Lecture Eight: Numerical Solution of Non-Linear Systems - Multivariate Newton-Raphson Algorithm

Solving Systems of Non-Linear Equations Using the Multivariate Newton-Raphson Algorithm

Let's assume we have the following system of non-linear equations:

$$f_1(x_1, x_2, ...x_k) = 0$$
$$f_2(x_1, x_2, ...x_k) = 0$$
$$.$$
$$.$$
$$.$$
$$f_k(x_1, x_2, ...x_k) = 0$$

where the above functions are real differentiable functions (we have k equations). Each function contains k variables.

To find a solution for the above system, we can use the multivariate version of the Newton-Raphson algorithm. In other words, we aim to find the values of $x_1, x_2, ...x_k$ that satisfy all the equations above being equal to zero at the same time.

To simplify understanding of the iterative procedure, suppose the system consists of only two equations, and each equation contains only two unknowns. Then we can write the system as follows:

$$f_1(x, \mathrm{y}) = 0$$
$$f_2(x, \mathrm{y}) = 0$$

The aim is to find the values of x, y that satisfy the above system being equal to zero.

Idea of the Algorithm: Let x_1, y_1 be the initial values for the unknowns x, y, respectively. It is preferable to choose these values such that the function values are close to zero when substituted with the initial values, i.e.,

$$f_1(x_1, y_1) \cong 0$$
$$f_2(x_1, y_1) \cong 0$$

In other words, it is preferable to choose the initial values so that they are close to the actual values of the roots.

After that, we plot the function f_1 (the graph of the function will be a surface, possibly irregular or non-planar, since it is a non-linear function). We then draw a plane surface $p_1 = (x, y)$ tangent to the surface of the function f_1 at the point $(x_1, y_1, f_1(x_1, y_1))$ (note that we have three coordinates in this case due to the presence of two variables in the function), such that this plane intersects the x-axis and y-axis at the points x_2, y_2

The same applies to the function f_2, where we draw the surface of the function and generate a plane surface $q_1(x, y)$ tangent to the surface of the function f_2 at the point $(x_1, y_1, f_2(x_1, y_1))$ such that this plane intersects the x-axis and y-axis at the points x_2, y_2.

The equations of the tangent surfaces can be written as follows:

$$p_1(x,y) = f_1(x_1,y_1) + (x - x_1)\frac{\partial f_1(x_1,y_1)}{\partial x} + (y - y_1)\frac{\partial f_1(x_1,y_1)}{\partial y}$$

$$q_1(x,y) = f_2(x_1,y_1) + (x - x_1)\frac{\partial f_2(x_1,y_1)}{\partial x} + (y - y_1)\frac{\partial f_2(x_1,y_1)}{\partial y}$$

where

$$\frac{\partial f_1(x,y)}{\partial x}, \frac{\partial f_1(x,y)}{\partial y}, \frac{\partial f_2(x,y)}{\partial x}, \frac{\partial f_2(x,y)}{\partial y}$$

are the first partial derivatives of the functions f_1, f_2 with respect to the unknowns x, y respectively.

Since the points x_2, y_2 are the intersection points of the tangents with the x-axis and y-axis, this means that the values of the tangent equations at the points x_2, y_2 equal zero. That is:

$$p_1(x_2,y_2) = f_1(x_1,y_1) + (x_2 - x_1)\frac{\partial f_1(x_1,y_1)}{\partial x} + (y_2 - y_1)\frac{\partial f_1(x_1,y_1)}{\partial y} = 0$$

$$\Rightarrow (x_2 - x_1)\frac{\partial f_1(x_1,y_1)}{\partial x} + (y_2 - y_1)\frac{\partial f_1(x_1,y_1)}{\partial y} = -f_1(x_1,y_1) \quad \text{...............(1)}$$

$$q_1(x_2,y_2) = f_2(x_1,y_1) + (x_2 - x_1)\frac{\partial f_2(x_1,y_1)}{\partial x} + (y_2 - y_1)\frac{\partial f_2(x_1,y_1)}{\partial y} = 0$$

$$\Rightarrow (x_2 - x_1)\frac{\partial f_2(x_1,y_1)}{\partial x} + (y_2 - y_1)\frac{\partial f_2(x_1,y_1)}{\partial y} = -f_2(x_1,y_1) \quad \text{...............(2)}$$

