NIDS Session 12

$\boxed{1}$ For $n = 2$ nodes the stability function of a collocation method is

$$R(z) = \frac{M^{(2)}(1) + z M^{(1)}(1) + z^2 M(1)}{M^{(2)}(0) + z M^{(1)}(0) + z^2 M(0)}$$

where $M(z) = \frac{1}{2}(z - c_1)(z - c_2)$

After simplification, we obtain

$$R(z) = \frac{1 + z\left(1 - \frac{1}{2}(c_1 + c_2)\right) + \frac{z^2}{2}(1 - c_1)(1 - c_2)}{1 - \frac{z}{2}(c_1 + c_2) + \frac{z^2}{2} c_1 c_2} = \frac{P(z)}{Q(z)}$$

For a method to be A-stable

$$\mathbb{C}^- := \{ z : \operatorname{Re}(z) \leq 0 \} \subseteq S := \{ z : |R(z)| \leq 1 \}$$

The method is A-stable if and only if

i) $|R(iy)| \leq 1 \quad \forall y \in \mathbb{R}$

ii) $R(z)$ is analytic in $\mathbb{C}^-$ (i.e. $R(z)$ has no poles in $\mathbb{C}^-$)

What are the poles of $R(z)$ (i.e. the zeros of $Q(z)$)?

$$Q(z) = \frac{1}{2} c_1 c_2 z^2 - \frac{z}{2}(c_1 + c_2) + 1$$

3 cases to consider:

- If $c_1 c_2 = 0$ $\quad \beta_1 = \dfrac{2}{c_1 + c_2} > 0$

- If $c_1 c_2 \neq 0$

$$\Delta = \frac{1}{4}(c_1 + c_2)^2 - 2c_1c_2 = \left(\frac{c_1 - c_2}{2}\right)^2 - c_1c_2$$

$$< \left(\frac{c_1 - c_2}{2}\right)^2$$

* If $\Delta < 0$

$$\beta_{1,2} = \frac{1}{c_1c_2}\left(\frac{c_1 + c_2}{2} \pm i\sqrt{-\Delta}\right)$$

$$\Rightarrow \mathrm{Re}(\beta_{1,2}) = \frac{c_1 + c_2}{2c_1c_2} > 0$$

* If $\Delta \geq 0$

$$\beta_{1,2} = \frac{1}{c_1c_2}\left(\frac{c_1 + c_2}{2} \pm \sqrt{\Delta}\right)$$

Since $\Delta < \left(\dfrac{c_1 + c_2}{2}\right)^2$

$$\frac{1}{c_1c_2}\left(\frac{c_1 + c_2}{2} - \sqrt{\Delta}\right) > 0$$

So again $\mathrm{Re}(\beta_{1,2}) > 0$

Hence, all poles are in $\mathbb{C}^+$

Now consider the first requirement $|R(iy)| \leq 1$

$$\Leftrightarrow \quad |P(iy)|^2 \leq |Q(iy)|^2$$

$$\Leftrightarrow \left(1 - \frac{y^2}{2}(1-c_1)(1-c_2)\right)^2 + \left(y\left(1 - \frac{1}{2}(c_1+c_2)\right)\right)^2$$

$$\leq \left(1 - \frac{y^2}{2} c_1 c_2\right)^2 + \left(\frac{y}{2}(c_1+c_2)\right)^2$$

Eventually, we obtain

$$\underbrace{\left(\frac{y^2}{2}\right)^2}_{\geq 0}\left(1 - (c_1+c_2)\right)\underbrace{\left((1-c_1)(1-c_2) + c_1 c_2\right)}_{\geq 0} \leq 0$$

The requirement is that $\quad 1 - (c_1 + c_2) \leq 0$

$$\Leftrightarrow \quad c_1 + c_2 \geq 1$$

Hence, the method is A-stable if and only if $c_1 + c_2 \geq 1$

$\boxed{2}$ We consider the system

$$\begin{cases} y'(t) = A y(t) \\ y(0) = y^0 \end{cases} \qquad A \in \mathbb{C}^{d \times d}$$

and we assume that $\quad Re(\langle y, Ay \rangle) \leq 0 \quad \forall y \in \mathbb{C}^d$

