Initial value problems

We consider the following problem

$$y'(t) = f(t, y(t)) \quad , \quad y(0) = y_0$$

with $f: \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$.

We want to approximate the solution y with the help of a numerical scheme.

First question: Does such a solution exist?
(Otherwise a numerical approximation does not make much sense!)

Theorem
Assume that the function $t \mapsto f(t,u)$ is continuous and fulfills the Lipschitz condition

$$\| f(t,u) - f(t,v) \| \leq L[f] \, \| u - v \|$$

for all $u, v \in \mathbb{R}^m$ with Lipschitz constant $L[f]$. Then there exists a unique solution y to the initial value problem on $[0,T]$ for every initial condition $y(0) = y_0$.

Background

Definition A vector norm $\| \cdot \| : X \to \mathbb{R}$ satisfies

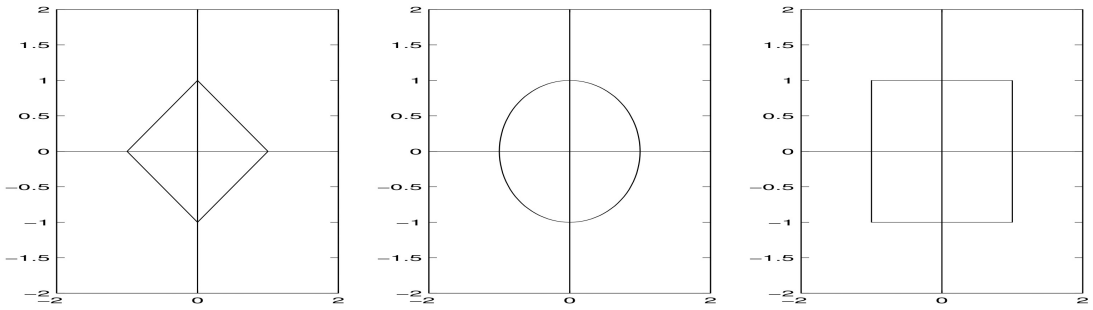
1) $\| u \| \geq 0$; $\quad \| u \| = 0 \Leftrightarrow u = 0$
2) $\| \alpha u \| = |\alpha| \, \| u \|$
3) $\| u \| - \| v \| \leq \| u \pm v \| \leq \| u \| + \| v \|$

A norm generalizes the notion of distance between points.

Important norms for vectors in $\mathbb{R}^m$

ℓ^p norms are defined by

$$\| x \|_p = \begin{cases} \left(\sum_{i=1}^{m} |x_i|^p \right)^{\frac{1}{p}} , & p \in [1, \infty) \\ \max_{i \in \{1, \dots, m\}} |x_i| , & p = \infty \end{cases}$$



$p=1$ $\quad$ $p=2$ (Euclidean norm) $\quad$ $p=\infty$

We can generalize vector norms to matrices:

Definition

Operator norm associated with the vector norm $\|\cdot\|$ is defined by

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

Properties

1) $\|Ax\| \leq \|A\| \, \|x\|$
2) $\|AB\| \leq \|A\| \, \|B\|$

Examples

vector norm	matrix norm
$\|x\|_1 = \sum_{i=1}^{m} \|x_i\|$	$\max_{j \in \{1,\dots,m\}} \sum_{i=1}^{m} \|a_{ij}\|$
$\|x\|_2 = \left(\sum_{i=1}^{m} \|x_i\|^2 \right)^{\frac{1}{2}}$	$\left(\rho[A^H A] \right)^{\frac{1}{2}}$
$\|x\|_\infty = \max_{i \in \{1,\dots,m\}} \|x_i\|$	$\max_{i \in \{1,\dots,m\}} \sum_{j=1}^{m} \|a_{ij}\|$

where ρ is the spectral radius of a matrix defined by

$$\rho[A] = \max |\lambda[A]| \quad \text{(maximal eigenvalue)}$$

Back to initial value problem

Important example for us: For $A \in \mathbb{R}^{m,m}$

$$y' = f(t,y) = Ay \quad , \quad y(0) = y_0$$

then f fulfills the Lipschitz condition

$$\|f(t,u) - f(t,v)\| = \|Au - Av\| \leq L[f]\,\|u - v\| \quad \text{for all } u,v \in \mathbb{R}^m$$

with the Lipschitz constant

$$L[f] = \max_{u \neq v} \frac{\|Au - Av\|}{\|u - v\|} \overset{w = u - v}{\uparrow} = \max_{y \neq 0} \frac{\|Aw\|}{\|w\|} = \|A\|$$

Thus, the matrix norm $\|A\|$ is the Lipschitz constant for $f(t,y) = Ay$.

This property is often not fulfilled for non linear problems.

