Num. Stability

Motivation:

For IVP: $y'(t) = -10 \cdot y(t)$, $0 < t < 1$. $y(0) = 1$. its sol. decays exponentially. But when we apply implicit / explicit Euler method. The num. sol. of implicit one looks good while the explicit one begins to oscillate and even explode when step length $h \uparrow$.

(1) model problem:

Fix $\lambda \in \mathbb{C}$. $y'(t) = \lambda y(t)$, $t_0 \le t \le T$. $y(t_0) = y_0$.

which has sol. $y(t) = y_0 e^{\lambda(t-t_0)}$. (*).

Note $|y(t)| = |y_0| e^{Re(\lambda)(t-t_0)}$. the limit behavior depend on $Re(\lambda) >, =, < 0$ when $t \to \infty$.

Rmk: Num. sol. should imitate its exp. decay behavior if $Re(\lambda) < 0$.

Def: A RK method is called absolutely

stable for a λh_k if it produces bdd approxi. $\sup_n |y_n| < \infty$ when applied on (*) for $Re(\lambda) < 0$. (h_k is k^{th} step length)

<u>e.g.</u> (Explicit Euler)

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

$$= \cdots = (1 + h\lambda)^n y_0.$$

We call $g(h\lambda) = 1 + h\lambda$ amplification. factor ($|g(h\lambda)| < 1 \Leftrightarrow |y_n| < |y_{n-1}|$)

So we require: $|g(h\lambda)| \leq 1$ for stability

<u>e.g.</u> (For $\lambda = -10$)

- $h = \frac{1}{20}$. it's good.
- $h = \frac{1}{6}$. oscillates but bdd.
- $h = \frac{1}{4}$. explodes.

<u>Rmk</u>: Absolutely stability only resures it's bdd but not good approxi.

Amplification factor examples:

i) Explicit Euler $g(\lambda h) = 1 + \lambda h$.

ii) Heun : $g(\lambda h) = \sum_0^2 (\lambda h)^k / k!$

iii) RK4 : $g(\lambda h) = \sum_0^4 (\lambda h)^k / k!$

Rmk : It's polynomial from Taylor expansion.

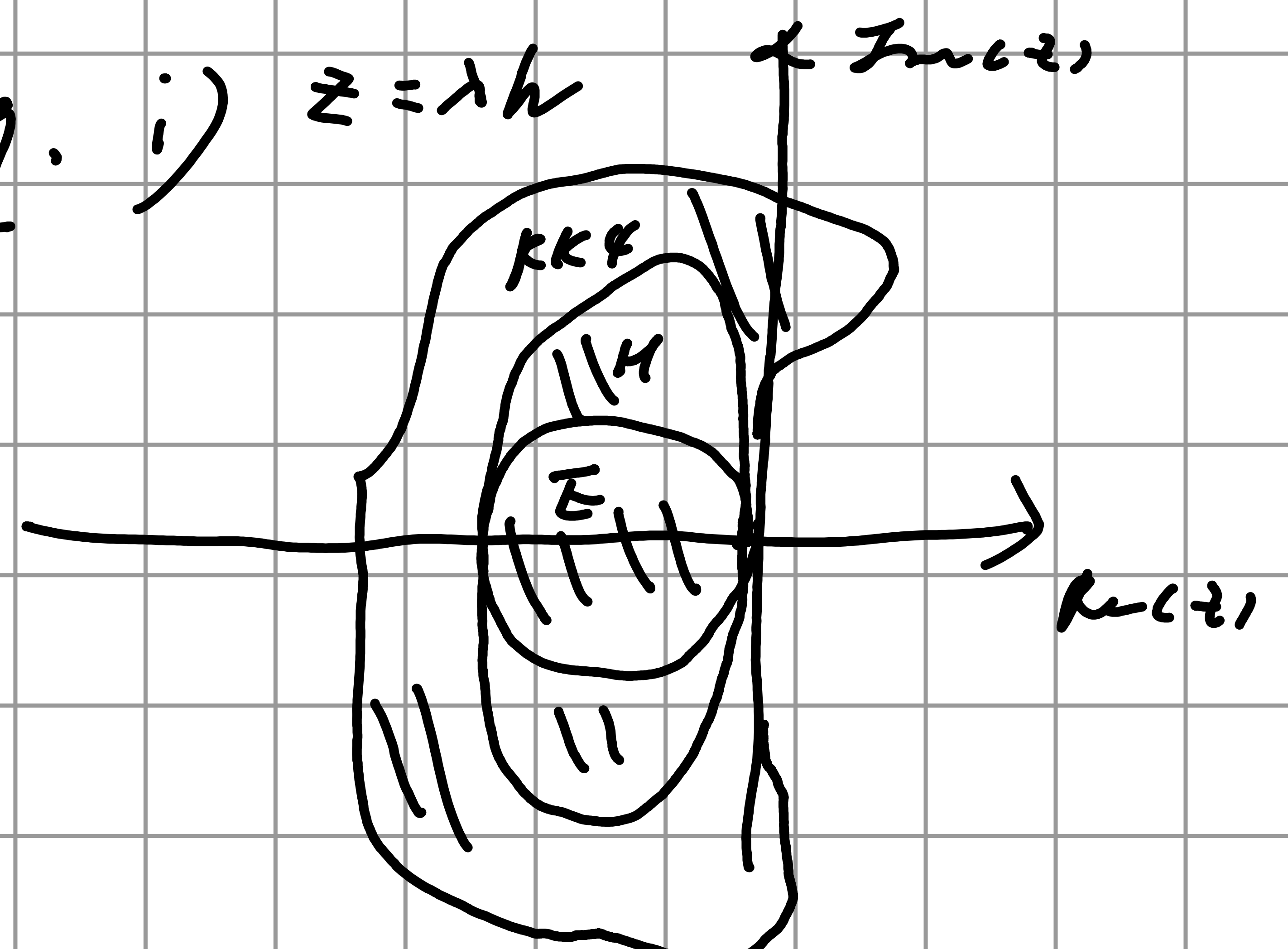
And $|g(\lambda h)| \uparrow$ when $|\lambda h| \uparrow$.

