university of groningen

Exam Numerical Methods

November 5th 2015 18.30-21.30

It is allowed to use a book (paper version only) and lecture notes, as well as a (graphical) pocket calculator. The use of electronic devices (tablet, laptop, mobile phone, etc.) is not allowed.

Always give a clear explanation of your answer. An answer without any computation will not be rewarded, so also copy the computations from your scratch paper.

Write your name and student number on each page!

Free points: $\boxed{10}$

Practica: $\boxed{18}$ For the 6 computer practica a maximum of 6*3=18 points can be earned.

1. For the intersection near $x = -3.18$ of $f_1(x) = \sin(x)$ and $f_2(x) = e^x$ we have to solve

$$\sin(x) = e^x$$

(a) $\boxed{3}$ Give the iteration formula when Newton's method is used for this problem.

(b) $\boxed{4}$ Someone uses the iterative method $x_{n+1} = \ln\{\sin(x_n)\}$, with $x_0 = -3$. Will this method be successful? Explain why.

(c) $\boxed{3}$ When the iterative method $x_{n+1} = x_n + \frac{1}{2}(\sin(x_n) - e^{x_n})$ is used, again with $x_0 = -3$, the first 4 iterations are given by

n	x_n
0	-3.00000000
1	-3.09545354
2	-3.14114215
3	-3.16298410
4	-3.17343886

Determine an error estimate for x_4

(d) $\boxed{4}$ Does the value $\alpha = \frac{1}{2}$ in (c) give optimal linear convergence? If not, determine the optimal value for α.

2. Consider the integral
$$I = \int_0^{\pi^2} \frac{\sin(\sqrt{x})}{\sqrt{4x}}\, dx$$

(a) $\boxed{4}$ (1) Suppose the Midpoint method is used, and one of the sub-intervals with length $\pi^2/8$ is centered around $x = \pi^2/4$. What is the Rectangular area for this segment?
(2) Give two reasons why the Midpoint method will not give optimal convergence.

By means of the substitution $u(x) = \frac{\sqrt{x}}{\pi}$, the integral is reformulated as $I = \int_0^1 \pi \sin(\pi u)\, du$.

With the Midpoint method the following results are obtained

n	$I(n)$
16	2.00321638
32	2.00080342
64	2.00020081
128	2.00005020
256	2.00001255

$I(n)$ is the approximation of the integral on a grid with n sub-intervals.

(b) $\boxed{4}$ (1) Compute the q-factor and explain that extrapolation is allowed.
(2) Compute an improved solution by means of extrapolation.

(c) $\boxed{6}$ (1) Give an error estimate for $I(256)$ based on $I(n)$ values.
(2) Also give the error estimate that follows from the theorem for the global error.
(3) Which estimate is the better one for this integral, theoretically? Explain why.

P.T.O.

3. Consider for x in [0 4] the differential equation $y'(x) = \alpha y^2 + (2x+1)$, with boundary condition $y(0) = 1$. The parameter α will be determined below.

   (a) $\boxed{5}$ Take $\alpha = -4$. Compute the solution with Heun's method (RK2) at $x = 1$ on a grid with $\Delta x = 0.5$.

   With a 2nd order method the solution for $\alpha = -1$ is determined on 2 grids with $\Delta x = 0.25$ and $\Delta x = 0.125$. The result at a selection of x locations is as follows

x_n	$\Delta x = 0.25$	$\Delta x = 0.125$
0.0	1.00000000	1.00000000
1.0	1.52303752	1.51971936
2.0	2.11450456	2.11867354
3.0	2.56470132	2.56793800
4.0	2.93866790	2.94122943

   (b) $\boxed{5}$ (1) Give an error estimate for the solution at $x = 4.0$ on the fine grid.
   (2) Which grid (i.e. how many segments) is probably required for an error of 1.0E-8?

