, biology, etc.
- Why use computational methods?
  - Many problems are very hard (if not impossible) to solve by hand, but numerical solutions can be (easily) obtained.
    - Lorenz attractor.
    - Bifurcation analysis (specially with higher co-dimensions).

Motivation

- Why study differential equations?
  - It is an important branch of mathematics and applied mathematics.
  - It is widely used in applications in engineering, physics, biology, etc.

- Why use computational methods?
  - Many problems are very hard (if not impossible) to solve by hand, but numerical solutions can be (easily) obtained.
    - Lorenz attractor.
    - Bifurcation analysis (specially with higher co-dimensions).
    - Fluid dynamics models, etc.

Motivation

- Why study differential equations?
  - It is an important branch of mathematics and applied mathematics.
  - It is widely used in applications in engineering, physics, biology, etc.
- Why use computational methods?
  - Many problems are very hard (if not impossible) to solve by hand, but numerical solutions can be (easily) obtained.
    - Lorenz attractor.
    - Bifurcation analysis (specially with higher co-dimensions).
    - Fluid dynamics models, etc.
- Why computer assisted proofs?

Motivation

- Why study differential equations?
  - It is an important branch of mathematics and applied mathematics.
  - It is widely used in applications in engineering, physics, biology, etc.
- Why use computational methods?
  - Many problems are very hard (if not impossible) to solve by hand, but numerical solutions can be (easily) obtained.
    - Lorenz attractor.
    - Bifurcation analysis (specially with higher co-dimensions).
    - Fluid dynamics models, etc.
- Why computer assisted proofs?
  - Numerical solutions may not be accurate.

Motivation

- Why study differential equations?
  - It is an important branch of mathematics and applied mathematics.
  - It is widely used in applications in engineering, physics, biology, etc.

- Why use computational methods?
  - Many problems are very hard (if not impossible) to solve by hand, but numerical solutions can be (easily) obtained.
    - Lorenz attractor.
    - Bifurcation analysis (specially with higher co-dimensions).
    - Fluid dynamics models, etc.
- Why computer assisted proofs?
  - Numerical solutions may not be accurate.
  - We want rigorous results (proofs).

Motivation

- Famous disasters caused by numerical errors:

Motivation

- Famous disasters caused by numerical errors:
  - The Patriot Missile Failure.

Motivation

- Famous disasters caused by numerical errors:
  - The Patriot Missile Failure.
  - Ariane 5 rocket failure.

Motivation

- Famous disasters caused by numerical errors:
  - The Patriot Missile Failure.
  - Ariane 5 rocket failure.
- A simple example (Henon map).

$$f(x, y) = (1 - ax^2 + y, bx)$$

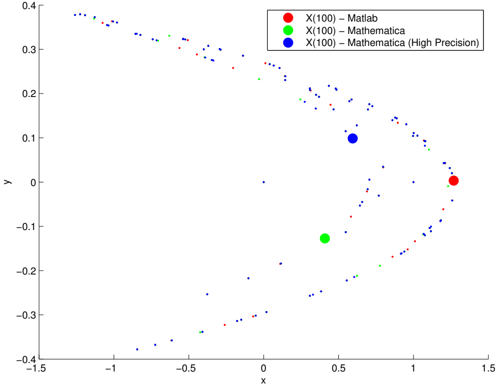
For $a = 1.4$, $b = 0.3$ and $X_0 = (0, 0)$, compute $X_n = f^n(X_0)$ for $n = 1, \ldots, 100$.

Motivation

- A simple example (Henon map).

$$f(x, y) = (1 - ax^2 + y, bx)$$

For $a = 1.4$, $b = 0.3$ and $X_0 = (0, 0)$, compute $X_n = f^n(X_0)$ for $n = 1, \ldots, 100$.



Computer Assisted Proofs

- What is a computer assisted proof?

Computer Assisted Proofs

- What is a computer assisted proof?
  - It is a proof that if performed with the help of a computer.

Computer Assisted Proofs

- What is a computer assisted proof?
  - It is a proof that if performed with the help of a computer.
- Some types of computer assisted proofs:

Computer Assisted Proofs

- What is a computer assisted proof?
  - It is a proof that if performed with the help of a computer.
- Some types of computer assisted proofs:
  - Automated proof assistants (Coq, ACL2, HOL4, etc).

Computer Assisted Proofs

- What is a computer assisted proof?
  - It is a proof that if performed with the help of a computer.
- Some types of computer assisted proofs:
  - Automated proof assistants (Coq, ACL2, HOL4, etc).
  - Proof by exhaustion (the four color theorem).

Computer Assisted Proofs

- What is a computer assisted proof?
  - It is a proof that if performed with the help of a computer.
- Some types of computer assisted proofs:
  - Automated proof assistants (Coq, ACL2, HOL4, etc).
  - Proof by exhaustion (the four color theorem).
  - Algebraic or exact computations (Mathematica, Computational Homology, etc).

Computer Assisted Proofs

- What is a computer assisted proof?
  - It is a proof that if performed with the help of a computer.
- Some types of computer assisted proofs:
  - Automated proof assistants (Coq, ACL2, HOL4, etc).
  - Proof by exhaustion (the four color theorem).
  - Algebraic or exact computations (Mathematica, Computational Homology, etc).
  - Validated (or Rigorous) Numerics (Interval Arithmetic).