Def : The stability region of a RK scheme is given by $SR = \{\lambda h \in \mathbb{C} \mid |g(h\lambda)| \leq 1\}$.

For $\lambda \in \mathbb{R}^1$, the stability interval is :

$RI = \{\lambda h \in \mathbb{R}^1 \mid |g(h\lambda)| \leq 1\}$ where $g(\lambda h)$ is amplify factor from applying scheme on (*).

e.g. i) $z = \lambda h$



a) Explicit Euler :

$RI = [-2, 0]$

b) Heun : $RI = [-2, 0]$

c) RK4 : $RI = [-2.78, 0]$.

ii) (Nonlinear case)

For IVP $y'(t) = \mu(y(t) - \cos(t)) - \sin(t)$

$y(0) = 1$. Note $\partial f / \partial y = \mu$. So μ will play the role of λ.

$\Rightarrow$ We need: $\mu \leq 0$. $-\mu h \leq 2.78$ for RK4.

<u>Def</u>: A method is called A-stable if

$\{z \in \mathbb{C} \mid \mathrm{Re}(z) \leq 0\} \subset SR$.

<u>Rmk</u>: Explicit scheme can't be A-stable

$\Rightarrow$ It motivates us to use the implicit method.

<u>e.g.</u> i) Implicit Euler: $g(\lambda h) = 1/(1-\lambda h)$.

ii) Implicit Trap.: $g(\lambda h) = \dfrac{1+\frac{1}{2}\lambda h}{1-\frac{1}{2}\lambda h}$.

<u>Rmk</u>: To chose step length. We need to consider: i) Stability (restriction on h)

ii) Accuracy: h small. accurate $\uparrow$

iii) Efficiency: h large. efficient $\uparrow$.

(2) <u>Stability func. for RK</u>:

Apply RK method on model problem: $(f = \lambda x)$

$k_i = \lambda(y_n + h \sum_j a_{ij} k_j)$, $y_{n+1} = y_n + h \underline{b}^T \underline{k}$

Set $\overline{k_i} = k_i / \lambda$. We rewrite it in:

$$\underline{\tilde{k}} = y_n \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + zA \cdot \underline{\tilde{k}}, \quad y_{n+1} = y_n + z\,\underline{b}^T \cdot \underline{\tilde{k}}$$

where we denote $z = h\lambda$.

$\Rightarrow \underline{\tilde{k}}(I_r - zA) = y_n \mathbb{1} \Rightarrow$ Solve $\underline{\tilde{k}} = (I_r - zA)^{-1} y_n \mathbb{1}$.

So $y_{n+1} = (1 + zb^T(I_r - zA)^{-1}\mathbb{1})y_n = g(z)y_n$

$g(z)$ is amplification factor we need.

1) Explicit RK:

$A = \begin{pmatrix} 0 & & 0 \\ & \ddots & \\ * & & 0 \end{pmatrix} \Rightarrow A^r = 0$. (i.e. r-nilpotent).

