DELFT UNIVERSITY OF TECHNOLOGY
Faculty of Electrical Engineering, Mathematics and Computer Science

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (CTB2400)

Thursday June 27 2024, 13:30-16:30

1. (a) Consider the test equation $y' = \lambda y$, then it follows that

$$\begin{aligned}
k_1 &= \lambda \Delta t w_n && (1)\\
k_2 &= \lambda \Delta t \left(w_n + \frac{1}{2}\lambda \Delta t w_n \right) && (2)\\
&= \left(\lambda \Delta t + \frac{1}{2}(\lambda \Delta t)^2 \right) w_n && (3)\\
k_3 &= \lambda \Delta t \left(w_n - \lambda \Delta t w_n + 2\left(\lambda \Delta t + \frac{1}{2}(\lambda \Delta t)^2 \right) w_n \right) && (4)\\
&= \left(\lambda \Delta t + (\lambda \Delta t)^2 + (\lambda \Delta t)^3 \right) w_n && (5)\\
w_{n+1} &= w_n + \alpha \lambda \Delta t w_n + \beta \left(\lambda \Delta t + \frac{1}{2}(\lambda \Delta t)^2 \right) w_n && (6)\\
&\quad + \gamma \left(\lambda \Delta t + (\lambda \Delta t)^2 + (\lambda \Delta t)^3 \right) w_n && (7)\\
&= \left(1 + (\alpha + \beta + \gamma)\lambda \Delta t + \left(\frac{1}{2}\beta + \gamma \right)(\lambda \Delta t)^2 + \gamma (\lambda \Delta t)^3 \right) w_n && (8)
\end{aligned}$$

Hence the amplification factor is given by

$$Q(\lambda \Delta t) = 1 + (\alpha + \beta + \gamma)\lambda \Delta t + \left(\frac{1}{2}\beta + \gamma \right)(\lambda \Delta t)^2 + \gamma (\lambda \Delta t)^3 . \quad (9)$$

(b) The local truncation error for the test equation $y' = \lambda y$ is given by

$$\tau_{n+1}(\Delta t) = \frac{e^{\lambda \Delta t} - Q(\lambda \Delta t)}{\Delta t} y_n. \quad (10)$$

The Taylor Series around 0 for $e^{\lambda \Delta t}$ is:

$$e^{\lambda \Delta t} = 1 + \lambda \Delta t + \frac{1}{2}(\lambda \Delta t)^2 + \frac{1}{6}(\lambda \Delta t)^3 + \mathcal{O}(\Delta t^4). \quad (11)$$

Hence, this gives

$$\begin{aligned}
e^{\lambda \Delta t} - Q(\lambda \Delta t) &= (1 - \alpha - \beta - \gamma)\lambda \Delta t + \left(\frac{1}{2} - \frac{1}{2}\beta - \gamma \right)(\lambda \Delta t)^2 && (12)\\
&\quad + \left(\frac{1}{6} - \gamma \right)(\lambda \Delta t)^3 + \mathcal{O}(\Delta t^4). && (13)
\end{aligned}$$

and hence $\tau_{n+1}(\Delta t) = \mathcal{O}(\Delta t^3)$ only if

$$\begin{aligned}
\alpha + \beta + \gamma &= 1, && (14)\\
\frac{1}{2}\beta + \gamma &= \frac{1}{2}, && (15)\\
\gamma &= \frac{1}{6}, && (16)\\
& && (17)
\end{aligned}$$

which have as solution

$$\alpha = \frac{1}{6}, \tag{18}$$
$$\beta = \frac{2}{3}, \tag{19}$$
$$\gamma = \frac{1}{6}. \tag{20}$$
$$\tag{21}$$

(c) Let $x_1 = y$ and $x_2 = y'$, then it follows that $y'' = x_2'$, and hence we get

$$\begin{aligned} &2x_2' + 2x_2 + x_1 = 2t, \\ &x_2 = x_1'. \end{aligned} \tag{22}$$

This expression is written as

$$\begin{aligned} &x_1' = x_2, \\ &x_2' = -\tfrac{1}{2}x_1 - x_2 + t. \end{aligned} \tag{23}$$

Finally, we get the following matrix–form:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}' = \begin{bmatrix} 0 & 1 \\ -\frac{1}{2} & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ t \end{bmatrix}. \tag{24}$$

