UNIVERSITY OF OSLO

Faculty of Mathematics and Natural Sciences

Examination in MAT3110/4110 — Introduction to numerical analysis

Day of examination: Tuesday, November 29th., 2022

Examination hours: 15:00 – 19:00

This problem set consists of 7 pages.

Appendices: Error formulas.

Permitted aids: None

Please make sure that your copy of the problem set is complete before you attempt to answer anything.

Problem 1

Let

$$u(x) = x^2 \quad v(x) = ax + b, \quad x \in [0, 1],$$

where a and b are constants. Determine a and b such that the L^2 distance $\|u - v\|_{L^2(0,1)}$ is minimal. Draw the graphs of u and v.

Løsningsforslag: We have

$$\begin{aligned}(u(x) - v(x))^2 &= (x^2 - (ax + b))^2 \\ &= x^4 - 2x^2(ax + b) + (ax + b)^2 \\ &= x^4 - 2ax^3 - 2bx^2 + a^2x^2 + 2abx + b^2.\end{aligned}$$

Hence

$$\begin{aligned}\|u - v\|_2^2 &= \int_0^1 x^4 - 2ax^3 - 2bx^2 + a^2x^2 + 2abx + b^2\,dx \\ &= \frac{1}{5} - \frac{1}{2}a - \frac{2}{3}b + \frac{1}{3}a^2 + ab + b^2 =: g(a, b).\end{aligned}$$

$$\frac{\partial g}{\partial a} = -\frac{1}{2} + \frac{2}{3}a + b, \quad \frac{\partial g}{\partial b} = -\frac{2}{3} + a + 2b.$$

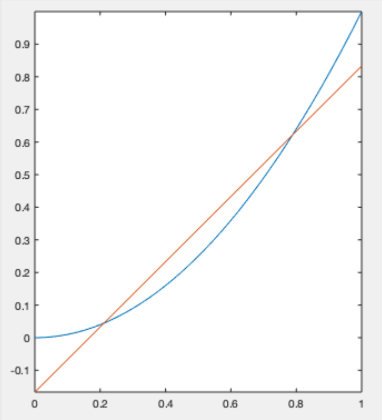
Solving for these to be zero,

$$\begin{aligned}\frac{2}{3}a + b &= \frac{1}{2} \\ a + 2b &= \frac{2}{3}\end{aligned}$$

gives

$$a = 1 \quad b = -\frac{1}{6}.$$

(Continued on page 2.)



Problem 2

Let $f : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(y) = \frac{c}{1+y^2}, \qquad \text{where } c \text{ is a positive constant.}$$

2a

Show that f is a contraction on $\mathbb{R}$ if $c < c_0$, where

$$c_0 = \frac{8\sqrt{3}}{9} \approx 1.54.$$

You don't have to show the "$\approx$".

Løsningsforslag: We need to show that $|f'(y)| < 1$.

$$f'(y) = -2c\frac{y}{(1+y^2)^2}.$$

To find the maximum we differentiate once more

$$\frac{d}{dy}\frac{y}{(1+y^2)^2} = \frac{(1+y^2)^2 - 4y^2(1+y^2)}{(1+y^2)^4} = \frac{(1-y^2)^2 - 4y^4}{(1+y^2)^4}.$$

Therefore $|f'(y)|$ will have max if $1 - y^2 = 2y^2$, i.e., $y = \pm 1/\sqrt{3}$. Therefore

$$\left|f'(y)\right| \le c\frac{18}{16\sqrt{3}} = \frac{c}{c_0}.$$

2b

Assume that $c = c_0/2$, we wish to find an approximation to the solution of $y = f(y)$, accurate to 17 decimal places. Explain how you could do this

(Continued on page 3.)

without using the "*abcd*-formula".

Løsningsforslag: f is a contraction with $L = 1/2$, we choose $y_0 = 0$, in a simple iteration; $y_{n+1} = f(y_n)$, which gives $0 < y_1 = c < 1$. We have the error formula

$$|y_n - y| \leq \frac{L^n}{1-L} |y_1 - y_0| = 2^{-(n-1)}.$$

We must choose n such that

$$2^{-(n-1)} \leq \frac{1}{2} \cdot 10^{-17} \quad \text{i.e.,} \quad 2^n \geq 10^{17},$$

which is satisfied for $n > 56.5$.

