CS 450 Numerical Analysis (22Fa): Homework 2

Xu Ke (3190110360)

Oct. 24, 2022

1 Theoretical Problems

1.1

Question-1: Consider a non-singular matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, use the definition of condition number to give a step-by-step proof showing that the matrix condition number is given by the following

$$\kappa(\boldsymbol{A}) = \|\boldsymbol{A}\| \cdot \|\boldsymbol{A}^{-1}\| \tag{1}$$

where $\|\boldsymbol{A}\|$ is the matrix norm. Furthermore, show that

$$\|\boldsymbol{A}\| \cdot \|\boldsymbol{A}^{-1}\| = \left(\max_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}\right) \cdot \left(\min_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}\right)^{-1}. \tag{2}$$

In addition, prove that $\kappa(\boldsymbol{A}) \geq 1$.

Considering y=Ax, then we can define the pertubated problem as, $y + \delta y = A(x + \delta x)$. Thus, we can get the condition number of matrix A, $\kappa(A)$:

$$\begin{aligned}
\kappa(A) &= \max \frac{relative forwarderror}{relativebackwarderror} = \max \frac{||\delta y||/||y||}{||\delta x||/||x||} \\
&= \max \frac{||A\delta x|| \cdot ||x||}{||\delta x|| \cdot ||y||} = \max \frac{||A\delta x||}{||\delta x||} \cdot \frac{||A^{-1}y||}{||y||} \\
&= \max \frac{||A|| \cdot ||\delta x||}{||\delta x||} \cdot \frac{||A^{-1}|| \cdot ||y||}{||y||} = \boxed{||A|| \cdot ||A^{-1}||}
\end{aligned} \tag{1}$$

According to the definition of operator norm, we will get $||A|| = \max_{x\in R^n, x\neq 0} \frac{||Ax||_2}{||x||_2}$, thus:

$$\begin{aligned}
||A^{-1}|| &= \max_{x\neq 0} \frac{||A^{-1}x||}{||x||} = \max_{x\neq 0} \frac{||A^{-1}Ax||}{||Ax||} = \max_{x\neq 0} \frac{||x||}{||Ax||} \\
&= ((\max_{x\neq 0} \frac{||x||}{||Ax||})^{-1})^{-1} = (\min_{x\neq 0} \frac{||Ax||}{||x||})^{-1}
\end{aligned} \tag{2}$$

$$\kappa(A) = ||A|| \cdot ||A^{-1}|| = \boxed{(\max_{x\neq 0} \frac{||Ax||_2}{||x||_2}) \cdot (\min_{x\neq 0} \frac{||Ax||_2}{||x||_2})^{-1}} \tag{3}$$

Furthermore, based on the inequality of matrix, we can easily get:

$$\kappa(A) = ||A|| \cdot ||A^{-1}|| \geq ||AA^{-1}|| = ||I|| = \boxed{1} \tag{4}$$

1.2

Question-2: Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ be an invertable matrix, and $\boldsymbol{x}$, $\boldsymbol{x} + \Delta\boldsymbol{x}$ be the solutions to the following systems

$$\boldsymbol{Ax} = \boldsymbol{b}, \tag{3}$$
$$(\boldsymbol{A} + \Delta\boldsymbol{A})(\boldsymbol{x} + \Delta\boldsymbol{x}) = \boldsymbol{b}. \tag{4}$$

Consider $\boldsymbol{b} \neq \boldsymbol{0}$, show the following:
(a) The following inequality holds

$$\frac{\|\Delta\boldsymbol{x}\|}{\|\boldsymbol{x} + \Delta\boldsymbol{x}\|} \leq \kappa(\boldsymbol{A})\frac{\|\Delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}. \tag{5}$$

(b) The following inequality holds

$$\frac{\|\Delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \kappa(\boldsymbol{A})\frac{\|\Delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}\left(\frac{1}{1 - \|\Delta\boldsymbol{A}\| \cdot \|\boldsymbol{A}^{-1}\|}\right). \tag{6}$$

1.2.1 (a)

Given $Ax = b$ and $(A + \Delta A)(x + \Delta x) = b$, then we have:

$$Ax = (A + \Delta A)(x + \Delta x) = Ax + A\Delta x + \Delta Ax + \Delta A\Delta x \tag{5}$$

Thus, we can get $-A\Delta x = \Delta A(x + \Delta x)$, which is equivalent to, $-A^{-1}A\Delta x = A^{-1}\Delta A(x + \Delta x)$ by multiplying A^{-1} at both sides, then we have:

$$||\Delta x|| = ||A^{-1}\Delta A(x + \Delta x)|| \leq ||A^{-1}|| \cdot ||\Delta A|| \cdot ||x + \Delta x|| \tag{6}$$

$$\frac{||\Delta x||}{||x + \Delta x||} \leq ||A^{-1}|| \cdot ||\Delta A|| = \kappa(A)\frac{||\Delta A||}{||A||} \tag{7}$$

Hence, we can conclude that $\frac{||\Delta x||}{||x+\Delta x||} \leq \kappa(A)\frac{||\Delta A||}{||A||}$.