Computer Assisted Proofs

- What is a computer assisted proof?
  - It is a proof that if performed with the help of a computer.
- Some types of computer assisted proofs:
  - Automated proof assistants (Coq, ACL2, HOL4, etc).
  - Proof by exhaustion (the four color theorem).
  - Algebraic or exact computations (Mathematica, Computational Homology, etc).
  - Validated (or Rigorous) Numerics (Interval Arithmetic).
- Some famous computer assisted proofs:

Computer Assisted Proofs

- What is a computer assisted proof?
  - It is a proof that if performed with the help of a computer.
- Some types of computer assisted proofs:
  - Automated proof assistants (Coq, ACL2, HOL4, etc).
  - Proof by exhaustion (the four color theorem).
  - Algebraic or exact computations (Mathematica, Computational Homology, etc).
  - Validated (or Rigorous) Numerics (Interval Arithmetic).
- Some famous computer assisted proofs:
  - The four color theorem.

Computer Assisted Proofs

- What is a computer assisted proof?
  - It is a proof that if performed with the help of a computer.
- Some types of computer assisted proofs:
  - Automated proof assistants (Coq, ACL2, HOL4, etc).
  - Proof by exhaustion (the four color theorem).
  - Algebraic or exact computations (Mathematica, Computational Homology, etc).
  - Validated (or Rigorous) Numerics (Interval Arithmetic).
- Some famous computer assisted proofs:
  - The four color theorem.
  - Kepler conjecture (optimal sphere packing).

Computer Assisted Proofs

- What is a computer assisted proof?
  - It is a proof that if performed with the help of a computer.
- Some types of computer assisted proofs:
  - Automated proof assistants (Coq, ACL2, HOL4, etc).
  - Proof by exhaustion (the four color theorem).
  - Algebraic or exact computations (Mathematica, Computational Homology, etc).
  - Validated (or Rigorous) Numerics (Interval Arithmetic).
- Some famous computer assisted proofs:
  - The four color theorem.
  - Kepler conjecture (optimal sphere packing).
  - Feigenbaum conjecture.

Computer Assisted Proofs

- What is a computer assisted proof?
  - It is a proof that if performed with the help of a computer.
- Some types of computer assisted proofs:
  - Automated proof assistants (Coq, ACL2, HOL4, etc).
  - Proof by exhaustion (the four color theorem).
  - Algebraic or exact computations (Mathematica, Computational Homology, etc).
  - Validated (or Rigorous) Numerics (Interval Arithmetic).
- Some famous computer assisted proofs:
  - The four color theorem.
  - Kepler conjecture (optimal sphere packing).
  - Feigenbaum conjecture.
  - Lorenz attractor (Smale's 14th problem).

Overview

- What we will discuss in this course?

Overview

- What we will discuss in this course?
  - The radii polynomial method (rigorous numerics) for ODEs.

Overview

- What we will discuss in this course?
  - The radii polynomial method (rigorous numerics) for ODEs.
  - We will not discuss methods to find numerical solutions.

Overview

- What we will discuss in this course?
  - The radii polynomial method (rigorous numerics) for ODEs.
  - We will not discuss methods to find numerical solutions.
- Some references:
  - Ordinary Differential Equations: A Constructive Approach, J.B. van den Berg, M. Gameiro, J.-P. Lessard, J.D. Mireles James, K. Mischaikow (in preparation).
  - Ordinary Differential Equations with Applications C. Chicone, Springer, 2006.
  - Dynamical Systems: Stability, Symbolic Dynamics, and Chaos, C. Robinson, CRC Press, 1998.
  - Introduction to Interval Analysis, R. Moore, R. Kearfott, M. Cloud, SIAM Press, 2009.

Overview

- Our goal is to study the dynamics of differential equations (ODEs).

Overview

- Our goal is to study the dynamics of differential equations (ODEs).
- The simplest building blocks of dynamics that we want to compute are:

Overview

- Our goal is to study the dynamics of differential equations (ODEs).
- The simplest building blocks of dynamics that we want to compute are:
  - Equilibrium solutions.

Overview

- Our goal is to study the dynamics of differential equations (ODEs).
- The simplest building blocks of dynamics that we want to compute are:
  - Equilibrium solutions.
  - Periodic solutions.

Overview

- Our goal is to study the dynamics of differential equations (ODEs).
- The simplest building blocks of dynamics that we want to compute are:
  - Equilibrium solutions.
  - Periodic solutions.
  - Solutions to IVP and BVP.

Overview

- Our goal is to study the dynamics of differential equations (ODEs).
- The simplest building blocks of dynamics that we want to compute are:
  - Equilibrium solutions.
  - Periodic solutions.
  - Solutions to IVP and BVP.
  - Invariant manifolds.

Overview

- Our goal is to study the dynamics of differential equations (ODEs).
- The simplest building blocks of dynamics that we want to compute are:
  - Equilibrium solutions.
  - Periodic solutions.
  - Solutions to IVP and BVP.
  - Invariant manifolds.
  - Connecting orbits.