Note that the equations (1) and (2) above can be transformed into matrix form as follows:

$$\begin{bmatrix} \dfrac{\partial f_1(x_1, \mathrm{y}_1)}{\partial x} & \dfrac{\partial f_1(x_1, \mathrm{y}_1)}{\partial y} \\ \dfrac{\partial f_2(x_1, \mathrm{y}_1)}{\partial x} & \dfrac{\partial f_2(x_1, \mathrm{y}_1)}{\partial y} \end{bmatrix} \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \end{bmatrix} = - \begin{bmatrix} f_1(x_1, \mathrm{y}_1) \\ f_2(x_1, \mathrm{y}_1) \end{bmatrix}$$

where the matrix J:

$$J(x_1, y_1) = \begin{bmatrix} \dfrac{\partial f_1(x_1, \mathrm{y}_1)}{\partial x} & \dfrac{\partial f_1(x_1, \mathrm{y}_1)}{\partial y} \\ \dfrac{\partial f_2(x_1, \mathrm{y}_1)}{\partial x} & \dfrac{\partial f_2(x_1, \mathrm{y}_1)}{\partial y} \end{bmatrix}$$

This is called the Jacobian matrix (or the Jacobian), which is a matrix of the first partial derivatives of the functions. Now, by multiplying both sides by the inverse of the Jacobian matrix, we obtain:

$$\begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \end{bmatrix} = - \begin{bmatrix} \dfrac{\partial f_1(x_1, \mathrm{y}_1)}{\partial x} & \dfrac{\partial f_1(x_1, \mathrm{y}_1)}{\partial y} \\ \dfrac{\partial f_2(x_1, \mathrm{y}_1)}{\partial x} & \dfrac{\partial f_2(x_1, \mathrm{y}_1)}{\partial y} \end{bmatrix}^{-1} \begin{bmatrix} f_1(x_1, \mathrm{y}_1) \\ f_2(x_1, \mathrm{y}_1) \end{bmatrix}$$

Note that:

$$\begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \end{bmatrix} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} - \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} - \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = -\begin{bmatrix} \dfrac{\partial f_1(x_1, y_1)}{\partial x} & \dfrac{\partial f_1(x_1, y_1)}{\partial y} \\ \dfrac{\partial f_2(x_1, y_1)}{\partial x} & \dfrac{\partial f_2(x_1, y_1)}{\partial y} \end{bmatrix}^{-1} \begin{bmatrix} f_1(x_1, y_1) \\ f_2(x_1, y_1) \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} - \begin{bmatrix} \dfrac{\partial f_1(x_1, y_1)}{\partial x} & \dfrac{\partial f_1(x_1, y_1)}{\partial y} \\ \dfrac{\partial f_2(x_1, y_1)}{\partial x} & \dfrac{\partial f_2(x_1, y_1)}{\partial y} \end{bmatrix}^{-1} \begin{bmatrix} f_1(x_1, y_1) \\ f_2(x_1, y_1) \end{bmatrix}$$

n general, for the non-linear solution, the formula is as follows:

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} x_{n-1} \\ y_{n-1} \end{bmatrix} - \begin{bmatrix} \dfrac{\partial f_1(x_{n-1}, y_{n-1})}{\partial x} & \dfrac{\partial f_1(x_{n-1}, y_{n-1})}{\partial y} \\ \dfrac{\partial f_2(x_{n-1}, y_{n-1})}{\partial x} & \dfrac{\partial f_2(x_{n-1}, y_{n-1})}{\partial y} \end{bmatrix}^{-1} \begin{bmatrix} f_1(x_{n-1}, y_{n-1}) \\ f_2(x_{n-1}, y_{n-1}) \end{bmatrix} \text{......(3)}$$

provided that the inverse of the matrix J exists.

