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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU/Minor AESB2210)
Thursday April 20th 2017, 18:30-21:30

1. (a) The local truncation error is given by

$$\tau_{n+1}(\Delta t) = \frac{y_{n+1} - z_{n+1}}{\Delta t}, \tag{1}$$

where z_{n+1} is computed by one step of the method starting from y_n, and we determine y_{n+1} by the use of a Taylor Series around t_n:

$$y_{n+1} = y_n + \Delta t y'(t_n) + \frac{\Delta t^2}{2} y''(t_n) + O(\Delta t^3). \tag{2}$$

We realize that

$$\begin{aligned} y'(t_n) & & &= f(t_n, y_n) \\ y''(t_n) &= \frac{df(t_n, y_n)}{dt} &= \frac{\partial f(t_n, y_n)}{\partial t} &+ \frac{\partial f(t_n, y_n)}{\partial y} y'(t_n) = \\ & &= \frac{\partial f(t_n, y_n)}{\partial t} &+ \frac{\partial f(t_n, y_n)}{\partial y} f(t_n, y_n). \end{aligned} \tag{3}$$

Hence, this gives

$$y_{n+1} = y_n + \Delta t y'(t_n) + \frac{\Delta t^2}{2} \left(\frac{\partial f(t_n, y_n)}{\partial t} + \frac{\partial f(t_n, y_n)}{\partial y} f(t_n, y_n) \right) + O(\Delta t^3). \tag{4}$$

For z_{n+1}, after substitution of the predictor–step for z^*_{n+1} into the corrector–step, and using the Taylor Series around (t_n, y_n)

$$z_{n+1} = y_n + \tfrac{\Delta t}{2} \left(f(t_n, y_n) + f(t_n + \Delta t, y_n + \Delta t f(t_n, y_n)) \right) =$$

$$y_n + \frac{\Delta t}{2} \left(f(t_n, y_n) + f(t_n, y_n) + \Delta t (\frac{\partial f(t_n, y_n)}{\partial t} + f(t_n, y_n) \frac{\partial f(t_n, y_n)}{\partial y}) + O(\Delta t^2) \right). \tag{5}$$

Then, it follows that

$$y_{n+1} - z_{n+1} = O(\Delta t^3), \text{ and hence } \tau_{n+1}(\Delta t) = \frac{O(\Delta t^3)}{\Delta t} = O(\Delta t^2). \tag{6}$$

(b) Consider the test–equation $y' = \lambda y$, then it follows that

$$\begin{aligned} w^*_{n+1} &= w_n + \lambda\Delta t w_n = (1+\lambda\Delta t)w_n, \\ w_{n+1} &= w_n + \frac{\Delta t}{2}(\lambda w_n + \lambda w^*_{n+1}) = \\ &= w_n + \frac{\Delta t}{2}(\lambda w_n + \lambda(w_n + \lambda\Delta t w_n)) = (1 + \lambda\Delta t + \frac{(\lambda\Delta t)^2}{2})w_n. \end{aligned} \tag{7}$$

Hence the amplification factor is given by

$$Q(\lambda\Delta t) = 1 + \lambda\Delta t + \frac{(\lambda\Delta t)^2}{2}. \tag{8}$$

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

$$\begin{aligned} x_2' + 2x_2 + 2x_1 &= \sin(t), \\ x_2 &= x_1'. \end{aligned} \tag{9}$$

This expression is written as

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= -2x_1 - 2x_2 + \sin(t). \end{aligned} \tag{10}$$

Finally, we get the following matrix–form:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \sin(t) \end{pmatrix}. \tag{11}$$

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

(d) The Modified Euler Method, applied to the system $\underline{x}' = A\underline{x} + \underline{f}$, gives

$$\begin{aligned} \underline{w}^*_1 &= \underline{w}_0 + \Delta t \left(A\underline{w}_0 + \underline{f}_0 \right), \\ \underline{w}_1 &= \underline{w}_0 + \tfrac{\Delta t}{2} \left(A\underline{w}_0 + \underline{f}_0 + A\underline{w}^*_1 + \underline{f}_1 \right). \end{aligned} \tag{12}$$

With the initial condition and $\Delta t = 0.1$, this gives

$$\underline{w}^*_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \frac{1}{10} \left(\begin{pmatrix} 0 & 1 \\ -2 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 1.2 \\ 1.4 \end{pmatrix}. \tag{13}$$

Then, the correction–step is given by

$$\begin{aligned} \underline{w}_1 &= \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \tfrac{1}{20} \left(\begin{pmatrix} 0 & 1 \\ -2 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -2 & -2 \end{pmatrix} \begin{pmatrix} 1.2 \\ 1.4 \end{pmatrix} + \begin{pmatrix} 0 \\ \sin(\frac{1}{10}) \end{pmatrix} \right) = \\ &= \begin{pmatrix} 1.17 \\ 1.445 \end{pmatrix} \end{aligned} \tag{14}$$

(e) 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 $-1 \pm i$. Using $\Delta t = 1$, it follows that

$$Q(\lambda \Delta t) = 1 + \lambda \Delta t + \frac{1}{2}\lambda^2 \Delta t^2 = 1 + (-1+i) + \frac{1}{2}(-1+i)^2 = 0 \qquad (15)$$

Herewith, it follows that $|Q(\lambda \Delta t)|^2 = 0 < 1$. Hence for $\Delta t = 1$, 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.