Overview - General idea

- Given a differential equation (ODE)

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

Overview - General idea

- Given a differential equation (ODE)

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

- Assume we want to find some specific type of solutions, e.g. periodic solution, solution to IVP, etc.

Overview - General idea

- Given a differential equation (ODE)

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

- Assume we want to find some specific type of solutions, e.g. periodic solution, solution to IVP, etc.
- We express the type of solution we want to find as a solution to a problem of the form $F(x) = 0$, for x in some appropriate space X.

Overview - General idea

- Given a differential equation (ODE)

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

- Assume we want to find some specific type of solutions, e.g. periodic solution, solution to IVP, etc.
- We express the type of solution we want to find as a solution to a problem of the form $F(x) = 0$, for x in some appropriate space X.
- We find an approximate (numerical) solution $\bar{x}$ to $F(x) = 0$, that is, $F(\bar{x}) \approx 0$.

Overview - General idea

- Given a differential equation (ODE)

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

- Assume we want to find some specific type of solutions, e.g. periodic solution, solution to IVP, etc.
- We express the type of solution we want to find as a solution to a problem of the form $F(x) = 0$, for x in some appropriate space X.
- We find an approximate (numerical) solution $\bar{x}$ to $F(x) = 0$, that is, $F(\bar{x}) \approx 0$.
- Next we want to prove that there is a true solution $\tilde{x}$ near $\bar{x}$ (a-posteriori existence theorem).

Overview - General idea

- Given a differential equation (ODE)

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

- Assume we want to find some specific type of solutions, e.g. periodic solution, solution to IVP, etc.
- We express the type of solution we want to find as a solution to a problem of the form $F(x) = 0$, for x in some appropriate space X.
- We find an approximate (numerical) solution $\bar{x}$ to $F(x) = 0$, that is, $F(\bar{x}) \approx 0$.
- Next we want to prove that there is a true solution $\tilde{x}$ near $\bar{x}$ (a-posteriori existence theorem).
- For this, we rewrite the problem $F(x) = 0$ as a fixed point problem of the form $T(x) = x$.

Overview - General idea

- Given a differential equation (ODE)

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

- Assume we want to find some specific type of solutions, e.g. periodic solution, solution to IVP, etc.
- We express the type of solution we want to find as a solution to a problem of the form $F(x) = 0$, for x in some appropriate space X.
- We find an approximate (numerical) solution $\bar{x}$ to $F(x) = 0$, that is, $F(\bar{x}) \approx 0$.
- Next we want to prove that there is a true solution $\tilde{x}$ near $\bar{x}$ (a-posteriori existence theorem).
- For this, we rewrite the problem $F(x) = 0$ as a fixed point problem of the form $T(x) = x$.
- Finally we try to prove that T is a contraction near $\bar{x}$.

Overview - General idea

- The main things we need to choose:

Overview - General idea

- The main things we need to choose:
- What is the map F and the space X? (problem dependent)

Overview - General idea

- The main things we need to choose:
- What is the map F and the space X? (problem dependent)
- How to find $\bar{x}$? (numerics)

Overview - General idea

- The main things we need to choose:
- What is the map F and the space X? (problem dependent)
- How to find $\bar{x}$? (numerics)
- How to construct the operator T? (Newton like operator)

Overview - General idea

- The main things we need to choose:
- What is the map F and the space X? (problem dependent)
- How to find $\bar{x}$? (numerics)
- How to construct the operator T? (Newton like operator)
- How to choose a region where T is a contraction? ($\overline{B_r(\bar{x})}$)

Overview - General idea

- The main things we need to choose:
- What is the map F and the space X? (problem dependent)
- How to find $\bar{x}$? (numerics)
- How to construct the operator T? (Newton like operator)
- How to choose a region where T is a contraction? ($\overline{B_r(\bar{x})}$)
  - How to choose r and prove that T is a contraction?

Overview - General idea

- The main things we need to choose:
- What is the map F and the space X? (problem dependent)
- How to find $\bar{x}$? (numerics)
- How to construct the operator T? (Newton like operator)
- How to choose a region where T is a contraction? ($\overline{B_r(\bar{x})}$)
  - How to choose r and prove that T is a contraction?
  - The *radii polynomial* method is a systematic way (based on the Newton-Kantorovich theorem) to choose r and prove that T is a contraction on $\overline{B_r(\bar{x})}$.

Zeros of Functions

- Given an ODE $\dot{x} = f(x)$, $x \in \mathbb{R}^n$ an equilibrium is a point x_0 satisfying

$$f(x_0) = 0.$$

Zeros of Functions

- Given an ODE $\dot{x} = f(x)$, $x \in \mathbb{R}^n$ an equilibrium is a point x_0 satisfying

$$f(x_0) = 0.$$

- We now focus on developing a method to prove existence of equilibria.

Zeros of Functions

- Given an ODE $\dot{x} = f(x)$, $x \in \mathbb{R}^n$ an equilibrium is a point x_0 satisfying
$$f(x_0) = 0.$$
- We now focus on developing a method to prove existence of equilibria.
- The problem is already of the form $f(x) = 0$. So we need to transform it into a fixed point problem of the form $T(x) = x$.

