Condition and Condition Number of a Matrix

In the solution of a linear system of equations or equivalently a matrix-vector equation of the form

$$A\underline{x} = \underline{b},$$

where A is a matrix, $\underline{b}$ is the known vector and $\underline{x}$ is the vector that is the solution of the process, we know that if the matrix A is singular then there is either a range of solutions or no solution, but in either case standard method for solving the equation – such as Gaussian elimination[1] or LU factorisation[2] - will fail. However, in cases when the matrix is *nearly* singular (within some concept of numerical accuracy in computing) then the expectation is that its automatic solution could also be problematic, in that small changes in the problem data can result in large changes in the solution. Given that the data that make up the matrix-vector system in most cases are prone to error, whether this arises for example from numerical approximation, measurement error or rounding error, then the sensitivity in the matrix-vector system will be expected to magnify the error and the significance of the solution that arises would be harmed.

In numerical analysis the concept of the *condition* of the matrix seeks to assess the quality of a matrix for the purposes of solving the linear system of equations above. The determination of a *condition number* is a quantification of the condition of the matrix; a matrix with a *low* condition number is deemed to be *well-conditioned* and now problems are envisaged in the solution and a matrix with a *very* high condition number is termed *ill-conditioned* and the 'solution' of the matrix-vector system has to be treated with care.

Let us consider an example. The following equation is 'close' to being singular in that the second row is similar to the first row:

$$\begin{pmatrix} 1 & 1 \\ 0.999 & 1.001 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

but the solution is strictly $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Let us now introduce an error of 0.000001 in $\underline{b}$

$$\begin{pmatrix} 1 & 1 \\ 0.999 & 1.001 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2.0000001 \\ 2 \end{pmatrix}$$

the solution is $\begin{pmatrix} 1.000501 \\ 0.999500 \end{pmatrix}$ (to 6 decimal places); an error of 0.000001 in the data results in an error of 0.0005 in the solution – a magnification of 500.

Similarly, if we introduce an error of 0.000001 in an element of A

$$\begin{pmatrix} 1 & 1.000001 \\ 0.999 & 1.001 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

the solution is $\begin{pmatrix} 0.999499 \\ 1.000500 \end{pmatrix}$; a similar magnification of the error to the previous example.

[1] Gaussian Elimination
[2] LU Factorisation

The magnification of the error through small changes in the input data can be illustrated by the method of plotting the graph in solving simultaneous equations[3]. When the two equations are similar, the lines that correspond to the equations lie close together, as shown in the following illustration.

In this case, the solution, where the lines cross, is not clear-cut. If there are small changes in the equations that define the lines then the apparent solution is disproportionately changed.

If the determinant of a matrix is zero then the matrix is singular. Hence, if the determinant is evaluated and a value of zero is obtained then that is an indication of problems in the ensuing solution of the linear system of equations. It may be thought therefore that a 'small' determinant indicates an ill-conditioned matrix. This is in fact not strictly the case; it is possible for a matrix to have a small determinant but still be well conditioned. For example a large diagonal matrix with all its diagonal elements equal to 0.1 has a 'small' determinant of 10^{-n} (where n is the dimension of the matrix), but for practical computing purposes it is clearly not significantly different from the identity matrix and does not present any problems when it is solved over.

The condition number relates the sensitivity of the relative error in the solution $\underline{x}$ to relative errors in the data $\underline{b}$. Let $\underline{\delta}$ represent the error in $\underline{b}$ and let $\underline{\varepsilon}$ be the corresponding error in the solution $\underline{x}$:

$$A(\underline{x} + \underline{\varepsilon}) = \underline{b} + \underline{\delta}.$$

Given that $A\underline{x} = \underline{b}$ then $A\underline{\varepsilon} = \underline{\delta}$ also follows from the equation above.

In the following analysis the condition number is defined using the concepts of vector norm[4] and matrix norm[5]. The relative error in the right hand side is $\frac{\|\underline{\delta}\|}{\|\underline{b}\|} = \frac{\|A\underline{\varepsilon}\|}{\|A\underline{x}\|}$ and the corresponding relative error in the solution is $\frac{\|\underline{\varepsilon}\|}{\|\underline{x}\|}$. Let us now consider the ratio of the relative error in the solution to the relative error in the input:

$$\frac{\|\underline{\varepsilon}\|\|\underline{b}\|}{\|\underline{x}\|\|\underline{\delta}\|} = \frac{\|A^{-1}\underline{\delta}\|\|A\underline{x}\|}{\|\underline{x}\|\|\underline{\delta}\|} \le \frac{\|A^{-1}\|\|\underline{\delta}\|\|A\|\|\underline{x}\|}{\|\underline{x}\|\|\underline{\delta}\|} = \|A^{-1}\|\|A\|,$$

where A^{-1} is the inverse[6] of the matrix A.

The quantity $\|A^{-1}\|\|A\|$ is termed the condition number of the matrix A and we often write:

$$\kappa(A) = \|A^{-1}\|\|A\|.$$

Note that the value of the condition number depends on the matrix norm. However, the purpose of the condition number is to provide an indication of the suitability of the matrix for solving over and to estimate the loss of significance. These purposes can be met with any of the different types of norm, and hence if there is a choice of matrix norm

[3] Graphical solution of simultaneous equations
[4] Vector Norm
[5] Matrix Norm
[6] Identity and Inverse Matrix

in the evaluation of the condition number then it is recommended that the most convenient or most efficient one is chosen.

Examples of determining the condition number of 2x2 and 3x3 matrices are given below. The Excel spreadsheet[7] contains these examples and allows experimentation with the elements of the matrix A to investigate its effect on condition number.

Example: The Condition Number of a 2x2 Matrix

For the matrix $A = \begin{pmatrix} 3 & -2 \\ -1 & 4 \end{pmatrix}$, $\|A\|_1 = 6$, $\|A\|_\infty = 5$, $\|A\|_F = \sqrt{30} = 5.4772$ to four decimal places.

The inverse $A^{-1} = \begin{pmatrix} 0.4 & 0.2 \\ 0.1 & 0.3 \end{pmatrix}$, $\|A\|_1 = 0.5$, $\|A\|_\infty = 0.6$, $\|A\|_F = 0.54772$ to four decimal places.

Hence the condition number in the 1-norm is $\kappa_1(A) = 6 \times 0.5 = 3$. Similarly $\kappa_\infty(A) = 5 \times 0.6 = 3$ and $\kappa_F(A) = 5.4772 \times 0.54772 = 3$.

This example and another example for the condition number of a 3x3 matrix can be found on the associated spreadsheet.

[7] Condition Number (Spreadsheet)