Numerical Solutions for DEs HW3

YANG, Ze (5131209043)

April 1, 2017

A Note to TA:
Hi, this is the senior student from Antai College who did not register for this course. I would like to do all the assignments for practice, but feel free to just skip my homework if you don't have time. Thank you again for allowing me to access the assignments and other class material! :)
- Ze

Problem 1. (Iserles 4.4) Determine all values of θ such that the theta method (1.13) is absolutely stable.

Proof. The theta method reads

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h[\theta \boldsymbol{f}(t_n, \boldsymbol{y}_n) + (1-\theta)\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1})], \qquad n = 0, 1, \ldots$$

To find the a-stability region ($\mathcal{D}$) we apply the method to scalar ode,

$$
\begin{aligned}
y_{n+1} &= y_n + h[\theta \lambda y_n + (1-\theta)\lambda y_{n+1}] \\
y_{n+1} &= \frac{1 + h\lambda\theta}{1 - h\lambda(1-\theta)} y_n = \frac{1 + z\theta}{1 - z(1-\theta)} y_n
\end{aligned}
\tag{1}
$$

Therefore, we want to find θ, such that $\mathbb{C}^- = \{z \in \mathbb{C} : \text{Re}(z) < 0\} \subseteq \{z \in \mathbb{C} : \left|\frac{1+z\theta}{1-z(1-\theta)}\right| < 1\} = \mathcal{D}$ (By the definition of a-stability). I.e. we want $|1 + z\theta| < |1 - z(1-\theta)|$ $(*)$ holds for all z that $\text{Re}(z) < 0$.

Write $z = x + yi$:

· $\theta = 1/2$, the method is a-stable, since it's the trapezoid method.
· $\frac{1}{2} < \theta \le 1$, we write $\theta = \frac{1}{2} + \delta$. Let $z = -1/\delta + 0i$. Then $LHS = 1/2\delta$, $RHS = |1 + 1/\delta(-1/2 - \delta)| = 1/2\delta$. Then the equation $(*)$ does not hold, so in this case the method is not a-stable.
· $0 \le \theta < \frac{1}{2}$, we want to show

$$|1 + \theta x + \theta yi| < |1 - (1-\theta)x - (1-\theta)yi|$$

for all $(x, y) \in \mathbb{C}^-$. Denote the complex number inside absolute value as r (right hand side) and l (left hand side). $\text{Im}(l) = \theta y$, $\text{Im}(l) = (1-\theta)y$, since $0 \le \theta < \frac{1}{2} \Rightarrow 1 - \theta > \theta > 0$. So $\text{Im}(l) > \text{Im}(r) > 0$, Hence $|\text{Im}(l)| > |\text{Im}(r)|$.
For the real parts, $\text{Re}(l) = 1 + x\theta$, $\text{Re}(r) = 1 - (1-\theta)x$.

1. $1 + x\theta \le 0$: since we must have $x < 0$. $(2\theta - 1) < 0 \Rightarrow (2\theta - 1)x > 0$, $2 + (2\theta - 1)x > 0 \Rightarrow 1 - (-\theta + 1)x > -1 - \theta x \ge 0$, i.e. $\text{Re}(r) > -\text{Re}(l) \ge 0$.
2. $1 + x\theta > 0$: $\theta x < -(1-\theta)x \Rightarrow 0 < 1 + \theta x < 1 - (1-\theta)x \Rightarrow \text{Re}(r) > \text{Re}(l) \ge 0$.

In either case we have $|\text{Re}(l)| > |\text{Re}(r)|$. So finally, we have $|r| < |l|$ for all $z \in \mathbb{C}^-$. And we conclude that for $\theta \in [0, 1/2]$, the method is a-stable.

$\square$

Problem 2. (Iserles 4.9) The two-step method

$$\boldsymbol{y}_{n+2} - \boldsymbol{y}_n = 2h\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}), \quad n = 0, 1, \ldots$$

is called the explicit midpoint rule.

    a. Denote by $w_1(z)$ and $w_2(z)$ the zeroes of the underlying function $\eta(z, \cdot)$, show that $w_1(z)w_2(z) \equiv -1$ for all $z \in \mathbb{C}$.

    b. Show that $\mathcal{D} = \emptyset$.

    c. We say that $\tilde{\mathcal{D}}$ is a *weak linear stability domain* of a numerical method if, when applied to the scalar linear equation, it produces a uniformly bounded solution sequence. (It is easy to see that $\tilde{\mathcal{D}} = \mathrm{cl}\mathcal{D}$ for most methods of interest.) Determine explicitly $\tilde{\mathcal{D}}$ for the method.