2. (a) The first order backward difference formula for the first derivative is given by

$$d'(t) \approx \frac{d(t) - d(t-h)}{h}.$$

Using $t = 20$, and $h = 10$ the approximation of the velocity is

$$\frac{d(20) - d(10)}{10} = \frac{100 - 40}{10} = 6 \text{ (m/s)}.$$

(b) Taylor polynomials are:

$$\begin{aligned} d(0) &= d(2h) - 2hd'(2h) + 2h^2 d''(2h) - \frac{(2h)^3}{6} d'''(\xi_0)\,, \\ d(h) &= d(2h) - hd'(2h) + \frac{h^2}{2} d''(2h) - \frac{h^3}{6} d'''(\xi_1)\,, \\ d(2h) &= d(2h). \end{aligned}$$

We know that $Q(h) = \frac{\alpha_0}{h} d(0) + \frac{\alpha_1}{h} d(h) + \frac{\alpha_2}{h} d(2h)$, which should be equal to $d'(2h) + O(h^2)$. This leads to the following conditions:

$$\begin{array}{rcrcrcl} \frac{\alpha_0}{h} & + & \frac{\alpha_1}{h} & + & \frac{\alpha_2}{h} & = & 0\,, \\ -2\alpha_0 & - & \alpha_1 & & & = & 1\,, \\ 2\alpha_0 h & + & \frac{1}{2}\alpha_1 h & & & = & 0\,. \end{array}$$

(c) The truncation error follows from the Taylor polynomials:

$$d'(2h) - Q(h) = d'(2h) - \frac{d(0) - 4d(h) + 3d(2h)}{2h} = \frac{\frac{8h^3}{6} d'''(\xi_0) - 4(\frac{h^3}{6} d'''(\xi_1))}{2h} = \frac{1}{3} h^2 d'''(\xi).$$

Using the new formula with $h = 10$ we obtain the estimate:

$$\frac{d(0) - 4d(10) + 3d(20)}{20} = \frac{0 - 4 \times 40 + 3 \times 100}{20} = 7 \text{ (m/s)}.$$

3. (a) **Newton-Raphson's method** is an iterative method to find $p \in \mathbb{R}$ such that $f(p) = 0$. Suppose $f \in C^2[a, b]$. Let $\bar{x} \in [a, b]$ be an approximation of the root p such that $f'(\bar{x}) \neq 0$, and suppose that $|p - \bar{x}|$ is small. Consider the first-degree Taylor polynomial about $\bar{x}$:

$$f(x) = f(\bar{x}) + (x - \bar{x})f'(\bar{x}) + \frac{(x - \bar{x})^2}{2} f''(\xi(x)), \tag{16}$$

in which $\xi(x)$ between x and $\bar{x}$. Using that $f(p) = 0$, equation (16) yields

$$0 = f(\bar{x}) + (p - \bar{x})f'(\bar{x}) + \frac{(p - \bar{x})^2}{2} f''(\xi(x)).$$

Because $|p - \bar{x}|$ is small, $(p - \bar{x})^2$ can be neglected, such that

$$0 \approx f(\bar{x}) + (p - \bar{x})f'(\bar{x}).$$

Note that the right-hand side is the formula for the tangent in $(\bar{x}, f(\bar{x}))$. Solving for p yields

$$p \approx \bar{x} - \frac{f(\bar{x})}{f'(\bar{x})}.$$

This motivates the Newton-Raphson method, that starts with an approximation p_0 and generates a sequence $\{p_n\}$ by

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}, \quad \text{for} \quad n \geq 1.$$

Remark 1 *One can also give a graphical derivation following Figure 4.2 from the book.*

(b) The first derivative of g equals

$$g'(x) = 1 - \frac{(f'(x))^2 - f(x)f''(x)}{(f'(x))^2} = \frac{f(x)f''(x)}{(f'(x))^2}.$$

Substitution of $f(x) = \sin(x)$, $f'(x) = \cos(x)$ and $f''(x) = -\sin(x)$ yields

$$g'(x) = -\frac{\sin^2(x)}{\cos^2(x)} = -\tan^2(x).$$

Since $\tan(-\pi/4) = -1$, $\tan(\pi/4) = 1$ and the tangent function is monotonically increasing on the interval $[-1, 1]$ any initial guess inside the interval $(-1, 1)$ will lead to a convergent iteration process.