(A)

In this exercise, I use the common definition

$\langle x, y \rangle = x^* y \qquad x^* = \bar{x}^T$ is the conjugate transpose

We have the norm

$$\|C\| = \sup_{\substack{u, v \in \mathbb{C}^d \\ \|u\| \leq 1 \\ \|v\| \leq 1}} |\langle u, Cv \rangle|$$

i)
$$\frac{d}{dt} \|y(t)\|^2 = \frac{d}{dt} \langle y(t), y(t) \rangle$$
$$= \langle \dot{y}(t), y(t) \rangle + \langle y(t), \dot{y}(t) \rangle$$
$$= \overline{\langle y(t), \dot{y}(t) \rangle} + \langle y(t), \dot{y}(t) \rangle$$
$$= 2 \operatorname{Re}(\langle y(t), A y(t) \rangle)$$
$$\leq 0 \quad \text{(by assumption (9))}$$

Conclusion: $\|y(t)\|^2$ is a decreasing function of time (and so is $\|y(t)\|$)

ii) Assume that A is normal, then there exist a unitary matrix of eigenvectors Q such that $A = Q \Lambda Q^*$ where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_d)$ is the diagonal matrix of eigenvalues

Consider $\langle y, Ay\rangle$ for $y = q_k$ (the k eigenvector)

$$\langle q_k, Aq_k\rangle = \langle q_k, \lambda_k q_k\rangle = \lambda_k \underbrace{\langle q_k, q_k\rangle}_{1}$$

$$\Rightarrow \mathrm{Re}(q_k, Aq_k) = \mathrm{Re}(\lambda_k) \leq 0 \quad \text{By assumption (1)}$$

$$\Rightarrow \lambda_k \in \mathbb{C}^- \quad \forall\, k = 1, \dots, d$$

iii) For a diagonalizable matrix A and a function $f: \mathbb{C} \to \mathbb{C}$ defined on the spectrum of $A = PDP^{-1}$ (meaning $f(\lambda_k)$ is well defined for all k)

Then $f(A)$ is defined as $f(A) = P f(D) P^{-1}$

$$f(D) := \mathrm{diag}(f(\lambda_1), \dots \; f(\lambda_d))$$

For the details, see Higham 2008 " Functions of matrices "

For a unitarily diagonalizable matrix A, the expression becomes $f(A) = P f(D) P^*$, because $P^{-1} = P^*$

Now, coming back to the exercise with $A = QDQ^*$, we obtain

$$\begin{aligned}
\| R(A) \| &= \| Q R(D) Q^* \| \\
&= \| R(D) \| \\
&= \max_{\lambda \in \Lambda(A)} | R(\lambda) |
\end{aligned}$$

$\Lambda(A) = \{\lambda_1, \dots \lambda_d\}$ the spectrum of A

$$\leq \sup_{z \in \mathbb{C}^-} |R(z)|$$

(v) Now consider a general matrix A

Denote $H := \frac{1}{2}(A^* + A)$ the Hermitian part of A
$(H^* = H)$

and $S := \frac{1}{2}(A - A^*)$ the skew Hermitian part
$(S^* = -S)$

Note $A = H + S$

Define $\mathcal{A} : \mathbb{C} \to \mathbb{C}^{d \times d}$ as

$$\mathcal{A}(w) = wH + S$$

Define $\varphi : \mathbb{C} \to \mathbb{C}$ as

$\varphi(w) = \langle u, R(\mathcal{A}(w))v \rangle$ for some fixed vectors $u, v \in \mathbb{C}^d$

$$\langle y, \mathcal{A}(w)y \rangle = w\langle y, Hy \rangle + \langle y, Sy \rangle$$
$$= w \underbrace{y^* H y}_{\in \mathbb{R}} + \underbrace{y^* S y}_{\in i\mathbb{R}}$$

Reminder: If H is Hermitian, for any $x \in \mathbb{C}^d$

$$x^* H x - \overline{x^* H x} = x^* H x - (x^* H x)^*$$
$$= x^* (H - H^*) x = 0$$