Here, we have $A = \begin{bmatrix} 0 & 1 \\ -\frac{1}{2} & -1 \end{bmatrix}$ and $f = \begin{bmatrix} 0 \\ t \end{bmatrix}$. The initial conditions are given by $\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

(d) To this extent, we determine the eigenvalues of the matrix A. Subsequently, these eigenvalues are substituted into the amplification factor. The eigenvalues of A are given by $-\frac{1}{2} \pm \frac{1}{2}i$. Using $\Delta t = 2$, it follows that

$$Q(\lambda \Delta t) = 1 + \lambda \Delta t + \frac{1}{2}\lambda^2 \Delta t^2 + \frac{1}{6}\lambda^3 \Delta t^3 \tag{25}$$
$$= 1 + (-1+i) + \frac{1}{2}(-1+i)^2 + \frac{1}{6}(-1+i)^3 \tag{26}$$
$$= \frac{1}{3} - \frac{1}{3}i. \tag{27}$$

Herewith, it follows that $|Q(\lambda \Delta t)|^2 = \frac{2}{9} < 1$. Hence for $\Delta t = 2$, it follows that the method applied to the given system is stable. Note that this conclusion holds for both the eigenvalues of A since they are complex conjugates.

(e) The given method, applied to the system $\underline{x}' = A\underline{x} + \underline{f}$, gives

$$\begin{cases} \underline{k}_1 = \Delta t \left(A\underline{w}_n + \underline{f}(t_n) \right) \\ \underline{k}_2 = \Delta t \left(A\left(\underline{w}_n + \frac{1}{2}\underline{k}_1\right) + \underline{f}\left(t_n + \frac{1}{2}\Delta t\right) \right) \\ \underline{k}_3 = \Delta t \left(A\left(\underline{w}_n - \underline{k}_1 + 2\underline{k}_2\right) + \underline{f}\left(t_n + \Delta t\right) \right) \\ \underline{w}_{n+1} = \underline{w}_n + \frac{1}{6}\left(\underline{k}_1 + 4\underline{k}_2 + \underline{k}_3\right) \end{cases} \tag{28}$$

With the initial condition and $\Delta t = 2$, this gives

$$\begin{cases} \underline{k}_1 & = & \begin{bmatrix} 2 \\ -3 \end{bmatrix} \\ \\ \underline{k}_2 & = & \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ \\ \underline{k}_3 & = & \begin{bmatrix} 12 \\ -5 \end{bmatrix} \\ \\ \underline{w}_1 & = & \begin{bmatrix} 8/3 \\ 1/3 \end{bmatrix} \end{cases} \tag{29}$$

2. (a) First, we check the boundary conditions:

$$u(0) = 0 - \frac{1-e^0}{1-e} = \frac{1-1}{1-e} = 0, \quad u(1) = 1 - \frac{1-e^1}{1-e} = 0. \tag{30}$$

Further, we have

$$u'(x) = 1 + \frac{e^x}{1-e}, \tag{31}$$

$$u''(x) = \frac{e^x}{1-e}. \tag{32}$$

Hence, we immediately see

$$-u''(x) + u'(x) = -\frac{e^x}{1-e} + 1 + \frac{e^x}{1-e} = 1. \tag{33}$$

Hence, the solution $u(x) = 1 - \frac{1-e^x}{1-e}$ satisfies the differential and the boundary conditions, and therewith $u(x)$ is the solution to the boundary value problem (uniqueness can be demonstrated in a straightforward way, but this was not asked for).

(b) The domain of computation, being $(0, 1)$, is divided into subintervals with mesh points, we set $x_j = j\Delta x$, where we use n unknowns, such that $x_{n+1} = (n+1)\Delta x = 1$. We are looking for a discretization with an error of second order, $O((\Delta x)^2)$. To this extent, we use the following central differences approximation at x_j:

$$u'(x_j) \approx \frac{u(x_{j+1}) - u(x_{j-1})}{2\Delta x}, \text{ for } j \in \{1, \ldots, n\}. \tag{34}$$

We note that the above formula can be derived formally by writing the derivative as

$$u'(x_j) = \frac{\alpha_0 u(x_{j-1}) + \alpha_1 u(x_j) + \alpha_2 u(x_{j+1})}{\Delta x}, \tag{35}$$

and solve α_0, α_1 and α_2 from checking the zeroth, first and second order derivatives of $u(x)$. Further, the second order derivative is approximated by

$$u''(x_j) \approx \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{(\Delta x)^2}. \tag{36}$$

