MA5233 Computational Mathematics

Lecture 16: Ordinary Differential Equations

Simon Etter

NUS National University of Singapore

2019/2020

Ordinary Differential Equations

Ordinary differential equation (ODE)

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

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

ODEs are also called *initial value problems*.

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

Example: Newton's law of motion $m\,\ddot{x} = F(x)$.
This can be written in above form 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

Picard-Lindelöf theorem

Assume $f : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz continuous, i.e. there exists a constant $L > 0$ such that for all $y_1, y_2 \in \mathbb{R}^n$ we have

$$\|f(y_2) - f(y_1)\| \leq L \|y_2 - y_1\|. \tag{1}$$

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

$$y(0) = y_0 \qquad \text{and} \qquad \dot{y}(t) = f\big(y(t)\big) \quad \text{for all } t \in [0, \infty). \tag{2}$$

Alternatively, assume f is locally Lipschitz continuous at y_0, i.e. there exist $L > 0$ and $\varepsilon > 0$ such that (1) holds for $y_1, y_2 \in \{y \mid \|y - y_0\| \leq \varepsilon\}$. Then, the solution to (2) may be defined only on $[0, T)$ for some $T < \infty$.

Examples

- $f(y) = y$ is Lipschitz with $L = 1$.
  Solution $y(t) = y_0 \exp(t)$ is defined on $[0, \infty)$.
- $f(y) = y^2$ is locally Lipschitz but not globally Lipschitz.
  Solution $y(t) = \frac{y_0}{1 - y_0 t}$ is defined only on $[0, \frac{1}{y_0})$.

Ordinary Differential Equations

Conditioning of initial value problems

Assume $f : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz continuous with constant L, and $(y_i : [0, T) \to \mathbb{R}^n)_{i \in \{1,2\}}$ satisfy the two ODEs

$$y_i(0) = s_i, \qquad \dot{y}_i = f\left(y_i\right).$$

Then,

$$\|y_2(t) - y_1(t)\| \leq e^{Lt} \, \|s_2 - s_1\| \qquad \text{for all } t \in [0, T).$$

Interpretation

The map $s \mapsto y(t)$ is Lipschitz continuous, but the Lipschitz constant deteriorates exponentially for $t \to \infty$.

Real-world consequences:

- Weather prediction is difficult for large t.
- Shooting a rocket to Mars requires course corrections.

Ordinary Differential Equations

Solving ODEs via quadrature

The solution to the ODE $\dot{y} = f(y)$ is given by

$$y(t) = y(0) + \int_0^t f\big(y(\tau)\big)\, d\tau.$$

Observation: ODEs can be solved via quadrature!

Problem: We don't know $f\big(y(\tau_k)\big)$ for quadpoints $\tau_k > 0$.

Solution 1: Use left-point rule: (This is known as Euler's method)

$$\tilde{y}(t) := y(0) + f\big(y(0)\big)\, t.$$

Solution 2: Use midpoint rule, and use left-point rule to estimate $y(\frac{t}{2})$:

$$\tilde{y}(t) := y(0) + f\big(\tilde{y}(\tfrac{t}{2})\big)\, t, \qquad \tilde{y}(\tfrac{t}{2}) := y(0) + f\big(y(0)\big)\, \tfrac{t}{2}.$$

Solution 3: Use trapezoidal rule, and use left-point rule to estimate $y(t)$:

$$\tilde{y}(t) := y(0) + \Big(f\big(y(0)\big) + f\big(\tilde{\tilde{y}}(t)\big)\Big)\, \tfrac{t}{2}, \qquad \tilde{\tilde{y}}(t) := y(0) + f\big(y(0)\big)\, t.$$

Ordinary Differential Equations

Solving ODEs via quadrature (continued)

Above schemes deliver poor accuracy since they use only few quadpoints.

Two ways to improve accuracy:

- Increase number of quadpoints $\to$ s-stage Runge-Kutta methods.
- Use composite quadrature.

There are limits to how far the first approach can be pushed.

All practical schemes use composite quadrature, which in this context amounts to the following.

- Assume we want to compute $\tilde{y}(T) \approx y(T)$.
- Introduce partition $0 = t_0 < t_1 < \ldots < t_m = T$.
- Use any of the schemes on previous slide to iteratively compute

$$y(0) = \tilde{y}(t_0) \to \tilde{y}(t_1) \to \ldots \to \tilde{y}(t_m) = \tilde{y}(T).$$

See `16_ordinary_differential_equations.jl`.

Ordinary Differential Equations

Abstract time-stepping scheme

A single step of Euler's / midpoint / trapezoidal rule can be interpreted as a function $\tilde{\Phi}_t : y(0) \to \tilde{y}(t)$.

Composite scheme is then given by

$$\tilde{y}(T) = \tilde{\Phi}_{t_m - t_{m-1}}\big(\dots \tilde{\Phi}_{t_2 - t_1}\big(\tilde{\Phi}_{t_1 - t_0}\big(y(0)\big)\big)\big).$$

$\tilde{\Phi}_t$ is an approximation to $\Phi_t : y(0) \mapsto y(t)$.

