Chapter 9
Adequacy of Solutions

After reading this chapter, you will be able to

- *Know the difference between ill conditioned and well conditioned system of equations*
- *Define the norm of a matrix*
- *Relate the norm of the matrix and of its inverse to the ill or well conditioning of the matrix, that is, how much trust can you have in the solution of the matrix.*

What does it mean by ill conditioned and well-conditioned system of equations?

A system of equations is considered to be **well conditioned** if a small change in the coefficient matrix or a small change in the right hand side results in a small change in the solution vector.

A system of equations is considered to be **ill conditioned** if a small change in the coefficient matrix or a small change in the right hand side results in a large change in the solution vector.

Example

Is this system of equations well conditioned?

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

Solution

The solution to the above set of equations is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Make a small change in the right hand side vector of the equations

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7.998 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -3.999 \\ 4.000 \end{bmatrix}$$

Make a small change in the coefficient matrix of the equations

$$\begin{bmatrix} 1.001 & 2.001 \\ 2.001 & 3.998 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3.994 \\ 0.001388 \end{bmatrix}$$

This system of equation "looks" ill conditioned as a small change in the coefficient matrix or the right hand side resulted in a large change in the solution vector.

Example

Is this system of equations well conditioned?

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

Solution

The solution to the above equations is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Make a small change in the right hand side vector of the equations.

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7.001 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1.999 \\ 1.001 \end{bmatrix}$$

Make a small change in the coefficient matrix of the equations.

$$\begin{bmatrix} 1.001 & 2.001 \\ 2.001 & 3.001 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2.003 \\ 0.997 \end{bmatrix}$$

This system of equation "looks" well conditioned as small changes in the coefficient matrix or the right hand side resulted in small changes in the solution vector.

So what if the system of equations is ill conditioning or well conditioning?

Well, if a system of equations is ill conditioned, we cannot trust the solution as much. Remember the velocity problem in Chapter 5. The values in the coefficient matrix are squares of time, etc. For example if instead of $a_{11} = 25$, you used $a_{11} = 24.99$, would you want it to make a huge difference in the solution vector. If it did, would you trust the solution?

Later we will see how much (quantifiable terms) we can trust the solution to a system of equations. Every invertible square matrix has a **condition number** and coupled with the **machine epsilon**, we can quantify how many significant digits one can trust in the solution.

To calculate condition number of an invertible square matrix, I need to know what norm of a matrix means. How is the norm of a matrix defined?

Just like the determinant, the norm of a matrix is a simple unique scalar number. However, norm is always positive and is defined for all matrices – square or rectangular; invertible or noninvertible square matrices.

One of the popular definitions of a norm is the row sum norm (also called uniform-matrix norm). For a mxn matrix [A], the row sum norm of [A] is defined as

$$\|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|$$

that is, find the sum of the absolute of the elements of each row of the matrix [A]. The maximum out of the 'm' such values is the row sum norm of the matrix [A].

Example

Find the row sum norm of the following matrix [A].

$$A = \begin{bmatrix} 10 & -7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{bmatrix}$$

Solution

$$\|A\|_\infty = \max_{1 \le i \le 3} \sum_{j=1}^{3} |a_{ij}|$$

$$\max[(|10| + |-7| + |0|), (|-3| + |2.099| + |6|), (|5| + |-1| + |5|)]$$

$$= \max[(10 + 7 + 0), (3 + 2.099 + 6), (5 + 1 + 5)]$$

$$= \max[17, 11.099, 11]$$

$$= 17.$$

How is norm related to the conditioning of the matrix?

Let us start answering this question using an example. Go back to the "ill conditioned" system of equations,

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

that gives the solution as

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Denoting the above set of equations as

$$[A][X] = [C]$$

$$\|X\|_\infty = 2$$

$$\|C\|_\infty = 7.999$$

Making a small change in the right hand side,

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7.998 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -3.999 \\ 4.000 \end{bmatrix}$$