So $(I_r - zA)^{-1} = I + zA + \cdots + (zA)^{r-1}$

i.e. $g(z)$ will be a polynomial of degree up to r. (So $|g(z)| \gg 1$. if $|z| \gg 1$)

2) Implicit RK:

$g(z)$ has form $g(z) = q(z)^{-1} \cdot p(z)$, st. p, q are both polynomials of degree $\le r$.

Def: One method called L-stable if it's A-stable and $\lim_{|z| \to \infty} |g(z)| = 0$.

Rmk: i) L-stable method can have better damping property (i.e. less oscillation when $|z|$ large)

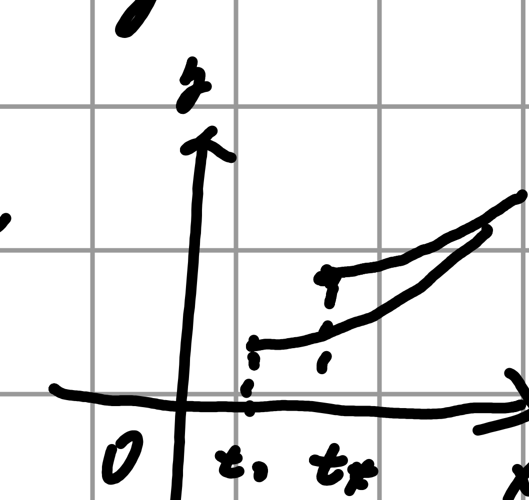
ii) Implicit Euler: $|g(z)| = 1/|1-z|$ is L-stable.

Implicit Trape.: $|g(z)| = \left|\frac{1+\frac{1}{2}z}{1-\frac{1}{2}z}\right|$ is A-stable but not L-stable.

(3) Application of Stab. Analy.:

Def. For IVP. $y' = f(t,y)$, $y(t_0) = y_0$, $t \geq t_0$.

i) Its global sol. is asymptotical stab. if any sol. v of perturbed problem $v'(t) = f(t, v(t))$, $v(t_*) = y(t_*) + w_*$, $t \geq t_*$ for some $t_* \geq t_0$ is also global where perturbation $\|w_*\| \leq \delta$ & $\|(v-y)(t)\| \to 0$ $(t \to \infty)$ holds.



ii) Let this IVP be solved by Lip. conti. one-step method $y_n = y_{n-1} + h_n F(h_n; t_n, y_{n-1}, y_n)$

$\tilde{y}_0 = y_0$. Its sol. (y_n) is called numerically stable if any discrete sol. $(v_n)_{n \geq n^*}$ of $v_n = v_{n-1} + h_n F(h_n, t_n, v_{n-1}, v_n)$

$v_{n^*} = y_{n^*} + w^*$. $n \geq n^*$. for some $t_{n^*} \geq t_0$

where perturbation $\|w^*\| \leq \delta$ satisfies:

$\|v_n - y_n\| \to 0 \ (n \to \infty)$

<u>Hypo</u>: All eigenvalues $\lambda(t)$ of $f_x(t, y(t))$ are assumed to satisfy: $Re(\lambda(t)) \leq 0$.

<u>Prove</u>: i) The method is asymptotically stable

ii) The method with $SR \subset \mathbb{C}$ is numerically stable if the step length h_n is choosen st. $h_n \lambda(t_n) \in SR$. $\forall n \geq 0$.

<u>Rmk</u>: i) $\lambda(t)$ will take role of λ in the model problem. (e.g. for $d=1 \Rightarrow \lambda h \leq \min f_y(t, y(t))$)

ii) It will be wrong if f_x is not diagonalizable. (counterexample exists)

i.e. no complex system of eigenvalues

proof:

i) General problem:

Write $w = v - y$. So $w'(t) = f(t, v(t)) - f(t, y(t))$

$= \int_0^1 \frac{\partial f}{\partial s}(t, y(t) + s w(t))\, ds$

$= \int_0^1 f_x(t, y(t) + s w(t))\, w(t)\, ds$

$\overset{\text{Taylor}}{=} f_x(t, y(t)) w(t) + O(\|w(t)\|^2)$.

1') Simplification: linearization

We get $w'(t) = f_x(t, y(t)) w(t)$.

$w(t_*) = w_*$ for $t \geq t_*$.

2') Simplification: localization.

Freeze $t = t_*$ locally on $y(t)$:

$w'(t) = f_x(t_*, y(t_*)) w(t)$. $w(t_*) = w_*$. $t \geq t_*$.