Proof. (a.) We have $\rho(w) = w^2 - 1$, $\sigma(w) = 2w$. So let $z = \lambda h$, $\eta(w, z) = \rho(w) - z\sigma(w) = w^2 - 2zw - 1$. Suppose $w_1(z)$ and $w_2(z)$ solves $\eta(\cdot, z) = 0$, then by Vieta's formula, we have $w_1(z)w_2(z) = \frac{-1}{1} = -1$. $\square$

(b.) We have $w_1(z) = \frac{2z + \sqrt{4z^2 + 4}}{2} = z + \sqrt{z^2 + 1}$. $w_2(z) = z - \sqrt{z^2 + 1}$. So $\mathcal{D} = \{z \in \mathbb{C} : |w_1(z)| < 1, |w_2(z)| < 1\}$. But

$$|w_1(z)| \cdot |w_2(z)| = |w_1(z)w_2(z)| = |-1| = 1$$

implies that $|w_1(z)| = 1/|w_2(z)|$. Hence either (1) $|w_1| = |w_2| = 1$, or (2) one of the two roots has modulus strictly smaller than 1, another one strictly greater than 1. And for all $z \in \mathbb{C}$, one of the two cases must be true. $\Rightarrow \forall z \in \mathbb{C}, z \notin \mathcal{D}$. $\mathcal{D} = \emptyset$.

(c) Apply to scalar linear equation $y' = \lambda y$,

$$y_{n+2} - y_n = 2h\lambda y_{n+1}$$

$\square$

Problem 3. (Iserles 5.8) Show that the *Hénon-Heiles system* (using the definition from wikipedia)

$$
\begin{aligned}
q_1' &= p_1 \\
p_1' &= -q_1 - 2\lambda q_1 q_2 \\
q_2' &= p_2 \\
p_2' &= -q_2 - \lambda(q_1^2 - q_2^2)
\end{aligned}
\tag{2}
$$

is Hamiltonian and identify explicitly the hamiltonian energy.

Proof. Let $\boldsymbol{q} = (q_1, q_2)$, $\boldsymbol{p} = (p_1, p_2)$. Assume the system is Hamiltonian, then

$$
\begin{aligned}
\nabla_{\boldsymbol{p}} H &= \dot{\boldsymbol{q}} = \begin{pmatrix} p_1 & p_2 \end{pmatrix} \\
\nabla_{\boldsymbol{q}} H &= -\dot{\boldsymbol{p}} = \begin{pmatrix} q_1 + 2\lambda q_1 q_2 & q_2 + \lambda(q_1^2 - q_2^2) \end{pmatrix}
\end{aligned}
\tag{3}
$$

Hence

$$H(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2}(p_1^2 + p_2^2) + f(\boldsymbol{q})$$

$$H(\boldsymbol{p}, \boldsymbol{q}) = \int (q_1 + 2\lambda q_1 q_2) dq_1 + g(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2}q_1^2 + \lambda q_2 q_1^2 + g(\boldsymbol{p}, \boldsymbol{q})$$

$$H(\boldsymbol{p}, \boldsymbol{q}) = \int [q_2 + \lambda(q_1^2 - q_2^2)] dq_2 + h(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2}q_2^2 + \lambda q_1^2 q_2 - \frac{1}{3}\lambda q_2^3 + h(\boldsymbol{p}, \boldsymbol{q})$$

Combine the first integrals above we have:

$$H(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2}(q_1^2 + q_2^2) + \lambda\left(q_1^2 q_2 - \frac{1}{3}q_2^3\right)$$

Check the partial derivatives of H, we find that it gives the system indeed. Hence the system is Hamiltonian, and the hamiltonian energy is given by $H(\boldsymbol{p}, \boldsymbol{q})$.

$\square$

Problem 4. The symplectic Euler method for the Hamiltonian system reads

$$\begin{cases} \boldsymbol{p}_{n+1} = \boldsymbol{p}_n - h\frac{\partial H(\boldsymbol{p}_{n+1}, \boldsymbol{q}_n)}{\partial \boldsymbol{q}} \\ \boldsymbol{q}_{n+1} = \boldsymbol{q}_n + h\frac{\partial H(\boldsymbol{p}_{n+1}, \boldsymbol{q}_n)}{\partial \boldsymbol{p}} \end{cases} \tag{4}$$

a. Show that this is a first order method.

b. Prove from basic principles that, as implied by its name, the method is indeed symplectic. (Hint: let $G = \nabla^2 H$, and write $G = [G_{11}, G_{12}; G_{21}, G_{22}]$).

c. Assuming that the Hamiltonian is separable, $H(\boldsymbol{p}, \boldsymbol{q}) = T(\boldsymbol{p}) + V(\boldsymbol{q})$, where T and V correspond to kinetic and potential energy respectively. Show that the method can be implemented explicitly.

d. Use symplectic Euler and explicit Euler to solve the problem of non-linear pendulum. The Hamiltonian is $H(p, q) = \frac{1}{2}p^2 - \cos(q)$, and the Hamiltonian equations are

$$\dot{p} = -\sin q, \quad \dot{q} = p$$

with initial condition $p(0) = 0$, $q(0) = 1$. Plot the error of the numerical methods in the Hamiltonian H.