   (c) $\boxed{4}$ (1) Give an improved solution (extrapolation) for the solution at $x = 4.0$.
   (2) What is the order of convergence of this extrapolation?

4. We will fit a (least-squares) straight line through the data

x_i	0	50	100
f_i	-1.5	-0.3	1.5

   (a) $\boxed{3}$ Apply a coordinate transformation, such that the $\hat{x}_i$ points are centered around $x = 0$ as follows: $\hat{x}_1 = -1$, $\hat{x}_2 = 0$, $\hat{x}_3 = 1$. Why is this necessary?

   (b) $\boxed{5}$ Determine the straight-line fit through the original data.

5. Consider the tri-diagonal system $A\vec{x} = \vec{b}$, with $A = \begin{pmatrix} 12 & 6 & 0 \\ -4 & 12 & 6 \\ 0 & 7 & -12 \end{pmatrix} \quad b = \begin{pmatrix} 108 \\ 36 \\ -18 \end{pmatrix}$

   (a) $\boxed{4}$ Determine the LU factorisation, with ones on the main diagonal of U.

   (b) $\boxed{4}$ When the Jacobi method is used, how many iterations are required to reduce the initial error in $\vec{x}_0 = (1\ 2\ 3)^T$ with a factor 100?

   (c) $\boxed{5}$ Compute $\vec{x}_1$, the result after 1 iteration with the initial vector $\vec{x}_0 = (0\ 0\ 0)^T$
   (1) when the Jacobi method is used.
   (2) when the SOR method is used with $\omega = 2$.

6. Consider for [0 4] the partial differential equation $\dfrac{\partial \phi}{\partial t} + \beta \dfrac{\partial \phi}{\partial x} = 10^{-4} \dfrac{\partial^2 \phi}{\partial x^2}$.

   For $\partial^2/\partial x^2$ the standard [1 -2 1]-formula is applied. The gradient term at the lhs can be discretised as $\dfrac{\partial \phi}{\partial x}$ at $x_i \approx \dfrac{\phi_i - \phi_{i-1}}{\Delta x}$ (method1) or $\dfrac{\partial \phi}{\partial x}$ at $x_i \approx \dfrac{\phi_{i+1} - \phi_i}{\Delta x}$ (method2)

   (a) $\boxed{4}$ For $\beta = 0$ the diffusion equation is obtained. When a time step of $\Delta t = 0.02$ is used for the explicit Euler method, which restriction follows for the number of segments in the domain [0 4]?

   (b) $\boxed{5}$ For $\beta > 0$ a different stability restriction is obtained for the explicit Euler method.
   (1) Determine this stability restriction when method 1 is used.
   (2) Is it more strict (for a fixed Δx) than for $\beta = 0$? Does it help if we use method 2?

Total: $\boxed{100}$

1 a) $f(x) = \sin(x) - e^x = 0$

$$x_{n+1} = x_n - \frac{(\sin(x_n) - e^{x_n})}{(\cos(x_n) - e^{x_n})}$$

b) $g(x) = \ln(\sin(x)) \Rightarrow g'(x) = \frac{1}{\sin(x)} \cdot \cos(x)$

$$\left.\begin{aligned} g'(-3) &\approx 7.02 \\ g'(-3.18) &\approx -26.02 \end{aligned}\right\} \Rightarrow |g'(p)| > 1$$

no convergence

Remark: $x_1 = \ln(\sin(-3))$ is complex number, no reason for break down

c) $K = \frac{x_4 - x_3}{x_3 - x_2} = 0.478655065$

$$\varepsilon_4 \leq \frac{K}{1-K} \left| (x_4 - x_3) \right| = 9.598\,6812\ E{-3}$$

d) $g(x) = x + \alpha(\sin(x) - e^x)$

$g'(x) = 1 + \alpha(\cos(x) - e^x)$

$g'(-3.18) = 1 + \alpha(-1.04084818) = 0$

$$\Rightarrow \alpha = \frac{-1}{-1.04084818} = 0.9607549$$

$\alpha = 1/2$ not optimal

2a) (1) $f(m)\cdot\Delta x = \dfrac{\sin\left(\sqrt{\pi^2/4}\right)}{\sqrt{\pi^2/4}} \cdot \dfrac{\pi^2}{8} = \pi/8$

