Ordinary differential equations

Initial value problem

Initial value problem (IVP)

Let $u(t)$ is an unknown function of a real argument $t \in [0, T)$.

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

and

$$u(t = 0) = u_0$$

Assume that $f(t, u)$ is smooth enough so that the solution exists and is unique.

Initial value problem (IVP)

Higher-order equations can be converted to systems of first-order equations:

$$\ddot{u} = f(t, u)$$

Define $w = \dot{u}$

$$\frac{d}{dt}\begin{bmatrix} u \\ w \end{bmatrix} = \begin{bmatrix} w \\ f(t, u) \end{bmatrix}$$

Numerical methods for IVP: discretization

Define a mesh

$$0 = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_N = T$$

Define a *mesh function* y_n

$$y_n \approx u(t_n)\,, \qquad n = 0, 1, \cdots, N$$

y_n satisfies a discretized form of the ODE.

Numerical methods for IVP: discretization

Since

$$\dot{u} = \lim_{\tau \to 0} \frac{u(t+\tau) - u(t)}{\tau} \,,$$

replace $\dot{u}$ with

$$\frac{y_{n+1} - y_n}{\tau}$$

where τ is the step size

$$\tau = t_{n+1} - t_n$$

(only consider a uniform mesh for simplicity).

Numerical methods for IVP: recurrence relations

Explicit Euler scheme

$$\frac{y_{n+1} - y_n}{\tau} = f(t_n, y_n)$$

for $n = 0, 1, \cdots, N - 1$.

Numerical methods for IVP: recurrence relations

Explicit Euler scheme

$$\frac{y_{n+1} - y_n}{\tau} = f(t_n, y_n)$$

for $n = 0, 1, \cdots, N-1$.

Initial condition: $y_0 = u(t=0) \equiv u_0$.

For $n \geqslant 1$:

$$y_{n+1} = y_n + \tau f(t_n, y_n)\ .$$

Numerical methods for IVP: discretization

Explicit Euler scheme

$$\frac{y_{n+1} - y_n}{\tau} = f(t_n, y_n)$$

Implicit Euler scheme

$$\frac{y_{n+1} - y_n}{\tau} = f(t_{n+1}, y_{n+1})$$

Symmetrized scheme

$$\frac{y_{n+1} - y_n}{\tau} = \frac{1}{2}\Big(f(t_n, y_n) + f(t_{n+1}, y_{n+1})\Big)$$

Analysis of the finite difference methods for ODEs

- Convergence
- Approximation order
- Stability

Convergence of a finite difference method

Consider a fixed $t \in [0, T)$.

Consider a sequence of uniform meshes with $\tau \to 0$, s.t.

$$t_n = n\tau = t \qquad \tau \to 0$$

(this requires $n \to \infty$)

Convergence of a finite difference method

A finite difference method is *convergent* at t if the *truncation error* tends to zero:

$$|y_n - u(t_n)| \to 0\,, \qquad \tau \to 0\,, \qquad t_n = t\ .$$

A method converges on $[0, T)$ if it converges for all $t \in [0, T)$.

A method has the p-th order of convergence if

$$|y_n - u(t_n)| = O(\tau^p)\,, \quad \tau \to 0$$

Truncation error

Consider the explicit Euler's scheme

$$\frac{y_{n+1} - y_n}{\tau} = f(t_n, y_n)$$

Derive the equation for the truncation error $z_n \equiv y_n - u(t_n)$.

Truncation error

Consider the explicit Euler's scheme

$$\frac{y_{n+1} - y_n}{\tau} = f(t_n, y_n)$$

Derive the equation for the truncation error $z_n \equiv y_n - u(t_n)$.

Using $y_n = u(t_n) + z_n$

$$\frac{z_{n+1} - z_n}{\tau} = f(t_n, u_n + z_n) - \frac{u_{n+1} - u_n}{\tau}$$

where $u_n \equiv u(t_n)$.