Example

The problem $y' = f(t,y) = y^2 \quad , \quad y(0) = y_0 > 0$

has the solution $y(t) = \frac{y_0}{1 - y_0 t}$. This function is only defined for $t < \frac{1}{y_0}$ and blows up at $= 1/y_0$. Therefore, we cannot expect a solution on every interval $[0,T]$.

Test: $$y'(t) = \left(y_0 (1 - y_0 t)^{-1}\right) = y_0 \cdot (-1) \cdot (1 - y_0 t)^{-2} \cdot (-y_0)$$
$$= y_0^2 (1 - y_0 t)^{-2} = (y(t))^2$$

i.e. $y'(t) = (y(t))^2$ and $y(0) = \frac{y_0}{1 - y_0 \cdot 0} = y_0$ ∎

The function $f(t,y) = y^2$ is not Lipschitz continuous on $\mathbb{R}$. So we cannot apply the Theorem.

To see that we show that there exists no $L[f]$ such that

$$\| f(t,u) - f(t,v) \| \leq L[f] \, \|u-v\| \quad \text{for all } u, v.$$

If such a $L[f]$ existed, the we can choose $u=0$ and $v = L[f]+1$ and find

$$\| f(t,0) - f(t, L[f]+1) \| = \| 0^2 - (L[f]+1)^2 \|$$
$$= \underbrace{(L[f]+1)}_{> L[f]} \; \underbrace{\| 0 - (L[f]+1) \|}_{\|u-v\|}$$
$$> L[f] \, \|u-v\|$$

Other possible examples that fit in the setting

Given a problem

$$y'' = f(t, y, y') \text{ with } y(0) = y_0, \; y'(0) = y_0'$$

we rewrite the equation into

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y \\ y' \end{pmatrix}$$

and get a first order system

$$x' = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} x_2 \\ f(t, x_1, x_2) \end{pmatrix} \text{ with } x(0) = \begin{pmatrix} y_0 \\ y_0' \end{pmatrix}.$$

Conclusion: We can rewrite a higher order initial value problem into an initial value problem of order 1 in a higher dimensional space. Therefore it is enough to study first order problems in the following.

Explicit (forward) Euler method (1768)

Idea: Replace y' in $y' = f(t, y)$ by finite difference approximation

$$y'(t_n) \approx \frac{y(t_n + h) - y(t_n)}{h}$$

Let y_n denote the numerical approximation to $y(t_n)$ in

$$\frac{y_{n+1} - y_n}{h} = f(t_n, y_n), \quad y_0 = y(t_0)$$

Rewritten this yields the explicit (forward) Euler method

$$\begin{cases} y_{n+1} = y_n + h f(t_n, y_n) \\ t_{n+1} = t_n + h \end{cases}$$

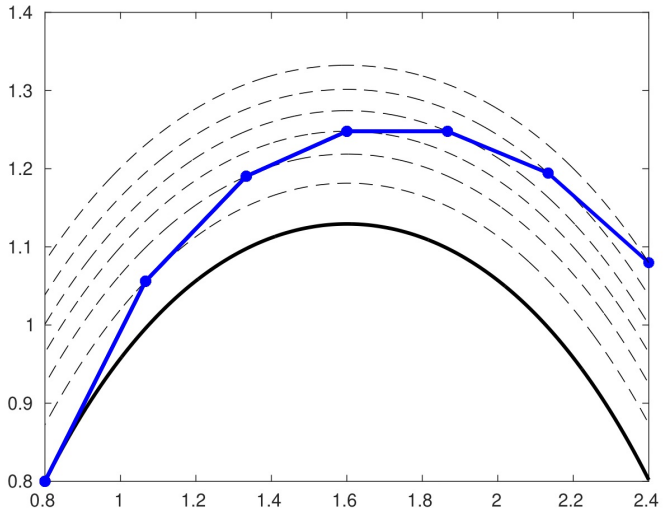
We can show that it is a useful method by expanding in Taylor series

$$\begin{aligned} y(t+h) &= y(t) + h y'(t) + \frac{h^2}{2!} y''(\xi) \\ &= y(t) + h f(t, y(t)) + O(h^2) \end{aligned}$$

Neglecting $O(h^2)$ we see

$$\begin{aligned} y(t+h) &\approx y(t) + h f(t, y(t)) \\ y_{n+1} &= y_n + h f(t_n, y_n) \quad \text{(explicit Euler)} \end{aligned}$$

Geometric interpretation:



