MADS-CLAUSEN-INSTITUTE, SDU
Christian Wolff

22/02/2023
contact: cwo@mci.sdu.dk

Computational Physics: Solutions to problem Set 3

$\boxed{1}$ Analysis of time integrators – Kepler problem

See solution programs

$\boxed{2}$ A simple matrix exponential integrator

a) In the problem, it is given that the matrix M can be represented in diagonalized form:

$$M = R^{-1}\Lambda R. \tag{1}$$

Assuming that for some n we can write $M^n = R^{-1}\Lambda^n R$ (which is definitely true for $n = 1$), we find:

$$\begin{aligned} M^{n+1} =& M^n M = \underbrace{R^{-1}\Lambda^n R}_{\text{assumption}} \cdot \underbrace{R^{-1}\Lambda R}_{\text{representation of } M} . && (2)\\ =& R^{-1}\Lambda^n \underbrace{R^{-1}R}_{=\mathbb{I}} \Lambda R && (3)\\ =& R^{-1}\Lambda^n \Lambda R = R^{-1}\Lambda^{n+1} R. && (4) \end{aligned}$$

So, by induction, every power of M can be written in this form.

We apply this to the Taylor expansion:

$$\exp(M) = \mathbb{I} + \sum_{n=1}^{\infty} \frac{M^n}{n!} = R^{-1}R + \sum_{n=1}^{\infty} \frac{R^{-1}\Lambda^n R}{n!} = R^{-1}\left[\mathbb{I} + \sum_{n=1}^{\infty} \frac{\Lambda^n}{n!}\right] R.$$

The matrices R^{-1} and R are already in the way that we want, so let's focus on the square bracket.

The matrix Λ is diagonal, so a power of Λ is a diagonal matrix with the powers of the diagonal elements:

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{pmatrix} \qquad \Rightarrow \qquad \Lambda^n = \begin{pmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N^n \end{pmatrix} \tag{5}$$

Applying this to the square brackets in the Taylor series, we find:

$$\mathbb{I} + \sum_{n=1}^{\infty} \frac{\Lambda^n}{n!} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} + \sum_{n=1}^{\infty} \begin{pmatrix} \lambda_1^n/n! & 0 & \cdots & 0 \\ 0 & \lambda_2^n/n! & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N^n/n! \end{pmatrix}. \tag{6}$$

Each diagonal entry of this sum is of the form

$$1 + \sum_{n=1}^{\infty} \frac{\lambda_i^n}{n!} = \exp(\lambda_i). \tag{7}$$

With this, we have shown the problem.

b) The first proof is simple:

$$\exp(M) = R^{-1} \exp(\Lambda) R = R^{-1} \exp(\Lambda/2) \exp(\Lambda/2) R \tag{8}$$
$$= R^{-1} \exp(\Lambda/2) R R^{-1} \exp(\Lambda/2) R = \exp(M/2) \exp(M/2) \tag{9}$$
$$= [\exp(M/2)]^2. \tag{10}$$

By reapeating this, we find:

$$\exp(M) = [\exp(M/2/2)]^{2\times 2}. \tag{11}$$
$$= [\exp(M/2/2/2)]^{2\times 2\times 2}. \tag{12}$$
$$\cdots \tag{13}$$
$$= [\exp(M/2^n)]^{2^n}. \tag{14}$$

The aim is now to estimate the error introduced by truncating a Taylor series. The full exponential is

$$\exp(M) = \mathbb{I} + \sum_{n=1}^{\infty} \frac{M^n}{n!}, \tag{15}$$

we truncate after N terms:

$$T_N(M) = \mathbb{I} + \sum_{n=1}^{N} \frac{M^n}{n!}, \tag{16}$$

the residue is therefore

$$R = \exp(M) - T_N(M) = \sum_{n=N+1}^{\infty} \frac{M^n}{n!} = M^{N+1} \Big[\sum_{n=0}^{\infty} \frac{M^n}{(n+N+1)!} \Big], \tag{17}$$