Truncation error

$$\frac{z_{n+1} - z_n}{\tau} = f(t_n, u_n + z_n) - \frac{u_{n+1} - u_n}{\tau} \\ + f(t_n, u_n) - f(t_n, u_n)$$

The r.h.s. is $\psi_n + \phi_n$, with

$$\psi_n = -\frac{u_{n+1} - u_n}{\tau} + f(t_n, u_n)$$

and

$$\phi_n = f(t_n, u_n + z_n) - f(t_n, u_n)$$

Truncation error

$$\frac{z_{n+1} - z_n}{\tau} = \psi_n + \phi_n$$

$$\phi_n = f(t_n, u_n + z_n) - f(t_n, u_n)$$

ϕ_n is identically zero if $f(t,u)$ is u-independent.
Otherwise, $\phi_n \propto z_n$:

$$\phi_n = z_n \left.\frac{\partial f}{\partial u}\right|_{(t_n, u_n + \theta z_n)}, \qquad |\theta| \leqslant 1.$$

Truncation error

$$\frac{z_{n+1} - z_n}{\tau} = \psi_n + \phi_n$$

Approximation error, a.k.a. *residual*:

$$\psi_n = -\frac{u_{n+1} - u_n}{\tau} + f(t_n, u_n)$$

If $y_n \equiv u_n$, $\psi_n = 0$.
If $\psi_n = O(\tau^p)$ as $\tau \to 0$, the *approximation order* is p.

Truncation error of the Euler's method

Expand u_{n+1} into the Taylor series around t_n :

$$\begin{aligned} u_{n+1} &\equiv u(t_n + \tau) \\ &= u_n + \dot{u}_n \tau + \ddot{u}_n \frac{\tau^2}{2} + \cdots \end{aligned}$$

The residual

$$\begin{aligned} \psi_n &= -\frac{u_{n+1} - u_n}{\tau} + f(t_n, u_n) \\ &= -\dot{u}_n + f(t_n, u_n) + \ddot{u}_n \frac{\tau}{2} + \cdots \\ &= O(\tau) \end{aligned}$$

Truncation error vs round-off error

- Small approximation errors require small τ.
- Roundoff errors increase for τ too small:

$$\varepsilon_r \sim \frac{1}{\tau}$$

There is a limit on how small we can make the total error.

Truncation error of the symmetrized Euler's method

Given the scheme

$$\frac{y_{n+1} - y_n}{\tau} = \frac{1}{2}\Big(f(t_n, y_n) + f(t_{n+1}, y_{n+1})\Big)$$

the residual is

$$\psi_n = -\frac{u_{n+1} - u_n}{\tau} + \frac{1}{2}\Big(f(t_n, u_n) + f(t_{n+1}, u_{n+1})\Big)$$

Expand u_{n+1} into the Taylor series around t_n :

$$u_{n+1} = u_n + \dot{u}_n\tau + \ddot{u}_n\frac{\tau^2}{2} + O(\tau^3)$$
$$\dot{u}_{n+1} = \dot{u}_n + \ddot{u}_n\tau + O(\tau^2)$$

Truncation error of the symmetrized Euler's method

The residual is

$$
\begin{aligned}
\psi_n &= -\frac{u_{n+1} - u_n}{\tau} + \frac{1}{2}\Big(f(t_n, u_n) + f(t_{n+1}, u_{n+1})\Big) \\
&= -\dot{u}_n - \ddot{u}\frac{\tau}{2} + O(\tau^2) + \frac{1}{2}\Big(\dot{u}_n + \dot{u}_n + \ddot{u}_n\tau + O(\tau^2)\Big) \\
&= O(\tau^2)
\end{aligned}
$$

The symmetrized Euler's scheme has the 2nd order of approximation.

Symmetrized Euler's method

- $\psi_n = O(\tau^2)$
- The method is implicit

Can we have an *explicit* 2nd order scheme?

Runge-Kutta methods

Predictor-corrector methods

IVP for $\dot{u} = f(t,u)$

Explicit Euler scheme

$$\frac{y_{n+1} - y_n}{\tau} = f(t_n, y_n)$$

Approximation order is linear, $\psi_n = O(\tau)$.

Symmetrized scheme

$$\frac{y_{n+1} - y_n}{\tau} = \frac{1}{2}\Big(f(t_n, y_n) + f(t_{n+1}, y_{n+1})\Big)$$

Approximation order is quadratic, $\psi_n = O(\tau^2)$.