Convergence

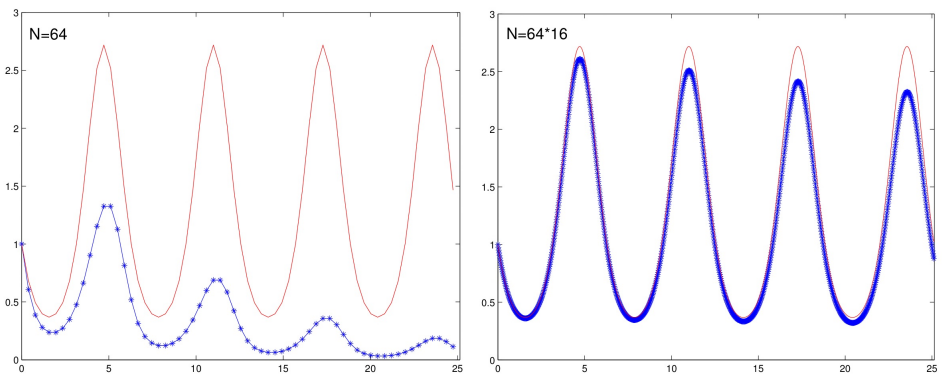
Definition

A method is convergent if for every initial value problem with a vector field f and every fixed $T = N \cdot h$ it holds that

$$\lim_{N \to \infty} \| y_{N,h} - y(T) \| = 0$$

This means that for N (number of steps) to infinity or h (step size) to zero the approximation becomes the exact solution (difference tends to zero.)

Example:



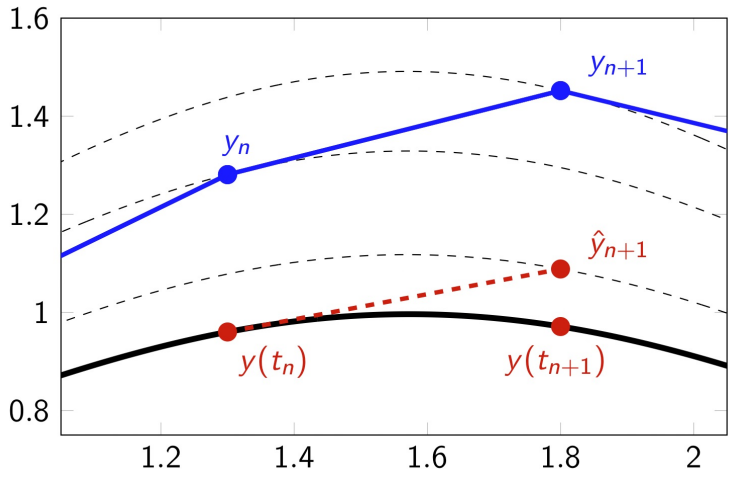
In the following, we analyse the error of a method, ie. we consider the difference between exact solution and numerical solution..

We look at two different types of errors:

global error: $e_{n+1} = y_{n+1} - y(t_{n+1})$

local error: $l_{n+1} = \hat{y}_{n+1} - y(t_{n+1})$

↖ solution of one step of the scheme where we start on the exact solution



Rule of thumb:

- We are most interested in the global error.
- The local error is easier to obtain (there exists an approach that helps to find it).
- Once we have the local error, we can use it to estimate the global error.

Local error

The local error is the error made after a single step given the exact solution

$$l_{n+1} = \hat{y}_{n+1} - y(t_{n+1})$$

For the explicit Euler scheme $\hat{y}_{n+1}$ is given by

$$\hat{y}_{n+1} = y(t_n) + f(t_n, y(t_n))$$

(insert exact solution in scheme)

We can interpret the local error as the residual in the equation

$$y(t_{n+1}) = y(t_n) + h\, f(t_n, \underbrace{y(t_n)}_{= y'(t_n)}) - l_{n+1}$$

(l_{n+1} ← residual)

If we expand in Taylor series for exact solution:

$$y(t_{n+1}) = y(t_n) + h\, y'(t_n) + \frac{h^2}{2} y''(t_n) + \dots$$

Thus, the local error is given by

$$l_{n+1} = -\frac{h^2}{2} y''(t_n) - \frac{h^3}{3!} y'''(t_n) + \dots \in O(h^2)$$

$$\leq C h^2$$

(for smooth function y'')

Summary:

- Insert exact solution in the numerical scheme
- Expand in Taylor series

Global error

The nth global error is the error made by the scheme after n steps. For the explicit Euler scheme we have:

$$\begin{aligned}
\|e_{n+1}\| &= \| y_{n+1} - y(t_{n+1}) \| \\
&= \| \underbrace{y_n + h f(t_n, y_n)}_{y_{n+1}\,(\text{scheme})} - \underbrace{(y(t_n) + h y'(t_n) - l_{n+1})}_{y(t_{n+1})\,(\text{scheme + local error})} \| \\
&\leq \|e_n\| + \|h f(t_n, y(t_n) + e_n) - h f(t_n, y(t_n))\| \\
&\qquad + \|l_{n+1}\| \\
&\leq \|e_n\| + hL\|e_n\| + \|l_{n+1}\|
\end{aligned}$$

Thus, the local error accumulates into the global error. We can express the error using the following lemma:

Lemma (Grönwall)