The numerical solution is accepted if one of the following two conditions is satisfied:

1. $|x_n - x_{n-1}| \le \varepsilon_1, |y_n - y_{n-1}| \le \varepsilon_2$ meaning that both absolute errors must be within the allowable error.
2. The absolute values of the functions at the proposed roots x_n, y_n are less than $\varepsilon_1, \varepsilon_2$ respectively; that is:

$$|f_1(x_n, y_n)| \leq \varepsilon_1 \;\&\; |f_2(x_n, y_n)| \leq \varepsilon_2$$

where $\varepsilon_1, \varepsilon_2$ represent the upper limits of the allowable absolute error in the estimation of x, y respectively.

The steps of the Multivariate Newton-Raphson algorithm can be summarized as follows:

1. Set the allowable absolute error for each unknown $\varepsilon_1, \varepsilon_2$.
2. Choose an initial value for each root x_1, y_1.
3. Set $n = 2$.
4. Calculate the root according to the following formula:

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} x_{n-1} \\ y_{n-1} \end{bmatrix} - \begin{bmatrix} \dfrac{\partial f_1(x_{n-1}, \mathrm{y}_{n-1})}{\partial x} & \dfrac{\partial f_1(x_{n-1}, \mathrm{y}_{n-1})}{\partial y} \\ \dfrac{\partial f_2(x_{n-1}, \mathrm{y}_{n-1})}{\partial x} & \dfrac{\partial f_2(x_{n-1}, \mathrm{y}_{n-1})}{\partial y} \end{bmatrix}^{-1} \begin{bmatrix} f_1(x_{n-1}, \mathrm{y}_{n-1}) \\ f_2(x_{n-1}, \mathrm{y}_{n-1}) \end{bmatrix}$$

5. If $|x_n - x_{n-1}| \leq \varepsilon_1, |y_n - y_{n-1}| \leq \varepsilon_2$ at the same time, meaning that both absolute errors must be within the allowable error, or if the absolute values of the functions at the proposed roots x_n, y_n are less than $\varepsilon_1, \varepsilon_2$ respectively; that is:

$$|f_1(x_n, y_n)| \leq \varepsilon_1 \;\&\; |f_2(x_n, y_n)| \leq \varepsilon_2$$

Then we say that x_n, y_n are the required roots.

6. Otherwise, set $n = n + 1$ and go to step 4.

The inverse of a 2×2 matrix:

If we have the matrix M :

$$M = \begin{bmatrix} a & \mathrm{b} \\ c & \mathrm{d} \end{bmatrix}$$

Then its inverse is given by the formula:

$$M^{-1} = \frac{1}{ad - cd} \begin{bmatrix} \mathrm{d} & \mathrm{-b} \\ -c & \mathrm{a} \end{bmatrix}$$

Although it is straightforward to find the inverse using this formula, it is always preferable to use a programming language to compute the inverse.

Example: Find the values of x, y that satisfy the following system of non-linear equations equal to zero:

$$f_1(x_1, \mathrm{y}_1) = x^2 + y^2 - 1$$
$$f_1(x_1, \mathrm{y}_1) = xy$$

Using $\varepsilon_1, \varepsilon_2 = 0.0001$ and initial values $x_1 = 0.5, y_1 = 0.1$

Solution:

Step 1: The upper limit for the allowable absolute error is $\varepsilon_1, \varepsilon_2 = 0.0001$

Step 2: The initial values are $x_1 = 0.5, y_1 = 0.1$.

Step 3: Set $n = 2$.

Step 4: Calculate the root values:

We need to compute the partial derivatives

$$\begin{bmatrix} \frac{\partial f_1(x_1, \mathrm{y}_1)}{\partial x} & \frac{\partial f_1(x_1, \mathrm{y}_1)}{\partial y} \\ \frac{\partial f_2(x_1, \mathrm{y}_1)}{\partial x} & \frac{\partial f_2(x_1, \mathrm{y}_1)}{\partial y} \end{bmatrix}$$

$$\frac{\partial f_1(x_1, \mathrm{y}_1)}{\partial x} = \frac{\partial}{\partial x}\left(x^2 + y^2 - 1\right) = 2x$$

$$\frac{\partial f_1(x_1, \mathrm{y}_1)}{\partial y} = \frac{\partial}{\partial x}\left(x^2 + y^2 - 1\right) = 2y$$

$$\frac{\partial f_2(x_1, \mathrm{y}_1)}{\partial x} = \frac{\partial}{\partial x}(xy) = y$$

$$\frac{\partial f_2(x_1, \mathrm{y}_1)}{\partial y} = \frac{\partial}{\partial x}(xy) = x$$

