MA3227 Numerical Analysis II

Lecture 15: Ordinary Differential Equations

Simon Etter

National University of Singapore

2019/2020

Ordinary Differential Equations

Introduction

Ordinary differential equations describe time evolution phenomena.
Examples: rocket trajectories, weather prediction, stock markets, etc.

Mathematically, these problems can be formulated as follows.

Given $f : \mathbb{R}^n \to \mathbb{R}^n$ and $y_0 \in \mathbb{R}^n$, find $y : [0, T] \to \mathbb{R}^n$ such that

$$y(0) = y_0 \qquad \text{and} \qquad \dot{y}(t) = f\big(y(t)\big) \quad \text{for all } t \in [0, T].$$

$\dot{y}(t) = \frac{dy}{dt}(t)$ is a common shorthand for the time derivative.

The final time T may arise in different ways:

- T is predetermined, e.g. $T = 1$ week for weather prediction.
- T depends on some event, e.g. $T =$ time when rocket lands.

Example ODEs

$$y(0) = y_0 \in \mathbb{R}, \qquad \dot{y}(t) = \lambda\, y(t) \quad \Longrightarrow \quad y(t) = y_0 \exp(\lambda t)$$

$$y(0) = y_0 \in \mathbb{R}, \qquad \dot{y}(t) = y(t)^2 \quad \Longrightarrow \quad y(t) = \frac{y_0}{1 - y_0 t}$$

Ordinary Differential Equations

Real-world ODE example: Newton's law of motion

Consider a point particle of mass $m > 0$ subject to an effective/total force $F \in \mathbb{R}^3$. The force typically depends on the location $x \in \mathbb{R}^3$ of the particle; thus we write $F(x)$.

Example: gravitation is stronger the closer you are to earth.

Newton's law states that the trajectory $x(t) \in \mathbb{R}^3$ (position as a function of time) of the above particle satisfies

$$m\,\ddot{x}(t) = F\big(x(t)\big).$$

This equation can be written in the form $\dot{y} = f(y)$ by setting

$$y = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}, \qquad f(y) = \begin{pmatrix} y[2] \\ \frac{1}{m}\,F(y[1]) \end{pmatrix}$$

such that

$$\dot{y} = \begin{pmatrix} \dot{x} \\ \ddot{x} \end{pmatrix} = \begin{pmatrix} \dot{x} \\ \frac{1}{m}\,F(x) \end{pmatrix} = \begin{pmatrix} y[2] \\ \frac{1}{m}\,F(y[1]) \end{pmatrix} = f(y).$$

Ordinary Differential Equations

Remark 1

Many textbooks allow $f(y,t)$ to also depend on the time variable t. I omit this generality because it is rarely needed in practice and complicates the notation.

Moreover, we can always reduce $\dot{x} = f(x,t)$ to $\dot{y} = \tilde{f}(y)$ by setting

$$y = \begin{pmatrix} t \\ x \end{pmatrix}, \qquad \tilde{f}(y) = \begin{pmatrix} 1 \\ f(y[2], y[1]) \end{pmatrix}.$$

Remark 2

ODEs are similar to PDEs in that the problem is to find a function $y(t)$ given an equation in terms of $y(t)$ and its derivatives which is to hold at every point in the domain.

Remark 2 (continued)

Historically/formally, the difference between ODEs and PDEs is that in an ODE, the unknown $y(t)$ depends on only a single variable while in a PDE the unknown $u(x_1, \dots, x_d)$ depends on several variables. The derivatives in an ODE are thus ordinary derivatives, while in a PDE they are partial derivatives.

In practice, the terms are hardly ever used in this way.

The defining property of an ODE is that the unknown $y(t)$ depends on a single variable, and fixed values for $y(t)$ and its derivatives are specified at a single point t_0. For this reason, ODEs are also called *initial value problems*.

The defining property of a PDE is that the unknown $u(x_1, \dots, x_d)$ may depend on one or more variables, and fixed values of $u(x_1, \dots, x_d)$ and its derivatives are specified at at least two points. For this reason, PDEs are also called *boundary value problems*.

