Adequacy of Solutions

Autar Kaw
Humberto Isaza

http://nm.MathForCollege.com
Transforming Numerical Methods Education for STEM Undergraduates

Adequacy of Solutions

http://nm.MathForCollege.com

Objectives

1. *differentiate between ill-conditioned and well-conditioned systems of equations,*
2. *define the norm of a matrix,*
3. *define the condition number of a square matrix,*
4. *relate the condition number to the ill or well conditioning of a system of equations, that is, determine how much trust you can trust the solution of a set of equations.*

Well-conditioned and ill-conditioned

What do you 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 1

Is this system of equations well-conditioned?

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

Example 1 (cont.)

Solution

The solution to the 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

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

Example 1 (cont.)

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 last systems of equation “looks” ill-conditioned because a small change in the coefficient matrix or the right hand side resulted in a large change in the solution vector.

Example 2

Is this system of equations well-conditioned?

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

Example 2 (cont.)

Solution

The solution to the previous 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

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

Example 2 (cont.)

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 because small changes in the coefficient matrix or the right hand side resulted in small changes in the solution vector.

Well-conditioned and ill-conditioned

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

Well, if a system of equations is ill-conditioned, we cannot trust the solution as much. Revisit the velocity problem, Example 5.1 in Chapter 5. The values in the coefficient matrix $[A]$ are squares of time, etc. For example, if instead of $a_{11} = 25$, you used $a_{11} = 24.99$, would you want this small change 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 in 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.

Condition number

To calculate the condition number of an invertible square matrix, I need to know what the 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, the norm is always positive and is defined for all matrices – square or rectangular, and invertible or noninvertible square matrices.

One of the popular definitions of a norm is the row sum norm (also called the uniform-matrix norm). For a $m \times n$ 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 value of the elements of each row of the matrix . The maximum out of the such values is the row sum norm of the matrix .

Example 3

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

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

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

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

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

$$= 17.$$

How is the 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

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

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

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

How is the norm related to the conditioning of the matrix? (cont.)

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

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

right hand side vector is found by

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

and the change in the solution vector is found by

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

How is the norm related to the conditioning of the matrix? (cont.)

then

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

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

and

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

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

then

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

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

How is the norm related to the conditioning of the matrix? (cont.)

The relative change in the norm of the solution vector is

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

$$= 2.9995$$

The relative change in the norm of the right hand side vector is

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

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

How is the norm related to the conditioning of the matrix? (cont.)

See the small relative change of 1.250×10^{-4} in the right hand side vector results in a large relative change in the solution vector as 2.9995.
In fact, the ratio between the relative change in the norm of the solution vector and the relative change in the norm of the right hand side vector is

$$\frac{\left\|\Delta X\right\|_\infty / \left\|X\right\|_\infty}{\left\|\Delta C\right\|_\infty / \left\|C\right\|_\infty} = \frac{2.9995}{1.250\times 10^{-4}}$$

$$= 23993$$

How is the norm related to the conditioning of the matrix? (cont.)

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 by

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

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

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

How is the norm related to the conditioning of the matrix? (cont.)

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 by

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

the change in the right hand side vector is then found by

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

How is the norm related to the conditioning of the matrix? (cont.)

and the change in the solution vector is

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

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

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

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

How is the norm related to the conditioning of the matrix? (cont.)

then

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

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

The relative change in the norm of solution vector is

$$\frac{\|\Delta X\|_\infty}{\|X\|_\infty} = \frac{0.001}{2}$$

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

The relative change in the norm of the right hand side vector is

$$\frac{\|\Delta C\|_\infty}{\|C\|_\infty} = \frac{0.001}{7}$$

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

How is the norm related to the conditioning of the matrix? (cont.)

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

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

$$\frac{\left\|\Delta X\right\|_{\infty}/\left\|X\right\|_{\infty}}{\left\|\Delta C\right\|_{\infty}/\left\|C\right\|_{\infty}}=\frac{5\times10^{-4}}{1.429\times10^{-4}}$$

$$=3.5$$

Properties of norms

What are some 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, $\|A + B\| \leq \|A\| + \|B\|$
4. For two matrices $[A]$ and $[B]$ that can be multiplied as $[A][B]$, $\|AB\| \leq \|A\|\|B\|$

Identifying well-conditioned and ill conditioned system of equations

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, will it tell us how many significant digits we could trust in the solution of a system of simultaneous linear equations?

Identifying well-conditioned and ill conditioned system of equations (cont.)

there is a relationship that exists between

$$\frac{\|\Delta X\|}{\|X\|} \text{ and } \frac{\|\Delta C\|}{\|C\|}$$

and between

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

these relationships are

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

Identifying well-conditioned and ill conditioned system of equations (cont.)

and

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

the above two inequalities show 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 $[A][X] = [C]$

Proof

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

that

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

Proof

let

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

then $[A]$ is changed to $[A']$ the $[X]$ will change to $[X']$ such that

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

Proof (cont.)

From Equations (1) and (2),

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

Denoting change in $[A]$ and $[X]$ matrices as $[\Delta A]$ and $[\Delta X]$, respectively

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

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

then

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

Proof (cont.)

Expanding the previous 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 the norm of multiplied matrices is less than the multiplication of the individual norms of the matrices,

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

Proof (cont.)

Multiplying both sides by $\|A\|$

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

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

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

the relative error in a solution vector is Cond (A) relative error in right hand side.
the possible relative error in the solution vector is $\le$ Cond (A) $\times \in_{mach}$

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

Example 4

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

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

Example 4 (cont.)

Solution

For

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

it can be show

$$[A]^{-1}=\begin{bmatrix}-3999 & 2000\\ 2000 & -1000\end{bmatrix}$$

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

$$\|A^{-1}\|_{\infty}=5999$$

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

$$=5.999\times 5999.4$$

$$=35990$$

Example 4 (cont.)c

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

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

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

$$Cond(A) \times \in_{mach} = 35990 \times 0.119209 \times 10^{-6}$$

$$= 0.4290 \times 10^{-2}$$

comparing it with 0.5×10^{-m}

$$0.5 \times 10^{-m} < 0.4290 \times 10^{-2}$$

$$m \leq 2$$

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

Example 5

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

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

Example 5 (cont.)

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

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

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

$$\begin{aligned} \text{Cond}\,(\text{A}) &= \|A\|_\infty \|A^{-1}\|_\infty \\ &= 5 \times 5 \\ &= 25 \end{aligned}$$

Example 5 (cont.)

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

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

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

$$Cond(A) \times \in_{mach} = 25 \times 0.119209 \times 10^{-6}$$

$$= 0.2980 \times 10^{-5}$$

Comparing it with 0.5×10^{-m}

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

$$m \leq 5$$

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

Key terms

Ill-Conditioned matrix
Well-Conditioned matrix
Norm
Condition Number
Machine Epsilon
Significant Digits