3') Simplification: diagonalization.

$\exists Q$. s.t. $A = f_x(t, y(t_*))$ have $QAQ^{-1} = D$

$= \operatorname{diag}\{\lambda_i\}$. $\lambda_i \in \mathbb{C}$.

Let $\tilde{w} = Qw$. $\Rightarrow$ $\tilde{w}'(t) = D\tilde{w}(t)$.

Which's reduced to model problem.

$Re\,\lambda_i \leq 0 \Rightarrow \widetilde{w_i}$ decays exponentially.

For Q regular (Q is invertible). We have:

$$\|W(t)\| = \|Q^{-1}\widetilde{W_i}(t)\| \leq C\|\widetilde{W_i}(t)\| \leq Ce^{\lambda(t-t_*)}\|\widetilde{W_*}\|$$

$$\leq Ce^{\lambda t}\|W_*\|. \quad \lambda = \max_i Re(\lambda_i).$$

ii) <u>Discrete problem</u>:

For g is amplification func. of RK method in (2) and recall it has form:

$y_n = g(hA)y_{n-1}$. Set $\widetilde{y_k} = Qy_k$. We have:

$\widetilde{y_n} = Qg(hA)Q^{-1}\widetilde{y_{n-1}} = g(hD)\widetilde{y_{n-1}}$ since g is rational func. (i.e. $p(x)/q(x)$).

<u>Algorithm</u>: To solve $y' = f(t,y)$:

a) Evaluate f_x at $(t_n, y(t_n))$.

b) Compute eigenvalues λ_i.

c) Choose h_n st. $h_n\lambda_i \in SR$. for $Re(\lambda_i) \leq 0$.

(4) <u>Stiff problem</u>:

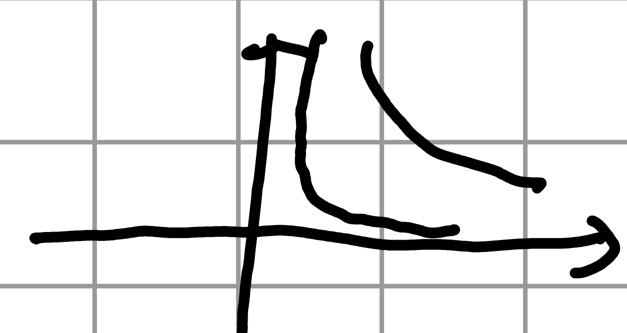
Consider the following IVP:

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} -50 & 49 \\ 49 & -50 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} \quad \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

$\Rightarrow y_1(t) = e^{-t} + e^{-99t}, \quad y_2(t) = e^{-t} - e^{-99t}$

But e^{-99t} decay much faster than e^{-t}.

So the behavior of s.l. will be dominant by e^{-t}.



But for num. stab.: We need to take e^{-99t} into account.

<u>Def</u>: i) Stiff problem is about components with different time scales

<u>Rmk</u>: Alternatively, a IVP is stiff if the step size needed to maintain abs. stab. of expl. Euler method is much smaller than the step size needed for accuracy.

ii) The IVP is stiff along s.l. traject if $\exists$ eigenvalues $\lambda(t)$ of $f_x(t, y(t))$

St. $K(t) = \max_{Re\lambda(t)<0} |Re\lambda(t)| \Big/ \min_{Re\lambda(t)<0} |Re\lambda(t)| \gg 1$

<u>Rmk</u>: i) We only consider $Re(\lambda(t))<0$.

ii) There's no exact math. def.

iii) It's characterized by having components that decay in totally different speed. And we want to solve both components over long time when step length h isn't so small.

iv) Impl. methods are often A-stable so h can be chosen large to be stable. ($\Rightarrow$ no stiff)

<u>Rmk</u>: When dealing stiff problem with some not small transient (initial error e_0) We'd like to use the method that is L-stable (e.g. For trapezoidal: its amplification factor is $1+\frac{h\lambda}{2} \Big/ 1-\frac{h\lambda}{2} \approx 1$ if h is small. So the transient will not

decay fast. While implicit Euler has $\frac{1}{1-h\lambda}$ is small so the transient decay decay fast then it performs better.)

(5) Example of Impl. RK:

① Gauss method:

i) It's based on Gaussian quadrature.

ii) It's A-stable.

iii) It has order 2s for s-stage method.

② Radau method:

i) It's L-stable.

ii) It has order 2s-1 for s-stage method

Rmk: For s=1. ①. ② both implicit Euler.