Part MATH-UA 252 — Numerical Analysis

Based on lectures by Jason Kaye
Homework done by Aaron Ma

Fall 2024

These notes are not endorsed by the lecturers, and I have modified them (often significantly) after lectures. They are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.

1 Problem 1

1.1 a

See the code

1.2 b

For Problem 4d:
The result now is $(1,1)^T$ and has an error of 0, the pivoting avoids the error generated by floating-point calculation, which leads to a wrong answer of $(0,1)^T$
For problem 4e:
The error is now 9.712478273807791e-11, which is similar to what we had before, indicating the result is probably not because of the rounding error generated during Gaussian Elimination

1.3 c

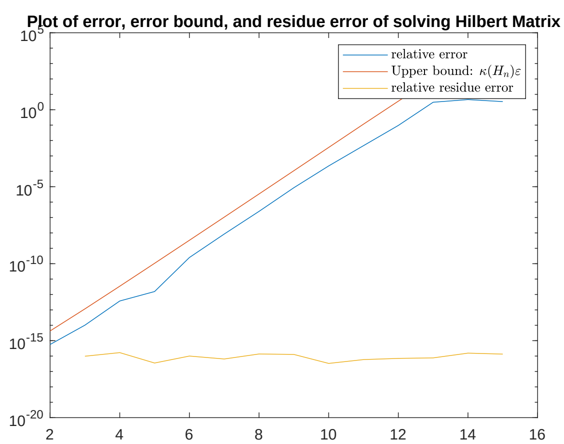

Figure 1: Plot of Error, Error bound, Residue Error of solving Hilbert Matrix

As the graph indicates, the relative residue error is always small, meaning the Gaussian Elimination is backward stable. However, the relative error increases as the size of the matrix increases, still, it is smaller than the upper bound.

1.4 d

According the the program, The solution generated with Gaussian Elimination with partial pivoting has a
relative error:

0.718045524835773
relative residue error:
0.396018360173385
This means that the Gaussian Elimination is not backward stable under this situation.
While the usual upper bound of the relative error, $\kappa(A) \cdot \varepsilon$, is 9.948098177737327e-15, smaller than the relative error, it is not suitable for this situation.
As the algorithm is not backward stable now, the relative error can be bigger than ε, which means that we can only estimate it with $\kappa(A) \cdot$(relative residue), which is 19.516441822920534 and is larger than the relative error. As a result, it does not violate what we learned in class.

1.5 e

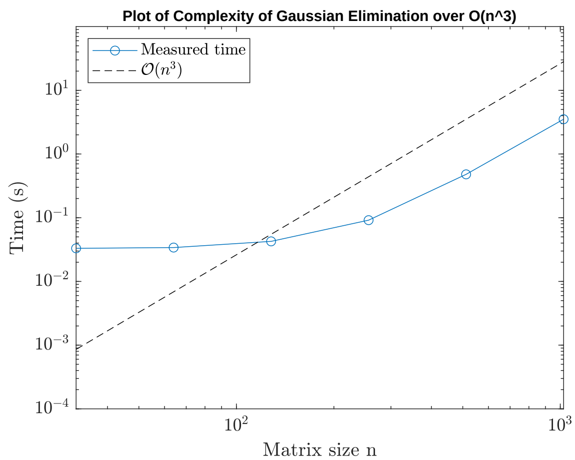

Figure 2: Plot of Complexity of Gaussian Elimination

As the graph shows, the slope of $\mathcal{O}(n^3)$ is greater than the measured time and as a result, the complexity is $\mathcal{O}(n^3)$