Yet $x^* H x - \overline{x^* H x} = 2i\,\mathrm{Im}(x^* H x) = 0$

$\Rightarrow x^* H x \in \mathbb{R}$

Similarly for a skew hermitian matrix S

$x^* S x + \overline{x^* S x} = 2\,\mathrm{Re}(x^* S x)$

$= x^*(S + S^*) x$

$= 0$

$\Rightarrow x^* S x \in i\mathbb{R}$ (purely imaginary number)

$\mathrm{Re}(\langle y, A(\omega) y\rangle) = \mathrm{Re}(\omega)\langle y, H y\rangle \quad (2)$

Recall that by assumption (1) $\mathrm{Re}(\langle y, A y\rangle) \leq 0$

and $\langle y, A y\rangle = \underbrace{\langle y, H y\rangle}_{\in \mathbb{R}} + \underbrace{\langle y, S y\rangle}_{\in i\mathbb{R}}$

$\Rightarrow \mathrm{Re}(\langle y, A y\rangle) = \langle y, H y\rangle \leq 0$

Conclusion: if $\mathrm{Re}(\omega) \geq 0$, then $\mathrm{Re}(\langle y, A(\omega) y\rangle) \leq 0$

(from (2))

v) General result: For a matrix $C \in \mathbb{C}^{d \times d}$

define $\mathcal{F}(C) = \left\{ \dfrac{\langle y, Cy \rangle}{\|y\|_2^2} : y \in \mathbb{C}^d, y \neq 0 \right\}$

called the field of values

Lemma: $\Lambda(C) \subseteq \mathcal{F}(C)$

Proof: Consider an eigenvalue / eigenvector pair of C

(λ, v) $\qquad Cv = \lambda v$

$$\Rightarrow \quad \langle v, Cv \rangle = \lambda \langle v, v \rangle$$

$$\Rightarrow \quad \lambda = \frac{\langle v, Cv \rangle}{\|v\|_2^2} \in \mathcal{F}(C)$$

Applying this lemma to $\mathcal{A}(\omega)$,

$\mathcal{F}(\mathcal{A}(\omega)) \subseteq \mathbb{C}^-$ (from question iv))

Consequently $\Lambda(\mathcal{A}(\omega)) \subseteq \mathbb{C}^-$

vi) Simply notice that $\mathcal{A}(ix)$ is skew Hermitian

Indeed $\mathcal{A}(ix) = ixH + S$

$$\Rightarrow \quad \mathcal{A}(ix)^* = -(ixH + S) = -\mathcal{A}(ix)$$

Since $\mathcal{A}(ix)$ is skew Hermitian, it is in particular normal

vii) For a general matrix A, we can define the matrix function $f(A)$ as

$f(A) = Z f(J) Z^{-1}$, where $A = Z J Z^{-1}$ is the Jordan canonical form of A

$f(J)$ is defined via the derivatives of f

for the details, see Higham 2008

Note $A(1) = A$

$$\|R(A)\| = \sup_{\substack{u,v \in \mathbb{C}^d \\ \|u\| \le 1 \\ \|v\| \le 1}} |\langle u, R(A(1)) v \rangle|$$

$$= \sup_{u,v} | \varphi(1) |$$

$$\le \sup_{x \in \mathbb{R}} \sup_{u,v} | \varphi(ix) |$$

$$= \sup_{x \in \mathbb{R}} \| R(A(ix)) \|$$

$$= \sup_{x \in \mathbb{R}} \max_{\lambda \in \Lambda(A(ix))} | R(\lambda) |$$

$$\le \sup_{Re(B) = 0} | R(B) |$$

$$\le \sup_{Re(B) \le 0} | R(B) |$$

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

$\Rightarrow \quad \| y_{n+1} \| \leq \| R(hA) \| \, \| y_n \|$

$$\leq \sup_{Re(z) \leq 0} | R(z) | \, \| y_n \|$$

Since by assumption, the numerical method is A stable

$$\sup_{Re(z) \leq 0} | R(z) | \leq 1$$

Therefore $\| y_{n+1} \| \leq \| y_n \|$ as required

$\boxed{3}$ The stability function of a collocation method is

$$R(z) = \frac{T^{(s)}(1) + z\, T^{(s-1)}(1) + \dots + z^s\, T(1)}{T^{(s)}(0) + z\, T^{(s-1)}(0) + \dots + z^s\, T(0)} = \frac{P(z)}{Q(z)}$$

where

$$T(t) = \frac{1}{s!} \prod_{i=1}^{s} (t - c_i)$$

. consider the Gauss method, all collocation points c_i are internal to $[0,1]$ / $c_i \in (0,1) \quad \forall i = 1 \dots s$