2c

Consider the initial value problem

$$y'(t) = g(y), \quad y(0) = 0, \quad g(y) = \frac{1}{1+y^2}. \tag{1}$$

We wish to find an approximation to $y(t)$ for $t \in [0, 1]$, and set $\Delta t = 1/N$ for some integer $N \geq 1$. We use the implicit scheme

$$y_{i+1} = y_i + \frac{\Delta t}{2} \left(g(y_i) + g(y_{i+1}) \right), \quad i = 0, 1, \ldots, N-1.$$

Using pseudo-code explain how you would implement this scheme. (Make sure your implementation is well-defined).

Løsningsforslag: Here is a Matlab version

```
N=50; T=1; dt=T/N;
y0=0;
g=@(y) 1./(1+y.^2);
y=zeros(1,N+1); t=linspace(0,T,N+1);
y(1)=y0;
tol=0.5*1e-16;
for i=2:N+1
    z=y(i-1)+0.5*dt*g(y(i-1));
    yn=y(i-1)+dt*g(y(i-1));  % initial guess
    ynn=z+0.5*dt*g(yn);
    n=1;
    while (abs(yn-ynn)>tol)&&(n<57)
        % fixpoint iteration to find y_{i}
        % know that (from a) 57 iterations suffice
        n=n+1;
        yn=ynn;
        ynn=z+0.5*dt*g(yn);
    end
    y(i)=ynn;
end
```

(Continued on page 4.)

2d

The truncation error τ_i is defined by

$$\tau_i = \frac{y(t_{i+1}) - y(t_i)}{\Delta t} - \frac{1}{2}\big(g(y(t_{i+1})) + g(y(t_i))\big)$$

where $y(t)$ is the exact solution of (1) and $t_i = i\Delta t$. Show that the truncation error satisfies

$$|\tau_i| \leq \frac{\Delta t^2}{12} M,$$

where

$$M = \left\| g'' g^2 + (g')^2 g \right\|_\infty .$$

What is the convergence order of this method?

Løsningsforslag: We get that

$$\tau_i = \frac{1}{\Delta t}\Big(\int_{t_i}^{t_{i+1}} g(y(t))\,dt - \frac{\Delta t}{2}\left(g(y(t_i)) + g(y(t_{i+1}))\right)\Big),$$

We recognise the error of the trapezoidal rule as

$$\int_{t_i}^{t_{i+1}} g(y(t))\,dt - \frac{\Delta t}{2}\left(g(y(t_i)) + g(y(t_{i+1}))\right) \leq \frac{\Delta t^3}{12} \max_t \left| \frac{d^2}{dt^2} g(y(t)) \right|$$

Now we have

$$\frac{d}{dt} g(y(t)) = g'g,$$
$$\frac{d^2}{dt^2} g(y(t)) = g''g^2 + (g')^2 g.$$

Since we know that the error of the method is bounded by $K \max_i |\tau_i| = \mathcal{O}(\Delta t^2)$ the method is second order.

Problem 3

Let L^n be an $n \times n$ lower triangular non-singular matrix.

3a

Show that $\left(L^2\right)^{-1}$ is lower triangular. If L^{n+1} is a general non-singular $(n+1) \times (n+1)$ lower triangular matrix written as

$$L^{n+1} = \begin{pmatrix} L^n & \mathbf{0} \\ \boldsymbol{r}^\top & \alpha \end{pmatrix},$$

where L_n is an $n \times n$ non-singular lower triangular matrix, $\boldsymbol{r}$ is an $n \times 1$ vector and $\alpha \neq 0$ is a number. Show that the inverse $\left(L^{n+1}\right)^{-1}$ is given by

$$\left(L^{n+1}\right)^{-1} = \begin{pmatrix} (L^n)^{-1} & \mathbf{0} \\ -\frac{1}{\alpha}\boldsymbol{r}^\top (L^n)^{-1} & \frac{1}{\alpha} \end{pmatrix}, \tag{2}$$

and conclude that $(L^n)^{-1}$ is lower triangular for any $n \geq 0$.

(Continued on page 5.)

Løsningsforslag: We compute

$$L^{n+1}\left(L^{n+1}\right)^{-1} = \begin{pmatrix} L^n & \mathbf{0} \\ \boldsymbol{r}^\top & \alpha \end{pmatrix} \begin{pmatrix} (L^n)^{-1} & \mathbf{0} \\ -\frac{1}{\alpha}\boldsymbol{r}^\top (L^n)^{-1} & \frac{1}{\alpha} \end{pmatrix} = \begin{pmatrix} I_n & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{pmatrix}.$$

If $(L^2)^{-1}$ is lower triangular, then $(L^3)^{-1}$ is lower triangular etc.

