UNIVERSITY OF BRISTOL

School of Mathematics

NUMERICAL ANALYSIS
MATH 30029W
(Paper code MATH–30029W)

December 2024   2 hour 30 minutes

Solutions

Do not turn over until instructed.

All questions are new.

1. (a) (*Similar to example in notes*)

(i) Perform row operation $R_2 \to R_2 - 100R_1$ gives

$$\begin{bmatrix} 0.01 & 1.6 & 32.1 \\ 0 & -159 & -3.19 \times 10^3 \end{bmatrix}$$

since using 3 digits, so $22 - 3210 = 3188$ rounds to -3190. Then $y = -3190/159 = 20.1$ after rounding and $0.01x = 32.1 - 32.2 = -0.1$ after rounding meaning that $x = -0.1/0.01 = -10$. This is way out.

(ii) Partial pivoting requires us to swap rows before eliminating. So

$$\begin{bmatrix} 1 & 0.6 & 22 \\ 0.01 & 1.6 & 32.1 \end{bmatrix}$$

and doing $R_2 \to R_2 - 0.01R_1$ gives

$$\begin{bmatrix} 1 & 0.6 & 22 \\ 0 & 1.59 & 31.9 \end{bmatrix}$$

after rounding. Now $y = 31.9/1.59 = 20.1$ after rounding and back substituting $x = 22 - 12.1 = 9.9$. This is much better.

(b) (*Follows standard method described in notes, in set homeworks and on past papers.*)
Perform LU decomposition in steps

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -1 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & \frac{3}{2} & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & \frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & \frac{3}{2} & -1 \\ 0 & 0 & 0 & \frac{5}{3} \end{bmatrix} = LU.$$

Solve in two steps: $L\mathbf{y} = \mathbf{b}$ then $U\mathbf{x} = \mathbf{y}$. So first

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & \frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Forward substitution easily gives $\mathbf{y} = (1, 0, 1, \frac{1}{3})^T$. Next

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & \frac{3}{2} & -1 \\ 0 & 0 & 0 & \frac{5}{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ \frac{1}{3} \end{bmatrix}.$$

Back substitution results in $\mathbf{x} = (\frac{7}{5}, \frac{2}{5}, \frac{4}{5}, \frac{1}{5})^T$.

(c) (*Unseen*)

The trick is to see that $B = PAP$ where

$$P = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

is such that $P^{-1} = P$. So $B\mathbf{z} = \mathbf{b}$ is $A(P\mathbf{z}) = P\mathbf{b} = \mathbf{b}$ which means $P\mathbf{z} = \mathbf{x}$ and so $\mathbf{z} = P\mathbf{x}$. Only now do we need the calculation of $\mathbf{x}$ from part (b). So $\mathbf{z} = (\frac{1}{5}, \frac{4}{5}, \frac{2}{5}, \frac{7}{5})^T$.

Continued...

2. (a) (*Similar to methods used in notes and homeworks.*)

Draw the two graphs, spot there is only one intersection, at $x = x^*$, say, which is obviously positive and less than 1. Solutions of $x^3 + x - 1 = 0$ are equivalent to solutions of $x = 1/(1 + x^2)$ (since $1 + x^2$ is non-vanishing) and we are done.

(b) (*Similar to examples in notes and homeworks*)

The map is defined by $x_{n+1} = g(x_n)$ where $g(x) = 1/(1 + x^2)$.

For $0 \leq x \leq 1$, $g(x) \in [\frac{1}{2}, 1] \subset [0, 1]$ since it is monotonically decreasing and takes its max/min values at $x = 0$, $x = 1$. Also

$$|g'(x)| = \left|\frac{-2x}{(1 + x^2)^2}\right| < 1, \qquad \text{for } 0 \leq x \leq 1$$

which requires some work to establish. For example,

$$g''(x) = \frac{6x^2 - 2}{(1 + x^2)^3}$$

implies there is a max/min in the interval $0 < x < 1$ at $x = 1/\sqrt{3}$ at which $|g'(1/\sqrt{3})| = 9/(8\sqrt{3}) < 1$. This is a maximum since $g'(0) = 0$ and $g'(1) = \frac{1}{2}$.

Hence, by the Fixed Point Theorem, there exists a unique fixed point $x^* \in (0, 1)$ s.t. all $x_0 \in [0, 1]$ will converge to x^*.

Finally, since $g'(x) \neq 0$ for $x \neq 0$ then $g'(x^*) \neq 0$ and so the scheme has first order convergence.

