Numerical Analysis : Recitations

Aakash Jog

2015-16

Contents

---

1 Instructor Information

Ron Levie

E-mail: ronlevie@post.tau.ac.il

2 Errors

Definition 1 (Error)**.** The absolute error in representation is defined as

$$e_x = x - \widetilde{x}$$

The relative error in representation is defined as

$$\delta = \frac{x - \widetilde{x}}{x}$$

Recitation 1 – Exercise 1.
The dimensions of a field are measured. The length is measured to be $\widetilde{x} = 800\text{m}$, with an absolute error bounded by 16. The width is measured to be $\widetilde{y} = 30\text{m}$, with an absolute error e_y, such that $|e_y| \leq 6$.

1. Find the approximate bounds for $|\delta_x|$ and $|\delta_y|$.

2. Find the bounds on the absolute error in the calculated area of the field.

Recitation 1 – Solution 1.
1.

$$\begin{aligned} |\delta_x| &= \frac{|e_x|}{|x|} \\ &\leq \frac{16}{|x|} \\ &\approx \frac{16}{800} \\ &= 0.02 \\ \therefore |\delta_x| &\leq 0.02 \end{aligned}$$

$$
\begin{aligned}
|\delta_y| &= \frac{|e_y|}{|y|} \\
&\leq \frac{6}{|y|} \\
&\approx \frac{6}{300} \\
&= 0.02 \\
\therefore |\delta_y| &\leq 0.02
\end{aligned}
$$

2. The measured area of the field is

$$
\begin{aligned}
\widetilde{A} &= \widetilde{x}\widetilde{y} \\
&= 800 \cdot 300 \\
&= 240000
\end{aligned}
$$

The maximum area of the field is

$$
\begin{aligned}
A_{\text{max}} &= (\widetilde{x} + e_{x\text{max}})(\widetilde{y} + e_{y_{\text{max}}}) \\
&= (800 + 16)(300 + 6) \\
&= 249696
\end{aligned}
$$

The maximum area of the field is

$$
\begin{aligned}
A_{\text{min}} &= (\widetilde{x} + e_{x\text{min}})(\widetilde{y} + e_{y_{\text{min}}}) \\
&= (800 - 16)(300 - 6) \\
&= 230496
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
|e_{xy}| &\leq (A_{\text{max}} - A_{\text{min}}) \\
&\leq 9696
\end{aligned}
$$

3.

$$
\begin{aligned}
|\delta_{xy}| &= \frac{|e_{xy}}{|xy|} \\
&\leq \frac{9696}{|xy|} \\
&\leq \frac{9696}{230496} \\
&\approx 0.042
\end{aligned}
$$

2.1 Propagation of Error

Recitation 1 – Exercise 2.

Let $\widetilde{x}$, $\widetilde{y}$ be approximations of x, y.

1. Find a formula for the absolute error in $x + y$ in terms of e_x and e_y.

2. Find a formula for δ_{x+y}, δ_{x-y} in terms of δ_x, δ_y, x, y.

3. Let $\delta = \max\{\delta_x, \delta_y\}$. Assuming $x, y > 0$, show

$$|\delta_{x-y}| \leq \frac{x+y}{|x-y|}\delta$$

Recitation 1 – Solution 2.

1.

$$\begin{aligned} e_{x+y} &= (x+y) - (\widetilde{x} + \widetilde{y}) \\ &= (x - \widetilde{x}) + (y - \widetilde{y}) \\ &= e_x + e_y \end{aligned}$$

2.

$$\begin{aligned} \delta_{x+y} &= \frac{e_{x+y}}{x+y} \\ &= \frac{e_x + e_y}{x+y} \\ &= \frac{x\delta_x + y\delta_y}{x+y} \end{aligned}$$

Similarly,

$$\begin{aligned} \delta_{x-y} &= \frac{e_{x-y}}{x-y} \\ &= \frac{e_x - e_y}{x-y} \\ &= \frac{x\delta_x - y\delta_y}{x-y} \end{aligned}$$

3.

$$\begin{aligned} |\delta_{x-y}| &= \left|\frac{x\delta_x - y\delta_y}{x-y}\right| \\ &\le \frac{|x||\delta_x| + |y||\delta_y|}{|x-y|} \\ &\le \frac{x\delta + y\delta}{|x-y|} \\ &= \frac{x+y}{|x-y|}\delta \end{aligned}$$

Recitation 1 – Exercise 3.
Find a formula for δ_{xy}, in terms of x, y, δ_x, δ_y.

Recitation 1 – Solution 3.

$$\begin{aligned} \delta_a &= \frac{a - \widetilde{a}}{a} \\ \therefore \widetilde{a} &= a(1 - \delta_a) \end{aligned}$$

Therefore,

$$\begin{aligned} \widetilde{x}\widetilde{y} &= (x(1-\delta_x))\left(y(1-\delta_y)\right) \\ &= xy(1 - \delta_x - \delta_y + \delta_x\delta_y) \end{aligned}$$

Also,

$$\widetilde{x}\widetilde{y} = xy(1 - \delta_{xy})$$

Therefore,

$$\delta_{xy} = \delta_x + \delta_y - \delta_x\delta_y$$

3 Interpolation by Polynomials

Theorem 1 (Existence and Uniqueness Theorem)**.** *There exists a unique polynomial $p_n(x)$ which approximates $f(x)$ between the sample points, i.e.*

$$|e_n(x)| = |f(x) - p_n(x)|$$

Recitation 2 – Exercise 1.
Find the interpolation polynomial for the data

x_i	$f(x_i)$
1	3
2	2
4	1

Recitation 2 – Solution 1.
Let

$$p(x) = a_0 + a_1 x + a_2 x^2$$

be the required interpolation polynomial.
Therefore,

$$\begin{aligned}
&p(1) = 3 \\
&\therefore 3 = a_0 + a_1 \cdot 1 + a_2 \cdot 1^2 \\
&p(2) = 2 \\
&\therefore 2 = a_0 + a_1 \cdot 2 + a_2 \cdot 2^2 \\
&p(4) = 1 \\
&\therefore 1 = a_0 + a_1 \cdot 4 + a_2 \cdot 4^2
\end{aligned}$$