Assume that a non-negative sequence (a_n) satisfies

$$a_{n+1} \leq b\, a_n + c h^p$$

with $a_0 = 0$ for $p > 0$. Then for all $n \in \mathbb{N}$

$$a_{n+1} \leq \frac{ch^p}{b-1}(b^{n+1} - 1).$$

Proof

$$\begin{aligned}
a_{n+1} &\leq b\, a_n + c h^p \\
&\leq b(b\, a_{n-1} + c h^p) + c h^p \\
&\leq b(b(b\, a_{n-2} + c h^p) + c h^p) + c h^p \\
&\leq \dots
\end{aligned}$$

$$\leq \underbrace{b^{n+1} a_0}_{=0} + \sum_{k=0}^{n} b^k ch^p$$

$$= ch^p \frac{b^{n+1}-1}{b-1}$$

$$= \frac{ch^p}{b-1}\left(b^{n+1}-1\right)$$

□

Theorem

If f fulfills a Lipschitz condition, then the explicit Euler method is convergent, i.e. the global error tends to zero as the step size tends to zero (or the number of steps tends to ∞). More precisely, $\|e_N\| \leq h\,C(T)$ for a step size h and a constant $C(T)$ that does not depend on h.

Proof

For a number N of steps, a final time T and a step size $h = \frac{T}{N}$, we consider the global error. For $n \in \{1, \dots, N\}$, we have

$$\|e_{n+1}\| = \|e_n + h\left(f(t_n, y(t_n)+e_n) - f(t_n, y(t_n))\right) + \ell_{n+1}\|$$

$$\leq \|e_n\| + hL[f]\,\|y(t_n)+e_n - y(t_n)\| + \|\ell_{n+1}\|$$

$$= (1+hL[f])\,\|e_n\| + \|\ell_{n+1}\|$$

Grönwall's inequality with $b = 1+hL[f]$ and $p = 2$ becomes

$$a_{n+1} \leq \frac{ch^2}{hL[f]}\left(\left(1+\frac{(n+1)hL[f]}{n+1}\right)^{n+1} - 1\right)$$

$$\leq \frac{ch}{L[f]}\left(\exp\left((n+1)hL[f]\right) - 1\right).$$

Thus, we find that

$$\|e_n\| \leq \frac{c}{L[f]} h \left(e^{TL[f]} - 1\right)$$

where

$$c = \max_n \frac{\|\ell_n\|}{h^2} \approx \max_t \frac{\|y''(t)\|}{2}$$

This shows that

$$\|e_n\| \leq h\, C(T) = h \max_t \frac{\|y''(t)\|}{2} \frac{e^{TL[f]} - 1}{L[f]}$$

($C(T)$: independent of h)

and therefore the method is convergent as $\|e_n\|$ tends to zero as the step size h tends to zero. □

Conclusion

1) local error can be obtained by inserting the exact solution into the scheme and an application of Taylor's expansion.
   (less important in applications but easier to obtain)
2) global error can be obtained by a recursion including the local error
   (important, use local error in derivation)

Question: Is the error bound really helpful?

Example:

$$y' = -100\,y \,, \quad y(0) = 1$$

Then $L[f] = 100$ and exact solution $y(t) = e^{-100t}$ with $y''(t) = 100^2\, e^{-100t}$. Thus, we obtain the error bound

$$\|e_n\| \leq h \max_t \frac{\|y''(t)\|}{2} \frac{e^{TL[f]} - 1}{L[f]}$$

$$= h \frac{100^2}{2} \cdot \frac{e^{100T} - 1}{100}$$

$$\approx h \; 50 \cdot e^{100T} \underset{T=1}{\approx} h \cdot 1{,}34 \cdot 10^{45}$$

What is the actual error of the scheme?

$$y_{n+1} = y_n + hf(t_n, y_n) = (1 - 100h)\, y_n = \dots = (1-100h)^{n+1} \underbrace{y(0)}_{=1}$$

For $T = 1$ and $\frac{1}{N} = h = \frac{1}{100}$, it follows that

$$\|e_n\| = \|y(1) - y_N\| = \left| e^{-100} - \left(1 - 100 \cdot \tfrac{1}{N}\right)^N \right| \leq 3{,}72 \cdot 10^{-44}$$

$$= h \; 3{,}72 \cdot 10^{-42}$$

The error estimate is useless in practice! But we will do better later!

Order of consistency

For an arbitrary scheme

$$y_{n+1} = \Phi_h(f, t_n, y_{n+1}, y_n, y_{n-1}, \dots)$$ (examples in following sections)

the local error is given by

$$l_{n+1} = y(t_{n+1}) - \Phi_h(f, t_n, y(t_{n+1}), y(t_n), y(t_{n-1}), \dots)$$

Definition

The order of consistency is p if

$$l_{n+1} = O(h^{p+1})$$

as $h \to 0$ for every analytic f.

Alternative

The order of consistency is p if the method is exact for all polynomials $y(t) = P(t)$ of degree p or less.