Terminology:

- $\tilde{\Phi}_t(y)$: numerical propagator.
- $\Phi_t(y)$: exact propagator.

Goal: error estimate

$$\|\tilde{y}(T) - y(T)\| = \mathcal{O}\big(f(m)\big)$$

assuming equispaced partition $\big(t_k := T\,\frac{k}{m}\big)_{k=0}^{m}$.

Ordinary Differential Equations

Error analysis for abstract time-stepping scheme

Assumptions on numerical propagator:

- Consistency: $\|\tilde{\Phi}_t(y) - \Phi_t(y)\| = \mathcal{O}\big(t^{p+1}\big)$ for some $p > 0$.
- Stability: $\|\tilde{\Phi}_t(y_2) - \tilde{\Phi}_t(y_1)\| \leq (1 + \tilde{L}\, t)\, \|y_2 - y_1\|$ for some $\tilde{L} > 0$.

Then,

$$\|\tilde{y}(T) - y(T)\| = \mathcal{O}\Big(m^{-p}\Big) \qquad \text{for } m \to \infty.$$

Observe: consistency & stability $\implies$ convergence.

Proof. For notational convenience, we set $\Delta t := \frac{T}{m}$ and

$$\Phi(y) := \Phi_{\Delta t}(y), \quad \tilde{\Phi}(y) := \tilde{\Phi}_{\Delta t}(y), \quad y_k := y(k\, \Delta t), \quad \tilde{y}_k := \tilde{y}(k\, \Delta t).$$

Ordinary Differential Equations

Proof (continued). We compute

$$
\begin{aligned}
\|\tilde{y}(T) - y(T)\| &= \|\tilde{\Phi}(\tilde{y}_{m-1}) - \Phi(y_{m-1})\| \\
&\leq \|\tilde{\Phi}(\tilde{y}_{m-1}) - \tilde{\Phi}(y_{m-1})\| + \|\tilde{\Phi}(y_{m-1}) - \Phi(y_{m-1})\| \\
&\leq (1 + \tilde{L}\Delta t)\, \|\tilde{y}_{m-1} - y_{m-1}\| + \mathcal{O}\big(\Delta t^{p+1}\big) \\
&\leq (1 + \tilde{L}\Delta t)^2\, \|\tilde{y}_{m-2} - y_{m-2}\| + \big(1 + (1 + \tilde{L}\,\Delta t)\big)\, \mathcal{O}\big(\Delta t^{p+1}\big) \\
&\leq \dots \\
&\leq (1 + \tilde{L}\,\Delta t)^m\, \|\tilde{y}_0 - y_0\| + \left(\sum_{k=0}^{m-1} (1 + \tilde{L}\,\Delta t)^k\right) \mathcal{O}\big(\Delta t^{p+1}\big) \\
&\leq 0 + (1 + \tilde{L}\,\Delta t)^{m-1}\, \mathcal{O}\big(\Delta t^{p}\big)
\end{aligned}
$$

Claim follows after observing that since $\Delta t = \frac{T}{m}$, we have

$$
\mathcal{O}\big(\Delta t^p\big) = \mathcal{O}\big(m^{-p}\big), \qquad (1 + \tilde{L}\,\Delta t)^{m-1} \leq \exp\big(\tilde{L}\,T\,\tfrac{m-1}{m}\big) \leq \exp\big(\tilde{L}\,T\big).
$$

Ordinary Differential Equations

Consistency of Euler's, midpoint and trapezoidal method

Assuming y (and equivalently f) has sufficiently many derivatives, the consistency error $\tilde{y}(t) - y(t)$ can be estimated using Taylor series.

- Euler's method:

$$\begin{aligned} \tilde{y}(t) &= y(0) + f\big(y(0)\big)\, t \\ y(t) &= y(0) + \quad \dot{y}(0)\, t + \mathcal{O}(t^2) \end{aligned}$$

  Since $\dot{y}(0) = f\big(y(0)\big)$, we have $\tilde{y}(t) - y(t) = \mathcal{O}(t^2)$.

- Midpoint method: $\tilde{y}(t) = y(0) + f\big(y(0) + f(y(0))\frac{t}{2}\big)\, t$

$$\begin{aligned} \tilde{y}(t) &= y(0) + f\big(y(0)\big)\, t + f'\big(y(0)\big)\, f\big(y(0)\big)\, \tfrac{t^2}{2} + \mathcal{O}(t^3) \\ y(t) &= y(0) + \quad \dot{y}(0)\, t + \qquad\qquad \ddot{y}(0)\, \tfrac{t^2}{2} + \mathcal{O}(t^3) \end{aligned}$$

  Since $\dot{y}(0) = f\big(y(0)\big)$ and $\ddot{y}(0) = f'\big(y(0)\big)\, \dot{y}(0) = f'\big(y(0)\big)\, f\big(y(0)\big)$, we have $\tilde{y}(t) - y(t) = \mathcal{O}(t^3)$.