Example: The 1d Poisson equation $-u''(x) = f(x)$ is an ODE in the formal sense because there is only a single independent variable x. However, it is usually called a PDE if we impose $u(0) = u(1) = 0$.

Ordinary Differential Equations

Outline

Our goal for the next few weeks is to develop numerical methods which can compute approximate point values $\tilde{y}(t) \approx y(t)$ for some finite number of time points $t \in [0, T]$.

We begin this journey by studying a few properties of the exact problem:

- Does a solution to $y(0) = y_0, \ \dot{y}(t) = f(y(t))$ exist?
- If yes, is the solution unique?
- If yes, how sensitive is the solution to perturbations in y_0?

These are fundamental questions which should always be addressed before attempting to solve any problem numerically. For illustration, the corresponding statements for linear systems are:

- $Ax = b$ has a unique solution x for all b if A is invertible.
- $Ax = b$ and $A(x + \Delta x) = b + \Delta b$ implies

$$\frac{\|\Delta x\|}{\|x\|} \le \|A\| \|A^{-1}\| \frac{\|\Delta b\|}{\|b\|}.$$

Ordinary Differential Equations

Def: Lipschitz continuity

A function $f : D \to \mathbb{R}^n$ with $D \subset \mathbb{R}^n$ is called Lipschitz continuous with Lipschitz constant $L > 0$ if for all $y_1, y_2 \in D$ we have

$$\big\| f(y_1) - f(y_2) \big\| \leq L \, \|y_1 - y_2\|.$$

Aside: contractions are Lipschitz-continuous functions with $L < 1$.

Picard-Lindelöf theorem (existence and uniqueness for ODEs)

Assume $f : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz continuous.

Then, there exists a unique differentiable $y : [0, \infty) \to \mathbb{R}^n$ such that

$$y(0) = y_0 \qquad \text{and} \qquad \dot{y}(t) = f\big(y(t)\big) \quad \text{for all } t \in [0, \infty).$$

Ordinary Differential Equations

Proof (not examinable).

Let Y be the space of all differentiable functions $y : [0, T] \to \mathbb{R}^n$ with $y(0) = y_0$ for some fixed $T > 0$, and consider the map

$$P : Y \to Y, \qquad y(t) \mapsto y_0 + \int_0^t f(y(\tau))\, d\tau.$$

It is clear that $y(t)$ solves the ODE if and only if $y = P(y)$. The idea of the proof is therefore to show that $P(y)$ has a unique fixed point for $T > 0$ small enough.

Once existence and uniqueness of a solution on $[0, T]$ is established, we can apply the same result on the interval $[T, 2T]$ using the now known $y(T)$ as initial value. Repeating this process indefinitely allows us to patch together a solution on $[0, \infty)$.

Ordinary Differential Equations

Proof (continued, not examinable).

The key step in the proof is therefore to show that $P(y)$ has a unique fixed point for a suitably chosen $T > 0$.

This follows from the Banach fixed-point theorem if we can show that

- $P(y)$ is a contraction, i.e. $\|P(y_1) - P(y_2)\| \le q \|y_1 - y_2\|$, and
- the space Y is complete (Cauchy sequences have a limit in Y).

Both of these conditions depend on the norm that we impose on Y. For this proof, we will use the norm

$$\|y\|_Y = \max_{t \in [0,T]} \|y(t)\|.$$

The completeness condition does not appear in the formulation of Banach's theorem provided in Lecture 14 because there we assumed that $g(x)$ is a function from $\mathbb{R}^N$ onto itself, and $\mathbb{R}^N$ is known to be complete.

Showing that Y is complete under the above norm is beyond our scope; instead, we will simply assume that it is. It then remains to determine $T > 0$ such that $P(y)$ is a contraction. We will do so on the next slide.

Ordinary Differential Equations

Proof (continued, not examinable).