Example (Explicit Euler)

Using Taylor expansion, we find

$$l_{n+1} = \underbrace{y(t_{n+1})}_{y(t_n) + h y'(t) + O(h^2)} - \big(y(t_n) + h \underbrace{f(t_n, y(t_n))}_{y'(t_n)}\big) = O(h^2) = O(h^{1+1})$$

Alternatively, we insert the monomials $y(t) = t^q$, $q = 0, 1, \dots$

- $y(t) = 1$, $f(t, y(t)) = y'(t) = 0$. Then

$$y(t_{n+1}) = 1$$

$$y(t_n) + h f(t, y(t_n)) = 1 + h \cdot 0 = 1$$

↺ coincide

- $y(t) = t$, $f(t, y(t)) = y'(t) = 1$. Then

$$y(t_{n+1}) = t_{n+1}$$
$$y(t_n) + h f(t_n, y(t_n)) = t_n + h \cdot 1 = t_{n+1}$$

↕ coincide

- $y(t) = t^2$, $f(t, y(t)) = y'(t) = 2t$. Then

$$y(t_{n+1}) = t_{n+1}^2 = (t_n + h)^2$$
$$y(t_n) + h f(t_n, y(t_n)) = t_n^2 + 2ht_n = (t_n + h)^2 - h^2$$

↘ don't coincide

By linearity it follows that

$$y(t_{n+1}) = y(t_n) + h f(t_n, y(t_n))$$

for all polynomials up to order 1.

We have shown that the explicit Euler method is consistent of order 1.

Inner products

Definition

A bilinear form $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{C}$ satisfying

1) $\langle u, u \rangle \geq 0$, $\langle u, u \rangle = 0 \Leftrightarrow u = 0$
2) $\langle u, v \rangle = \overline{\langle v, u \rangle}$
3) $\langle u, \alpha v \rangle = \alpha \langle u, v \rangle$, $\alpha \in \mathbb{C}$
4) $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$

Important: An inner product generates a norm

$$\langle u, u \rangle = \|u\|^2$$

Examples

1) Scalar product in $\mathbb{R}^n$: $\langle u, v \rangle_{\mathbb{R}^n} = u^T v$
   Euclidean norm: $\|u\|_2^2 = \sum_{i=1}^{n} |u_i|^2$

2) General inner product in $\mathbb{C}^n$: $\langle u, v \rangle_G = u^{\#} G v$
   for G symmetric, positiv definite matrix

3) Inner product for $L^2[0,1]$: $\langle u, v \rangle_{L^2} = \int_0^1 \overline{u(x)} v(x) dx$
   Norm for L^2: $\|u\|_{L^2}^2 = \int_0^1 |u(x)|^2 dx$

Theorem (Cauchy-Schwarz inequality)

$$|\langle u, v \rangle| \leq \|u\| \, \|v\|$$

Definition (strong accretivity/dissipativity condition)

A mapping $A : X \to X$ fulfills a strong accretivity condition, if there exists $\mu > 0$ such that

$$\langle Au - Av, u - v \rangle \geq \mu \|u - v\|^2$$

A is called strongly dissipative $\Leftrightarrow$ $-A$ is strongly accretive,

i.e. there exists $\tilde{\mu} > 0$ such that

$$\langle Au - Av, u - v \rangle \leq -\tilde{\mu} \|u - v\|^2$$

Example:

Let $A \in \mathbb{R}^{n,n}$ be a symmetric positive def. matrix. Then there exist n linearly independent, orthonormal eigenvectors $v_1, \dots, v_n$ with there corresponding eigenvalues $\lambda_1, \dots, \lambda_n$.

First, note that due to the linearity

$$\begin{aligned} &\langle Au_1 - Au_2, u_1 - u_2 \rangle \\ =&\langle A(u_1 - u_2), u_1 - u_2 \rangle \\ =&\langle Au, u \rangle \qquad \text{for } u := u_1 - u_2. \end{aligned}$$

As the eigenvectors build a basis of $\mathbb{R}^n$, every u can be writen as

$$u = \alpha_1 v_1 + \dots + \alpha_n v_n \quad \text{for } \alpha_1, \dots, \alpha_n \in \mathbb{R}.$$

It then follows that

$$\begin{aligned} &\langle Au, u \rangle \\ =&\langle A(\alpha_1 v_1 + \dots + \alpha_n v_n), \alpha_1 v_1 + \dots + \alpha_n v_n \rangle \\ =&\sum_{i,j=1}^{n} \alpha_i \alpha_j \langle Av_i, v_j \rangle \\ =&\sum_{i,j=1}^{n} \alpha_i \alpha_j \lambda_i \langle v_i, v_j \rangle \\ =&\sum_{i=1}^{n} \alpha_i^2 \lambda_i \langle v_i, v_i \rangle \\ \geq&\min_{i \in \{1,\dots,n\}} \lambda_i \sum_{i,j=1}^{n} \langle \alpha_i v_i, \alpha_j v_j \rangle = \min_{i \in \{1,\dots,n\}} \lambda_i \|u\|^2 \end{aligned}$$