Therefore,

$$\begin{pmatrix} 1 & 1 & 1^2 \\ 1 & 2 & 2^2 \\ 1 & 4 & 4^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

Therefore, solving,

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \frac{13}{3} \\ -\frac{3}{2} \\ \frac{1}{6} \end{pmatrix}$$

Therefore,

$$p(x) = \frac{13}{3} - \frac{3}{2}x + \frac{1}{6}x^2$$

Definition 2. Let the sample points be $x_0, \ldots, x_{n-1}$.
Lagrange polynomials are n polynomials of degree $n-1$, each of which is 0 at all sample points, except one, at which it is 1.

$$l_i(x) = \frac{(x-x_0)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_{n-1})}{(x_i-x_0)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_{n-1})}$$

Recitation 2 – Exercise 2.
Find the interpolation polynomial for the data

x_i	$f(x_i)$
1	3
2	2
4	1

using Lagrange polynomials.

Recitation 2 – Solution 2.
Let

$$x_0 = 1$$
$$x_1 = 2$$
$$x_3 = 4$$

Therefore,

$$\begin{aligned} l_0(x) &= \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} \\ &= \frac{(x-2)(x-4)}{(1-2)(1-4)} \\ l_1(x) &= \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} \\ &= \frac{(x-1)(x-4)}{(2-1)(2-4)} \\ l_2(x) &= \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} \\ &= \frac{(x-1)(x-2)}{(4-1)(4-2)} \end{aligned}$$

Therefore,

$$\begin{aligned} p(x) &= f(x_0)l_0(x) + f(x_1)l_1(x) + f(x_2)l_2(x) \\ &= 3l_0(x) + 2l_1(x) + l_2(x) \end{aligned}$$

Recitation 3 – Exercise 1.
Given

x_i	$f(x_i)$
0	0
$\frac{\pi}{4}$	$\frac{\sqrt{2}}{2}$
$\frac{\pi}{2}$	1

Find the interpolating polynomial in Newton's form.

Recitation 3 – Solution 1.
The interpolating polynomial is

$$p_2(x) = A_0 + A_1(x - x_0) + A_2(x - x_0)(x - x_1)$$

where

$$A_k = f[x_0, \ldots, x_k]$$

Therefore,

$$\begin{aligned} \therefore f[0] &= f(0) \\ &= 0 \\ \therefore f\left[\frac{\pi}{4}\right] &= f\left(\frac{\pi}{4}\right) \\ &= \frac{\sqrt{2}}{2} \\ \therefore f\left[\frac{\pi}{2}\right] &= f\left(\frac{\pi}{2}\right) \\ &= 1 \end{aligned}$$

Therefore,

$$f\left[0, \frac{\pi}{4}\right] = \frac{f\left[\frac{\pi}{4}\right] - f[0]}{\frac{\pi}{4} - 0}$$
$$= \frac{\frac{\sqrt{2}}{2} - 0}{\frac{\pi}{4}}$$
$$f\left[\frac{\pi}{4}, \frac{\pi}{2}\right] = \frac{f\left[\frac{\pi}{2}\right] - f\left[\frac{\pi}{4}\right]}{\frac{\pi}{2} - \frac{\pi}{4}}$$

Therefore,

$$f\left[0, \frac{\pi}{4}, \frac{\pi}{2}\right] = \frac{f\left[\frac{\pi}{4}, \frac{\pi}{2}\right] - f\left[0, \frac{\pi}{4}\right]}{\frac{\pi}{2} - 0}$$
$$= \frac{8(1 - \sqrt{2})}{\pi^2}$$

Therefore,

$$A_0 = 0$$
$$A_1 = \frac{2\sqrt{2}}{\pi}$$
$$A_2 = \frac{8(1 - \sqrt{2})}{\pi^2}$$

Therefore,

$$p_2(x) = \frac{2\sqrt{2}}{\pi}x + \frac{8(1 - \sqrt{2})}{\pi^2}(x)\left(x - \frac{\pi}{4}\right)$$

Recitation 3 – Exercise 2.
$\sin\left(\frac{\pi}{3}\right)$ was approximated using Newton's method, at sample points 0, $\frac{\pi}{4}$, $\frac{\pi}{2}$, to be

$$p_2\left(\frac{\pi}{3}\right) = \frac{2\sqrt{2}}{3} + \frac{8(1 - \sqrt{2})}{36}$$
$$= 0.8507$$

Find the bounds on the error in this approximation.