We have

$$
\begin{aligned}
\left\|P(y_1)(t) - P(y_2)(t)\right\| &= \left\| \int_0^t \big(f(y_1(\tau)) - f(y_2(\tau))\big)\, d\tau \right\| \\
&\le \int_0^t \left\|f(y_1(\tau)) - f(y_2(\tau))\right\| d\tau \\
&\le L \int_0^t \left\|y_1(\tau) - y_2(\tau)\right\| d\tau \\
&\le Lt \max_{\tau\in[0,t]} \left\|y_1(\tau) - y_2(\tau)\right\| \\
&\le LT \max_{\tau\in[0,T]} \left\|y_1(\tau) - y_2(\tau)\right\|.
\end{aligned}
$$

and hence

$$
\|P(y_1) - P(y_2)\|_Y \le LT \, \|y_1 - y_2\|_Y .
$$

Thus, for $T < \frac{1}{L}$ we have that $P(y)$ is a contraction and therefore has a unique fixed point by the Banach fixed-point theorem.

This completes the proof of the Picard-Lindelöf theorem.

Ordinary Differential Equations

Local version of Picard-Lindelöf theorem

Assume $f : \mathbb{R}^n \to \mathbb{R}^n$ is locally Lipschitz continuous, i.e. for every $y_0 \in \mathbb{R}^n$ there exist $\delta, L > 0$ such that for all $y_1, y_2 \in \mathbb{R}^n$ with $\|y_i - y_0\| \leq \delta$ we have

$$\|f(y_1) - f(y_2)\| \leq L \|y_1 - y_2\|.$$

Then, there exists a $T > 0$ and a unique differentiable $y : [0, T) \to \mathbb{R}^n$ such that

$$y(0) = y_0 \qquad \text{and} \qquad \dot{y}(t) = f\big(y(t)\big) \quad \text{for all } t \in [0, T).$$

Ordinary Differential Equations

Proof. The idea of the proof is again to show that the map

$$P : Y \to Y, \qquad y(t) \mapsto y_0 + \int_0^t f(y(\tau))\, d\tau$$

with $Y = \{ y : [0, T] \to \mathbb{R}^n \mid y(0) = y_0, \dot{y} = f(y) \}$ is a contraction and therefore has a unique fixed point for T small enough. This can be done as follows.

Let δ, L be the local Lipschitz radius and constant, respectively, which as mentioned in the above statement may depend on the initial value y_0.

As before, we must choose $T < \frac{1}{L}$ to ensure that $P(y)$ is a contraction.

In addition, we must also choose T small enough such that we can guarantee that $y(t)$ stays within a radius δ of y_0 where we have the Lipschitz bound.

Working out the corresponding bound on T is highly technical; hence I omit the details. It is intuitively clear that $\|y(t) - y_0\| \le \delta$ can always be achieved by choosing T small enough.

Proof (continued).

The above gives us existence and uniqueness of a solution on $[0, T]$ for some $T > 0$. As before, we can then repeat the argument to extend the solution to an interval $[T, T']$ with $T' > T$. However, unlike before we can no longer guarantee that $T' = 2T$ will work because the extension length T now depends on the initial value y_0.

In particular, it is possible that as we proceed along the time axis, the extension length decreases fast enough such that the sum of all extension lengths is finite.

This is the reason why local Lipschitz continuity of $f(y)$ only guarantees that a solution exists on $[0, T)$ but not necessarily on $[0, \infty)$.

The following example demonstrates that solutions to ODEs may indeed have finite domains.

Example

- $f(y) = \lambda\, y$ is globally Lipschitz with $L = \lambda$.
  Hence the solution $y(t) = y_0 \exp(\lambda\, t)$ is defined on $[0, \infty)$.
- $f(y) = y^2$ is only locally Lipschitz since $|f'(y)| = 2\,|y| \to \infty$ for $y \to \infty$. Hence the solution $y(t) = \frac{y_0}{1 - y_0 t}$ is defined only on $[0, \frac{1}{y_0})$.

Ordinary Differential Equations

Discussion

The Picard-Lindelöf theorem shows that if $f(y)$ is Lipschitz continuous, then $\dot{y} = f(y)$ has a unique solution.