$\langle Av_i, v_j \rangle = \lambda_i v_i$

$= 1$ if $i = j$ as v_i, v_j are
$= 0$ if $i \neq j$ orthonormal

$$\Rightarrow \quad \mu = \min \lambda_i \qquad (\mu \text{ is the smallest eigenvalue})$$

Implicit methods

Standard example:

Implicit (backward) Euler

$$y_{n+1} = y_n + h\, f(t_{n+1}, y_{n+1})$$

This means that y_{n+1} depends on itself. In practice we need to solve a nonlinear function. This creates extra costs but one can use larger steps for many examples (stiff problems, later)

Consistency of the method

Using Taylor expansion we find

$$\ell_{n+1} = y(t_{n+1}) - \big(y(t_n) + h \underbrace{f(t_{n+1}, y(t_{n+1}))}_{y'(t_{n+1})}\big)$$

$$= y(t_n) + h y'(t_n) + O(h^2) - y(t_n) - h\, y'(t_{n+1})$$

$$= h\, y'(t_n) + O(h^2) - h\,(y'(t_n) + O(h)) \in O(h^2) + h\, O(h) \in O(h^2)$$

Other examples for implicit schemes:

-) trapezoidal rule

$$y_{n+1} = y_n + \frac{h}{2}\left(f(t_n, y_n) + f(t_{n+1}, y_{n+1})\right)$$

Exercise: consistent of order 2

-) implicit midpoint rule

$$y_{n+1} = y_n + h\, f\left(\frac{t_n + t_{n+1}}{2}, \frac{y_n + y_{n+1}}{2}\right)$$

Exercise: consistent of order 2

·) theta method

$$y_{n+1} = y_n + h\left(\theta f(t_{n+1}, y_{n+1}) + (1-\theta) f(t_n, y_n)\right)$$

$\theta \in [0,1]$

for $\theta = 0$: explicit Euler

$\theta = \frac{1}{2}$: trap. rule

$\theta = 1$: implicit Euler

Using the strong dissipativity condition, we can provide an error bound for the implicit Euler method with an improved error constant compared to our first error bound for the explicit Euler method.

We consider the equation:

$$y'(t) = Ay(t) + f(t), \quad y(0) = y_0.$$

where A fulfills a strong dissipativity condition with $\mu > 0$

Scheme: $y_{n+1} = y_n + h A y_{n+1} + h f(t_{n+1})$

Local error: $y(t_{n+1}) = y(t_n) + h A y(t_{n+1}) + h f(t_{n+1}) - l_{n+1}$

Global error: $e_n = y_n - y(t_n)$

We obtain:

$$\begin{aligned} y_{n+1} - y(t_{n+1}) &= y_n + h A y_{n+1} + h f(t_{n+1}) \\ &\quad - y(t_n) + h A y(t_{n+1}) + h f(t_{n+1}) + l_{n+1} \\ &= e_n + h A (y_{n+1} - y(t_{n+1})) + l_{n+1} \end{aligned}$$

$$\Leftrightarrow \quad e_{n+1} - e_n = h A e_{n+1} + l_{n+1}$$

We "test" (multiply) with e_{n+1}:

$$\underbrace{\langle e_{n+1} - e_n, e_{n+1}\rangle}_{(*)} = \underbrace{h\langle A e_{n+1}, e_{n+1}\rangle}_{(**)} + \underbrace{\langle l_{n+1}, e_{n+1}\rangle}_{(***)}$$

$$\begin{aligned}(*) &= \tfrac{1}{2}\langle e_{n+1} - e_n, e_{n+1}\rangle + \tfrac{1}{2}\langle e_{n+1} - e_n, e_{n+1} - e_n\rangle \\ &\quad + \tfrac{1}{2}\langle e_{n+1} - e_n, e_n\rangle \\ &= \tfrac{1}{2}\|e_{n+1}\|^2 - \tfrac{1}{2}\langle e_n, e_{n+1}\rangle + \tfrac{1}{2}\|e_{n+1} - e_n\|^2 \\ &\quad + \tfrac{1}{2}\langle e_{n+1}, e_n\rangle - \tfrac{1}{2}\|e_n\|^2 \\ &= \tfrac{1}{2}\left(\|e_{n+1}\|^2 - \|e_n\|^2 + \|e_{n+1} - e_n\|^2\right)\end{aligned}$$

$(**) \leq -h\mu\|e_{n+1}\|^2$ (strongly dissipative)

$(***) \leq \|l_{n+1}\|\,\|e_{n+1}\|$ (Cauchy Schwarz)

$$= \frac{1}{\sqrt{h\mu}}\|l_{n+1}\|\sqrt{h\mu}\,\|e_{n+1}\| \leq \frac{1}{2}\frac{1}{h\mu}\|l_{n+1}\|^2 + \frac{1}{2}h\mu\|e_{n+1}\|^2$$

