Assignment 1

Name: 陈子蔚 SID: 12332440

1

As for 0, $\forall i, d_i = 0 \land s = 0 \land E = 0$. Only 1 case.

As for signed floating numbers, $0 < d_0 \le \beta - 1 \land 1 \le 0 \le \beta - 1, 0 < i \le p - 1 \land L \le E \le U$.

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

Totally, $(\beta - 1)\beta^{p-1}(U - L + 1) + 1$ cases.

The smallest fraction part is 1. The smallest exponent part is L. The smallest positive floating number is $1 \times \beta^L = \beta^L$.

The largest fraction part is $(\beta - \beta^{-(p-1)})$. The largest exponent part is U. The largest positive floating number is $(\beta - \beta^{-(p-1)}) \times \beta^U = \beta^{U+1}(1 - \beta^{-p})$.

2

a

$$\epsilon = \sin(x + h) - \sin(x) \tag{1}$$

b

$$\epsilon_r = \frac{\sin(x + h) - \sin(x)}{\sin(x)} \tag{2}$$

c

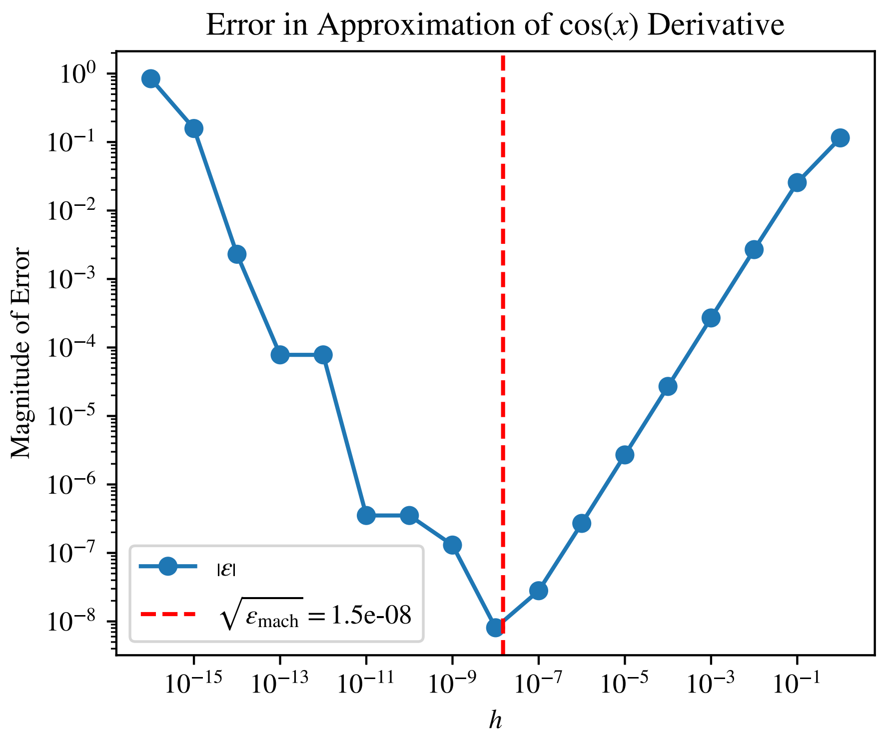
$$\begin{aligned}
\kappa &= \frac{\left|\frac{\sin(x+h)-\sin(x)}{\sin(x)}\right|}{\left|\frac{h}{x}\right|} && (3)\\
&= \left|\frac{\frac{\sin(x+h)-\sin(x)}{\sin(x)}}{\frac{h}{x}}\right| && (4)\\
&= \left|\frac{\sin(x+h)-\sin(x)}{h} \cdot \frac{x}{\sin(x)}\right| && (5)\\
&\simeq \left|\frac{x\cos(x)}{\sin(x)}\right| && (6)\\
&= \left|\frac{x}{\tan(x)}\right| && (7)
\end{aligned}$$

