DELFT UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING, MATHEMATICS AND COMPUTER SCIENCE

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS
(CTB2400)
Tuesday July 12 2022, 13:30-16:30

1. (a) The test equation is given by

$$y' = \lambda y.$$

Application of the method to the test equation gives

$$w_{n+1} = w_n + \frac{1}{2}\lambda\Delta t w_n + \frac{1}{2}\lambda\Delta t w_{n+1}.$$

This is equivalent to

$$\left(1 - \frac{1}{2}\lambda\Delta t\right) w_{n+1} = \left(1 + \frac{1}{2}\lambda\Delta t\right) w_n.$$

Hence the amplification factor is given by

$$Q(\lambda\Delta t) = \frac{1 + \frac{1}{2}\lambda\Delta t}{1 - \frac{1}{2}\lambda\Delta t}.$$

(b) The local truncation error for the test equation is given as

$$\tau_{n+1}(\Delta t) = \frac{e^{\lambda\Delta t} - Q(\lambda\Delta t)}{\Delta t} y_n. \tag{1}$$

A Taylor expansion of $e^{\lambda\Delta t}$ around $\lambda\Delta t = 0$ yields

$$e^{\lambda\Delta t} = 1 + \lambda\Delta t + \frac{1}{2}(\lambda\Delta t)^2 + \frac{1}{6}(\lambda\Delta t)^3 + \mathcal{O}(\Delta t^4). \tag{2}$$

A Taylor expansion of $Q(\lambda\Delta t)$ around $\frac{1}{2}\lambda\Delta t = 0$ yields

$$\begin{aligned} Q(\lambda\Delta t) &= \frac{1 + \frac{1}{2}\lambda\Delta t}{1 - \frac{1}{2}\lambda\Delta t} \\ &= \left(1 + \frac{1}{2}\lambda\Delta t\right)\left(1 + \frac{1}{2}\lambda\Delta t + \left(\frac{1}{2}\lambda\Delta t\right)^2 + \left(\frac{1}{2}\lambda\Delta t\right)^3 + \mathcal{O}(\Delta t^4)\right) \\ &= 1 + \lambda\Delta t + \frac{1}{2}(\lambda\Delta t)^2 + \frac{1}{4}(\lambda\Delta t)^3 + \mathcal{O}(\Delta t^4). \end{aligned} \tag{3}$$

Equations (2) and (3) are substituted into relation (1) to obtain

$$\tau_{n+1} = -\frac{1}{12} y_n \lambda^3 \Delta t^2 + \mathcal{O}(\Delta t^3),$$

hence

$$T = -\frac{1}{12} y_n \lambda^3.$$

(c) The characteristic equation of A is given by

$$\begin{aligned} & \det(A-\lambda I) = 0 \\ \Rightarrow \quad & \begin{vmatrix} -1-\lambda & 2 & -2 \\ 0 & -2-\lambda & -2 \\ 0 & 2 & -2-\lambda \end{vmatrix} = 0 \\ \Rightarrow \quad & (-1-\lambda)\begin{vmatrix} -2-\lambda & -2 \\ 2 & -2-\lambda \end{vmatrix} = 0 \\ \Rightarrow \quad & (-1-\lambda)\left((-2-\lambda)^2+4\right) = 0. \end{aligned}$$

The eigenvalues of A are calculated from this as $\lambda_1 = -1$ and $\lambda_2 = \overline{\lambda_3} = -2+2i$. Because λ_2 and λ_3 are each other complex conjugates, stability is governed by λ_1 and λ_2.

For $\lambda_1 = -1$ and $\Delta t = 1$ we obtain

$$\begin{aligned} Q(\lambda_1 \Delta t) &= Q(-1) \\ &= \frac{1+\frac{1}{2}(-1)}{1-\frac{1}{2}(-1)} \\ &= \frac{\frac{1}{2}}{\frac{3}{2}} \\ &= \frac{1}{3}, \end{aligned}$$

and therefore

$$|Q(\lambda_1 \Delta t)| = \frac{1}{3} \leq 1. \tag{4}$$

For $\lambda_2 = -2+2i$ and $\Delta t = 1$ we obtain

$$\begin{aligned} Q(\lambda_2 \Delta t) &= Q(-2+2i) \\ &= \frac{1+\frac{1}{2}(-2+2i)}{1-\frac{1}{2}(-2+2i)} \\ &= \frac{i}{2-i)} \\ &= -\frac{1}{5}+\frac{2}{5}i, \end{aligned}$$

and therefore

$$|Q(\lambda_2 \Delta t)| = \sqrt{\frac{1}{25}+\frac{4}{25}} = \sqrt{\frac{1}{5}} \leq 1. \tag{5}$$

From (4) and (5) it follows that the method applied to the given IVP is stable for $\Delta t = 1$.