We next study how sensitive this solution is to perturbations in the initial value y_0, i.e. we aim to provide a bound on $\|y_1(t) - y_2(t)\|$ assuming that $y_1(t), y_2(t)$ both satisfy the same ODE $\dot{y}_i = f(y_i)$ but differ in the initial value $y_{1;0}$, $y_{2;0}$.

This question is important because in real-world applications, errors in y_0 are unavoidable (e.g. measurement, rounding or numerical approximation errors). Understanding sensitivities gives us a way to quantify how much we can trust our mathematical predictions.

The key tool to answer this question is Gronwall's inequality introduced on the next slide.

Ordinary Differential Equations

Gronwall's inequality

$$y(0) = y_0 \in \mathbb{R}, \qquad \dot{y}(t) \leq \lambda\, y(t) \qquad \Longrightarrow \qquad y(t) \leq y_0 \exp(\lambda t).$$

Proof. Consider

$$z(t) = \exp(-\lambda t)\, y(t).$$

Then, $z(0) = y_0$ and

$$\begin{aligned} \dot{z}(t) &= -\lambda \exp(-\lambda t)\, y(t) + \exp(-\lambda t)\, \dot{y}(t) \\ &\leq -\lambda \exp(-\lambda t)\, y(t) + \exp(-\lambda t)\, \lambda\, y(t) \\ &= 0. \end{aligned}$$

We conclude that $z(t) = \exp(-\lambda t)\, y(t) \leq y_0$ and thus

$$y(t) \leq y_0 \exp(\lambda t).$$

Ordinary Differential Equations

Thm: stability / sensitivity of ODEs

Assume $f : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz continuous with Lipschitz constant L, and assume $y_1, y_2 : [0, T) \to \mathbb{R}^n$ satisfy $\dot{y}_i = f(y_i)$. Then,

$$\|y_2(t) - y_1(t)\| \leq \exp(Lt) \, \|y_2(0) - y_1(0)\| \qquad \text{for any } t < T.$$

Proof (not examinable). We have $\frac{d}{dt}\|y(t)\| \leq \|\dot{y}(t)\|$ for any differentiable $y : \mathbb{R} \to \mathbb{R}^N$ since by the triangle inequality we have

$$\tfrac{\|y(t+h)\| - \|y(t)\|}{h} \leq \tfrac{\|y(t+h) - y(t)\|}{h}.$$

We can therefore bound

$$\begin{aligned} \frac{d}{dt}\|y_2(t) - y_1(t)\| &\leq \|\dot{y}_2(t) - \dot{y}_1(t)\| \\ &= \|f(y_2(t)) - f(y_1(t))\| \\ &\leq L \, \|y_2(t) - y_1(t)\| \end{aligned}$$

from which the claim follows by Gronwall's inequality.

Ordinary Differential Equations

Discussion of sensitivity result

The bound on the previous slide is a two-sided coin:

- Good news: error at time t is proportional to error at time 0.
- Bad news: constant of proportionality is $\exp(Lt)$.

The exponential function has a noticeable two-stage character:

- For $t \lesssim \frac{1}{L}$, $\exp(Lt)$ is fairly small.
- For $t \gtrsim \frac{1}{L}$, $\exp(Lt)$ grows *very* quickly.

In practice, this means that there is often a characteristic time-scale $\frac{1}{L}$ beyond which numerical simulations become unreliable.

For example, we can predict the weather fairly accurately for the next 2-3 days, but predictions beyond one week are virtually impossible.

Ordinary Differential Equations

Summary

- Picard-Lindelöf theorem: $\dot{y} = f(y)$ has a unique solution on $[0, \infty)$ if f is globally Lipschitz continuous. If $f(y)$ is only locally Lipschitz continuous, then the solution may exist only on some finite interval.
- Stability estimate: Assume $\dot{y}_i(t) = f(y_i(t))$ and $f(y)$ has Lipschitz constant L. Then,

$$\|y_1(t) - y_2(t)\| \leq \|y_1(0) - y_2(0)\| \exp(Lt).$$