(2) f singular at $x=0$ $\left(f''(0) \to \pm\infty \right)$

interval $[0, \pi^2] \to \Delta x$ not in full accuracy

b) $q = \dfrac{I_{64} - I_{128}}{I_{128} - I_{256}} = 4.0002656 \mathrel{\hat{=}} 2^2$

extrapolation o.k.

$T_2 = I_{256} + \dfrac{I_{256} - I_{128}}{3} = 2$

c) (2) $\varepsilon_{256} = \dfrac{1-0}{24}\left(\dfrac{1}{256}\right)^2 M$ $\rightarrow$ $1.971325 \cdot 10^{-5}$

$$\left.\begin{aligned} f(x) &= \pi \sin(\pi x) \\ f'(x) &= \pi^2 \cos(\pi x) \\ f''(x) &= -\pi^3 \sin(\pi x) \end{aligned}\right\} \Rightarrow M = \pi^3$$

(1) $\varepsilon_{256} = \left| \dfrac{I_{256} - I_{128}}{3} \right| = 1.2549999 \cdot 10^{-5}$

(3) Since M "belongs" to one x-location and not the full interval, E_n based on I_n more accurate

3 a) $f(x,y) = -4y^2 + 2x + 1$

$k_1 = 0.5\, f(0,1) = -3/2$
$k_2 = 0.5\, f(0.5,\ 1 + -3/2) = 0.5\, f(0.5, -\frac{1}{2}) = 1/2$

$y_1 = 1 + \frac{1}{2}\left(-\frac{3}{2} + \frac{1}{2}\right) = 1/2$

$k_1 = 0.5\, f(0.5,\ 1/2) = 1/2$
$k_2 = 0.5\, f(1,\ 1/2 + 1/2) = 0.5\, f(1,1) = -1/2$

$y_2 = 1/2 + 1/2\left(\frac{1}{2} + -1/2\right) = 1/2$

b) 1) $\varepsilon \approx \frac{1}{3}(2.94122943 - 2.93866790) = 8.5384333\,E{-4}$

2) 2^{nd} order $\Rightarrow$ grid $p \cdot \Delta x \rightsquigarrow$ error $p^2 \cdot \varepsilon$

$p^2 \cdot \varepsilon = 1.0\,E{-8} \Rightarrow p = 0.0034222$

grid $p \cdot \Delta x \Rightarrow \Delta x \approx 4.277\,E{-4}$
segments 11689

c) $\frac{4}{3}\, 2.94122943 - \frac{1}{3}\, 2.93866790 = 2.94208327$

$O(\Delta x^4)$

4. a) $\hat{x} = \frac{x-50}{50}$ prevents unbalanced matrix

$\Leftrightarrow$ inaccurate solution

b) $M_0 = 3$ $\quad F_0 = -0.3$

$M_1 = 0$ $\quad F_1 = 1.5 + 0 + 1.5 = 3$

$M_2 = 2$

$3a + 0b = -0.3 \qquad a = -0.1$

$0a + 2b = 3 \qquad b = 1.5$

$y = -0.1 + 1.5\hat{x} = -0.1 + 1.5\left(\frac{x-50}{50}\right)$

$= -1.6 + 0.03x$

5a)

$$\begin{pmatrix} 12 & 6 & 0 \\ -4 & 12 & 6 \\ 0 & 7 & -12 \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 & 0 \\ -4 & \alpha_2 & 0 \\ 0 & 7 & \alpha_3 \end{pmatrix} \begin{pmatrix} 1 & \gamma_1 & 0 \\ 0 & 1 & \gamma_2 \\ 0 & 0 & 1 \end{pmatrix}$$