Zeros of Functions

- Given an ODE $\dot{x} = f(x)$, $x \in \mathbb{R}^n$ an equilibrium is a point x_0 satisfying

$$f(x_0) = 0.$$

- We now focus on developing a method to prove existence of equilibria.
- The problem is already of the form $f(x) = 0$. So we need to transform it into a fixed point problem of the form $T(x) = x$.
- *How to define T? (Recall that we want it to be a contraction)*

Zeros of Functions

- Given an ODE $\dot{x} = f(x)$, $x \in \mathbb{R}^n$ an equilibrium is a point x_0 satisfying

$$f(x_0) = 0.$$

- We now focus on developing a method to prove existence of equilibria.
- The problem is already of the form $f(x) = 0$. So we need to transform it into a fixed point problem of the form $T(x) = x$.
- *How to define T? (Recall that we want it to be a contraction)*
    - The first idea is to take $T(x) = x - Df(x)^{-1} f(x)$

Zeros of Functions

- Given an ODE $\dot{x} = f(x)$, $x \in \mathbb{R}^n$ an equilibrium is a point x_0 satisfying
$$f(x_0) = 0.$$
- We now focus on developing a method to prove existence of equilibria.
- The problem is already of the form $f(x) = 0$. So we need to transform it into a fixed point problem of the form $T(x) = x$.
- *How to define T? (Recall that we want it to be a contraction)*
  - The first idea is to take $T(x) = x - Df(x)^{-1} f(x)$
  - $Df(x)^{-1}$ is too hard to compute in general!

Zeros of Functions

- Given an ODE $\dot{x} = f(x)$, $x \in \mathbb{R}^n$ an equilibrium is a point x_0 satisfying
$$f(x_0) = 0.$$
- We now focus on developing a method to prove existence of equilibria.
- The problem is already of the form $f(x) = 0$. So we need to transform it into a fixed point problem of the form $T(x) = x$.
- *How to define T? (Recall that we want it to be a contraction)*
  - The first idea is to take $T(x) = x - Df(x)^{-1} f(x)$
  - $Df(x)^{-1}$ is too hard to compute in general!
  - An easier option is $T(x) = x - Df(\bar{x})^{-1} f(x)$

Zeros of Functions

- Given an ODE $\dot{x} = f(x)$, $x \in \mathbb{R}^n$ an equilibrium is a point x_0 satisfying

$$f(x_0) = 0.$$

- We now focus on developing a method to prove existence of equilibria.
- The problem is already of the form $f(x) = 0$. So we need to transform it into a fixed point problem of the form $T(x) = x$.
- *How to define T? (Recall that we want it to be a contraction)*
  - The first idea is to take $T(x) = x - Df(x)^{-1}f(x)$
  - $Df(x)^{-1}$ is too hard to compute in general!
  - An easier option is $T(x) = x - Df(\bar{x})^{-1}f(x)$
  - Still too hard!

Zeros of Functions

- Given an ODE $\dot{x} = f(x)$, $x \in \mathbb{R}^n$ an equilibrium is a point x_0 satisfying

$$f(x_0) = 0.$$

- We now focus on developing a method to prove existence of equilibria.
- The problem is already of the form $f(x) = 0$. So we need to transform it into a fixed point problem of the form $T(x) = x$.
- *How to define T? (Recall that we want it to be a contraction)*
  - The first idea is to take $T(x) = x - Df(x)^{-1}f(x)$
  - $Df(x)^{-1}$ is too hard to compute in general!
  - An easier option is $T(x) = x - Df(\bar{x})^{-1}f(x)$
  - Still too hard!
  - So we take is $T(x) = x - Af(x)$, where $A \approx Df(\bar{x})^{-1}$.

Contraction Mapping Theorem

Given a function $T\colon X \to X$, where X is a topological space, an element $x \in X$ is a fixed point of T, if $T(x) = x$.

Contraction Mapping Theorem

Given a function $T: X \to X$, where X is a topological space, an element $x \in X$ is a fixed point of T, if $T(x) = x$.

Definition

Let $(X, \mathbf{d})$ denote a metric space. A function $T: X \to X$ is a contraction if there is a number $\kappa \in [0, 1)$, called a contraction constant, such that

$$\mathbf{d}(T(x), T(y)) \leq \kappa \mathbf{d}(x, y)$$

for all $x, y \in X$.

Contraction Mapping Theorem

Given a function $T: X \to X$, where X is a topological space, an element $x \in X$ is a fixed point of T, if $T(x) = x$.

Definition

Let $(X, \mathbf{d})$ denote a metric space. A function $T: X \to X$ is a contraction if there is a number $\kappa \in [0, 1)$, called a contraction constant, such that

$$\mathbf{d}(T(x), T(y)) \le \kappa \mathbf{d}(x, y)$$

for all $x, y \in X$.

Theorem (Contraction Mapping Theorem)

Let $(X, \mathbf{d})$ be a complete metric space. If $T: X \to X$ is a contraction with contraction constant κ, then there exists a unique fixed point $\tilde{x} \in X$ of T. Furthermore, $\tilde{x}$ is globally attracting, and for any $x \in X$,

$$\mathbf{d}(T^n(x), \tilde{x}) \le \frac{\kappa^n}{1 - \kappa} \mathbf{d}(T(x), x).$$

Contraction Mapping Theorem

Proof.