(c) (*Again, familiar type of example.*)

Here we are presented with $x_{n+1} = g(x_n)$ with

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

Assuming a fixed point $x^* = (1 - x^*)/(x^*)^2$ rearranges to the original cubic. So same fixed point. Now

$$g'(x) = \frac{-2 + x}{x^3}$$

whose size is greater than 1 for all $0 < x < 1$. Hence $|g'(x^*)| > 1$ and the scheme cannot converge to x^* apart from if $x_0 = x^*$.

(d) (*Again familiar type of example: it is just Newton's method.*)

Here we are presented with $x_{n+1} = g(x_n)$ with

$$g(x) = \frac{2x^3 + 1}{3x^2 + 1}.$$

Assuming a fixed point $x^*(3(x^*)^2 + 1) = 2(x^*)^3 + 1$ rearranges to the original cubic. So same fixed point. Now

$$g'(x) = \frac{6x^2(3x^2 + 1) - 6x(2x^3 + 1)}{(3x^2 + 1)^2} = \frac{6x(x^3 + x - 1)}{(3x^2 + 1)^2}$$

and so $g'(x^*) = 0$. Since $g(x)$ is continuous, there is a non-vanishing region around $x = x^*$ where $|g'(x)| < 1$ and this means the scheme will converge for x_0 sufficiently close to x^*.

(e) (*Unrelated to the rest of this question, but related to Q4. Standard methods, similar examples in notes and homeworks*).

We Taylor expand approximation about x_0 and equate with $f'(x_0)$ thus:

$$\begin{aligned} f'(x_0) \;\approx\; & \alpha(f + 2hf' + 2h^2 f'' + \tfrac{8}{6}h^3 f'''(\xi_1)) \\ & +\beta(f + hf' + \tfrac{1}{2}h^2 f'' + \tfrac{1}{6}h^3 f'''(\xi_2)) \\ & +\gamma f \end{aligned}$$

for $\xi_1 \in (x_0, x_0 + 2h)$ and $\xi_2 \in (x_0, x_0 + h)$. Then we match coefficients so that

$$\alpha + \beta + \gamma = 0, \qquad 2h\alpha + h\beta = 1, \qquad 4\alpha + \beta = 0.$$

This gives

$$\alpha = -1/2h, \qquad \beta = 4/2h, \qquad \gamma = -3/2h$$

so that

$$f'(x_0) \approx \frac{-f(x_0 + 2h) + 4f(x_0 + h) - 3f(x_0)}{2h}$$

and the error is $O(h^2)$.

Continued...

3. (a) (*Simple numerical methods following definitions in notes*)
Denoting $f(x) = 1/(1+x^2)$ as the integrand we have

$$T_1 = \frac{1}{2}(f(0) + f(1)) = \frac{3}{4} \quad (= 0.75)$$

and

$$S_2 = \frac{1}{6}(f(0) + 4f(1/2) + f(1)) = 47/60 \quad (= 0.783333).$$

(b) (*More simple numerical methods following definitions in notes*)

$$T_2 = \frac{1}{4}(f(0) + 2f(1/2) + f(1)) = 0.775$$

and

$$T_4 = \frac{1}{8}(f(0) + 2f(1/4) + 2f(1/2) + 2f(3/4) + f(1)) = 0.7827941.$$

(c) (*Need to remember Romberg iterations, otherwise straightforward*)

$$T_2^{(1)} = \frac{4T_2 - T_1}{3} = 0.7833333$$

(same as S_2 since Romberg once on the same number of sub-intervals is the same as Simpson). And

$$T_4^{(1)} = \frac{4T_4 - T_2}{3} = 0.7853921.$$

Then we have

$$T_4^{(2)} = \frac{16T_4^{(1)} - T_2^{(1)}}{15} = 0.7855294.$$

We imagine that $T_4^{(2)}$ is the most accurate since the error associated with this iterate is $O(h^6)$ compared with $T_4^{(1)}$ which is $O(h^4)$. However, we see that $T_4^{(1)}$ is closest to the exact value, which is surprising and unexplained.

(d) (*Follows methods in class, unseen example but simple.*)
(i) First $\phi_0(x) = 1$ since this is a polynomial of degree 0 satisfying $\phi_0(0) = 1$. Next, we let $\phi_1(x) = A_1 x + B_1$ and $B_1 = 1$ plus

$$\langle \phi_1, \phi_0 \rangle = 0 = A_1 \int_0^1 x \, dx + 1.$$

Then $A_1 = -2$ and $\phi_1(x) = -2x + 1$.
Next let $\phi_2(x) = A_2 x^2 + B_2 x + 1$ and require

$$\langle \phi_2, \phi_0 \rangle = 0 = A_2 \int_0^1 x^2 \, dx + B_2 \int_0^1 x \, dx + \int_0^1 dx = A_2/3 + B_2/2 + 1$$

and

$$\langle \phi_2, \phi_1 \rangle = 0 = A_2 \int_0^1 x^2(1-2x) \, dx + B_2 \int_0^1 x(1-2x) \, dx + \int_0^1 (1-2x) \, dx$$
$$= A_2(-1/6) + B_2(-1/6).$$

So $A_2 = -B_2$ and $A_2 = 6$, $B_2 = -6$ which gives the required result.