$\alpha_1 = 12 \qquad \gamma_1 = 6/12 = 1/2$

$\alpha_2 = 12 - (-4) \cdot \frac{1}{2} = 14 \qquad \gamma_2 = 6/14 = 3/7$

$\alpha_3 = -12 - 7 \cdot \frac{3}{7} = -15$

$$\begin{pmatrix} 12 & 0 & 0 \\ -4 & 14 & 0 \\ 0 & 7 & -15 \end{pmatrix} \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1 & 3/7 \\ 0 & 0 & 1 \end{pmatrix}$$

b) row1 $6/12$

row2 $10/12 \longrightarrow \mu = 5/6$

row3 $7/12$

$\left(\frac{5}{6}\right)^{25} = 0.01048\ldots$

$\left(\frac{5}{6}\right)^{26} = 0.008735\ldots$

$\Longrightarrow$ 26 iterations

c) Jacobi $x_1 = \frac{1}{12}(108 - 6.0) = 9$

$x_2 = \frac{1}{12}(36 - (-4).0 - 6.0) = 3$

$x_3 = \frac{1}{-12}(-18 - 7.0) = 3/2$

SOR $\hat{x}_1 = \frac{1}{12}(108 - 6.0) = 9 \qquad x_1 = 2\hat{x}_1 - x_1 = 18$

$\hat{x}_2 = \frac{1}{12}(36 - (-4).18 - 6.0) = 9 \qquad x_2 = 2\hat{x}_2 - x_2 = 18$

$\hat{x}_3 = \frac{1}{-12}(-18 - 7.18) = 12 \qquad x_3 = 2\hat{x}_3 - x_3 = 24$

6a) $R = \dfrac{\Delta t\, D}{\Delta x^2} \leq \dfrac{1}{2}$

$$\frac{0.02 \; 10^{-4}}{\Delta x^2} \leq 1/2 \implies \Delta x \geq 2\; 10^{-3}$$

$$\implies \# \text{ segments} \leq \frac{4}{2\; 10^{-3}} = 2000$$

b) $$\phi_i^{n+1} = \phi_i^n + \Delta t\left(10^{-4}\, \frac{\phi_{i+1}^n - 2\phi_i^n + \phi_{i-1}^n}{\Delta x^2} - \beta\, \frac{\phi_i^n - \phi_{i-1}^n}{\Delta x}\right)$$

$$\text{OR } -\beta\, \frac{\phi_{i+1}^n - \phi_i^n}{\Delta x}$$

positive coeff meth 1: ϕ_{i+1} $\dfrac{\Delta t\, 10^{-4}}{\Delta x^2} \geq 0$ no problem

ϕ_{i-1} $\dfrac{\Delta t\, 10^{-4}}{\Delta x^2} + \dfrac{\beta\, \Delta t}{\Delta x}$ no problem

ϕ_i $1 - \dfrac{2\Delta t\, 10^{-4}}{\Delta x^2} - \dfrac{\beta \Delta t}{\Delta x} \geq 0$

$\hookrightarrow \Delta t$ has to be smaller if $\beta > 0$

meth 2: ϕ_{i+1} $\dfrac{\Delta t\, 10^{-4}}{\Delta x^2} - \dfrac{\beta \Delta t}{\Delta x} \geq 0$

$\hookrightarrow \Delta t$ has to be smaller

ϕ_{i-1} $\dfrac{\Delta t\, 10^{-4}}{\Delta x^2} + 0 \geq 0$ no problem

ϕ_i $1 - \dfrac{2\Delta t\, 10^{-4}}{\Delta x^2} + \dfrac{\beta\, \Delta t}{\Delta x} \geq 0$

$\hookrightarrow \Delta t$ can be larger if $\beta > 0$

method 1: Δt smaller $\to$ so more strict

method 2: depends on values of Δx and β if it helps