

ILL-Conditioned Systems

The solutions of some linear systems (that can be represented by systems of linear equations) are more sensitive to round-off error than others. For some linear systems a small change in one of the values of the coefficient matrix or the right-hand side vector causes a large change in the solution vector.

Consider the following system of two linear equations in two unknowns.

$$\begin{bmatrix} 400 & -201 \\ -800 & 401 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 200 \\ -200 \end{bmatrix}$$

The system can be solved by using previously covered methods and the solution is $x_1 = -100$ and $x_2 = -200$

Now, let us make a slight change in one of the elements of the coefficient matrix. Change A_{11} from 400 to 401 and see how this small change affects the solution of the following.

$$\begin{bmatrix} 401 & -201 \\ -800 & 401 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 200 \\ -200 \end{bmatrix}$$

This time the solution is $x_1 = 40000$ and $x_2 = 79800$

With a modest change in one of the coefficients one would expect only a small change in the solution. However, in this case the change in solution is quite significant. It is obvious that in this case the solution is very sensitive to the values of the coefficient matrix A .

When the solution is highly sensitive to the values of the coefficient matrix A or the right-hand side constant vector b , the equations are called to be ill-conditioned. Ill-conditioned systems pose particular problems where the coefficients or constants are estimated from experimental results or from a mathematical model. Therefore, we cannot rely on the solutions coming out of an ill-conditioned system. The problem is then how do we know when a system of linear equations is ill-conditioned. To do that we have to first define *vector* and *matrix norms*.

Vector and Matrix Norms

A norm of a vector is a measure of its length or magnitude. There are, however, several ways to define a vector norm. For the purpose of this discussion we will use a computationally simple formulation of a vector norm in the following manner.

$$\|x\| \triangleq \max_{k=1}^n \{x_k\} \quad (1)$$

The notation $\max\{ \}$ denotes the maximum element of the set. The formulation shown by Eqn. (1) is also called the *infinity norm* of the vector x . Note the following properties of infinity norm.

$$\begin{aligned}\|x\| &> 0 \text{ for } x \neq 0 \\ \|x\| &= 0 \text{ for } x = 0 \\ \|\alpha x\| &= \|\alpha\| \cdot \|x\| \\ \|x + y\| &\leq \|x\| + \|y\|\end{aligned}\tag{2}$$

The last property is called the *triangle inequality*.

Now we need to consider the notion of a matrix norm. A matrix norm can be defined in terms of a vector norm in the following manner.

$$\|A\| \triangleq \max_{k=1}^n \left\{ \sum_{j=1}^n |A_{kj}| \right\}$$

Note that the expression for $\|A\|$ involves summing the absolute values of elements in the rows of A .

Consider the following linear algebraic system.

$$\begin{bmatrix} 2 & 3 & -7 \\ 5 & 4 & -2 \\ 7 & -3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ -7 \\ 11 \end{bmatrix}$$

We have $\|b\| = \max\{3, 7, 11\} = 11$ and $\|A\| = \max\{12, 11, 16\} = 16$

The matrix norm satisfies all the properties of a vector norm and, in addition, a matrix norm has the following important property.

$$\|Ax\| \leq \|A\| \cdot \|x\|\tag{3}$$

Condition Number

Let us investigate first, how a small change in the b vector changes the solution vector. x is the solution of the original system and let $x + \Delta x$ is the solution when b changes from b to $b + \Delta b$.

The we can write

$$A(x + \Delta x) = b + \Delta b$$

$$\text{or, } Ax + A\Delta x = b + \Delta b$$

But because $Ax = b$, it follows that $A\Delta x = \Delta b$.

$$\text{i.e., } \Delta x = A^{-1}\Delta b$$

By using the relationship shown in Eqn. (3) we can write that

$$\|A^{-1}\Delta b\| \leq \|A^{-1}\| \cdot \|\Delta b\|$$

$$\text{i.e., } \|\Delta x\| \leq \|A^{-1}\| \cdot \|\Delta b\| \quad (4)$$

Again using Eqn. (3) to the original system, $Ax = b$ we can write that

$$\|Ax\| \leq \|A\| \cdot \|x\|$$

$$\text{i.e., } \|b\| \leq \|A\| \cdot \|x\|$$

$$\text{or, } \|A\| \cdot \|x\| \geq \|b\| \quad (5)$$

Divide Eqn. (4) by Eqn. (5)