d

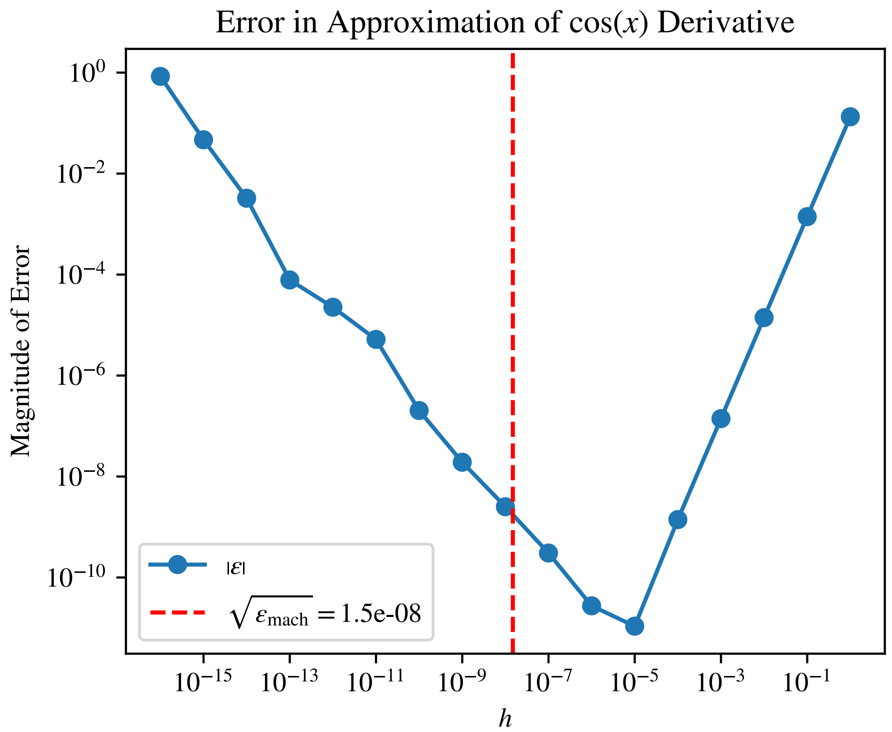
The problem is highly sensitive where the condition number is extremely large, which typically occurs when $\tan(x)$ is close to 0.

3

a



b



4

b

A tolerance such that it is less than the machine epsilon was used, which is a reasonable estimate for when to stop the series summation.

Output

```
# x_list: [1, -1, 5, -5, 10, -10, 15, -15, 20, -20]
19
19
37
37
53
53
68
68
83
83
# max order
[(1, 2.7182818284590455, 2.718281828459045, 4.440892098500626e-16),
(-1, 0.36787944117144245, 0.36787944117144233, 1.1102230246251565e-16),
(5, 148.4131591025766, 148.4131591025766, 0.0),
(-5, 0.006737946999084642, 0.006737946999085467, 8.248610128269718e-16),
(10, 22026.465794806714, 22026.465794806718, 3.637978807091713e-12),
(-10, 4.5399929670419935e-05, 4.5399929762484854e-05, 9.206491905638936e-
14),
(15, 3269017.3724721107, 3269017.3724721107, 0.0),
(-15, 3.059100025508472e-07, 3.059023205018258e-07, 7.682049021416746e-
12),
(20, 485165195.40979046, 485165195.4097903, 1.7881393432617188e-07),
(-20, 6.147561848704381e-09, 2.061153622438558e-09, 4.086408226265824e-
09)]
# [(x, my_exp, builtin_exp, error)]
```

5

$$\|x\|_1 = \sum_{1\leq i\leq n} |x_i| \tag{8}$$

$$\|x\|_2 = \sqrt{\sum_{1\leq i\leq n} |x_i|^2} \tag{9}$$

$$\|x\|_\infty = \max_{1\leq i\leq n} |x_i| \tag{10}$$

Notice that,

$$\forall i, |x_i| \leq \sum_{1\leq i\leq n} |x_i| = \|x\|_1 \tag{11}$$

As for the 1st inequality,

$$\|x\|_2^2 = \sum_{1\le i\le n} |x_i|^2 = \sum_{1\le i\le n} |x_i|\cdot|x_i| \tag{12}$$

$$\le \sum_{1\le i\le n} |x_i|\cdot\|x\|_1 = \|x\|_1 \cdot \sum_{1\le i\le n} |x_i| = \|x\|_1^2 \tag{13}$$

Since $\|x\|_2 \ge 0 \wedge \|x\|_1 \ge 0$,

$$\|x\|_2 \le \|x\|_1 \tag{14}$$

Then,

$$\|x\|_1 = \sum_{1\le i\le n} |x_i| = x^\top[1,\cdots,1]^\top \le \|x\|_2 \cdot \|[1,\cdots,1]^\top\|_2 = \sqrt{n}\|x\|_2 \tag{15}$$