IVP for $\dot{u} = f(t,u)$

Explicit Euler scheme

$$\frac{y_{n+1} - y_n}{\tau} = f(t_n, y_n)$$

Approximation order is linear, $\psi_n = O(\tau)$.

Symmetrized scheme

$$\frac{y_{n+1} - y_n}{\tau} = \frac{1}{2}\Big(f(t_n, y_n) + f(t_{n+1}, y_{n+1})\Big)$$

Approximation order is quadratic, $\psi_n = O(\tau^2)$.

Can we have an explicit quadratic method?

A two-step Runge-Kutta method

Use an intermediate value $y_{n+1/2}$:

$$t_n \longrightarrow t_{n+1/2} \longrightarrow t_{n+1}$$

Given y_n at $t = t_n$, make a "half-step"

$$\frac{y_{n+1/2} - y_n}{\tau/2} = f(t_n, y_n)$$

and then

$$\frac{y_{n+1} - y_n}{\tau} = f(t_n + \frac{\tau}{2}, y_{n+1/2})$$

Approximation error of the 2-step RK method

The finite-difference scheme is ($f_n \equiv f(t_n, y_n)$)

$$\frac{y_{n+1} - y_n}{\tau} = f(t_n + \frac{\tau}{2}, y_n + \frac{\tau}{2} f_n)$$

The corresponding residual is

$$\psi_n = -\frac{u_{n+1} - u_n}{\tau} + f(t_n + \frac{\tau}{2}, u_n + \frac{\tau}{2} f_n)$$

Approximation error of the 2-step RK method

At $\tau \to 0$,

$$f(t_n + \frac{\tau}{2}, u_n + \frac{\tau}{2} f_n) = f_n + \frac{\tau}{2}\Big(\frac{\partial f_n}{\partial t} + f_n \frac{\partial f_n}{\partial u}\Big) + O(\tau^2)$$

Differentiate the original ODE, $\dot{u} = f(t, u(t))$:

$$\ddot{u} = \frac{\partial f}{\partial t} + \dot{u}\frac{\partial f}{\partial u} = \frac{\partial f}{\partial t} + f\frac{\partial f}{\partial u}$$

Approximation error of the 2-step RK method

At $\tau \to 0$,

$$\begin{aligned}\psi_n &= -\frac{u_{n+1} - u_n}{\tau} + f(t_n + \frac{\tau}{2}, u_n + \frac{\tau}{2} f_n) \\ &= -\dot{u}_n - \frac{\tau}{2}\ddot{u}_n + \left(\dot{u}_n + \frac{\tau}{2}\ddot{u}_n\right) + O(\tau^2)\end{aligned}$$

So that the approximation order is indeed quadratic.

Predictor-corrector interpretation

Predictor step:

$$\frac{y_{n+1/2} - y_n}{\tau/2} = f(t_n, y_n)$$

The approximation order is $O(\tau)$.

Corrector step: refine the prediction,

$$\frac{y_{n+1} - y_n}{\tau} = f(t_n + \frac{\tau}{2}, y_{n+1/2})$$

so that the result is $O(\tau^2)$-accurate.

Higher-order Runge-Kutta methods

The two-stage method can be identically reformulated as follows.
Given y_n at $t = t_n$, compute

$$k_1 = f(t_n, y_n)$$

$$k_2 = f(t_n + \frac{\tau}{2}, y_n + \frac{\tau}{2} k_1)$$

Then,

$$\frac{y_{n+1} - y_n}{\tau} = k_2 \ .$$

Higher-order Runge-Kutta methods

Given y_n at $t = t_n$, compute

$$
\begin{aligned}
k_1 &= f(t_n, y_n) \\
k_2 &= f(t_n + a_2\tau, y_n + \tau b_{21} k_1) \\
k_3 &= f(t_n + a_3\tau, y_n + \tau b_{31} k_1 + \tau b_{32} k_2) \\
&\dots
\end{aligned}
$$

Finally,

$$
\frac{y_{n+1} - y_n}{\tau} = \sum_{j=1}^{m} \sigma_j k_j \ .
$$

The coefficients, a_j, b_{jl} and σ_j, are chosen to maximize the order of approximation.