Denoting the above set of equations by

$$[A][X'] = [C']$$

and the change in right hand side vector

$$[\Delta \mathrm{C}] = [\mathrm{C}'] - [\mathrm{C}]$$

and the change in the solution vector as

$$[\Delta \mathrm{X}] = [\mathrm{X}'] - [\mathrm{X}]$$

then

$$[\Delta \mathrm{C}] = \begin{bmatrix} 4.001 \\ 7.998 \end{bmatrix} - \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

$$= \begin{bmatrix} 0.001 \\ -0.001 \end{bmatrix}$$

and

$$[\Delta \mathrm{X}] = \begin{bmatrix} -3.999 \\ 4.000 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -5.999 \\ 3.000 \end{bmatrix}$$

then

$$\|\Delta \mathrm{C}\|_\infty = 0.001$$

$$\|\Delta X\|_\infty = 5.999$$

Relative change in the norm of the solution vector is

$$\frac{\|\Delta X\|_\infty}{\|X\|_\infty} = \frac{5.999}{2}$$

$$= 2.9995$$

Relative change in the norm of the right hand side vector is

$$\frac{\|\Delta C\|_\infty}{\|C\|_\infty} = \frac{0.001}{7.999}$$

$$= 1.250 \times 10^{-4}$$

See the small relative change in the right hand side vector of 1.250 x 10^{-4} results in a large relative change in the solution vector as 2.995.

In fact, the ratio between the relative change in the norm of the solution vector to the relative change in the norm of the right hand side vector is

$$\frac{\|\Delta X\|_\infty / \|X\|_\infty}{\|\Delta C\|_\infty / \|C\|_\infty}$$

$$= \frac{2.9995}{1.250 \times 10^{-4}}$$

$$= 23957$$

Let us now go back to the “well-conditioned” system of equations.

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Denoting the system of equations as

$$[A][X] = [C]$$

$$\|X\|_\infty = 2$$

$$\|C\|_\infty = 7$$

Making a small change in the right hand side vector

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7.001 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1.999 \\ 1.001 \end{bmatrix}$$

Denoting the above set of equations as

$$[A][X'] = [C']$$

and the change in the right hand side vector

$$[\Delta C] = [C'] - [C]$$

and the change in the solution vector as

$$[\Delta X] = [X'] - [X]$$

then

$$[\Delta C] = \begin{bmatrix} 4.001 \\ 7.001 \end{bmatrix} - \begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 0.001 \\ 0.001 \end{bmatrix}$$

and

$$[\Delta X] = \begin{bmatrix} 1.999 \\ 1.001 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.001 \\ 0.001 \end{bmatrix}$$

then

$$\|\Delta C\|_\infty = 0.001$$

$$\|\Delta X\|_\infty = 0.001$$

Relative change in the norm of solution vector is

$$\frac{\|\Delta X\|_\infty}{\|X\|_\infty}$$

$$= \frac{0.001}{2}$$

$$= 5 \times 10^{-4}$$

Relative change in the norm of the right hand side vector is

$$\frac{\|\Delta C\|_\infty}{\|C\|_\infty}$$

$$= \frac{0.001}{7}$$

$$= 1.429 \times 10^{-4}$$

See a small relative change in the right hand side vector norm of 1.429 x 10^{-4} results in a small relative change in the solution vector norm of 5 x 10^{-4}.

In fact, the ratio between the relative change in the norm of the solution vector to the relative change in the norm of the right hand side vector is

$$\frac{\|\Delta X\|_\infty / \|X\|_\infty}{\|\Delta C\|_\infty / \|C\|_\infty}$$

$$= \frac{5 \times 10^{-4}}{1.429 \times 10^{-4}}$$

$$= 3.5$$

What are some of the properties of norms?

1. For a matrix [A], $\|A\| \geq 0$

2. For a matrix [A] and a scalar k, $\|kA\| = |k|\|A\|$

3. For two matrices [A] and [B] of same order,

$$\|A+B\| \le \|A\| + \|B\|$$

4. For two matrices [A] and [B] that can be multiplied as [A] [B],

$$\|AB\| \le \|A\|\,\|B\|$$

Is there a general relationship that exists between $\|\Delta X\|/\|X\|$ and $\|\Delta C\|/\|C\|$ or between $\|\Delta X\|/\|X\|$ and $\|\Delta A\|/\|A\|$? If so, it could help us identify well-conditioned and ill conditioned system of equations.

If there is such a relationship, will it help us quantify the conditioning of the matrix, that is, tell us how many significant digits we could trust in the solution of a system of simultaneous linear equations?

