MAE5002 Advanced Numerical Analysis Spring2024

Homework 1：Answers

1.

1)Total number of normalized floating-point number:

There are β-1 possible choices for the leading digit d_0, β choices for p-1 fractions, and $U + L - 1$ possible values for the exponent. And because the number could be zero, 1 should be added. There is a positive and negative sign in front of values which leads double of the total numbers.

Therefore, the total numbers of normalized floating-point number is

$$2(\beta-1)\beta^{p-1}(U+L-1)+1$$

2) Smallest positive normalized number:

For the smallest positive normalized number, the leading digit is 1 and the remaining digits of the mantissa are 0, and the exponent is smallest possible value

$$UFL = \beta^L$$

3) Largest floating-point number:

Largest floating-point number has $\beta - 1$ as the value for each digit of the mantissa and the largest possible value for the exponent.

$$OFL = (\beta - 1)(1 + \frac{1}{\beta} + \cdots + \frac{1}{\beta^{p-1}})\beta^U = \beta^{U+1}\left(1 - \frac{1}{\beta^p}\right)$$

2 Exercise 1.4

a) Absolute value = $\sin(x + h) - \sin(x)$

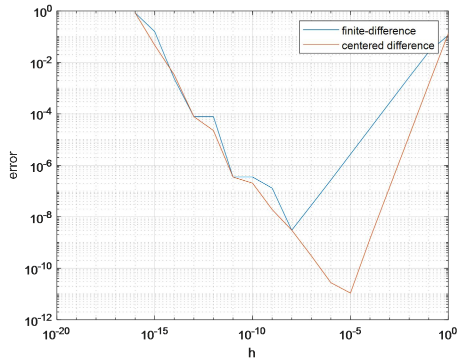
b) Relative value = $\dfrac{\sin(x+h)-\sin(x)}{\sin(x)}$

c) Condition number = $\left|\dfrac{\dfrac{\sin(x+h)-\sin(x)}{\sin(x)}}{\dfrac{x+h-x}{x}}\right| = \left|\dfrac{x\cos(x)}{\sin(x)}\right| = \left|\dfrac{x}{\tan(x)}\right|$

d) When condition number is much larger or smaller than 1, i.e. x is near any integer multiple of π except 0, the problem is highly sensitive.

3 Programming

Code: refer to Q3 in zip file



As the figure shows, there is a minimum value error =10^{-8} for the magnitude of the error when h= 10^{-8}, which corresponds to the rule in Example 1.3. While for centered difference approximation, there is a minimum value for the magnitude of the error when $h = 10^{-5}$.

4.

stopping criterion:

(1) x^n/n! < epsilon

(2) [(x^n/n!) / (1+…+x^(n-1)/(n-1)!)] < epsilon

Both (1) and (2) is right, it does not make contributions to the series sum.The answer (2) is more reasonable.More details can be found in Q4.mlx

5.

(1)

$x = (x_1, \ldots, x_n)^T \quad ||x||_\infty = \max_{0 \le i \le n} |x_i| = |x_j|$

For :$||x||_2^2 = \sum_{i=1}^{n} |x_i|^2 \le (\sum_{i=1}^{n} |x_i|)^2 = ||x||_1^2$

So : $||x||_2 \le ||x||_1$

For : $||x||_1^2 = (\sum_{i=1}^{n} |x_i|)^2 \le n\sum_{i=1}^{n} |x_i|^2 = n * ||x||_2^2$

So: $||x||_1 \le \sqrt{n}\ ||x||_2$

So: $||x||_2 \le ||x||_1 \le \sqrt{n}\ ||x||_2$

(2)

For : $||x||_\infty^2 = |x_j|^2 \le \sum_{i=1}^{n} |x_i|^2 = ||x||_2^2$

So : $||x||_\infty \le ||x||_2$

For : $||x||_2^2 = \sum_{i=1}^{n} |x_i|^2 \le \sum_{j=1}^{n} |x_j|^2 \le n\,|x_j|^2 = n||x||_\infty^2$

So : $||x||_2 \le \sqrt{n}\ ||x||_\infty$

So : $||x||_\infty \le ||x||_2 \le \sqrt{n}\ ||x||_\infty$

6.

(a)

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 3 & 2 \end{bmatrix} \xrightarrow[c3-c1]{c2-c1} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow{c3-2*c2} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

So: A is singular

(b)

$$[A|b] = \begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & 2 & 1 & 4 \\ 1 & 3 & 2 & 6 \end{bmatrix} \xrightarrow[c3-c1]{c2-c1} \begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 2 \\ 0 & 2 & 2 & 4 \end{bmatrix} \xrightarrow{c3-2*c2} \begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

So: when $b = [2\ 4\ 6]^T$, there are infinite solutions.

(c)

$$||A||_1 = \max_j \sum_{i=1}^{n} |a_{ij}| = 6$$

$$||A||_\infty = \max_i \sum_{j=1}^{n} |a_{ij}| = 6$$

(d)

For A is singular,so the condition number of A is infinite.

7.8.

More details can be found in Q7.mlx and Q8.mlx