Recitation 3 – Solution 2.

$$\begin{aligned}|e_n(x)| &= |f(x) - p_n(x)| \\ &\leq \left| \frac{f^{(n+1)}(c)}{(n+1)!} \prod_{j=0}^{n} (x - x_j) \right|\end{aligned}$$

where $c \in [\min\{x_0, \ldots, x_n, x\}, \max\{x_0, \ldots, x_n, x\}]$.
Therefore,

$$\begin{aligned}|e_2(x)| &\leq \left| \frac{\sin^{(3)}(c)}{3!} \prod_{j=0}^{3} (x - x_j) \right| \\ \therefore \left| e_2\left(\frac{\pi}{3}\right) \right| &\leq \left| \frac{\sin^{(3)}(c)}{3!} \prod_{j=0}^{3} \left(\frac{\pi}{3} - x_j\right) \right| \\ &\leq \left| \frac{\sin^{(3)}(c)}{3!} \left(\frac{\pi}{3} - 0\right)\left(\frac{\pi}{3} - \frac{\pi}{4}\right)\left(\frac{\pi}{3} - \frac{\pi}{2}\right) \right| \\ &\leq \left| \frac{-\cos(c)}{6} \frac{\pi^3}{(3)(12)(6)} \right| \\ &\leq \left| \frac{-\cos(c)\pi^3}{1296} \right|\end{aligned}$$

Therefore, as $|\cos(c)|$ is bounded by 0 and 1,

$$\begin{aligned}\left| e_2\left(\frac{\pi}{3}\right) \right| &\leq \left| \frac{\pi^3}{1296} \right| \\ &< 0.0242\end{aligned}$$

Recitation 4 – Exercise 1.

Find Hermite's interpolating polynomial for the sample points 1, 1, e, for the function $f(x) = \ln(x)$.

Recitation 4 – Solution 1.

$$f[x_0, \ldots, x_k] = \begin{cases} \frac{f[x_1,\ldots,x_k] - f[x_0,\ldots,x_{k-1}]}{x_k - x_0} & ; \quad x_k \neq x_0 \\ \frac{f^{(k)}(x_0)}{k!} & ; \quad x_k = x_0 \end{cases}$$

Therefore,

$$\begin{aligned}f[1] &= 0 \\ f[1] &= 0 \\ f[e] &= 1\end{aligned}$$

Therefore,

$$
\begin{aligned}
f[1,1] &= \frac{f'(1)}{1!} \\
&= \left.\frac{1}{x}\right|_{x=1} \\
&= 1 \\
f[1,e] &= \frac{1-0}{e-1} \\
&= \frac{1}{e-1}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
f[1,1,e] &= \frac{\frac{1}{e-1}-1}{e-1} \\
&= \frac{2-e}{(e-1)^2}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
p_2(x) &= f[1] + f[e](x-1) + \frac{2-e}{(e-1)^2}(x-1)(x-1) \\
&= 0 + 1(x-1) + \frac{2-e}{(e-1)^2}(x-1)^2
\end{aligned}
$$

4 Fixed Point Iterations and Root Finding

Recitation 5 – Exercise 1.
Show that

$$
\begin{aligned}
e_n &= \alpha - x_n \\
&\approx -\frac{f(x_n)}{f'(x_n)}
\end{aligned}
$$

Recitation 5 – Solution 1.
By Lagrange's Mean Value Theorem, $\exists c \in (a,b)$, such that

$$
f'(c) = \frac{f(b)-f(a)}{b-a}
$$

Let

$$\begin{aligned} b &= x_n \\ a &= \alpha \end{aligned}$$

Therefore,

$$\frac{f(x_n) - f(\alpha)}{x_n - \alpha} = f'(c_n)$$

where $c_n \in (\alpha, x_n)$.
Therefore,

$$\begin{aligned} -e_n &= x_n - \alpha \\ &= \frac{f(x_n)}{f'(c_n)} \end{aligned}$$

Therefore, as $\lim_{n\to\infty} x_n = 2$, for $n \to \infty$,

$$c_n = x_n$$

Therefore,

$$e_n = -\frac{f(x_n)}{f'(x_n)}$$

Recitation 5 – Exercise 2.
Let

$$f(x) = e^{-x} - \frac{1}{2}$$

1. Show that f has a root in $[0, 1]$.

2. Show that Newton's method converges to the root α of f, and that α is unique.

Recitation 5 – Solution 2.

1.

$$\begin{aligned} f(0) &= e^0 - \frac{1}{2} \\ &= \frac{1}{2} \\ f(1) &= \frac{1}{e} - \frac{1}{2} \\ &< \frac{1}{2.7} - \frac{1}{2} \\ &< 0 \end{aligned}$$

Therefore, by the intermediate value theorem, $\exists \alpha$ such that $f(\alpha) = 0$. Hence, f has a root in $[0, 1]$.

2.

$$\begin{aligned} g(x) &= x - \frac{f(x)}{f'(x)} \\ &= x + \frac{e^{-x} - \frac{1}{2}}{e^{-x}} \\ &= x + 1 - \frac{1}{2}e^x \end{aligned}$$

Therefore,

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

Therefore, as the extrema of g are in $[0, 1]$, $g : [0, 1] \to [0, 1]$.
Similarly, $g'(x)$ is decreasing.
Hence, by the fixed point theorem, as $\lim_{n\to\infty} x_n = \alpha$, α is unique.

Theorem 2. *For the method $x_{n+1} = g(x_n)$, if $\alpha = g(\alpha)$ and $|g'(\alpha)| < 1$, then $\exists$ a neighbourhood $(\alpha - \varepsilon, \alpha + \varepsilon) = \mathcal{N}$, of α, such that for any $x_0 \in \mathcal{N}$,*

$$\lim_{n\to\infty} x_n = \alpha$$

Definition 3 (Rate of convergence)**.** For a converging iterative method, p is called the rate of convergence if $\exists c \neq 0$, such that

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^p} = c$$

which is equivalent to

$$|e_{n+1}| = (c + \mathrm{o}(1))\,|e_n|^p$$

where o(1) is a sequence whose limit is 0.

Theorem 3. *Let $p \in \mathbb{N}$. If $g(\alpha) = \alpha$, and for $1 \leq k < p$,*

$$g^{(k)}(\alpha) = 0$$

and

$$g^{(p)}(\alpha) \neq 0$$

then ,the rate of convergence if p.