Proof. (a.) We examine the truncation error $\boldsymbol{T}_n^{[1,2]}/h$ for $\boldsymbol{q}$ and $\boldsymbol{p}$ respectively. Denote $\boldsymbol{f} = \dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{q}}$, $\boldsymbol{g} = \dot{\boldsymbol{q}} = \frac{\partial H}{\partial \boldsymbol{p}}$.

$$\begin{aligned} \boldsymbol{T}_n^{[1]}/h &= \frac{\boldsymbol{p}(t_{n+1}) - \boldsymbol{p}(t_n)}{h} - \boldsymbol{f}(\boldsymbol{p}(t_{n+1}), \boldsymbol{q}(t_n)) \\ &= \frac{1}{h}[\dot{\boldsymbol{p}}(t_n)h + \frac{1}{2}\ddot{\boldsymbol{p}}(t_n)h^2 + O(h^3)] - (\boldsymbol{f}(t_n) + \boldsymbol{f}'(t_n)h + O(h^2)) \\ &= -\frac{1}{2}\boldsymbol{f}'(t_n)h + O(h^2) \end{aligned} \tag{5}$$

The calculation is same for $\boldsymbol{T}_n^{[2]}/h$, we obtain $\boldsymbol{T}_n^{[2]}/h = -\frac{1}{2}\boldsymbol{g}'(t_n)h + O(h^2)$. Hence we conclude that the method is of order 1. $\square$

(b.) Let

$$\boldsymbol{G} = \nabla^2 H = \begin{pmatrix} \frac{\partial^2 H}{\partial \boldsymbol{p}^2} & \frac{\partial^2 H}{\partial \boldsymbol{p} \partial \boldsymbol{q}} \\ \frac{\partial^2 H}{\partial \boldsymbol{q} \partial \boldsymbol{p}} & \frac{\partial^2 H}{\partial \boldsymbol{q}^2} \end{pmatrix} =: \begin{pmatrix} \boldsymbol{G}_{pp} & \boldsymbol{G}_{qp}^\top \\ \boldsymbol{G}_{qp} & \boldsymbol{G}_{qq} \end{pmatrix}$$

Then calculate the partial derivatives from the method's formula:

$$\begin{cases} \frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{p}_n} &= \boldsymbol{I} - h\boldsymbol{G}_{qp}\frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{p}_n} \\ \frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{q}_n} &= -h\boldsymbol{G}_{qq} - h\boldsymbol{G}_{qp}\frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{q}_n} \\ \frac{\partial \boldsymbol{q}_{n+1}}{\partial \boldsymbol{p}_n} &= h\boldsymbol{G}_{pp}\frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{p}_n} \\ \frac{\partial \boldsymbol{q}_{n+1}}{\partial \boldsymbol{q}_n} &= \boldsymbol{I} + h\boldsymbol{G}_{pq} + h\boldsymbol{G}_{pp}\frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{q}_n} \end{cases} \tag{6}$$

I.e.

$$\begin{cases} (\boldsymbol{I} + h\boldsymbol{G}_{qp})\frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{p}_n} = \boldsymbol{I} \\ (\boldsymbol{I} + h\boldsymbol{G}_{qp})\frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{q}_n} = -h\boldsymbol{G}_{qq} \\ -h\boldsymbol{G}_{pp}\frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{p}_n} + \frac{\partial \boldsymbol{q}_{n+1}}{\partial \boldsymbol{p}_n} = \boldsymbol{O} \\ -h\boldsymbol{G}_{pp}\frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{q}_n} + \frac{\partial \boldsymbol{q}_{n+1}}{\partial \boldsymbol{q}_n} = \boldsymbol{I} + h\boldsymbol{G}_{pq} \end{cases} \tag{7}$$

rewrite in matrix form:

$$\begin{pmatrix} \boldsymbol{I} + h\boldsymbol{G}_{qp} & \boldsymbol{O} \\ -h\boldsymbol{G}_{pp} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{p}_n} & \frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{q}_n} \\ \frac{\partial \boldsymbol{q}_{n+1}}{\partial \boldsymbol{p}_n} & \frac{\partial \boldsymbol{q}_{n+1}}{\partial \boldsymbol{q}_n} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I} & -h\boldsymbol{G}_{qq} \\ \boldsymbol{O} & \boldsymbol{I} + h\boldsymbol{G}_{pq} \end{pmatrix} \tag{8}$$