Since we approximate the derivatives at the point x_j, we use Taylor series expansion about x_j, to obtain:

$$\begin{aligned} u(x_{j+1}) &= u(x_j + \Delta x) = u(x_j) + \Delta x u'(x_j) + \frac{(\Delta x)^2}{2} u''(x_j) + \frac{(\Delta x)^3}{6} u'''(x_j) + O((\Delta x)^4), \\ u(x_{j-1}) &= u(x_j - \Delta x) = u(x_j) - \Delta x u'(x_j) + \frac{(\Delta x)^2}{2} u''(x_j) - \frac{(\Delta x)^3}{6} u'''(x_j) + O((\Delta x)^4), \end{aligned} \tag{37}$$

This gives

$$\begin{aligned} &-\frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{(\Delta x)^2} + \frac{u(x_{j+1}) - u(x_{j-1})}{2\Delta x} = -u''(x_j) + u'(x_j) \\ &+ \frac{O((\Delta x)^3)}{2\Delta x} + \frac{O((\Delta x)^4)}{(\Delta x)^2} = -u''(x_j) + u'(x_j) + O((\Delta x)^2). \end{aligned} \tag{38}$$

Hence the error is second order, that is $O((\Delta x)^2)$. Next, we neglect the truncation error, and set $w_j := u(x_j)$ to get

$$-\frac{w_{j+1} - 2w_j + w_{j-1}}{(\Delta x)^2} + \frac{w_{j+1} - w_{j-1}}{2\Delta x} = 1, \text{ for } j \in \{1, \ldots, n\}. \tag{39}$$

At the boundaries, we see for $j = 1$ and $j = n$, upon substituting $w_0 = 0$ and $w_{n+1} = 0$, respectively:

$$\begin{aligned} -\frac{w_2 - 2w_1 + 0}{(\Delta x)^2} + \frac{w_2 - 0}{2\Delta x} &= 1, \\ -\frac{0 - 2w_n + w_{n-1}}{(\Delta x)^2} + \frac{0 - w_{n-1}}{2\Delta x} &= 1. \end{aligned} \tag{40}$$

This can be rewritten more neatly as follows:

$$\begin{aligned} \frac{-w_2 + 2w_1}{(\Delta x)^2} + \frac{w_2}{2\Delta x} &= 1, \\ \frac{2w_n - w_{n-1}}{(\Delta x)^2} - \frac{w_{n-1}}{2\Delta x} &= 1. \end{aligned} \tag{41}$$

(c) Next, we use $\Delta x = 1/4$, then, from equations (39) and (41), one obtains the following system

$$32w_1 - 14w_2 = 1 \tag{42}$$

$$-18w_1 + 32w_2 - 14w_3 = 1 \tag{43}$$

$$-18w_2 + 32w_3 = 1 \tag{44}$$

3. (a) The equation that needs to be solved is

$$f(p_0) + \frac{f(p_1) - f(p_0)}{p_1 - p_0}(p_2 - p_0) = 0.$$

Solving this equation gives the steps:

$$\begin{aligned} & \frac{f(p_1) - f(p_0)}{p_1 - p_0}(p_2 - p_0) = -f(p_0), \\ \Rightarrow \quad & p_2 - p_0 = -\frac{p_1 - p_0}{f(p_1) - f(p_0)} f(p_0), \\ \Rightarrow \quad & p_2 = p_0 - \frac{p_1 - p_0}{f(p_1) - f(p_0)} f(p_0). \end{aligned}$$

We write the above as one quotient:

$$\begin{aligned} p_2 &= \frac{f(p_1) - f(p_0)}{f(p_1) - f(p_0)} p_0 - \frac{p_1 - p_0}{f(p_1) - f(p_0)} f(p_0), \\ \Rightarrow \quad p_2 &= \frac{p_0 f(p_1) - p_1 f(p_0)}{f(p_1) - f(p_0)}. \end{aligned} \tag{45}$$

Now we have two options:

A. Rewrite the above formula to the form given in the exercise, with $n = 2$, and conclude the formula for K_1;

B. Fill in the formula for K_1 into the formula for p_n, with $n = 2$, given in the exercise and show this results in the same formula.

Note: Only one of the options has to be present within your answer and earns at most 1/2 point.