Recitation 6 – Exercise 1.
Consider the following iteration for calculating $\alpha = r^{\frac{1}{3}}$, where $r > 0$.

$$g(x) = Ax + Brx^{-2} + Cr^2x^{-5}$$
$$x_{n+1} = g(x_n)$$

where $A, B, C \in \mathbb{R}$.

1. Find A, B, C, such that the method converges to $r^{\frac{1}{3}}$ with maximum rate of convergence.

2. What is the rate of convergence?

Recitation 6 – Solution 1.
1. For the method to converge to $r^{\frac{1}{3}}$, $r^{\frac{1}{3}}$ must be a fixed point of g. Therefore,

$$g\left(r^{\frac{1}{3}}\right) = Ar^{\frac{1}{3}} + Brr^{-\frac{2}{3}} + Cr^2r^{-\frac{5}{3}}$$
$$\therefore r^{\frac{1}{3}} = Ar^{\frac{1}{3}} + Br^{\frac{1}{3}} + Cr^{\frac{1}{3}}$$

For the rate of convergence to be maximum,

$$g'\left(r^{\frac{1}{3}}\right) = 0$$
$$\therefore A - 2B - 5C = 0$$

Also, for the rate of convergence to maximum,

$$g''\left(r^{\frac{1}{3}}\right) = 0$$
$$\therefore 6B + 30C = 0$$

Therefore, solving,

$$A = \frac{5}{9}$$
$$B = \frac{5}{9}$$
$$C = -\frac{1}{9}$$

Therefore, the rate of convergence is greater than 2.

2.

$$g'''(x) = -24Brx^{-5} - 210Cr^2x^{-8}$$

Therefore,

$$\begin{aligned} g'''\left(r^{\frac{1}{3}}\right) &= \frac{40}{3}e^{-\frac{2}{3}} \\ &\neq 0 \end{aligned}$$

Therefore the rate of convergence is 3.

5 LU Decomposition and Norms

Recitation 7 – Exercise 1.
Let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \\ \frac{1}{2} & 0 & 3 \end{pmatrix}$$

1. Find the PLU decomposition, i.e. the LU decomposition with pivoting, of A.

2. Represent P as a permutation vector.

3. Use the decomposition to solve $Ax = b$ for

$$b = \begin{pmatrix} 5 \\ 4 \\ 7 \end{pmatrix}$$

Recitation 7 – Solution 1.

1.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \\ 0.5 & 0 & 3 \end{pmatrix}$$

$$\xrightarrow[m_{21}=1, m_{31}=0.5]{R_2 \to R_2 - 1R_1, R_3 \to R_3 - \frac{1}{2}R_2} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & -1 & 1.5 \end{pmatrix}$$

$$\xrightarrow[m_{21}=0.5, m_{31}=1]{R_3 \leftrightarrow R_2} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 1.5 \\ 0 & 0 & 1 \end{pmatrix}$$

Therefore,

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

$$L = \begin{pmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$$U = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 1.5 \\ 0 & 0 & 1 \end{pmatrix}$$

2.

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

Therefore, the corresponding permutation vector is

$$V = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

3. Using V,

$$B \to \begin{pmatrix} 5 \\ 7 \\ 4 \end{pmatrix}$$

Therefore,

$$\begin{aligned} Ax &= b \\ \therefore LUx &= b \end{aligned}$$

Let

$$Ux = y$$

Therefore,

$$\begin{aligned} Ly &= b \\ \therefore \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} &= \begin{pmatrix} 5 \\ 7 \\ 4 \end{pmatrix} \end{aligned}$$

Therefore,

$$\begin{aligned} y_1 &= 5 \\ 0.5y_1 + y_2 &= 7 \\ y_1 + y_3 &= 4 \end{aligned}$$

Therefore, solving,

$$\begin{aligned} y_1 &= 5 \\ y_2 &= 4.5 \\ y_3 &= -1 \end{aligned}$$

Therefore,

$$Ux = y$$
$$\therefore \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 1.5 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 4.5 \\ -1 \end{pmatrix}$$

Therefore,

$$\begin{aligned} x_3 &= -1 \\ -x^2 + 1.5x_3 &= 4.5 \\ x_1 + 2x_2 + 3x_3 &= 5 \end{aligned}$$

Therefore, solving,

$$\begin{aligned} x_1 &= 20 \\ x_2 &= -6 \\ x_3 &= -1 \end{aligned}$$

6 Condition Number

Definition 4 (Condition number)**.** The condition number of a matrix A, with respect to a particular norm is defined as

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$$

Theorem 4. *Let*

$$Ax = B$$

be a matrix equation.
Then,

$$\frac{1}{\text{cond}(A)} \frac{\|e_b\|}{\|b\|} \leq \frac{\|e_x\|}{\|x\|} \leq \text{cond}(A) \frac{\|e_b\|}{\|b\|}$$

and the inequality is tight, i.e. there exist $\overline{x}$, $\overline{e_x}$, $\overline{b}$, $\overline{e_b}$, *such that there is an equality, i.e. the bounds are the best bounds possible.*

Recitation 8 – Exercise 1.
Consider the system

$$Ax = b$$

where

$$A = \begin{pmatrix} 100 & 99 \\ 99 & 98 \end{pmatrix}$$
$$B = \begin{pmatrix} 1.005 \\ 0.995 \end{pmatrix}$$

The accurate solution is

$$x = \begin{pmatrix} 0.015 \\ -0.005 \end{pmatrix}$$

Consider two approximations of the solution

$$\widetilde{x}_1 = \begin{pmatrix} -0.182 \\ 0.194 \end{pmatrix}$$
$$\widetilde{x}_2 = \begin{pmatrix} -19.685 \\ 19.895 \end{pmatrix}$$

1. Find the absolute error and the relative error in the RHS, in the infinity norm.
2. Fins the relative error, in the infinity norm, in x, assuming that x is known.
3. Can we conclude that a small relative error in the RHS implies a small relaive error in the LHS?
4. How can this problem be determined without knowing the actual values of x?