$$\frac{\|\Delta x\|}{\|A\| \cdot \|x\|} \leq \frac{\|A^{-1}\| \cdot \|\Delta b\|}{\|b\|}$$

$$\text{i.e., } \frac{\|\Delta x\|}{\|x\|} \leq \frac{\|A\| \cdot \|A^{-1}\| \cdot \|\Delta b\|}{\|b\|} \text{ or, } \frac{\|\Delta x\|}{\|x\|} \leq \|A\| \cdot \|A^{-1}\| \cdot \frac{\|\Delta b\|}{\|b\|}$$

$$\text{or, } \frac{\|\Delta x\|}{\|x\|} \leq K(A) \cdot \frac{\|\Delta b\|}{\|b\|} \quad (6)$$

where $K(A)$ is called the condition number of the matrix A and is defined as

$$K(A) \triangleq \|A\| \cdot \|A^{-1}\| \text{ provided } A \text{ is nonsingular.}$$

$K(A)$ is a measure of the relative sensitivity of the solution to changes in the right-hand side vector b . Eqn. (6) gives the upper bound of the relative change

Now, let us investigate what happens if a small change is made in the coefficient matrix A . Consider A is changed to $A + \Delta A$ and the solution changes from x to $x + \Delta x$. $(A + \Delta A)(x + \Delta x) = b$. It can be shown that the changes in the solution can be expressed in the following manner.

$$\frac{\|\Delta x\|}{\|x + \Delta x\|} \leq K(A) \frac{\|\Delta A\|}{\|A\|}$$

When the condition number $K(A)$ becomes large, the system is regarded as being *ill-conditioned*. Matrices with condition numbers near 1 are said to be *well-conditioned*.

In our previous example

$$A = \begin{bmatrix} 400 & -201 \\ -800 & 401 \end{bmatrix} \text{ and the corresponding inverse } A^{-1} = \begin{bmatrix} -1.002 & -0.502 \\ -2 & -1 \end{bmatrix}$$

$$\|A\| = 1201 \text{ and } \|A^{-1}\| = 3$$

$$\text{Condition number } K(A) \triangleq \|A\| \cdot \|A^{-1}\| = 3603$$

Scaling

Large condition numbers can also arise from equations that are in need of scaling.

Consider the following coefficient matrix which corresponds to one 'regular' equation and one 'large' equation.

$$A = \begin{bmatrix} 1 & -1 \\ 1000 & 1000 \end{bmatrix}$$

In this case the inverse of the matrix is:

$$A^{-1} = \begin{bmatrix} 0.5 & 0.0005 \\ -0.5 & 0.0005 \end{bmatrix}$$

$$\|A\| = \max\{2, 2000\} = 2000 \text{ and } \|A^{-1}\| = \max\{0.5005, 0.5005\}. \text{ The condition number is:}$$

$$K(A) = 2000 \times 0.5005 = 1001$$

Scaling (called Equilibration) can be used to reduce the condition number for a system that is poorly scaled. If each row of A is scaled by its largest element, then the new A and its inverse become

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad A^{-1} = \begin{bmatrix} 0.5 & 0.5 \\ -0.5 & 0.5 \end{bmatrix}$$

The condition number of the scaled system is 2.

We have just mentioned that when the condition number $K(A)$ becomes large, the system is regarded as being *ill-conditioned*. But, how large does $K(A)$ have to be before a system is regarded as *ill-conditioned*? There is no clear threshold. However, to assess the effects of ill-conditioning, a rule of thumb can be used. For a system with condition number $K(A)$, expect a loss of roughly $\log_{10} K(A)$ decimal places in the accuracy of the solution. Therefore for the system with

$$A = \begin{bmatrix} 400 & -201 \\ -800 & 401 \end{bmatrix} \text{ whose condition number } K(A) \text{ is } 3603 \text{ expect to lose 3 decimal places in accuracy.}$$

IEEE standard double precision numbers have about 16 decimal digits of accuracy, so if a matrix has a condition number of 10^{10} , you can expect only six digits to be accurate in the

answer. An important aspect of conditioning is that, in general, residuals $\mathbf{R} = \mathbf{Ax} - \mathbf{b}$ are reliable indicators of accuracy only if the problem is well-conditioned.

THE END