There is a relationship that exists between

$\frac{\|\Delta X\|}{\|X\|}$ and $\frac{\|\Delta C\|}{\|C\|}$, and between

$\frac{\|\Delta X\|}{\|X\|}$ and $\frac{\|\Delta A\|}{\|A\|}$.

These relationships are

$$\frac{\|\Delta X\|}{\|X+\Delta X\|} \le \|A\|\,\|A^{-1}\|\frac{\|\Delta C\|}{\|C\|}.$$

$$\frac{\|\Delta X\|}{\|X\|} \le \|A\|\,\|A^{-1}\|\frac{\|\Delta A\|}{\|A\|}$$

Looking at the above two inequalities, it shows that the relative change in the norm of the right hand side vector or the coefficient matrix can be amplified by as much as $\|A\|\,\|A^{-1}\|$.

This number $\|A\|\,\|A^{-1}\|$ is called the **condition number** of the matrix and coupled with the machine epsilon, we can quantify the accuracy of the solution of [A][X] = [C].

Prove for $[A][X]=[C]$

$$\frac{\|\Delta X\|}{\|X+\Delta X\|} \leq \|A\| \|A^{-1}\| \frac{\|\Delta A\|}{\|A\|}.$$

Proof

Let

$$[A][X] = [C] \qquad (1)$$

Then if [A] is changed to $[A']$, then[X] will change to $[X']$, such that

$$[A'][X'] = [C] \qquad (2)$$

From equations (1) and (2),

$$[A][X] = [A'][X']$$

Denoting change in [A] and [X] matrices as

$$[\Delta A] = [A'] - [A]$$

$$[\Delta X] = [X'] - [X]$$

then

$$[A][X] = ([A] + [\Delta A])([X] + [\Delta X])$$

Expanding the above expression

$$[A][X] = [A][X] + [A][\Delta X] + [\Delta A][X] + [\Delta A][\Delta X]$$

$$0 = [A][\Delta X] + [\Delta A]([X] + [\Delta X])$$

$$-[A][\Delta X] = [\Delta A]([X] + [\Delta X])$$

$$[\Delta X] = -[A]^{-1}[\Delta A]([X] + [\Delta X])$$

Applying the theorem of norms that norm of multiplied matrices is less than the multiplication of the individual norms of the matrices,

$$\|\Delta X\| \leq \|A^{-1}\| \|\Delta A\| \|X + \Delta X\|$$

Multiplying both sides by $\|A\|$

$$\|A\| \|\Delta X\| \leq \|A\| \|A^{-1}\| \|\Delta A\| \|X + \Delta X\|$$

$$\frac{\|\Delta X\|}{\|X+\Delta X\|} \leq \|A\| \|A^{-1}\| \frac{\|\Delta A\|}{\|A\|}$$

How do I use the above theorems to find how many significant digits are correct in my solution vector?

Relative error in solution vector is <= Cond (A)* relative error in right hand side.

Possible relative error in the solution vector is <= Cond (A) * $\in_{mach}$

Hence Cond (A)*$\in_{mach}$ should give us the number of significant digits at least correct in our solution by comparing it with $0.5\text{x}10^{-m}$.

Example

How many significant digits can I trust in the solution of the following system of equations?

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

Solution

$$[A] = \begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix}$$

$$[A]^{-1} = \begin{bmatrix} -3999.31 & 2000.1 \\ 2000.1 & -1000.1 \end{bmatrix}$$

$$\|A\|_\infty = 5.999$$

$$\|A^{-1}\|_\infty = 5999.4$$

$$\begin{aligned} \text{Cond}(A) &= \|A\|_\infty \|A^{-1}\|_\infty \\ &= 5.999 \times 5999.4 \\ &= 35990 \end{aligned}$$

Assuming single precision with 24 bits used in the mantissa for real numbers, the machine epsilon

$$\begin{aligned} \in_{mach} &= 2^{1-24} \\ &= 0.119209 \times 10^{-6} \end{aligned}$$

$$Cond(A) * \in_{mach}$$

$$= 35990 \times 0.119209 \times 10^{-6}$$

$$= 0.4290 x 10^{-2}$$

Comparing it with 0.5×10^{-m}

$0.5\times10^{-m} < 0.4290\times10^{-2}$

$m\leq2$

So two significant digits are at least correct in the solution vector.