Recitation 8 – Solution 1.
1.

$$\begin{aligned} A\widetilde{x}_1 &= \begin{pmatrix} 100 & 99 \\ 99 & 98 \end{pmatrix} \begin{pmatrix} -0.182 \\ 0.194 \end{pmatrix} \\ &= \begin{pmatrix} 1.006 \\ 0.994 \end{pmatrix} \\ &= \widetilde{b}_1 \end{aligned}$$

Therefore,

$$\begin{aligned} e_{b_1} &= b - \widetilde{b}_1 \\ &= \begin{pmatrix} -0.001 \\ 0.001 \end{pmatrix} \end{aligned}$$

Therefore,

$$\|e_{b_1}\|_\infty = 0.001$$

Therefore,

$$\begin{aligned} \delta_{b_1} &= \frac{\|e_{b_1}\|_\infty}{\|b\|_\infty} \\ &= \frac{0.001}{1.005} \\ &\approx 10^{-3} \end{aligned}$$

Similarly,

$$\begin{aligned} \|e_{b_2}\|_\infty &= 0.1 \\ \delta_{b_2} &\approx 10^{-1} \end{aligned}$$

2.

$$\begin{aligned} e_1 &= x - \widetilde{x}_1 \\ &= \begin{pmatrix} 0.197 \\ -0.199 \end{pmatrix} \end{aligned}$$

Therefore,

$$\begin{aligned} \delta_{x_1} &= \frac{\|e_1\|_\infty}{\|x\|_\infty} \\ &= \frac{0.199}{0.015} \\ &\approx 13 \end{aligned}$$

Similarly,

$$\delta_{x_2} = 1326$$

3. Therefore, even though the relative error in B, i.e. the RHS is small, the error in x, i.e. in the LHS is huge. Hence, a small relative error in the RHS does not imply a small relative error in the LHS.

4.

$$
\begin{aligned}
A &= \begin{pmatrix} 100 & 99 \\ 99 & 98 \end{pmatrix} \\
\therefore A^{-1} &= \begin{pmatrix} -98 & 99 \\ 99 & -100 \end{pmatrix}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\|A\|_\infty &= 199 \\
\left\|A^{-1}\right\|_\infty &= 199
\end{aligned}
$$

Therefore,

$$\operatorname{cond}(A) = 199^2$$

Therefore,

$$\frac{\|e_x\|}{\|x\|} \leq 199^2 \frac{\|e_b\|}{\|b\|}$$

and the inequality is tight.
Therefore, as the inequality is tight, it is possible that an error in the RHS can be multiplied by 199^2 in the LHS.
Therefore if the condition number is large, then such a problem might exist.

Realistically, if A^{-1} can be calculated, x can be calculated x accurately.

7 Iterative Methods for Systems of Linear Equations

Algorithm 1 Jacobi Method

1: Find lower triangular L, diagonal D, and upper triangular U, such that $A = L + D + U$
2: $C \leftarrow D^{-1}$
3: $B_J \leftarrow \left(I - D^{-1}A\right) = -D^{-1}(L + U)$
4: $d_J \leftarrow D^{-1}b$
5: $x^{(n+1)} \leftarrow Bx^{(n)} + d$

Algorithm 2 Gauss-Seidel Method

Find lower triangular L, diagonal D, and upper triangular U, such that $A = L + D + U$

$C \leftarrow (L + D)^{-1}$

$B_{GS} \leftarrow \Big(I - (L + D)^{-1}A\Big) = -(L + D)^{-1}U$

$d_{GS} \leftarrow (L + D)^{-1}b$

$x^{(n+1)} \leftarrow Bx^{(n)} + d$

Theorem 5 (Sufficient condition for convergence of iterative method for systems of linear equations)**.** *Let $\|\cdot\|$ be a norm. If $\|B\| < 1$, then the method*

$$x^{(n+1)} = Bx^{(n)} + d$$

converges for any initial condition $x^{(0)}$.

Theorem 6 (Necessary condition for convergence of iterative method for systems of linear equations)**.** *The method*

$$x^{(n+1)} = Bx^{(n)} + d$$

converges for any $x^{(0)}$, if and only if

$$\rho(B) < 1$$

Recitation 9 – Exercise 1.
Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Consider the system

$$Ax = \alpha$$

1. Write B for Jacobi and Gauss-Seidel methods.

2. Find, for each method, a sufficient condition for convergence based on the ∞ norm.

3. Find, for each method, an equivalent condition for convergence.

4. Does the Jacobi method converge if and only if the Gauss-Seidel method converges?

5. Find a matrix A such that the above equivalent condition is met, but the above sufficient condition is not.

6. Let

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

and

$$\|B_J\|_\infty = \frac{1}{2}$$
$$\|B_{GS}\|_\infty = \frac{1}{2}$$

If

$$x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

then how many iterations are required in order to guarantee a relative error $\frac{\|e^{(k)}\|_\infty}{\|x\|_\infty}$ less than 10^{-6}?

Recitation 9 – Solution 1.