$\Rightarrow \quad T(1) \neq 0$ and $T(0) \neq 0$

$\Rightarrow \quad P(z) \in \mathbb{P}^s \quad Q(z) \in \mathbb{P}^s$

Since the given method has order $2s$, then

$R = R_{s,s}$ by uniqueness of Padé approximations

- consider now the Radau method

$c_1 > 0$ but $c_s = 1$

$\Rightarrow \quad T(1) = 0$ but $T(0) \neq 0$

$\Rightarrow \quad P(z) \in \mathbb{P}^{s-1}, \quad Q(z) \in \mathbb{P}^{s}$

Since Radau has order $2s-1$

$R = R_{s-1,s}$

- Finally for the Lobatto method:

$c_1 = 0$ and $c_s = 1$

$\Rightarrow \quad T(1) = T(0) = 0$

$\Rightarrow \quad P(z) \in \mathbb{P}^{s-1} \quad Q(z) \in \mathbb{P}^{s-1}$

Since the Lobatto $2s-2$

$R = R_{s-1,s-1}$

$\boxed{4}$ We consider a numerical method of order p

i) $y(t_0+h) - \Phi_h(y_0) = C(y_0)\, h^{p+1} + O(h^{p+2})$

We replace h by $2h$

$$y(t_0 + 2h) - \underbrace{\Phi_{2h}(y_0)}_{\omega} = 2^{p+1} C(y_0) h^{p+1} + O(h^{p+2})$$

ii) Denote $y_2 = (\Phi_h \circ \Phi_h)(y_0) = \Phi_h(y_1)$

$$\varphi_h(y_1) - \Phi_h(y_1) = C(y_1) h^{p+1} + O(h^{p+2})$$

Informal : $C(y_1) = C(y_0 + O(h))$

$$= C(y_0) + O(h)$$

Thus $\varphi_h(y_1) - \Phi_h(y_1) = C(y_0) h^{p+1} + O(h^{p+2})$

iii) $\varphi_{2h}(y_0) - \varphi_h(y_1)$

$$= \varphi_h(y(t_1)) - \varphi_h(y_1)$$

$$= \varphi_h(y_1) + \varphi_h'(y_1)(y(t_1) - y_1) + \dots - \varphi_h(y_1) \quad (A)$$

Note that

$$\varphi_h(y_0) = y(t_0 + h) = y_0 + h y'(t_0) + O(h^2)$$

$$\Rightarrow \varphi_h'(y_0) = I + O(h)$$

Thus (1) becomes

$$y(t_1) - y_1 + O(h^{p+2})$$

$$= C(y_0) h^{p+1} + O(h^{p+2})$$

iv) from ii) and iii)

$$\phi_h(y_1) - y_2 = \varphi_{2h}(y_0) - C(y_0)h^{p+1} + O(h^{p+2}) - y_2$$

$$= C(y_0)h^{p+1} + O(h^{p+2})$$

$$\Rightarrow \varphi_{2h}(y_0) - y_2 = 2C(y_0)h^{p+1} + O(h^{p+2})$$

v) $$y(t_0 + 2h) = \omega + 2^{p+1} C(y_0) h^{p+1} + O(h^{p+2})$$

$$= y_2 + 2C(y_0)h^{p+1} + O(h^{p+2})$$

$$\Rightarrow 2C(y_0)h^{p+1} = \frac{y_2 - \omega}{2^p - 1} + O(h^{p+2})$$

$$y(t_0 + 2h) = y_2 + \frac{y_2 - \omega}{2^p - 1} + O(h^{p+2})$$

$$= \underbrace{\frac{2^p y_2 - \omega}{2^p - 1}}_{= \beta_2} + O(h^{p+2})$$