Higher-order Runge-Kutta methods

In practice,

- The most popular RK methods have $m = 4$ and are $O(\tau^4)$.
- Meshes are not uniform, step size is adaptive.

Absolute and conditional stability

Stiff systems of ODEs

Asymptotic stability of ODEs

Consider two IVPs

$$\begin{cases} \dot{u} = f(t, u) \\ u(0) = u_0 \end{cases} \qquad\qquad \begin{cases} \dot{w} = f(t, w) \\ w(0) = w_0 \end{cases}$$

The ODE is *asymptotically stable* if

$$|u(t) - w(t)| \to 0\,, \qquad t \to \infty$$

A motivating example

Consider the ODE

$$\dot{u} = \lambda u\,, \quad t \geqslant 0\,, \quad u(0) = 1$$

with $\lambda < 0$.

The solution is

$$u(t) = e^{\lambda t}\,,$$

which is monotonically decreasing to zero at $t \to \infty$.

A motivating example (a.k.a. the *model equation*)

For any $\tau > 0$, we have (the *asymptotic stability*)

$$|u(t+\tau)| < |u(t)|$$

We expect a similar condition,

$$|y_{n+1}| < |y_n|\,, \qquad n = 0, 1, \cdots$$

to hold for a numeric solution, y_n, of a discretized equation.

Explicit Euler's scheme for the model equation

Consider the explicit Euler's scheme

$$\frac{y_{n+1} - y_n}{\tau} = \lambda y_n \,, \qquad n = 0, 1, \cdots$$

Equivalently, $y_{n+1} = (1 + \tau\lambda) y_n$.

$|y_{n+1}| < |y_n|$ for $n \to \infty$ iff

$$|1 + \lambda\tau| < 1 \qquad \Leftrightarrow \qquad 0 < \tau < \frac{2}{|\lambda|}$$

Explicit Euler's scheme is *conditionally stable*.

Implicit Euler's scheme for the model equation

Consider the implicit Euler's scheme

$$\frac{y_{n+1} - y_n}{\tau} = \lambda y_{n+1}\,, \qquad n = 0, 1, \cdots$$

Equivalently, $y_{n+1} = \dfrac{1}{1-\tau\lambda} y_n\,.$

So that $|y_{n+1}| < |y_n|$ for all $\tau > 0$ and $\lambda < 0$.

Implicit Euler's scheme is *absolutely stable*.

Stiff systems of ODEs

Consider the system of independent ODEs

$$\begin{cases} \dot{u}_1 + \lambda_1 u_1 = 0 \\ \dot{u}_2 + \lambda_2 u_2 = 0 \end{cases}$$

The solution is

$$\begin{cases} u_1(t) = u_1(0)e^{-\lambda_1 t} \\ u_2(t) = u_2(0)e^{-\lambda_2 t} \end{cases}$$

Let $\lambda_2 \gg \lambda_1$.

For $t \gg 1/\lambda_2$, the solution is dominated by $u_1(t)$, but numerically, the step size is limited by $\tau < 1/\lambda_2$.

Stiff systems of ODEs

For a general system of linear equations,

$$\dot{\mathbf{u}} = \mathbf{A}\mathbf{u}$$

Assuming $\mathbf{A}$ can be diagonalized: $\mathbf{\Lambda} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ is diagonal. Then use $\mathbf{u} = \mathbf{Q}\mathbf{w}$, and

$$\dot{\mathbf{w}} = \mathbf{\Lambda}\mathbf{w}$$

The system is *stiff* if

$$s = \frac{\max_k |\operatorname{Re} \lambda_k|}{\min_k |\operatorname{Re} \lambda_k|} \gg 1$$

Stiff systems of ODEs

- If the $\mathbf{A}$ matrix is t-dependent, so is the stiffness ratio, $s(t)$
- Nonlinear systems: linearize, consider a local stiffness ratio $s(t)$
- Stiffness is not a precise term, there is no hard cutoff for s.
- For stiff systems, implicit methods work better.