1.

$$\begin{aligned} B_J &= -D^{-1}(L+U) \\ &= -\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}^{-1} \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{a} & 0 \\ 0 & -\frac{1}{d} \end{pmatrix} \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -\frac{b}{a} \\ -\frac{c}{d} & 0 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} B_{GS} &= -(L+D)^{-1}U \\ &= -\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{a} & 0 \\ \frac{c}{ad} & -\frac{1}{d} \end{pmatrix} \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -\frac{b}{a} \\ 0 & \frac{bc}{ad} \end{pmatrix} \end{aligned}$$

2. A sufficient condition for convergence, for the Jacobi method is

$$\begin{aligned} \|B_J\|_\infty &< 1 \\ \therefore \max\left\{\left|\frac{b}{a}\right|, \left|\frac{c}{d}\right|\right\} &< 1 \end{aligned}$$

A sufficient condition for convergence, for the Gauss-Seidel method is

$$\begin{aligned} \|B_{GS}\|_\infty &< 1 \\ \therefore \max\left\{\left|\frac{b}{a}\right|, \left|\frac{bc}{ad}\right|\right\} &< 1 \end{aligned}$$

3. An equivalent condition for convergence, for the Jacobi method is

$$\begin{aligned} \rho(B_J) &< 1 \\ \therefore \max\{\lambda_1\} &< 1 \\ \therefore \sqrt{\left|\frac{bc}{ad}\right|} &< 1 \end{aligned}$$

Similarly,

$$\begin{aligned} \rho(B_{GS}) &< 1 \\ \therefore \left|\frac{bc}{ad}\right| &< 1 \end{aligned}$$

4.

$$\begin{aligned} \rho(B_j) &< 1 \\ \iff \sqrt{\left|\frac{bc}{ad}\right|} &< 1 \\ \iff \left|\frac{bc}{ad}\right| &< 1 \\ \iff \rho(B_{GS}) &< 1 \end{aligned}$$

Therefore, the Jacobi method converges if and only if the Gauss-Seidel method converges.

5. For the above equivalent condition to be met,

$$\begin{aligned} & \left|\frac{bc}{ad}\right| < 1 \\ \iff & |bc| < |ad| \end{aligned}$$

For the above sufficient condition to be not met,

$$\begin{aligned} & \left|\frac{b}{a}\right| > 1 \\ \iff & b > a \end{aligned}$$

Therefore, if

$$\begin{aligned} a &= 1 \\ b &= 2 \\ c &= \frac{1}{4} \\ d &= 1 \end{aligned}$$

the matrix A converges.

6.

$$\begin{aligned} e^{(k)} &= x - x^{(k)} \\ &= B^k e^{(0)} \end{aligned}$$

Therefore,

$$\begin{aligned} \left\|e^{(k)}\right\| &= \left\|B^k e^{(0)}\right\| \\ &\leq \left\|B^k\right\| \left\|e^{(0)}\right\| \\ &\leq \|B\|^k \|x - 0\| \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{\left\|e^{(k)}\right\|_\infty}{\|x\|_\infty} &\leq {\|B\|_\infty}^k \\ &= 2^{-k} \end{aligned}$$

Therefore, for the required accuracy,

$$\begin{aligned} {\|B\|_\infty}^k &< 10^{-6} \\ \therefore 2^{-k} &< 10^{-6} \end{aligned}$$

Therefore, $k = 20$ is sufficient for the required accuracy.

8 Numerical Differentiation

Recitation 10 – Exercise 1.
Let $f \in C^4$, with samples of f at $-h$, $2h$, and a sample of f' at $-h$ given.

1. Calculate Hermite's interpolation polynomial and derive the formula for the error.
2. Find an approximation of $f'(h)$ by differentiating the sum of the interpolation polynomial and the error.

Recitation 10 – Solution 1.
1. Let

$$\begin{aligned} x_0 &= -h \\ x_1 &= -h \\ x_2 &= 2h \end{aligned}$$

Therefore,

$$\begin{aligned} f[x_0] &= f(-h) \\ f[x_1] &= f(-h) \\ f[x_2] &= f(2h) \end{aligned}$$

Therefore,

$$\begin{aligned} f[x_0, x_1] &= f'(-h) \\ f[x_1, x_2] &= \frac{f(2h) - f(-h)}{3h} \end{aligned}$$

Therefore,

$$f[x_0, x_1, x_2] = \frac{f(2h) - f(-h) - 3hf'(-h)}{9h^2}$$

Therefore,

$$\begin{aligned} p_2(x) &= f[x_0] + f[x_0, x_1](x - x_0)^2 + f[x_0, x_1, x_2](x - x_0)(x - x_2) \\ &= f(-h) + f'(-h)(x + h) + \frac{f(2h) - f(-h) - 3hf'(-h)}{9h^2}(x + h)^2 \end{aligned}$$

Therefore,

$$\begin{aligned} \psi(x) &= \prod(x - x_i) \\ &= (x + h)^2(x - 2h) \end{aligned}$$

Therefore,

$$\begin{aligned} f(x) &= p_2(x) + f[-h,-h,2h,x]\psi(x) \\ &= p_2(x) f[-h,-h,2h,x](x+h)^2(x-2h) \end{aligned}$$

Therefore, the error is

$$e(x) = f[-h,-h,2h,x](x+h)^2(x-2h)$$

2.

$$\begin{aligned} f'(h) &\approx {p_2}'(h) \\ \therefore {p_2}'(x) &= f'(-h) + 2\frac{f(2h) - f(-h) - 3hf'(-h)}{9h^2}(x+h) \\ \therefore {p_2}'(h) &= f'(-h) + 4\frac{f(2h) - f(-h) - 3hf(-h)}{9h} \\ \therefore f'(h) &\approx f'(-h) + 4\frac{f(2h) - f(-h) - 3hf(-h)}{9h} \end{aligned}$$

The error is

$$\begin{aligned} f(x) &= p_2(x) + e(x) \\ \therefore f'(x) &= {p_2}'(x) + e'(x) \\ \therefore f'(h) &= {p_2}'(h) + e'(h) \end{aligned}$$