We can give an upper bound for the residual norm (i.e. the absolute error of the Taylor approximation) using the triangle inequality:

$$||R|| = ||M^{N+1}|| \cdot || \sum_{n=0}^{\infty} \frac{M^n}{(n+N+1)!} || \leq ||M^{N+1}|| \cdot \sum_{n=0}^{\infty} \frac{||M||^n}{(n+N+1)!} \leq ||M^{N+1}|| \cdot \Big[1 + \sum_{n=1}^{\infty} ||M||^n \Big]. \tag{18}$$

We now use the infinte series

$$\sum_{n=1} x^n = \frac{1}{1-x} \qquad \text{for} \qquad |x| < 1. \tag{19}$$

Therefore for $||M|| < 1$, we can estimate:

$$||R|| \leq ||M^{N+1}|| \frac{2 - ||M||}{1 - ||M||} = ||M||^N \underbrace{\frac{(2 - ||M||)||M||}{1 - ||M||}}_{=F}. \tag{20}$$

If we can show that $F < 1$, then we have a robust bound for the error. So we write (using $x = ||M||$ to save space):

$$F = \frac{(2-x)x}{1-x} < 1 \qquad \Leftrightarrow \qquad 2x - x^2 < 1 - x \qquad \Leftrightarrow \qquad x^2 - 3x + 1 > 0. \tag{21}$$

With zeros $x_{1/2} = \frac{1}{2}(3 \pm \sqrt{5})$, this inequality is true if $x > \frac{1}{2}(3 + \sqrt{5})$ or if $x < \frac{1}{2}(3 - \sqrt{5})$.

Putting everything together, we can guarantee for the error

$$||R|| < ||M||^N, \tag{22}$$

if the matrix M is small enough, specifically if $||M|| < \frac{1}{2}(3 - \sqrt{5}) \approx 0.38$. This is not the tightest possible error bound, but it will do for our purposes. Note that in practice $M = \tau A$ with some time interval τ, so $||M||$ can always be reduced by reducing τ.

We now come to the question how Equation (14) is useful for time-integration. In other words: is it benefitial to divide the matrix M by two, compute the Taylor expansion and then square the result or is it better to compute the truncated Taylor expansion directly with M?

For this, we must compare the two errors $||R_1||$ and $||R_2||$ of the residuals

$$R_1 = \exp(M) - T_N(M) \tag{23}$$
$$R_2 = \exp(M) - [T_N(M/2)]^2. \tag{24}$$

We can write:

$$T_N(M) = \exp(M) - R_1 \tag{25}$$
$$T_N(M/2) = \exp(M/2) - R_3, \tag{26}$$

where we know from the scaling law $||R_3|| \leq 2^{-4}||R_1|| = ||R_1||/16$. We now insert this in the expression for R_2:

$$R_2 = \exp(M) - [\exp(M/2) - R_3]^2 = \exp(M/2)R_3 + R_3 \exp(M/2) - R_3^2. \tag{27}$$

If $||\exp(M)|| < 1$, i.e. for oscillatory or decaying problems, we can therefore estimate:

$$||R_2|| \leq ||\exp(M/2)R_3|| + ||R_3 \exp(M/2)|| + ||R_3||^2 \leq 2||R_3|| + ||R_3||^2 \leq ||R_1||/8 + ||R_1||^2/256. \tag{28}$$

So using Equation (14) will reduce the error roughly by a factor of 8.

c) see example program.

d) see example program.

e) If we truncate the Taylor expansion after the first order instead of after th fourth order, we obtain Euler's method.

f) The stability contour is given by the stability function $R(z)$ that performs one time step:

$$\tilde{y}(t_{n+1}) = R(z)\tilde{y}(t_n). \tag{29}$$

In our case, that is exactly the approximated time-evolution operator. It is:

$$R(z) = 1 + z\left(1 + \frac{z}{2}\left(1 + \frac{z}{3}\left(1 + \frac{z}{4}\right)\right)\right). \tag{30}$$

The area of stability is the area of the complex plane, where $|R(z)| < 1$. One easy way to find this is to plot the function $R(z)$ as demonstrated in the program *stability_contour.m*. The stability contour looks just like that of a standard 4th order Runge-Kutta.