And solve this linear system:

$$\boldsymbol{\Phi} = \begin{pmatrix} \frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{p}_n} & \frac{\partial \boldsymbol{p}_{n+1}}{\partial \boldsymbol{q}_n} \\ \frac{\partial \boldsymbol{q}_{n+1}}{\partial \boldsymbol{p}_n} & \frac{\partial \boldsymbol{q}_{n+1}}{\partial \boldsymbol{q}_n} \end{pmatrix} = \begin{pmatrix} (\boldsymbol{I} + h\boldsymbol{G}_{qp})^{-1} & -h\boldsymbol{G}_{qq}(\boldsymbol{I} + h\boldsymbol{G}_{qp})^{-1} \\ h\boldsymbol{G}_{pp}(\boldsymbol{I} + h\boldsymbol{G}_{qp})^{-1} & \boldsymbol{I} + h\boldsymbol{G}_{pq} - h^2\boldsymbol{G}_{pp}\boldsymbol{G}_{qq}(\boldsymbol{I} + h\boldsymbol{G}_{qp})^{-1} \end{pmatrix} \tag{9}$$

So:

$$\mathbf{\Phi}^\top \mathbf{J} = \begin{pmatrix} (\mathbf{I} + h\mathbf{G}_{pq})^{-1} & h\mathbf{G}_{pp}(\mathbf{I} + h\mathbf{G}_{pq})^{-1} \\ -h\mathbf{G}_{qq}(\mathbf{I} + h\mathbf{G}_{pq})^{-1} & \mathbf{I} + h\mathbf{G}_{qp} - h^2\mathbf{G}_{pp}\mathbf{G}_{qq}(\mathbf{I} + h\mathbf{G}_{pq})^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{O} & \mathbf{I} \\ -\mathbf{I} & \mathbf{O} \end{pmatrix}$$
$$= \begin{pmatrix} -h\mathbf{G}_{pp}(\mathbf{I} + h\mathbf{G}_{pq})^{-1} & (\mathbf{I} + h\mathbf{G}_{pq})^{-1} \\ -\mathbf{I} - h\mathbf{G}_{qp} + h^2\mathbf{G}_{pp}\mathbf{G}_{qq}(\mathbf{I} + h\mathbf{G}_{pq})^{-1} & -h\mathbf{G}_{qq}(\mathbf{I} + h\mathbf{G}_{pq})^{-1} \end{pmatrix}$$
(10)

And then it is clear that $(\mathbf{\Phi}^\top \mathbf{J})\mathbf{\Phi} = \mathbf{J}$. So the method is indeed symplectic. $\square$

(c) If the method is separable, $H(\boldsymbol{p}, \boldsymbol{q}) = T(\boldsymbol{p}) + V(\boldsymbol{q})$, then it becomes:

$$\begin{cases} \boldsymbol{p}_{n+1} = \boldsymbol{p}_n - h\frac{\partial V(\boldsymbol{q}_n)}{\partial \boldsymbol{q}} \\ \boldsymbol{q}_{n+1} = \boldsymbol{q}_n + h\frac{\partial T(\boldsymbol{p}_{n+1})}{\partial \boldsymbol{p}} \end{cases}$$
(11)

is an explicit method. Denote $\nabla_{\boldsymbol{q}} V = \boldsymbol{f}$ and $\nabla_{\boldsymbol{p}} T = \boldsymbol{g}$, and suppose we know them explicitly. The method is implemented as follows:

Given $\boldsymbol{q}_0, \boldsymbol{p}_0$.
For $k = 1$ to n:

$$\boldsymbol{p}_k \leftarrow \boldsymbol{p}_{k-1} - h\boldsymbol{f}(\boldsymbol{q}_k)$$
$$\boldsymbol{q}_k \leftarrow \boldsymbol{q}_{k-1} + h\boldsymbol{g}(\boldsymbol{p}_k) = \boldsymbol{q}_{k-1} + h\boldsymbol{g}(\boldsymbol{p}_{k-1} - h\boldsymbol{f}(\boldsymbol{q}_k))$$

which is explicit, and resembles the Runge-Kutta method. $\square$

(d) $\square$

Problem 5. Implement implicit Euler, implicit midpoint method and trapezoid method, compare their rates of convergence and execution time.

Solution. See the attached code.