Therefore,

$$\begin{aligned} e(x) &= f[-h,-h,2h]\psi(x) \\ \therefore e'(x) &= f[-h,-h,2h,x,x]\psi(x) + f[-h,-h,2h,x]\psi'(x) \\ \therefore e'(h) &= f[-h,-h,2h,h,h]\psi(h) + f[-h,-h,2h,h]\psi'(h) \end{aligned}$$

Therefore, substituting $\psi(h)$ and $\psi'(h)$, for $c \in [-h, 2h]$,

$$\begin{aligned} e'(h) &= f[-h,-h,2h,h,h]\left(-4h^3\right) \\ &= -\frac{f^{(4)}(c)}{4!}4h^3 \\ &= \mathrm{O}\left(h^3\right) \end{aligned}$$

$f[x_0, \ldots, x_n] = \frac{f^{(n)}(c)}{n!}$

Recitation 11 – Exercise 1.
Let $f \in C^4$. Using Taylor expansion, find an approximation of $f''(a)$ of the form

$$f''(a) = Af(a-h) + Bf(a) + Cf(a+h) + e$$

where e is of the form

$$e = kf^{(n)}(c)h^m$$

1. Choose A, B, C, such that m is maximized.

2. Assume that for some $M > 0$,

$$\left| f''''(x) \right| \le M$$

   Assume that the value s of f are known up to an error bounded by ε. The Differentiation formula above is calculated using the noisy data. Bound the overall error in the approximation of $f''(a)$.

3. What is the optimal h for the error bound?

Recitation 11 – Solution 1.
1. For $c_1 \in (a, x)$,

$$\begin{aligned} f(x) = \quad & f(a) \\ & + f'(a)(x-a) \\ & + \frac{1}{2} f''(a)(x-a)^2 \\ & + \frac{1}{6} f'''(a)(x-a)^3 \\ & + \frac{1}{24} f''''(c_1) \end{aligned}$$

If $x = a - h$, $x - a = -h$. Therefore,

$$f(a-h) = \quad f(a) - hf'(a) + \frac{h^2}{2} f''(a) - \frac{h^3}{6} f'''(a) + \frac{h^4}{24} f''''(c_1)$$

If $x = a + h$, $x - a = h$. Therefore,

$$f(a-h) = \quad f(a) + hf'(a) + \frac{h^2}{2} f''(a) + \frac{h^3}{6} f'''(a) + \frac{h^4}{24} f''''(c_2)$$

If $x = a$, $h = 0$. Therefore,

$$f(a) = f(a)$$

Therefore,

$$\begin{aligned} Af(a-h) + Bf(a) + Cf(a+h) + e = \quad & Af(a) + Bf(a) + Cf(a) \\ & - Ahf'(a) + Chf'(a) \\ & + A\frac{h^2}{2}f''(a) + C\frac{h^2}{2}f''(a) \\ & + \dots \end{aligned}$$

Therefore,

$$f''(a) = Af(a-h) + Bf(a) + Cf(a+h) + e$$

Therefore, comparing,

$$\begin{aligned} A + B + C &= 0 \\ -A + C &= 0 \\ \frac{h^2}{2}A + \frac{h^2}{2}C &= 1 \end{aligned}$$

Therefore, solving,

$$\begin{aligned} A &= \frac{1}{h^2} \\ B &= -\frac{2}{h^2} \\ C &= \frac{1}{h^2} \end{aligned}$$

Therefore,

$$f''(a) \approx \frac{f(a+h) - 2f(a) + f(a+h)}{h^2}$$

Therefore,

$$\begin{aligned}
\frac{f(a+h)-2f(a)+f(a+h)}{h^2} &= \frac{f(a)-2f(a)+f(a)}{h^2} \\
&+ \frac{-hf'(a)-0+hf'(a)}{h^2} \\
&+ f''(a) \\
&\frac{\frac{h^3}{6}f'''(a)-\frac{h^3}{6}f'''(a)}{h^2} \\
&\frac{\frac{h^4}{24}f''''(c_1)+\frac{h^4}{24}f''''(c_2)}{h^2} \\
&= f''(a)+\frac{2h^2}{24}\frac{f''''(c_1)+f''''(c_2)}{2}
\end{aligned}$$

Therefore, by the intermediate value theorem, for $c \in [c_1, c_2]$,

$$\frac{f''''(c_1)+f''''(c_2)}{2} = f''''(c)$$

Therefore,

$$\frac{f(a+h)-2f(a)+f(a+h)}{h^2} = f''(a)+\frac{h^2}{12}f''''(c)$$

Therefore,

$$e = \frac{h^2}{12}f''''(c)$$

2. Let the error in f be η. Therefore,

$$f(x) = \widetilde{f}(x)+\eta(x)$$

Therefore,

$$\begin{aligned}
f''(a) &\approx \frac{\widetilde{f}(a-h)-2\widetilde{f}(a)+\widetilde{f}(a+h)}{h^2} \\
&= \frac{f(a-h)-2f(a)+f(a+h)}{h^2} - \frac{\eta(a-h)-2\eta(h)+\eta(a+h)}{h^2} \\
&= f''(a)+e-\frac{\eta(a-h)-2\eta(a)+\eta(a+h)}{h^2}
\end{aligned}$$

Let

$$E = e-\frac{\eta(a-h)-2\eta(a)+\eta(a+h)}{h^2}$$

Therefore, by triangle inequality,

$$\begin{aligned}|E| &\leq |e| + \left|\frac{\eta(a-h)}{h^2}\right| + \left|\frac{\eta(a)}{h^2}\right| + \left|\frac{\eta(a+h)}{h^2}\right| \\ &\leq \frac{h^2}{12}M + \frac{4\varepsilon}{h^2}\end{aligned}$$

3. Let

$$k(h) = \frac{h^2}{12}M + \frac{4\varepsilon}{h^2}$$

Therefore, minimizing,

$$h = \left(\frac{48\varepsilon}{M}\right)^{\frac{1}{4}}$$

Recitation 12 – Exercise 1.
Find an integration rule of the form

$$\int_0^h f(x)\,\mathrm{d}x \approx Af(0) + Bf'(0) + Cf(h)$$

where $f \in C^3$. Find the error formula, and find a composite integration method based on the integration rule, for $\int_a^b f(x)\,\mathrm{d}x$.
How many sample points are needed to guarantee an error of absolute value less than or equal to $\frac{10^{-9}}{72}$, with $f(x) = \sin x$, and $[a, b] = [0, 1]$.