Thus,

$$\|x\|_2 \le \|x\|_1 \le \sqrt{n}\|x\|_2 \tag{16}$$

As for the 2nd inequality, suppose $j = \arg\max_{0\le i\le n} |x_i|$,

$$\|x\|_\infty^2 = x_j^2 \le \sum_{1\le i\le n, i\ne j} {x_i}^2 + x_j^2 = \sum_{1\le i\le n} {x_i}^2 = \|x\|_2^2 \tag{17}$$

That is,

$$\|x\|_\infty \le \|x\|_2 \tag{18}$$

Notice that,

$$\forall i, |x_i| \le |x_j| \tag{19}$$

Then,

$$\|x\|_2 = \sqrt{\sum_{1\le i\le n} |x_i|^2} \le \sqrt{\sum_{1\le i\le n} |x_j|^2} = \sqrt{n|x_j|^2} = \sqrt{n}|x_j| = \sqrt{n}\|x\|_\infty \tag{20}$$

6

a&b

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 1 & 2 & 1 & 4 \\ 1 & 3 & 2 & 6 \end{array}\right] \xrightarrow{r_2 - r_1, r_3 - r_1} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 2 \\ 0 & 2 & 2 & 4 \end{array}\right] \xrightarrow{r_2 - \frac{1}{2} r_3} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right] \tag{21}$$

There are only 2 pivots in A and $b \in c(A)$ since $x = [1,1,1]^\top$ is a possible solution.

Thus, A is singular and there are infinite solutions to $Ax = b$.

c

$$\|A\|_1 = \max_{1\le j\le n} \|A_{(:,j)}\|_1 = 6 \tag{22}$$

$$\|A\|_\infty = \max_{1\le i\le n} \|A_{(i,:)}\|_1 = 6 \tag{23}$$

d

Since A is singular,

$$\text{cond}(A) = \infty \tag{24}$$

7

a

Note that L here is the L^{-1} in the slide.

$$U = L_m P_m \cdots L_1 P_1 A = LPA \tag{25}$$
$$P = P_m \cdots P_1 \tag{26}$$
$$L = L_m P_m \cdots L_1 P_1 P^\top \tag{27}$$

b

$$Ax = b \Leftrightarrow LPAx = Ux = LPb \tag{28}$$
$$A[x_1, \cdots, x_n] = [b_1, \cdots, b_n] \Leftrightarrow I = [e_1, \cdots, c_n] = AA^{-1} \tag{29}$$

c

$$\text{cond}(A) = \|A\| \|A^{-1}\| \tag{30}$$

Output

```
# P
[[0. 1. 0.]
 [0. 0. 1.]
 [1. 0. 0.]]
# L
[[ 1.    0.    0.  ]
 [-1.    1.    0.  ]
 [ 0.25 -0.5   1.  ]]
# U
[[4.  4.  2. ]
 [0.  2.  2. ]
 [0.  0.  0.5]]
# A_inv
[[ 1.   1.  -1. ]
 [-2.  -1.   1.5]
 [ 2.   0.5 -1. ]]
# A_inv @ A
[[1. 0. 0.]
 [0. 1. 0.]
 [0. 0. 1.]]
# cond_1
60.0
# cond_inf
63.0
```

8

The code of `LPU_factorization` and `Ux_equal_b_solution` in 7 is also used here. Moreover, the package `scipy` is used to check the correctness.

Output

```
[-28.28427125  20.          10.         -30.          14.14213562
  20.           0.         -30.           7.07106781  25.
  20.         -35.35533906  25.        ]
[ 0.00000000e+00  0.00000000e+00 -3.55271368e-15 -1.77635684e-15
  0.00000000e+00  0.00000000e+00  8.88178420e-16  0.00000000e+00
  0.00000000e+00  0.00000000e+00  8.88178420e-16  4.44089210e-15
  7.10542736e-15]
# solutions by the manual program
[-28.28427125  20.          10.         -30.          14.14213562
  20.           0.         -30.           7.07106781  25.
  20.         -35.35533906  25.        ]
[ 0.00000000e+00  0.00000000e+00 -3.55271368e-15 -1.77635684e-15
  0.00000000e+00  0.00000000e+00  8.88178420e-16  0.00000000e+00
  0.00000000e+00  0.00000000e+00  8.88178420e-16  4.44089210e-15
  0.00000000e+00]
# solutions by scipy
```