- Given $x_0 \in X$

$$\mathbf{d}(x_{n+1}, x_n) = \mathbf{d}(T(x_n), T(x_{n-1})) \leq \kappa \mathbf{d}(x_n, x_{n-1}).$$

Contraction Mapping Theorem

Proof.

- Given $x_0 \in X$

$$\mathbf{d}(x_{n+1}, x_n) = \mathbf{d}(T(x_n), T(x_{n-1})) \leq \kappa \mathbf{d}(x_n, x_{n-1}).$$

- Hence, by induction

$$\mathbf{d}(x_{n+1}, x_n) \leq \kappa^n \mathbf{d}(x_1, x_0).$$

Contraction Mapping Theorem

Proof.

- Given $x_0 \in X$

$$\mathbf{d}(x_{n+1}, x_n) = \mathbf{d}(T(x_n), T(x_{n-1})) \leq \kappa \mathbf{d}(x_n, x_{n-1}).$$

- Hence, by induction

$$\mathbf{d}(x_{n+1}, x_n) \leq \kappa^n \mathbf{d}(x_1, x_0).$$

- Applying the triangle inequality for $n < m$

$$\begin{aligned} \mathbf{d}(x_n, x_m) &\leq \sum_{j=n}^{m-1} \mathbf{d}(x_{j+1}, x_j) \leq \sum_{j=n}^{m-1} \kappa^j \mathbf{d}(x_1, x_0) \\ &\leq \kappa^n \left(\sum_{k=0}^{\infty} \kappa^k \right) \mathbf{d}(x_1, x_0) \leq \kappa^n \frac{1}{1-\kappa} \mathbf{d}(x_1, x_0). \end{aligned}$$

Contraction Mapping Theorem

Proof.

- Given $x_0 \in X$

$$\mathbf{d}(x_{n+1}, x_n) = \mathbf{d}(T(x_n), T(x_{n-1})) \leq \kappa \mathbf{d}(x_n, x_{n-1}).$$

- Hence, by induction

$$\mathbf{d}(x_{n+1}, x_n) \leq \kappa^n \mathbf{d}(x_1, x_0).$$

- Applying the triangle inequality for $n < m$

$$\begin{aligned}
\mathbf{d}(x_n, x_m) &\leq \sum_{j=n}^{m-1} \mathbf{d}(x_{j+1}, x_j) \leq \sum_{j=n}^{m-1} \kappa^j \mathbf{d}(x_1, x_0) \\
&\leq \kappa^n \left(\sum_{k=0}^{\infty} \kappa^k \right) \mathbf{d}(x_1, x_0) \leq \kappa^n \frac{1}{1-\kappa} \mathbf{d}(x_1, x_0).
\end{aligned}$$

- This implies that $\{x_n\}$ is a Cauchy sequence.

Contraction Mapping Theorem

Proof.

- Hence there exists $\tilde{x} \in X$ such that

$$\lim_{n \to \infty} x_n = \tilde{x}.$$

$\square$

Contraction Mapping Theorem

Proof.

- Hence there exists $\tilde{x} \in X$ such that

$$\lim_{n \to \infty} x_n = \tilde{x}.$$

- By continuity of T,

$$\tilde{x} = \lim_{n \to \infty} x_n = \lim_{n \to \infty} T(x_{n-1}) = T\left(\lim_{n \to \infty} x_{n-1}\right) = T(\tilde{x}).$$

$\square$

Contraction Mapping Theorem

Proof.

- Hence there exists $\tilde{x} \in X$ such that

$$\lim_{n\to\infty} x_n = \tilde{x}.$$

- By continuity of T,

$$\tilde{x} = \lim_{n\to\infty} x_n = \lim_{n\to\infty} T(x_{n-1}) = T\left(\lim_{n\to\infty} x_{n-1}\right) = T(\tilde{x}).$$

- Assume that $\tilde{y}$ is another fixed point of T, that is $T(\tilde{y}) = \tilde{y}$ and $\mathbf{d}(\tilde{y}, \tilde{x}) > 0$. Then

$$\mathbf{d}(\tilde{y}, \tilde{x}) = \mathbf{d}(T(\tilde{y}), T(\tilde{x})) \leq \kappa \mathbf{d}(\tilde{y}, \tilde{x}),$$

which is a contradiction.

$\square$

Contraction Mapping Theorem

Proof.

- Note that $x_0 \in X$ was arbitrary. Hence we have

$$\begin{aligned}\mathbf{d}(T^n(x), \tilde{x}) &\leq \mathbf{d}(T^n(x), T^m(x)) + \mathbf{d}(T^m(x), \tilde{x}) \\ &\leq \frac{\kappa^n}{1-\kappa}\mathbf{d}(T(x), x) + \mathbf{d}(T^m(x), \tilde{x}).\end{aligned}$$

  for any $x \in X$, and all $m > n$. Taking the limit as $m \to \infty$ yields the bound claimed.