(c) It follows from the linearization of the function $\mathbf{f}$ about the iterate $\mathbf{x}_{n-1}$ that

$$f_1(\mathbf{p}) \approx f_1(\mathbf{p}^{(n-1)}) + \frac{\partial f_1}{\partial p_1}(\mathbf{p}^{(n-1)})(p_1 - p_1^{(n-1)}) + \ldots + \frac{\partial f_1}{\partial p_m}(\mathbf{p}^{(n-1)})(p_m - p_m^{(n-1)}),$$

$$\vdots$$

$$f_m(\mathbf{p}) \approx f_m(\mathbf{p}^{(n-1)}) + \frac{\partial f_m}{\partial p_1}(\mathbf{p}^{(n-1)})(p_1 - p_1^{(n-1)}) + \ldots + \frac{\partial f_m}{\partial p_m}(\mathbf{p}^{(n-1)})(p_m - p_m^{(n-1)}).$$

Defining the Jacobian matrix of $\mathbf{f}(\mathbf{x})$ by

$$\mathbf{J}(\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \ldots & \frac{\partial f_1}{\partial x_m}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}) & \ldots & \frac{\partial f_m}{\partial x_m}(\mathbf{x}) \end{pmatrix},$$

the linearization can be written in the more compact form

$$\mathbf{f}(\mathbf{p}) \approx \mathbf{f}(\mathbf{p}^{(n-1)}) + \mathbf{J}(\mathbf{p}^{(n-1)})(\mathbf{p} - \mathbf{p}^{(n-1)}).$$

The next iterate, $\mathbf{p}^{(n)}$, is obtained by setting the linearization equal to zero:

$$\mathbf{f}(\mathbf{p}^{(n-1)}) + \mathbf{J}(\mathbf{p}^{(n-1)})(\mathbf{p}^{(n)} - \mathbf{p}^{(n-1)}) = 0, \tag{17}$$

which can be rewritten as

$$\mathbf{J}(\mathbf{p}^{(n-1)})\mathbf{s}^{(n)} = -\mathbf{f}(\mathbf{p}^{(n-1)}), \tag{18}$$

where $\mathbf{s}^{(n)} = \mathbf{p}^{(n)} - \mathbf{p}^{(n-1)}$. The new approximation equals $\mathbf{p}^{(n)} = \mathbf{p}^{(n-1)} + \mathbf{s}^{(n)}$. Finally, Newton-Raphson's formula for general nonlinear problems reads:

$$\mathbf{p}^{(n)} = \mathbf{p}^{(n-1)} - \mathbf{J}^{-1}(\mathbf{p}^{(n-1)})\mathbf{f}(\mathbf{p}^{(n-1)}). \tag{19}$$

(d) First, we rewrite the system into the form

$$\begin{aligned} f_1(w_1, w_2) &= 0, \\ f_2(w_1, w_2) &= 0, \end{aligned} \tag{20}$$

by setting

$$\begin{aligned} f_1(w_1, w_2) &:= 18w_1 - 9w_2 + (w_1)^2, \\ f_2(w_1, w_2) &:= -9w_1 + 18w_2 + (w_2)^2 - 9. \end{aligned} \tag{21}$$

We denote the Jacobi-matrix by $J(w_1, w_2)$. At the first step we compute

$$\underline{w}^{(1)} = \underline{w}^{(0)} - J(\underline{w}^{(0)})^{-1}F(\underline{w}^{(0)}), \tag{22}$$

where $\underline{w} = [w_1\ w_2]^T$. Note that

$$J(\underline{w}^{(0)}) = \begin{pmatrix} 18 + 2w_1^{(0)} & -9 \\ -9 & 18 + 2w_2^{(0)} \end{pmatrix}. \tag{23}$$

Using $w_1^{(0)} = w_2^{(0)} = 0$ we obtain:

$$J(\underline{w}^{(0)}) = \begin{pmatrix} 18 & -9 \\ -9 & 18 \end{pmatrix}. \tag{24}$$

This implies that

$$J(\underline{w}^{(0)})^{-1} = \frac{1}{18^2 - 81} \begin{pmatrix} 18 & 9 \\ 9 & 18 \end{pmatrix}. \tag{25}$$

Furthermore

$$F(\underline{w}^{(0)}) = \begin{pmatrix} 0 \\ -9 \end{pmatrix}, \tag{26}$$

so

$$\underline{w}^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \frac{1}{18^2 - 81} \begin{pmatrix} 18 & 9 \\ 9 & 18 \end{pmatrix} \begin{pmatrix} 0 \\ -9 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}. \tag{27}$$