1.2.2 (b)

From (a) above, we have already proven that $\frac{||\Delta x||}{||x+\Delta x||} \leq \kappa(A)\frac{||\Delta A||}{||A||}$. Thus in order to prove $\frac{||\Delta x||}{||x||} \leq \kappa(A)\frac{||\Delta A||}{||A||}(\frac{1}{1-||\Delta A||\cdot||A^{-1}||})$, we only need to show:

$$\frac{||x+\Delta x||}{||x||} \leq \frac{1}{1-||\Delta A||\cdot||A^{-1}||} \tag{8}$$

That is to say, we need to prove validity of:

$$||\Delta x + x|| \leq ||x|| + ||A^{-1}||\cdot||\Delta A||\cdot||\Delta x + x|| \tag{9}$$

Since we know $\Delta A(x+\Delta x) + A\Delta x = 0$, which implies:

$$\Delta A A^{-1}(x+\Delta x) + \Delta x = 0 \tag{10}$$

Thus, we will get:

$$||\Delta x + x|| - ||x|| \leq ||\Delta x|| = ||\Delta A A^{-1}(x+\Delta x)|| \leq ||A^{-1}||\cdot||\Delta A||\cdot||x+\Delta x|| \tag{11}$$

Hence, the inequality above has been proved, $||\Delta x + x|| \leq ||x|| + ||A^{-1}||\cdot||\Delta A||\cdot||\Delta x + x||$ holds. Thus, we can conclude that $\frac{||\Delta x||}{||x||} \leq \kappa(A)\frac{||\Delta A||}{||A||}(\frac{1}{1-||\Delta A||\cdot||A^{-1}||})$.

1.3

Question-3: Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix. Show that the function

$$\|\boldsymbol{x}\|_{\boldsymbol{A}} = \sqrt{\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}} \tag{7}$$

defines a norm on $\mathbb{R}^n$ (i.e., it satisfies the three defining properties of a norm). This vector norm is said to be *induced* by $\boldsymbol{A}$.

Moreover, what if $\boldsymbol{A}$ is positive semi-definite, does it still hold?

According to the definition of norm, $||x||_A$ can be a norm if and only if it satisfies (1) $||x||_A \geq 0$ and $||x||_A = 0$ if and only if $x = 0$, (2) $||\alpha x||_A = ||\alpha||\cdot||x||_A$ and (3) $||x_1 + x_2||_A \leq ||x_1||_A + ||x_2||_A$.

First, since A is a positive definite matrix, we must have:

$$||x||_A = \sqrt{x^TAx} = 0 \leftrightarrow x^TAx = 0 \leftrightarrow x = 0 \tag{12}$$

Second, replacing x with αx, we will get:

$$||\alpha x||_A = \sqrt{(\alpha x)^TA(\alpha x)} = \sqrt{\alpha^2x^TAx} = \alpha\sqrt{x^TAx} = \alpha||x||_A \tag{13}$$

Third, replacing x with $x_1 + x_2$, we will get:

$$\begin{aligned}
||x_1 + x_2||_A = \sqrt{(x_1 + x_2)^T A (x_1 + x_2)} = \sqrt{(x_1^T + x_2^T) A (x_1 + x_2)} \\
= \sqrt{x_1^T A x_1 + 2x_1^T A x_2 + x_2^T A x_2} \leq \sqrt{x_1^T A x_1 + 2\sqrt{x_1^T A x_1 x_2^T A x_2} + x_2^T A x_2} \\
= \sqrt{(\sqrt{x_1^T A x_1} + \sqrt{x_2^T A x_2})^2} = ||x_1||_A + ||x_2||_A
\end{aligned} \tag{14}$$

Considering that $||x||_A$ satisfies all three properties above, we can conclude that $||x||_A = \sqrt{x^T A x}$ defines a norm. Furthermore, if A is a positive semi-definite, $||x||_A$ will no longer be a norm since $||x||_A = 0$ doesn't necessarily spring from $x = 0$, thus dissatisfying property (1) mentioned above.

1.4

Question-4: Prove the following Sherman-Morrison formula, in which $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ and $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$

$$\left(\boldsymbol{A} - \boldsymbol{u}\boldsymbol{v}^T\right)^{-1} = \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{u}\left(1 - \boldsymbol{v}^T\boldsymbol{A}^{-1}\boldsymbol{u}\right)^{-1}\boldsymbol{v}^T\boldsymbol{A}^{-1}. \tag{8}$$

Furthermore, prove the following Woodbury formula

$$\left(\boldsymbol{A} - \boldsymbol{U}\boldsymbol{V}^T\right)^{-1} = \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{U}\left(\boldsymbol{I} - \boldsymbol{V}^T\boldsymbol{A}^{-1}\boldsymbol{U}\right)^{-1}\boldsymbol{V}^T\boldsymbol{A}^{-1}. \tag{9}$$