2 Problem 2

2.1 a

$$\begin{aligned}
Ax &= b \\
(A+\delta)\hat{x} &= \hat{b} \\
\therefore (A+\delta)(x-\hat{x}) &= b + \delta x - \hat{b} \\
x - \hat{x} &= (A+\delta)^{-1}(b + \delta x - \hat{b} \\
\frac{||x-\hat{x}||}{||x||} &\leq \frac{||(A+\delta)^{-1}||||b-\hat{b}+\delta x||}{||x||} \\
&\leq \frac{||(A+\delta)^{-1}||||b-\hat{b}||}{||x||} + \frac{||(A+\delta)^{-1}||||\delta x||}{||x||} \\
&= ||(A+\delta)^{-1}|| \cdot \frac{||b-\hat{b}||}{||b||} \cdot \frac{||b||}{||x||} + ||\delta||||(A+\delta)^{-1}|| \\
&\leq ||(A+\delta)^{-1}|| \cdot ||A|| \cdot (\frac{||b-\hat{b}||}{||b||} + \frac{||\delta||}{||A||})
\end{aligned}$$

2.2 b

$$\begin{aligned}
(A+\delta)^{-1} &= (I + A^{-1}\delta)^{-1}A^{-1} \\
&= (\sum_{k=0}^{\infty}(-1)^k(A^{-1}\delta)^k)A^{-1} \\
&= (I + \sum_{k=1}^{\infty}(-1)^k(A^{-1}\delta)^k)A^{-1} \\
&= A^{-1} + (\sum_{k=1}^{\infty}(-1)^k(A^{-1}\delta)^k)A^{-1} \\
\therefore ||(A+\delta)^{-1}|| &\leq ||A^{-1}|| + ||A^{-1}||\sum_{k=1}^{\infty}||A^{-1}\delta||^k \\
&= ||A^{-1}||(1 + \sum_{k=1}^{\infty}||A^{-1}\delta||^k)
\end{aligned}$$

2.3 c

$$\begin{aligned}
||(A+\delta)^{-1}|| &\leq ||A^{-1}||(1+\sum_{k=1}^{\infty}||A^{-1}\delta||^{k}) \\
&= ||A^{-1}||(1+\frac{||A^{-1}\delta||}{1-||A^{-1}\delta||}) \\
&\leq ||A^{-1}||(1+\frac{||A^{-1}||||\delta||}{1-||A^{-1}||||\delta||})
\end{aligned}$$

2.4 d

$$\begin{aligned}
||(A+\delta)^{-1}|| &\leq ||A^{-1}||(1+\sum_{k=1}^{\infty}||A^{-1}\delta||^{k}) \\
&\leq ||A^{-1}||(1+\frac{||A^{-1}||||\delta||}{1-||A^{-1}||||\delta||}) \\
&= ||A^{-1}||\frac{1}{1-||A^{-1}||||\delta||} \\
&= ||A^{-1}||\frac{1}{1-\frac{\kappa(A)}{||A||}||\delta||} \\
\therefore \frac{||x-\hat{x}||}{||x||} &\leq \frac{\kappa(A)}{1-\kappa(A)||\delta||/||A||}(\frac{||b-\hat{b}||}{||b||}+\frac{||\delta||}{||A||})
\end{aligned}$$

3 Problem 3

3.1 a

As both $sin(x)$ and $cos(x)$ is 2π-periodic, $f(x)$ is 2π-periodic.

$$\begin{aligned} f(-x) &= cos^3(sin(-x)) \\ &= cos^3(sin(x)) \\ &= f(x) \end{aligned}$$

Therefore it is also even

3.2 b

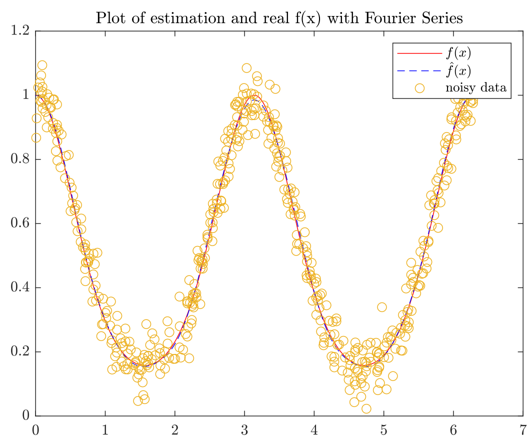

Figure 3: 3bEstimation and real function of f(x)

The max error is 1.465563893286781e-02