($|ab| \leq \frac{1}{2}a^2 + \frac{1}{2}b^2$)

Altogether, we have

$$\tfrac{1}{2}\left(\|e_{n+1}\|^2 - \|e_n\|^2 + \|e_{n+1} - e_n\|^2\right) \leq \left(-h\mu + \tfrac{1}{2}h\mu\right)\|e_{n+1}\|^2 + \frac{1}{2h\mu}\|l_{n+1}\|^2$$

$$\Leftrightarrow (1 + h\mu)\|e_{n+1}\|^2 + \underbrace{\|e_{n+1} - e_n\|^2}_{\geq 0} \leq \frac{1}{h\mu}\|l_{n+1}\|^2 + \|e_n\|^2$$

$$\Leftrightarrow \|e_{n+1}\|^2 \leq \frac{1}{1+h\mu}\left(\|e_n\|^2 + \frac{1}{h\mu}\|l_{n+1}\|^2\right)$$

$$\leq \frac{1}{1+h\mu}\|e_n\|^2 + \frac{c^2h^3}{(1+h\mu)\mu}$$

($\|l_{n+1}\| \leq ch^2$)

Then it follows that

$$\|e_{n+1}\|^2 \leq \frac{1}{1+h\mu}\|e_n\|^2 + \frac{c^2 h^3}{(1+h\mu)\,2\mu}$$

Reminder:

Lemma (Grönwall)

Assume that a non-negative sequence (a_n) satisfies

$$a_{n+1} \leq b\, a_n + \tilde{c} h^p$$

with $a_0 = 0$ for $p > 0$. Then for all $n \in \mathbb{N}$

$$a_{n+1} \leq \frac{\tilde{c} h^p}{b-1}\left(b^{n+1} - 1\right).$$

For $b = \frac{1}{1+h\mu}$, we similarly find that

$$\begin{aligned} a_{n+1} &\leq \frac{\tilde{c}\, h^p}{\frac{1-(1+h\mu)}{1+h\mu}}\left((1+h\mu)^{-n-1} - 1\right) \\ &= \frac{\tilde{c} h^{p-1}}{-\mu}(1+h\mu)\underbrace{\left((1+h\mu)^{-n-1} - 1\right)}_{<0} \\ &= \frac{\tilde{c} h^{p-1}}{\mu}(1+h\mu)\left(1 - (1+h\mu)^{-n-1}\right) \\ &< \frac{\tilde{c} h^{p-1}}{\mu}(1+h\mu) \end{aligned}$$

Thus, we obtain

$$\tilde{c} = \frac{c^2}{(1+h\mu)\,2\mu}$$

$$\|e_{n+1}\|^2 \le \frac{c^2 h^2}{(1+h\mu)\,2\mu}\,(1+h\mu)\cdot\frac{1}{\mu}$$

$$= \frac{c^2 h^2}{2\mu^2}$$

and in particular

$$\|e_{n+1}\| \le \frac{c\,h}{\sqrt{2}\,|\mu|}.$$

strong dissipativity

For previous example $y'(t) = -100\,y(t)$

$$A = -100, \quad \|A\| = 100, \quad \mu = 100$$

Then we avoid the e^{100} and obtain

$$\|e_{n+1}\| \le \max_t \frac{\|y''(t)\|}{2}\,\frac{h}{\sqrt{2}\cdot 100}$$

$$= \frac{100^2}{2}\,\frac{h}{\sqrt{2}\,100} = \frac{50}{\sqrt{2}}\,h$$

Comparison: For the explict Euler we proved with a more naive analysis

$$\|e_n\| \le h\;\; 50\cdot e^{100T} \underset{T=1}{\approx} h\cdot 1{,}34\cdot 10^{45}.$$

The error constant became much smaller!

Convergence order

Definition

The order of convergence is p if for every fixed $T = N \cdot h$ and every Lipschitz vector field f

$$\| y_{N,h} - y(T) \| = O(h^p)$$

as $h \to 0$.

Note

The convergence order is p if the global error is in $O(h^p)$ while the consistency order is p if the local error is in $O(h^{p+1})$

We want: A link between convergence and consistency: Stability

Theorem (Lax Principle)

Consistency + Stability = Convergence

Thus, we study stability in the following.

Stability

In general that means that the numerical solution stays bounded, i.e. $\| \Phi_h(y_{n+1}) \| \leq C$ for all $h > 0$. Here, we consider the linear test equation

$$y' = \lambda y \,, \quad y(0) = 1, \quad t \geq 0 \,, \quad \lambda \in \mathbb{C}.$$

As $y(t) = e^{\lambda t}$, it follows that

$$|y(t)| \leq C \iff \mathrm{Re}(\lambda) \leq 0$$

Mathematical stability:

Bounded solutions if $\mathrm{Re}(\lambda) \leq 0$

Question: • When does a numerical method have this property?