$\square$

Zeros of Functions

Theorem

Let X and Y be vector spaces. Let $U \subset X$ and $V \subset Y$. Consider $f : U \to V$. Assume that $A : Y \to X$ is an injective linear map. Let $T : U \to X$ be defined by

$$T(x) = x - Af(x).$$

Then, $T(\tilde{x}) = \tilde{x}$ if and only if $f(\tilde{x}) = 0$.

Definition

Let $f \in C^1(U, \mathbb{R}^N)$ where $U \subset \mathbb{R}^N$ is an open set. A point $\tilde{x} \in U$ is a *nondegenerate zero* of f if $f(\tilde{x}) = 0$ and $Df(\tilde{x})$ is invertible.

Radii Polynomials in Finite Dimensions

Theorem (Radii polynomials in finite dimensions)

Consider $f \in C^1(\mathbb{R}^N, \mathbb{R}^N)$*. Let* $\bar{x} \in \mathbb{R}^N$ *and* $A \in M_N(\mathbb{R})$*. Let* Y_0 *and* Z_0 *be non-negative constants and* $Z_2 : (0, \infty) \to [0, \infty)$ *a non-negative function satisfying*

$$\begin{aligned} \|Af(\bar{x})\| &\leq Y_0 \\ \|\mathrm{I} - ADf(\bar{x})\| &\leq Z_0 \\ \|A[Df(c) - Df(\bar{x})]\| &\leq Z_2(r)r, \quad \textit{for all } c \in \overline{B_r(\bar{x})} \textit{ and all } r > 0. \end{aligned}$$

Define

$$p(r) = Z_2(r)r^2 - (1 - Z_0)r + Y_0.$$

If there exists $r_0 > 0$ *such that* $p(r_0) < 0$*, then there exists a unique* $\tilde{x} \in \overline{B_{r_0}(\bar{x})}$ *satisfying* $f(\tilde{x}) = 0$ *and* $Df(\tilde{x})$ *is invertible.*

Radii Polynomials in Finite Dimensions

Definition

Since in practice the bound $Z_2(r)$ is usually chosen to be a polynomial with nonnegative coefficients, the function $p(r)$ is called a radii polynomial.

Radii Polynomials in Finite Dimensions

Proof.

- The assumption that $p(r_0) < 0$ is equivalent to

$$\kappa := Z_0 + Z_2(r_0)r_0 < 1 - \frac{Y_0}{r_0} \leq 1.$$

$\square$

Radii Polynomials in Finite Dimensions

Proof.

- The assumption that $p(r_0) < 0$ is equivalent to

$$\kappa := Z_0 + Z_2(r_0)r_0 < 1 - \frac{Y_0}{r_0} \leq 1.$$

- Define the mapping $T(x) = x - Af(x)$. Given $c \in \overline{B_{r_0}(\bar{x})}$,

$$\begin{aligned} \|DT(c)\| &= \|\mathrm{I} - ADf(c)\| \\ &\leq \|\mathrm{I} - ADf(\bar{x})\| + \|A[Df(\bar{x}) - Df(c)]\| \\ &\leq Z_0 + Z_2(r_0)r_0 = \kappa < 1. \end{aligned}$$

$\square$

Radii Polynomials in Finite Dimensions

Proof.

- The assumption that $p(r_0) < 0$ is equivalent to

$$\kappa := Z_0 + Z_2(r_0)r_0 < 1 - \frac{Y_0}{r_0} \leq 1.$$

- Define the mapping $T(x) = x - Af(x)$. Given $c \in \overline{B_{r_0}(\bar{x})}$,

$$\begin{aligned} \|DT(c)\| &= \|\mathrm{I} - ADf(c)\| \\ &\leq \|\mathrm{I} - ADf(\bar{x})\| + \|A[Df(\bar{x}) - Df(c)]\| \\ &\leq Z_0 + Z_2(r_0)r_0 = \kappa < 1. \end{aligned}$$

- In particular, $\|\mathrm{I} - ADf(\bar{x})\| < 1$ and hence $ADf(\bar{x})$ is invertible (Exercise). This implies that A is invertible.

$\square$

Radii Polynomials in Finite Dimensions

Proof.

- We now show that T maps $\overline{B_{r_0}(\bar{x})}$ into itself. Let $x \in \overline{B_{r_0}(\bar{x})}$. Apply the Mean Value Theorem inequality to obtain

$$
\begin{aligned}
\|T(x) - \bar{x}\| &\leq \|T(x) - T(\bar{x})\| + \|T(\bar{x}) - \bar{x}\| \\
&\leq \sup_{c \in \overline{B_{r_0}(\bar{x})}} \|DT(c)\| \|x - \bar{x}\| + \|Af(\bar{x})\| \\
&\leq (Z_0 + Z_2(r_0) r_0) r_0 + Y_0 < r_0,
\end{aligned}
$$

where the last inequality follows from the assumption that $p(r_0) < 0$.

Radii Polynomials in Finite Dimensions

Proof.