Recitation 12 – Solution 1.

$$\begin{aligned}f[x_0] &= f(0) \\ f[x_1] &= f(0) \\ f[x_2] &= f(h)\end{aligned}$$

Therefore,

$$\begin{aligned}f[x_0, x_1] &= f'(0) \\ f[x_1, x_2] &= \frac{f(h) - f(0)}{h}\end{aligned}$$

Therefore,

$$f[x_0, x_1, x_2] = \frac{f(h) - f(0)}{h^2} - \frac{f'(0)}{h}$$

Therefore,

$$\begin{aligned} p_2(x) &= f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\ &= f(0) + f'(0)x + \left(\frac{f(h) - f(0)}{h^2} - \frac{f'(0)}{h} \right) x^2 \end{aligned}$$

Therefore,

$$\begin{aligned} \int_0^h f(x)\,\mathrm{d}x \approx \int_0^h p_2(x)\,\mathrm{d}x \\ &\approx f(0)h + \frac{f'(0)}{2}h^2 + \left(\frac{f(h) - f(0)}{h^2} - \frac{f'(0)}{h} \right) \frac{h^3}{3} \\ &\approx \frac{2}{3}hf(0) + \frac{1}{6}h^2 f'(0) + \frac{1}{3}hf(h) \end{aligned}$$

Therefore,

$$f(x) = p_2(x) + f[0, 0, h, x](x - 0)(x - 0)(x - h)$$

Therefore,

$$\int_0^h f(x) = \int_0^h p_2(x) + \int_0^h f[0, 0, h, x](x - 0)(x - 0)(x - h)$$

Therefore,

$$E = \int_0^h f[0, 0, h, x](x - 0)(x - 0)(x - h)$$

In the interval $(0, h)$, $(x - h)$ is always negative, and x^2 is always positive. Therefore, by the mean value theorem for integrals,

$$E = f[0, 0, h, d] \int_0^h x^2(x - h)\,\mathrm{d}x$$

where $d \in (0, h)$. Therefore,

$$E = \frac{f^{(3)}(c)}{3!} \left(\frac{h^4}{4} - \frac{h^4}{3} \right)$$

where $c \in (0, h)$. Therefore,

$$E = -\frac{h^4}{72} f^{(3)}(c)$$

Let the interval (a, b) be divided in n intervals by $a = x_0 < \cdots < x_n = b$, where

$$x_k = a + hk$$

where

$$h = \frac{b - a}{n}$$

Therefore,

$$\int_a^b f(x)\,\mathrm{d}x = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(x)\,\mathrm{d}x$$

Let

$$x = y + x_k$$

Therefore,

$$\begin{aligned}
\int_a^b f(x)\,\mathrm{d}x &= \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(x)\,\mathrm{d}x \\
&= \sum_{k=0}^{n-1} \int_0^h f(y + x_k)\,\mathrm{d}y \\
&= \sum_{k=0}^{n-1} \frac{2}{3}hf(x_k) + \frac{1}{6}h^2 f'(x_k) + \frac{1}{3}hf(h + x_{k+1}) + E_k \\
&= \sum_{k=0}^{n-1} \frac{2}{3}hf(x_k) + \frac{1}{6}h^2 f'(x_k) + \frac{1}{3}hf(h + x_{k+1}) - \frac{h^4}{72}\left.\frac{\partial^3 f(y + x_k)}{\partial y^3}\right|_{y=c} \\
&= \sum_{k=0}^{n-1} \frac{2}{3}hf(x_k) + \frac{1}{6}h^2 f'(x_k) + \frac{1}{3}hf(h + x_{k+1}) - \frac{h^4}{72}f^{(3)}(c_k) \\
&= \quad \frac{2}{3}hf(x_0) + \sum_{k=1}^{n-1} hf(x_k) + \sum_{k=0}^{n-1} \frac{1}{6}h^2 f'(x_k) + \frac{1}{3}hf(x_n) \\
&\quad - \frac{n \sum\limits_{k=0}^{n-1} \frac{h^4}{72} f^{(3)}(c_k)}{n} \\
&= \quad \frac{2}{3}hf(x_0) + \sum_{k=1}^{n-1} hf(x_k) + \sum_{k=0}^{n-1} \frac{1}{6}h^2 f'(x_k) + \frac{1}{3}hf(x_n) - (b-a)\frac{h^3}{72}f^{(3)}(c)
\end{aligned}$$

Therefore,

$$\begin{aligned}
|e| &\leq \left| -n\frac{h^4}{72}f^{(3)}(c) \right| \\
&\leq \left| -(b-a)\frac{h^3}{72}f^{(3)}(c) \right| \\
&\leq 1 \cdot \frac{h^3}{72} \cdot 1 \\
\therefore \frac{10^{-9}}{72} &\geq \frac{h^3}{72}
\end{aligned}$$

Therefore,

$$h \leq 10^{-3}$$

Therefore,

$$n \geq 10^3$$