• Does $\mathrm{Re}(\lambda) \leq 0$ imply numerical stability?

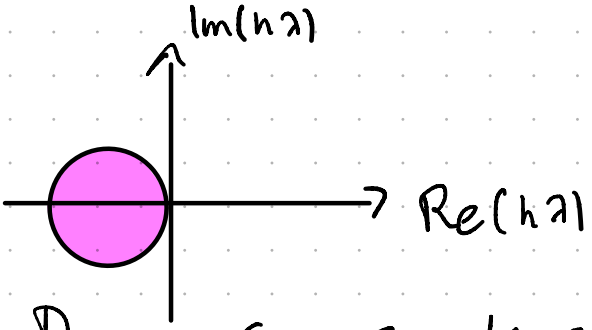
Definition (Stability region)

The stability region D of a method is the set of all $h\lambda \in \mathbb{C}$ such that $|y_n| \leq \tilde{K}$ when the method is applied to the linear test equation.

Example

• Explicit Euler method: $y_{n+1} = (1 + h\lambda) y_n$.

y_n remains bounded if and only if $|1 + h\lambda| \leq 1$

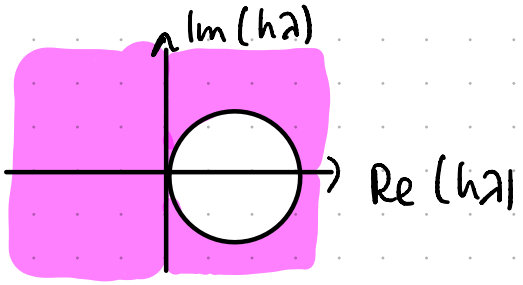


$$D_{Euler} = \{z \in \mathbb{C} : |1+z| \leq 1\}$$

·) Implicit Euler method: $y_{n+1} = y_n + h\lambda y_{n+1}$

$$\Leftrightarrow y_{n+1} = (1-h\lambda)^{-1} y_n$$

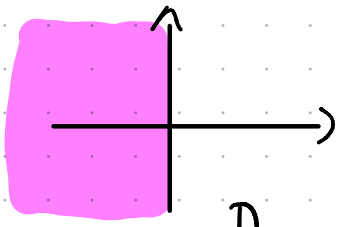
y_n remains bounded if and only if $|1-h\lambda|^{-1} \leq 1$



$$D_{impEuler} = \{z \in \mathbb{C} : |1-z|^{-1} \leq 1\}$$

·) Trapezoidal rule: $y_{n+1} = y_n + \frac{1}{2}h\lambda y_n + \frac{1}{2}h\lambda y_{n+1}$

$$\Leftrightarrow y_{n+1} = (1-\tfrac{h\lambda}{2})^{-1}(1+\tfrac{h\lambda}{2})\, y_n$$



$$D_{trap} = \{z \in \mathbb{C} : \left|\frac{1+\frac{z}{2}}{1-\frac{z}{2}}\right| \leq 1\}$$

Definition (A-stability)

A method is called A-stable if $\mathbb{C}^- = \{z \in \mathbb{C} : \mathrm{Re}\, z \leq 0\} \subseteq D$.

Examples:

·) Exp Euler is not A-stable.

·) Imp Euler is A-stable.

·) trapezoidal rule is A-stable.

Idea

If the original problem is stable, then an A-stable method will replicate this behaviour numerically.

Note (Study question)

The linear test equation is relevant for other problems:

Assume that $y' = Ay$ is diagonalizable with $AT = T\Lambda$.

Then $u = T^{-1}y$ satisfies $u' = \Lambda u$.

Exp Euler for $y' = Ay$:

$$y_{n+1} = (I + hA)\, y_n$$

Putting $u_n = T^{-1} y_n$ leads to

$$u_{n+1} = (I + h\Lambda)\, u_n \quad \text{i.e. independent eq } (u_{n+1})_k = (1 + h\lambda_k)(u_n)_k$$

Thus: T diagonalizes the differential eq and it's discretization. λ can then be interpreted as an eigenvalue of A.

We need $|1 + h\lambda| \leq 1$ for all eigenvalues λ.

Stiffness

There is no very accurate definition, but one can say a problem is stiff when explicit methods have severe step size restrictions and implicit methods have not and perform better.

Example
$$\begin{cases} y' = \lambda(y - \sin t) + \cos t \\ y(0) = 0 \end{cases}$$

Solution $y(t) = \sin t$ for every $\lambda \in \mathbb{R}$.

For an explicit method the stability region is bounded and depends on λ. For explicit Euler, we need for example $|1+h\lambda| \leq 1$. Thus, for $\lambda \ll -1$ the step size h has to be choosen very small! (i.e. many steps of the method are needed.)
No such problem for an A-stable method (h does not depend on λ).