- We now show that T maps $\overline{B_{r_0}(\bar{x})}$ into itself. Let $x \in \overline{B_{r_0}(\bar{x})}$. Apply the Mean Value Theorem inequality to obtain

$$\begin{aligned} \|T(x) - \bar{x}\| &\leq \|T(x) - T(\bar{x})\| + \|T(\bar{x}) - \bar{x}\| \\ &\leq \sup_{c \in \overline{B_{r_0}(\bar{x})}} \|DT(c)\| \|x - \bar{x}\| + \|Af(\bar{x})\| \\ &\leq (Z_0 + Z_2(r_0) r_0) r_0 + Y_0 < r_0, \end{aligned}$$

where the last inequality follows from the assumption that $p(r_0) < 0$.

- Hence, $T(\overline{B_{r_0}(\bar{x})}) \subset \overline{B_{r_0}(\bar{x})}$.

Radii Polynomials in Finite Dimensions

Proof.

- We now show that T maps $\overline{B_{r_0}(\bar{x})}$ into itself. Let $x \in \overline{B_{r_0}(\bar{x})}$. Apply the Mean Value Theorem inequality to obtain

$$\begin{aligned} \|T(x) - \bar{x}\| &\leq \|T(x) - T(\bar{x})\| + \|T(\bar{x}) - \bar{x}\| \\ &\leq \sup_{c \in \overline{B_{r_0}(\bar{x})}} \|DT(c)\| \|x - \bar{x}\| + \|Af(\bar{x})\| \\ &\leq (Z_0 + Z_2(r_0) r_0) r_0 + Y_0 < r_0, \end{aligned}$$

  where the last inequality follows from the assumption that $p(r_0) < 0$.
- Hence, $T(\overline{B_{r_0}(\bar{x})}) \subset \overline{B_{r_0}(\bar{x})}$.
- To see that T is a contraction on $\overline{B_{r_0}(\bar{x})}$, let $a, b \in \overline{B_{r_0}(\bar{x})}$ and apply the Mean Value Theorem to obtain

$$\|T(a) - T(b)\| \leq \sup_{c \in \overline{B_{r_0}(\bar{x})}} \|DT(c)\| \|a - b\| \leq \kappa \|a - b\|.$$

Radii Polynomials in Finite Dimensions

Proof.

- Since $\kappa < 1$, T is a contraction.

$\square$

Radii Polynomials in Finite Dimensions

Proof.

- Since $\kappa < 1$, T is a contraction.
- The contraction mapping theorem yields the existence of a unique $\tilde{x} \in \overline{B_{r_0}(\bar{x})}$ such that $T(\tilde{x}) = \tilde{x} - Af(\tilde{x}) = \tilde{x}$. The invertibility of A yields the existence of a unique solution $\tilde{x} \in \overline{B_{r_0}(\bar{x})}$ of $f(x) = 0$.

$\square$

Radii Polynomials in Finite Dimensions

Existence Interval

Consider a radii polynomial $p(r) = Z_2(r)r^2 - (1 - Z_0)r + Y_0$ for a function $f\colon \mathbb{R}^n \to \mathbb{R}^n$ and assume that $Z_2(r)$ is a polynomial with non-negative coefficients, or more generally a non-decreasing function of r. Observe that if there exists $r_0 > 0$ such that $p(r_0) < 0$, then there exists an interval $(r_-, r_+) \subset (0, \infty)$ containing r_0 over which the inequality is satisfied. The maximal such interval is called the existence interval for the radii polynomial p and denoted by

$$\mathrm{EI}(p).$$

Since $\tilde{x}$ is the unique zero of f in $\overline{B_r(\bar{x})}$ for all $r \in \mathrm{EI}(p)$, r_- provides tight bounds on the location of $\tilde{x}$, while r_+ provides information about the domain of isolation of $\tilde{x}$. In particular, if the existence interval for the radii polynomials is nonempty, then one can present an explicit domain $U \subset \mathbb{R}^n$ in which there exists a unique zero of f.

Example 1

- Consider $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f(x) = \begin{pmatrix} 4x_1^2 + x_2 - \lambda \\ x_1 + x_2^2 - 1 \end{pmatrix}$$

Example 1

- Consider $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f(x) = \begin{pmatrix} 4x_1^2 + x_2 - \lambda \\ x_1 + x_2^2 - 1 \end{pmatrix}$$

- Then

$$Df(\bar{x}) = \begin{pmatrix} 8\bar{x}_1 & 1 \\ 1 & 2\bar{x}_2 \end{pmatrix}$$

Example 1

- Consider $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f(x) = \begin{pmatrix} 4x_1^2 + x_2 - \lambda \\ x_1 + x_2^2 - 1 \end{pmatrix}$$

- Then

$$Df(\bar{x}) = \begin{pmatrix} 8\bar{x}_1 & 1 \\ 1 & 2\bar{x}_2 \end{pmatrix}$$

- We take $Y_0 := \|Df(\bar{x})^{-1} f(\bar{x})\|$ and $Z_0 := \|\mathrm{I} - ADf(\bar{x})\|$.