(ii) To get x_j need to solve $6x^2 - 6x + 1 = 0$ and this gives

$$x_{1,2} = \frac{1}{2} \pm \frac{1}{6}\sqrt{3}.$$

Then the weights are given by

$$w_1 = \int_0^1 \frac{x - x_2}{x_1 - x_2}\, dx = \frac{3}{\sqrt{3}} \left(\frac{1}{2} - \frac{1}{2} + \frac{1}{6}\sqrt{3} \right) = \frac{1}{2}$$

and

$$w_2 = \int_0^1 \frac{x - x_1}{x_2 - x_1}\, dx = -\frac{3}{\sqrt{3}} \left(\frac{1}{2} - \frac{1}{2} - \frac{1}{6}\sqrt{3} \right) = \frac{1}{2}.$$

Using these values we have

$$I \approx (f(1/2 + \sqrt{3}/6) + f(1/2 - \sqrt{3}/6))/2 = 0.7868852.$$

Continued...

4. (a) (*Bookwork*)

Euler is

$$y_{i+1} = y_i + hf(y_i, t_i), \qquad i = 0, 1, \ldots$$

with $y_0 = \alpha$.

(b) (*Bookwork/done in notes*)

The local truncation error is defined as $\tau_{i+1} = y(t_{i+1}) - y_{i+1}$ where $y(t_{i+1})$ is the exact solution at $t = t_{i+1}$ and y_{i+1} is the numerical solution at the same time assuming the $y_i = y(t_i)$ is exact.

For Euler

$$\begin{aligned}\tau_{i+1} &= y(t_i + h) - y(t_i) - hf(y_i, t_i) \\ &= y(t_i) + hy'(t_i) + \frac{h^2}{2}y''(t_i) + \ldots - y(t_i) - hy'(t_i) \\ &= \frac{h^2}{2}y''(t_i) + \ldots\end{aligned}$$

So local truncation error is $O(h^2)$.

(c) (*Unseen example. Similar seen on notes/homeworks*)

(i) Solution is easy: $y(t) = t^2/2 + t$.

(ii) First, y_i^h satisfies

$$y_{i+1}^h - y_i^h = 0$$

and has solution $y_i^h = A$ for a constant A. For the particular solution use ansatz given and substitute into Euler with $f = t_i + 1$

$$ah^2(i+1)^2 - ai^2h^2 + b(i+1)h - bih = h(ih + 1)$$

which simplifies to

$$ah^2(2i + 1) + bh = ih^2 + h$$

which means $a = 1/2$ and $b = 1 - ah = 1 - h/2$. Thus

$$y_i = A + t_i^2/2 + t_i(1 - h/2)$$

is the general solution, and using $y_0 = 0$ gives $A = 0$ so that

$$y_i = t_i^2/2 + t_i(1 - h/2)$$

is the discrete Euler solution.

(iii) At $t = t_N = Nh = 1$

$$E_N = y(1) - y_N = \frac{3}{2} - \left(\frac{1}{2} + 1(1 - h/2)\right) = h/2.$$

The global error is one order less than the local truncation error which is what we expect from the Euler method.

(d) (*Unseen example but nothing outrageous*).

(i) Now we have

$$\begin{aligned}
\tau_{i+1} &= y(t_i + h) - 4y(t_i) + 3y(t_{i-1}) + 2hf(y_{i-1}, t_{i-1}) \\
&= y(t_i) + hy'(t_i) + \frac{h^2}{2}y''(t_i) + \frac{h^3}{6}y'''(t_i)\dots \\
&\quad -4y(t_i) + 3(y(t_i) - hy'(t_i) + \frac{h^2}{2}y''(t_i) - \frac{h^3}{6}y'''(t_i)\dots) \\
&\quad +2h(y'(t_i) - hy''(t_i) + \frac{h^2}{2}y'''(t_i)\dots) \\
&= \frac{2}{3}h^3y'''(t_i) + \dots
\end{aligned}$$

and the local truncation error is $O(h^3)$.

(ii) For stability, we consider $h \to 0$ and seek solutions to the resulting homogeneous difference equation

$$y_{i+1} - 4y_i + 3y_{i-1} = 0$$

Using $y_i = Az^i$ gives $z^2 - 4z + 3 = 0$ whose roots are $z = 3$ and $z = 1$. Since one of the roots is s.t. $|z| > 1$ the scheme is unstable (root condition theorem).

End of examination.