3b

Let L^k be as above, and $\boldsymbol{y} \in \mathbb{R}^k$, $k \geq 2$. How many operations (additions, subtractions, multiplications or divisions) do you need to compute the product $\boldsymbol{y}^\top L^k$ as efficiently as possible?

Løsningsforslag: Since L^k is lower triangular, $L^k_{ji} = 0$ for $i > j$. We have that for $i = 1, \ldots, k$

$$\left(\boldsymbol{y}^\top L^k\right)_i = \sum_{j=1}^{k} y_j L^k_{ji} = \sum_{j=i}^{k} y_j L^k_{ji}.$$

For each i we have $k - i + 1$ multiplications and $k - i$ additions. This gives a total of

$$\sum_{i=1}^{k} 2k - 2i + 1 = 2k^2 - 2\frac{k(k+1)}{2} + k = k^2 \quad \text{operations.}$$

3c

Based on (2), formulate an algorithm which, given a non-singular lower triangular $n \times n$ matrix A, computes its inverse B.

Løsningsforslag: In Matlab

```
for i=1:n
   B(i,i)=1/A(i,i);
   for j=1:i-1
      B(i,j:i-1)=-B(i,i)*A(i,j:i-1)*B(j:i-1,j);
   end
end
```

with inner loop spelled out

```
for i=1:n
   B(i,i)=1/A(i,i);
   for j=1:i-1
      B(i,j)=0;
      for k=j:i-1
         B(i,j)=B(i,j)+A(i,k)*B(k,j);
      end
      B(i,j)=-B(i,j)*B(i,i);
   end
end
```

(Continued on page 6.)

Problem 4

Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of independent random variables, uniformly distributed in the interval $[0, 1]$.

4a

Explain why

$$\lim_{M\to\infty} \frac{4}{M}\sum_{i=1}^{M}\sqrt{1-X_i^2} = \pi \quad \text{almost surely.}$$

Løsningsforslag: We have that

$$\frac{4}{M}\sum_{i=1}^{M}\sqrt{1-X_i^2}$$

is the Monte-Carlo approximation to the integral

$$4\int_0^1 \sqrt{1-x^2}\,dx = \pi.$$

We know that Monte-Carlo approximations converge almost surely.

4b

How large must you choose M if you want to be 90% sure that

$$\left|\frac{4}{M}\sum_{i=1}^{M}\sqrt{1-X_i^2} - \pi\right| < 0.05?$$

Løsningsforslag: We have $\max f - \min f = 4$. We choose M so that

$$\frac{4}{0.05^2 M} \leq 0.1 \quad \Rightarrow \quad M \geq 16\,000.$$

THE END

(Continued on page 7.)

Some useful error formulas

Simple iteration, finding $\xi = f(\xi)$:

$$|x_n - \xi| \leq \frac{L^n}{1-L}\left|x_0 - x_1\right|, \quad 1 > L \geq \left|f'(x)\right|.$$

$$\left.\begin{aligned} &\text{Trapezoid rule:} && \left|E_1(f) - \int_a^b f(x)\,dx\right| \leq \frac{(b-a)^3}{12}M_2, \\ &\text{Simpson's rule:} && \left|E_2(f) - \int_a^b f(x)\,dx\right| \leq \frac{(b-a)^5}{2880}M_4, \end{aligned}\right\} \quad M_k = \max\left|f^{(k)}\right|.$$

Monte Carlo integration:

$$\mathcal{E}_M(f) \leq \frac{\left(\max_{\boldsymbol{x}\in[0,1]^d}\{f(\boldsymbol{x})\} - \min_{\boldsymbol{x}\in[0,1]^d}\{f(\boldsymbol{x})\}\right)}{2\sqrt{M}},$$

and

$$\text{Prob}\left(\left|I_M(f) - \int_{[0,1]^d} f(\boldsymbol{x})\,d\boldsymbol{x}\right| \geq \varepsilon\right) \leq \frac{\left(\max_{\boldsymbol{x}\in[0,1]^d}\{f(\boldsymbol{x})\} - \min_{\boldsymbol{x}\in[0,1]^d}\{f(\boldsymbol{x})\}\right)^2}{4\varepsilon^2 M}.$$

Polynomial interpolation:

$$|p_n(x) - f(x)| \leq \frac{M_{n+1}}{(n+1)!}\prod_{i=0}^{n}|x - x_i|.$$

Error for a one-step method; $y_{n+1} = y_n + \Delta t\Phi(y_n, \cdots)$ for the ODE $y' = f(y)$:

$$|e_n| \leq \frac{\max_i|\tau_i|}{L_\Phi}(e^{L_\Phi t_n} - 1).$$