Option A: We can rewrite Equation (45) to:

$$\begin{aligned} p_2 &= \frac{p_0 f(p_1) - p_1 f(p_0)}{f(p_1) - f(p_0)}, \\ \Rightarrow \quad p_2 &= \frac{p_0 f(p_1) - p_1 f(p_0) - p_1 f(p_1) + p_1 f(p_1)}{f(p_1) - f(p_0)}, \\ \Rightarrow \quad p_2 &= \frac{p_1(f(p_1) - f(p_0)) - (p_1 - p_0) f(p_1)}{f(p_1) - f(p_0)}, \\ \Rightarrow \quad p_2 &= p_1 - \frac{p_1 - p_0}{f(p_1) - f(p_0)} f(p_1), \end{aligned}$$

which is indeed of the form given in the exercise. Therefore, K_1 indeed has the formula

$$K_1 = \frac{f(p_1) - f(p_0)}{(p_1 - p_0)}.$$

Option B: The formula of the exercise, with $n = 2$ and the given formula for K_1 is:

$$p_2 = p_1 - \frac{p_1 - p_0}{f(p_1) - f(p_0)} f(p_1).$$

We write the above as one quotient:

$$\begin{aligned} p_2 &= \frac{f(p_1) - f(p_0)}{f(p_1) - f(p_0)} p_1 - \frac{p_1 - p_0}{f(p_1) - f(p_0)} f(p_1), \\ \Rightarrow \; p_2 &= \frac{p_0 f(p_1) - p_1 f(p_0)}{f(p_1) - f(p_0)}. \end{aligned}$$

The above equation is equal to Equation (45). Therefore, K_1 indeed has the formula

$$K_1 = \frac{f(p_1) - f(p_0)}{p_1 - p_0}.$$

(b) *Note: Every miscalculation in the calculation of K_1 gives a subtraction of $1/4$ point, with at most $1/2$ point being subtracted.*

Given that $p_0 = 1$ and $p_1 = 2$, we first calculate K_1, using the values from the given table:

$$\begin{aligned} K_1 &= \frac{f(p_1) - f(p_0)}{p_1 - p_0}, \\ &= \frac{f(2) - f(1)}{2 - 1}, \\ &= f(2) - f(1), \\ &= 2 - (-1), \\ &= 3. \end{aligned}$$

Note: Every miscalculation in the calculation of p_2 gives a subtraction of $1/4$ point, with at most $1/2$ point being subtracted.

Note: The value of p_2 should be consistent with your value for K_1.

Now p_2 can be calculated with the Secant method, with $n = 2$ and the values from the given table:

$$\begin{aligned} p_2 &= p_1 - \frac{f(p_1)}{K_1}, \\ &= 2 - \frac{f(2)}{3}, \\ &= 2 - \frac{2}{3}, \\ &= \frac{4}{3}. \end{aligned}$$

(c) The formula for K_2 is given by

$$K_2 = \frac{f(p_2) - f(p_1)}{p_2 - p_1}.$$

Motivation could be a repetition of the derivation of K_1. It is also sufficient if a motivation is given that all indices are increased by 1. No motivation gives a subtraction of $1/2$ point.

Note: Every miscalculation in the calculation of K_2 gives a subtraction of $1/4$ point, with at most $3/4$ point being subtracted.

Note: The value of K_2 should be consistent with your formula for K_2.

This formula gives

$$\begin{aligned}
K_2 &= \frac{f(p_2) - f(p_1)}{p_2 - p_1}, \\
&= \frac{f(\frac{4}{3}) - f(2)}{\frac{4}{3} - 2}, \\
&= \frac{\left(-\frac{2}{9}\right) - 2}{-\frac{2}{3}}, \\
&= \frac{-\frac{20}{9}}{-\frac{2}{3}}, \\
&= \frac{10}{3},
\end{aligned}$$

Note: Every miscalculation in the calculation of p_3 gives a subtraction of $^1/_4$ point, with at most $^3/_4$ point being subtracted.
Note: The value of p_3 should be consistent with your value for K_2.
and finally

$$\begin{aligned}
p_3 &= p_2 - \frac{f(p_2)}{K_2}, \\
&= \frac{4}{3} - \frac{f(\frac{4}{3})}{\frac{10}{3}}, \\
&= \frac{4}{3} - \frac{-\frac{2}{9}}{\frac{10}{3}}, \\
&= \frac{4}{3} - \frac{1}{15}, \\
&= \frac{7}{5}.
\end{aligned}$$