(d) First note that holds

$$\mathbf{w}_0 = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}.$$

We can show that

$$A\mathbf{w}_0 + \mathbf{b} = \mathbf{0}. \tag{6}$$

The given value for $\mathbf{w}_1$ is exactly equal to $\mathbf{w}_0$, so we also have as a direct consequence:

$$A\mathbf{w}_1 + \mathbf{b} = \mathbf{0}. \tag{7}$$

(6), (7) and the values for $\mathbf{w}_0$ and $\mathbf{w}_1$ can be substituted in the method, which leads to

$$\begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix},$$

which is mathematically correct. Therefor $\mathbf{w}_1$ as given is indeed the approximation of the exact solution at time $t = 1$.

Alternative solution: $\mathbf{w}_1$ can also be calculated explicitly be direct application of the method, which has the following calculations:

$$\begin{aligned}
& & \mathbf{w}_0 &= \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}, \\
&\text{Method:} & \mathbf{w}_1 &= \mathbf{w}_0 + \frac{1}{2}\left(A\mathbf{w}_0 + \mathbf{b} + A\mathbf{w}_1 + \mathbf{b}\right), \\
&\Rightarrow & \left(I - \frac{1}{2}A\right)\mathbf{w}_1 &= \left(I + \frac{1}{2}A\right)\mathbf{w}_0 + \mathbf{b}, \\
&\Rightarrow & \begin{bmatrix} 3/2 & -1 & 1 \\ 0 & 2 & 1 \\ 0 & -1 & 2 \end{bmatrix}\mathbf{w}_1 &= \begin{bmatrix} 11/2 \\ 1 \\ 7 \end{bmatrix}, \\
&\Rightarrow & \mathbf{w}_1 &= \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}.
\end{aligned}$$

No points will be given if a different method is used or a different system of differential equations is solved.

2. (a) The linear Lagrangian interpolation polynomial, with nodes a and b, is given by

$$p_1(x) = \frac{x-b}{a-b}f(a) + \frac{x-a}{b-a}f(b).$$

We approximate $f(x)$ by $p_1(x)$ in the integral $\int_a^b f(x)\mathrm{d}x$, from which follows:

$$\begin{aligned}
\int_a^b f(x)\,\mathrm{d}x &\approx \int_a^b p_1(x)\,\mathrm{d}x \\
&= \int_a^b \left\{\frac{x-b}{a-b}f(a) + \frac{x-a}{b-a}f(b)\right\}\,\mathrm{d}x \\
&= \left[\frac{1}{2}\frac{(x-b)^2}{a-b}f(a)\right]_a^b + \left[\frac{1}{2}\frac{(x-a)^2}{b-a}f(b)\right]_a^b \\
&= \frac{1}{2}(b-a)(f(a)+f(b)).
\end{aligned}$$

This is the Trapezoidal rule.

(b) The magnitude of the error of the numerical integration over interval $[a, b]$ is given by

$$\begin{aligned}
\left|\int_a^b f(x)\,\mathrm{d}x - \int_a^b p_1(x)\,\mathrm{d}x\right| &= \left|\int_a^b (f(x) - p_1(x))\,\mathrm{d}x\right| \\
&= \left|\int_a^b \frac{1}{2}(x-a)(x-b)f''(\xi(x))\,\mathrm{d}x\right| \\
&\le \frac{1}{2}\max_{x\in[a,b]}|f''(x)|\int_a^b |(x-a)(x-b)|\,\mathrm{d}x \\
&= \frac{1}{12}(b-a)^3\max_{x\in[a,b]}|f''(x)|.
\end{aligned}$$

(c) The composite Trapezoidal rule for $\displaystyle\int_0^1 x^2\,\mathrm{d}x$ with $h = 1/4$ is given by

$$\begin{aligned}
\frac{1}{h}\left(\frac{1}{2}x_0^2 + \left(\sum_{j=2}^{3} x_j^2\right) + \frac{1}{2}x_4^2\right) &= \frac{1}{4}\left(\frac{1}{2}0^2 + \frac{1}{4}^2 + \frac{1}{2}^2 + \frac{3}{4}^2 + \frac{1}{2}1^2\right) \\
&= \frac{11}{32} = 0.34375.
\end{aligned}$$

(d) Since $\displaystyle\int_0^1 x^2\,\mathrm{d}x = \frac{1}{3}$ the absolute value of the truncation error is:

$$\left|\frac{1}{3} - \frac{22}{64}\right| = \frac{1}{96} = 0.01041\overline{6}.$$

3. (a) Using central differences for the second order derivative at a node $x_j = j\Delta x$ gives

$$y''(x_j) \approx \frac{y_{j+1} - 2y_j + y_{j-1}}{\Delta x^2} =: Q(\Delta x). \tag{8}$$

Here, $y_j := y(x_j)$. Next, we will prove that this approximation is second order accurate, that is $|y''(x_j) - Q(\Delta x)| = \mathcal{O}(\Delta x^2)$.
Using Taylor's Theorem around $x = x_j$ gives

$$\begin{aligned} y_{j+1} &= y(x_j + \Delta x) = y(x_j) + \Delta x y'(x_j) + \tfrac{\Delta x^2}{2} y''(x_j) + \tfrac{\Delta x^3}{3!} y'''(x_j) + O(\Delta x^4) \\ y_{j-1} &= y(x_j - \Delta x) = y(x_j) - \Delta x y'(x_j) + \tfrac{\Delta x^2}{2} y''(x_j) - \tfrac{\Delta x^3}{3!} y'''(x_j) + O(\Delta x^4) \end{aligned} \tag{9}$$