Example

How many significant digits can I trust in the solution of the following system of equations?

$$\begin{bmatrix}1 & 2\\ 2 & 3\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}4\\ 7\end{bmatrix}$$

Solution

For

$$[A]=\begin{bmatrix}1 & 2\\ 2 & 3\end{bmatrix}$$

it can be shown

$$[A]^{-1}=\begin{bmatrix}-3 & 2\\ 2 & -1\end{bmatrix}$$

Then

$\|A\|_\infty = 5$,

$\|A^{-1}\|_\infty = 5$.

Cond (A) = $\|A\|_\infty \|A^{-1}\|_\infty$

$=5\times5$

$=25$

Assuming single precision with 24 bits of mantissa for real numbers, the machine epsilon

$\in_{mach} = 2^{1-24}$

$= 0.119209\times10^{-6}$

$Cond(A)*\in_{mach}$

$= 25\times0.119209\times10^{-6}$

$= 0.2980x10^{-5}$

Comparing it with 0.5×10^{-m}

$$0.5 \times 10^{-m} < 0.2980 \times 10^{-5}$$

$$m \leq 5$$

So five significant digits are at least correct in the solution vector.

Key Terms

Ill conditioned	Well conditioned
Norm	Condition number
Machine epsilon	Significant digits

Homework

1. **The adequacy of the solution of simultaneous linear equations depends on**

    A. Condition number

    B. Machine epsilon

    C. Product of condition number and machine epsilon

    D. Norm of the matrix.

2. **If a system of equations [A] [X] = [C] is ill conditioned, then**

    A. det(A) = 0

    B. Cond(A) = 1

    C. Cond(A) is large.

    D. $\|A\|$ is large.

If Cond(A) = 10^4 and $\in_{mach}$ = 0.119 x 10^{-6}, then in [A] [X] = [C], at least these many significant digits are correct in your solution,

A. 3

B. 2

C. 1

D. 0

3.

Make a small change in the coefficient matrix to

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

and find

$$\frac{\|\Delta X\|_\infty / \|X\|_\infty}{\|\Delta A\|_\infty / \|A\|_\infty}$$

Is it a large or small number? How is this number related to the condition number of the matrix?

4.

Make a small change in the coefficient matrix to

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

and find

$$\frac{\|\Delta X\|_\infty / \|X\|_\infty}{\|\Delta A\|_\infty / \|A\|_\infty}$$

Is it a large or a small number? Compare your results with the previous problem. How is this number related to the condition number of the matrix?

5. Prove

$$\frac{\|\Delta X\|}{\|X\|} \leq \|A\| \|A^{-1}\| \frac{\|\Delta C\|}{\|C\|}$$

6.

For

$$[A]=\begin{bmatrix}10 & -7 & 0\\ -3 & 2.099 & 6\\ 5 & -1 & 5\end{bmatrix}$$

gives

$$[A]^{-1}=\begin{bmatrix}-0.1099 & -0.2333 & 0.2799\\ -0.2999 & -0.3332 & 0.3999\\ 0.04995 & 0.1666 & 6.664\times10^{-5}\end{bmatrix}$$

a) What is the condition number of [A]?

b) How many significant digits can we at least trust in the solution of $[A][X]=[C]$ **if**

$\in_{mach}=0.1192\times10^{-6}$

c)Without calculating the inverse of the matrix [A], can you estimate the condition number of [A] using the theorem in Problem#6?

Answer: a) $\|A\|=17$

$\|A^{-1}\|=1.033$

Cond (A) = 17.56

b) 5

c) Try different values of right hand side of C= $[\pm1 \quad \pm1 \quad \pm1]^T$ ***with signs chosen randomly. Then*** $\|A^{-1}\| \le \|X\|$ ***obtained from solving equation set [A] [X] = [C] as*** $\|C\|=1$.

7.

Prove that the Cond(A) >=1.

Hint: We know that

$\|A\,B\| \le \|A\|\,\|B\|$

then if $[B]=[A]^{-1}$,

$\|A\,A^{-1}\| \le \|A\|\,\|A^{-1}\|$

$\|I\| \le \|A\|\,\|A^{-1}\|$

$1 \le \|A\|\,\|A^{-1}\|$

$\|A\|\,\|A^{-1}\| \ge 1$

Cond (A) ≥ 1.