In order to prove $(A - uv^T)^{-1} = A^{-1} + A^{-1}u(1 - v^T A^{-1} u)^{-1} v^T A^{-1}$ holds, we just need to show that $(A - uv^T)(A^{-1} + A^{-1}u(1 - v^T A^{-1}u)^{-1} v^T A^{-1}) = I$ holds:

$$\begin{aligned}
(A - uv^T)(A^{-1} + A^{-1}u(1 - v^T A^{-1}u)^{-1} v^T A^{-1}) \\
= I + u(1 - v^T A^{-1}u)^{-1} v^T A^{-1} - uv^T A^{-1} - uv^T A^{-1} u (1 - v^T A^{-1} u)^{-1} v^T A^{-1} \\
= I + \frac{uv^T A^{-1} - uv^T A^{-1} uv^T A^{-1}}{1 - v^T A^{-1} u} - uv^T A^{-1} = I + (uv^T A^{-1} - uv^T A^{-1}) = I
\end{aligned} \tag{15}$$

Thus, the Sherman-Morrison Formula $(A - uv^T)^{-1} = A^{-1} + A^{-1}u(1 - v^T A^{-1}u)^{-1} v^T A^{-1}$ proved. Then the Woodbury formula would just be the generalization to rank k modification:

$$\begin{aligned}
(A - UV^T)(A^{-1} + A^{-1}U(I - V^T A^{-1}U)^{-1} V^T A^{-1}) \\
= I + U(I - V^T A^{-1} U)^{-1} V^T A^{-1} - UV^T A^{-1} - UV^T A^{-1} U (I - V^T A^{-1} U)^{-1} V^T A^{-1} \\
= I + UV^T A^{-1} (1 - V^T A^{-1} U)^{-1} (I - UV^T A^{-1}) - UV^T A^{-1} = I + UV^T A^{-1}(I - I) = I
\end{aligned} \tag{16}$$

Thus, the Woodbury formula $(A - UV^T)^{-1} = A^{-1} + A^{-1}U(I - V^T A^{-1} U)^{-1} V^T A^{-1}$ proved.

1.5

Question-5: Let

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1+\epsilon \\ 1-\epsilon & 1 \end{bmatrix}. \tag{10}$$

(a) What is the determinant of $\boldsymbol{A}$?

(b) In floating-point arithmetic, for what range of values of ϵ will the computed value of the determinant be zero?

(c) What is the LU factorization of $\boldsymbol{A}$?

(d) In floating-point arithmetic, for what range of value of ϵ will be the computed value of U (in the LU factorization) be singular?

1.5.1 (a)

According to the definition of determinant, we can get:

$$det(A) = 1 \cdot 1 - (1+\epsilon)(1-\epsilon) = \boxed{\epsilon^2} \tag{17}$$

1.5.2 (b)

Suppose the machine epsilon of the floating-point system to be ϵ_{mach}, then the computed value of the determinant will be zero, if it is smaller than the machine epsilon, thus:

$$\boxed{-\sqrt{\epsilon_{mach}} < \epsilon < \sqrt{\epsilon_{mach}}} \tag{18}$$

1.5.3 (c)

Based on the rule of LU factorization, we will get:

$$L = \boxed{\begin{bmatrix} 1 & 0 \\ 1-\epsilon & 1 \end{bmatrix}} \tag{19}$$

$$U = \boxed{\begin{bmatrix} 1 & 1+\epsilon \\ 0 & \epsilon^2 \end{bmatrix}} \tag{20}$$

1.5.4 (d)

If U is singular in floating-point arithmetic, we know that ϵ^2 entry should be zero. Similarly as (b), then the computed value of the ϵ^2 will be zero if it is smaller than the machine epsilon, thus:

$$\boxed{-\sqrt{\epsilon_{mach}} < \epsilon < \sqrt{\epsilon_{mach}}} \tag{21}$$

2 Programming Problems

2.1 (a)

[1]:
```
import numpy as np
import numpy.linalg as la
import scipy.linalg as sla
```

[2]:
```
# Programming Problem 1
A = np.array([
    [2,4,-2],
    [4,9,-3],
    [-2,-1,7]
])
b = np.array([
    [2],
    [8],
    [10]
])
x1 = sla.solve(A,b)
x1
```

[2]:
```
array([[-7.00000000e+00],
       [ 4.00000000e+00],
       [-3.88578059e-16]])
```

Considering the computational mechanism of Python, the answer should be $\boxed{\text{[-7,4,0]}}$ instead.

2.2 (b)

[3]:
```
# Programming Problem 2
c = np.array([
    [4],
    [8],
    [-6]
])
(lu,piv) = sla.lu_factor(A)
x2 = sla.lu_solve((lu,piv),c)
x2
```

[3]:
```
array([[-1.],
       [ 1.],
       [-1.]])
```

So the answer would be $\boxed{[-1, 1, -1]}$ as shown above.