Substitution of these expressions into $Q(\Delta x)$ gives

$$|y''(x_j) - Q(\Delta x)| = \mathcal{O}(\Delta x^2).$$

This leads to the following discretisation formula for internal grid nodes:

$$\frac{-w_{j-1} + 2w_j - w_{j+1}}{\Delta x^2} + (x_j + 1)w_j = x_j^3 + x_j^2 - 2. \tag{10}$$

Here, w_j represents the numerical approximation of the solution y_j. To deal with the boundary $x = 0$, we use a virtual node at $x = -\Delta x$, and we define $y_{-1} := y(-\Delta x)$. Then, using central differences at $x = 0$ gives

$$0 = y'(0) \approx \frac{y_1 - y_{-1}}{2\Delta x} =: Q_b(\Delta x). \tag{11}$$

Using Taylor's Theorem, gives

$$\begin{aligned} Q_b(\Delta x) &= \\ &= \frac{y(0) + \Delta x y'(0) + \frac{\Delta x^2}{2} y''(0) + O(\Delta x^3)}{2\Delta x} \\ &- \frac{y(0) - \Delta x y'(0) + \frac{\Delta x^2}{2} y''(0) + O(\Delta x^3)}{2\Delta x} \\ &= y'(0) + \mathcal{O}(\Delta x^2). \end{aligned}$$

Again, we get an error of $\mathcal{O}(\Delta x^2)$.

(b) With respect to the numerical approximation at the virtual node, we get

$$\frac{w_1 - w_{-1}}{2\Delta x} = 0 \quad \Leftrightarrow \quad w_{-1} = w_1. \tag{12}$$

The discretisation at $x = 0$ is given by

$$\frac{-w_{-1} + 2w_0 - w_1}{\Delta x^2} + w_0 = -2. \tag{13}$$

Substitution of equation (12) into the above equation, yields

$$\frac{2w_0 - 2w_1}{\Delta x^2} + w_0 = -2. \tag{14}$$

Subsequently, we consider the boundary $x = 1$. To this extent, we consider its neighbouring point x_{n-1} and substitute the boundary condition $w_n = y(1) = y_n = 1$ into equation (10) to obtain

$$\begin{aligned} & \frac{-w_{n-2} + 2w_{n-1}}{\Delta x^2} + (x_{n-1} + 1)w_{n-1} && (15) \\ = \ & x_{n-1}^3 + x_{n-1}^2 - 2 + \frac{1}{\Delta x^2} && (16) \\ = \ & (1 - \Delta x)^3 + (1 - \Delta x)^2 - 2 + \frac{1}{\Delta x^2}. && (17) \end{aligned}$$

This concludes our discretisation of the boundary conditions. In order to get a symmetric discretisation matrix, one divides equation (14) by 2.

Next, we use $\Delta x = 1/3$. From equations (10, 14, 17) we obtain the following system

$$\begin{aligned} 9\frac{1}{2}w_0 - 9w_1 &= -1 \\ -9w_0 + 19\frac{1}{3}w_1 - 9w_2 &= -\frac{50}{27} \\ -9w_1 + 19\frac{2}{3}w_2 &= \frac{209}{27}. \end{aligned}$$

(c) The Gershgorin circle theorem states that the eigenvalues of a square matrix $\mathbf{A}$ are located in the complex plane in the union of circles

$$|z - a_{ii}| \leq \sum_{\substack{j \neq i \\ j=1}}^{n} |a_{ij}| \quad \text{where} \quad z \in \mathbb{C} \tag{18}$$

For the $n \times n$ matrix given in part (c) we have

- For $i = 1$:

$$\left| z - \left(\frac{2}{(\Delta x)^2} + 1\right) \right| \leq \frac{1}{(\Delta x)^2} \quad \Rightarrow \quad |\lambda|_{\min} \geq 1 + \frac{1}{(\Delta x)^2} \tag{19}$$

- For $i = 2 \ldots n - 1$:

$$\left| z - \left(\frac{2}{(\Delta x)^2} + 1\right) \right| \leq \frac{2}{(\Delta x)^2} \quad \Rightarrow \quad |\lambda|_{\min} \geq 1 \tag{20}$$

- For $i = n$:

$$\left| z - \left(\frac{2}{(\Delta x)^2} + 1\right) \right| \leq \frac{1}{(\Delta x)^2} \quad \Rightarrow \quad |\lambda|_{\min} \geq 1 + \frac{1}{(\Delta x)^2} \tag{21}$$

Hence, a lower bound for the smallest eigenvalue is 1. For a symmetric matrix $\mathbf{A}$ we have

$$\|\mathbf{A}^{-1}\| = \frac{1}{|\lambda|_{\min}} \leq 1 \tag{22}$$

This proves that the finite-difference scheme is stable, e.g., with constant $C = 1$.