Example 1

- Consider $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f(x) = \left(\begin{array}{c} 4x_1^2 + x_2 - \lambda \\ x_1 + x_2^2 - 1 \end{array} \right)$$

- Then

$$Df(\bar{x}) = \begin{pmatrix} 8\bar{x}_1 & 1 \\ 1 & 2\bar{x}_2 \end{pmatrix}$$

- We take $Y_0 := \|Df(\bar{x})^{-1} f(\bar{x})\|$ and $Z_0 := \|\mathrm{I} - ADf(\bar{x})\|$.
- To determine Z_2 it is useful to use a change of variables. Given $c \in \overline{B_r(\bar{x})}$, set $b := c - \bar{x} \in \overline{B_r(0)}$. Then

$$\begin{aligned}
&(Df(c) - Df(\bar{x})) = (Df(\bar{x} + b) - Df(\bar{x})) \\
&= \left[\begin{pmatrix} 8\bar{x}_1 + 8b_1 & 1 \\ 1 & 2\bar{x}_2 + 2b_2 \end{pmatrix} - \begin{pmatrix} 8\bar{x}_1 & 1 \\ 1 & 2\bar{x}_2 \end{pmatrix} \right] = \begin{pmatrix} 8b_1 & 0 \\ 0 & 2b_2 \end{pmatrix}
\end{aligned}$$

Example 1

- Hence

$$A\left(Df(c) - Df(\bar{x})\right) = \frac{1}{16\bar{x}_1\bar{x}_2 - 1}\begin{pmatrix} 16\bar{x}_2 b_1 & -2b_2 \\ -8b_1 & 16\bar{x}_1 b_2 \end{pmatrix}.$$

Example 1

- Hence

$$A\left(Df(c) - Df(\bar{x})\right) = \frac{1}{16\bar{x}_1\bar{x}_2 - 1}\begin{pmatrix} 16\bar{x}_2 b_1 & -2b_2 \\ -8b_1 & 16\bar{x}_1 b_2 \end{pmatrix}.$$

- Choosing the sup-norm we have

$$\begin{aligned}\|A\left(Df(c) - Df(\bar{x})\right)\|_\infty &= \max\left\{\frac{16|\bar{x}_2 b_1| + 2|b_2|}{|16\bar{x}_1\bar{x}_2 - 1|}, \frac{8|b_1| + 16|\bar{x}_1 b_2|}{|16\bar{x}_1\bar{x}_2 - 1|}\right\} \\ &\leq \max\left\{\frac{16|\bar{x}_2| + 2}{|16\bar{x}_1\bar{x}_2 - 1|}, \frac{8 + 16|\bar{x}_1|}{|16\bar{x}_1\bar{x}_2 - 1|}\right\} r\end{aligned}$$

where the inequality arises from the fact that $\|b\|_\infty = \max\{|b_1|, |b_2|\} \leq r$.

Example 1

- Hence
$$A\,(Df\,(c) - Df\,(\bar{x})) = \frac{1}{16\bar{x}_1\bar{x}_2 - 1}\begin{pmatrix} 16\bar{x}_2 b_1 & -2b_2 \\ -8b_1 & 16\bar{x}_1 b_2 \end{pmatrix}.$$
- Choosing the sup-norm we have
$$\begin{aligned} \|A\,(Df\,(c) - Df\,(\bar{x}))\,\|_\infty &= \max\left\{\frac{16|\bar{x}_2 b_1| + 2|b_2|}{|16\bar{x}_1\bar{x}_2 - 1|}, \frac{8|b_1| + 16|\bar{x}_1 b_2|}{|16\bar{x}_1\bar{x}_2 - 1|}\right\} \\ &\leq \max\left\{\frac{16|\bar{x}_2| + 2}{|16\bar{x}_1\bar{x}_2 - 1|}, \frac{8 + 16|\bar{x}_1|}{|16\bar{x}_1\bar{x}_2 - 1|}\right\} r \end{aligned}$$
where the inequality arises from the fact that $\|b\|_\infty = \max\{|b_1|, |b_2|\} \leq r$.
- Then we can set
$$Z_2 := \max\left\{\frac{16|\bar{x}_2| + 2}{|16\bar{x}_1\bar{x}_2 - 1|}, \frac{8 + 16|\bar{x}_1|}{|16\bar{x}_1\bar{x}_2 - 1|}\right\}.$$

Example 1

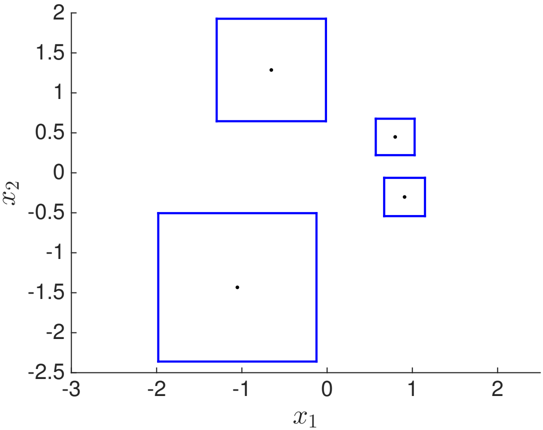


Figure: Largest existence and uniqueness enclosures for each equilibrium of f for $\lambda = 3$. For each $i = 1, 2, 3, 4$, the radius around $\bar{x}^{(i)}$ is the largest value of $I^{(i)}$. The smallest enclosure is too small to represent graphically.