Thus, the Jacobian matrix (or the matrix of partial derivatives) is given by:

$$\begin{bmatrix} \frac{\partial f_1(x_1, \mathrm{y}_1)}{\partial x} & \frac{\partial f_1(x_1, \mathrm{y}_1)}{\partial y} \\ \frac{\partial f_2(x_1, \mathrm{y}_1)}{\partial x} & \frac{\partial f_2(x_1, \mathrm{y}_1)}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x & 2\mathrm{y} \\ \mathrm{y} & \mathrm{x} \end{bmatrix}$$

And its inverse is given by:

$$\begin{bmatrix} 2x & 2y \\ y & x \end{bmatrix}^{-1} = \frac{1}{2x^2 - 2y^2} \begin{bmatrix} x & -2y \\ -y & 2x \end{bmatrix}$$

Now, we calculate the values of the roots according to equation (3):

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} - \begin{bmatrix} \frac{\partial f_1(x_1, y_1)}{\partial x} & \frac{\partial f_1(x_1, y_1)}{\partial y} \\ \frac{\partial f_2(x_1, y_1)}{\partial x} & \frac{\partial f_2(x_1, y_1)}{\partial y} \end{bmatrix}^{-1} \begin{bmatrix} f_1(x_1, y_1) \\ f_2(x_1, y_1) \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} - \frac{1}{2x_1^2 - 2y_1^2} \begin{bmatrix} x_1 & -2y_1 \\ -y_1 & 2x_1 \end{bmatrix} \begin{bmatrix} x_1^2 + y_1^2 - 1 \\ x_1 y_1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.5 \\ 0.1 \end{bmatrix} - \frac{1}{2(0.5)^2 - 2(0.1)^2} \begin{bmatrix} 0.5 & -2(0.1) \\ -0.1 & 2(0.5) \end{bmatrix} \begin{bmatrix} 0.5^2 + 0.1^2 - 1 \\ 0.5(0.1) \end{bmatrix}$$

$$= \begin{bmatrix} 0.5 \\ 0.1 \end{bmatrix} - \frac{1}{0.48} \begin{bmatrix} 0.5 & -0.2 \\ -0.1 & 1 \end{bmatrix} \begin{bmatrix} -0.74 \\ 0.05 \end{bmatrix}$$

$$= \begin{bmatrix} 0.5 \\ 0.1 \end{bmatrix} - \begin{bmatrix} 1.0417 & -0.4167 \\ -0.2083 & 2.0833 \end{bmatrix} \begin{bmatrix} -0.74 \\ 0.05 \end{bmatrix}$$

$$= \begin{bmatrix} 0.5 \\ 0.1 \end{bmatrix} - \begin{bmatrix} -0.7917 \\ 0.2583 \end{bmatrix} = \begin{bmatrix} 1.2917 \\ -0.1583 \end{bmatrix}$$

Step Five: Note that the absolute errors for the unknowns x, y are as follows:

$$|x_2 - x_1| = |1.29717 - 0.5| = 0.7917 > 0.0001$$

$$|y_2 - y_1| = |-0.1583 - 0.1| = 0.2583 > 0.0001$$

It is clear that both errors are greater than the permissible error; therefore, a new iteration is necessary. Continuing the solution, we find that the roots of the above system are $x = 1, y = 0$, as shown in the table below:

n	x_n	y_n	$\lvert x_n - x_{n-1}\rvert$	$\lvert y_n - y_{n-1}\rvert$
1	0.5000	0.1000	--	---
2	1.2917	-0.1583	0.7917	0.2583
3	1.0389	-0.0310	0.2528	0.1273
4	1.0011	-0.0011	0.0378	0.0299
5	1.0000	0.0000	0.0011	0.0011
6	1.0000	0.0000	0.0000	0.0000

Note: The student can apply the test for the absolute value of the function at each proposed value of the roots. Specifically, applying the following test:

$$|f_1(x_n, y_n)| \le \varepsilon_1 \text{ and } |f_2(x_n, y_n)| \le \varepsilon_2$$