- Trapezoidal method: analogous. Result is $\tilde{y}(t) - y(t) = \mathcal{O}(t^3)$.

Ordinary Differential Equations

Stability of Euler's, midpoint and trapezoidal method

Assuming $f(y)$ is Lipschitz continuous, $\|f(y_2) - f(y_1)\| \leq L \|y_2 - y_1\|$, we can estimate the stability as follows.

- Euler's method: $\tilde{\Phi}_t(y) = y + f\big(y\big)\, t$.

$$\begin{aligned} \|\tilde{\Phi}_t(y_2) - \tilde{\Phi}_t(y_1)\| &\leq \|y_2 - y_1\| + t\, \|f(y_2) - f(y_1)\| \\ &\leq (1 + tL)\, \|y_2 - y_1\|. \end{aligned}$$

- Midpoint method: $\tilde{\Phi}_t(y) = y + f\big(y + f(y)\, \frac{t}{2}\big)\, t$.

$$\begin{aligned} \|\tilde{\Phi}_t(y_2) - \tilde{\Phi}_t(y_1)\| &\leq \|y_2 - y_1\| + t\, \|f(y_2 + f(y_2)\, \tfrac{t}{2}) - f(y_1 + f(y_1)\, \tfrac{t}{2})\| \\ &\leq (1 + tL)\, \|y_2 - y_1\| + \tfrac{t^2}{2}\, L\, \|f(y_2) - f(y_1)\| \\ &\leq (1 + tL + (tL)^2)\, \|y_2 - y_1\|. \end{aligned}$$

- Trapezoidal method: analogous.

Conclusion

Euler: error $= \mathcal{O}\big(m^{-1}\big)$, midpoint & trapezoidal: error $= \mathcal{O}\big(m^{-2}\big)$.

Ordinary Differential Equations

Runge-Kutta methods

Assume we have quadrature points x_i and a sequence of quadrature weights w_{ij} with $i \in \{0, \dots, s\}$ and $j \in \{0, \dots, i-1\}$ such that

$$x_0 = 0, \qquad x_s = 1,$$

and

$$\int_0^{x_i} f(x)\, dx \approx \sum_{j=0}^{i} w_{ij}\, f(x_j) \qquad \text{for all } i \in \{0, \dots, s\}.$$

Then, we can compute an approximate solution to $\dot{y} = f(y)$ through

$$\tilde{y}(t) := y(0) + t \sum_{j=0}^{s} w_{sj}\, f_j \approx y(t)$$

where

$$f_i := f\left(y(0) + t \sum_{j=0}^{i} w_{ij} f_j \right) \approx f\big(y(t\, x_i)\big)$$

Algorithms of this form are known as *s-stage Runge-Kutta methods*.

Ordinary Differential Equations

Butcher's tableau

The parameters x_i and w_{ij} of Runge-Kutta methods can be conveniently represented in a *Butcher's tableau*:

$$
\begin{array}{c|ccccc}
0 & & & & & \\
x_1 & w_{10} & & & & \\
x_2 & w_{20} & w_{21} & & & \\
\vdots & \vdots & \vdots & \ddots & & \\
x_{s-1} & w_{s-1,0} & w_{s-1,1} & \cdots & w_{s-1,s-2} & \\
\hline
 & w_{s,0} & w_{s,1} & \cdots & w_{s,s-2} & w_{s,s-1}
\end{array}
$$

Examples

Euler

$$
\begin{array}{c|c}
0 & \\
\hline
 & 1
\end{array}
$$

Midpoint

$$
\begin{array}{c|cc}
0 & & \\
\frac{1}{2} & \frac{1}{2} & \\
\hline
 & 0 & 1
\end{array}
$$

Trapezoidal

$$
\begin{array}{c|cc}
0 & & \\
1 & 1 & \\
\hline
 & \frac{1}{2} & \frac{1}{2}
\end{array}
$$

Error analysis for Runge-Kutta schemes

The analysis from slides 8-11 works for arbitrary Runge-Kutta method.

In particular, it can be used to determine x_i, w_{ij} such that consistency error is $\mathcal{O}\left(t^{p+1}\right)$ with p as large as possible.

Bad: calculations get tedious very quickly for increasing s.

Good: others have done the work for us:

`https://en.wikipedia.org/wiki/List_of_Runge-Kutta_methods`.

Interesting observation: minimal number of stages s to achieve error $\mathcal{O}\left(m^{-p}\right)$ grows faster than p.

p	1	2	3	4	5	6	7	8
$\min s$	1	2	3	4	6	7	9	11

Ordinary Differential Equations

References and further reading

- E. Suli and D. F. Mayers. *An Introduction to Numerical Analysis.* Cambridge University Press (2003), `doi:10.1017/CBO9780511801181`

Can be accessed online for free via the library website!