Solution to System of Nonlinear Functions

Solution to System of Nonlinear Functions

First consider solving the following system of nonlinear equations:

$$\begin{cases} f_1(x_1, x_2) = 0, \\ f_2(x_1, x_2) = 0. \end{cases}$$

Suppose $(x_1^{(k)}, x_2^{(k)})$ is an approximation to the solution of the system above, and we try to compute $h_1^{(k)}$ and $h_2^{(k)}$ such that $(x_1^{(k)} + h_1^{(k)}, x_2^{(k)} + h_2^{(k)})$ satisfies the system.

Taylor's theorem for two variables,

system of nonlinear equations:

$$\begin{cases} f_1(x_1, x_2) = 0, \\ f_2(x_1, x_2) = 0. \end{cases}$$

Taylor's theorem for two variables,

$$0 = f_1(x_1^{(k)} + h_1^{(k)}, x_2^{(k)} + h_2^{(k)}) \approx f_1(x_1^{(k)}, x_2^{(k)}) + h_1^{(k)} \frac{\partial f_1}{\partial x_1}(x_1^{(k)}, x_2^{(k)}) + h_2^{(k)} \frac{\partial f_1}{\partial x_2}(x_1^{(k)}, x_2^{(k)})$$

$$0 = f_2(x_1^{(k)} + h_1^{(k)}, x_2^{(k)} + h_2^{(k)}) \approx f_2(x_1^{(k)}, x_2^{(k)}) + h_1^{(k)} \frac{\partial f_2}{\partial x_1}(x_1^{(k)}, x_2^{(k)}) + h_2^{(k)} \frac{\partial f_2}{\partial x_2}(x_1^{(k)}, x_2^{(k)})$$

Put this in matrix form

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_1^{(k)}, x_2^{(k)}) & \frac{\partial f_1}{\partial x_2}(x_1^{(k)}, x_2^{(k)}) \\ \frac{\partial f_2}{\partial x_1}(x_1^{(k)}, x_2^{(k)}) & \frac{\partial f_2}{\partial x_2}(x_1^{(k)}, x_2^{(k)}) \end{bmatrix} \begin{bmatrix} h_1^{(k)} \\ h_2^{(k)} \end{bmatrix} + \begin{bmatrix} f_1(x_1^{(k)}, x_2^{(k)}) \\ f_2(x_1^{(k)}, x_2^{(k)}) \end{bmatrix} \approx \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The matrix

$$J(x_1^{(k)}, x_2^{(k)}) \equiv \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_1^{(k)}, x_2^{(k)}) & \frac{\partial f_1}{\partial x_2}(x_1^{(k)}, x_2^{(k)}) \\ \frac{\partial f_2}{\partial x_1}(x_1^{(k)}, x_2^{(k)}) & \frac{\partial f_2}{\partial x_2}(x_1^{(k)}, x_2^{(k)}) \end{bmatrix}$$

is called the Jacobian matrix.

$$J(x_1^{(k)}, x_2^{(k)}) \begin{bmatrix} h_1^{(k)} \\ h_2^{(k)} \end{bmatrix} = - \begin{bmatrix} f_1(x_1^{(k)}, x_2^{(k)}) \\ f_2(x_1^{(k)}, x_2^{(k)}) \end{bmatrix}$$

Set $h_1^{(k)}$ and $h_2^{(k)}$ be the solution of the linear system

$$\begin{bmatrix} h_1^{(k)} \\ h_2^{(k)} \end{bmatrix} = - \quad J^{-1}(x_1^{(k)}, x_2^{(k)}) \begin{bmatrix} f_1(x_1^{(k)}, x_2^{(k)}) \\ f_2(x_1^{(k)}, x_2^{(k)}) \end{bmatrix}$$

then

$$\begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{bmatrix} = \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \end{bmatrix} + \begin{bmatrix} h_1^{(k+1)} \\ h_2^{(k+1)} \end{bmatrix}$$

Quick look at Generalized version of NRM

In general, we solve the system of n nonlinear equations $f_i(x_1, \cdots, x_n) = 0, i = 1, \ldots, n$. Let

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad F(X) = \begin{bmatrix} f_1(X) \\ f_2(X) \\ \vdots \\ f_n(X) \end{bmatrix}.$$

The problem can be formulated as solving

$$F(X) = 0, \quad F : \mathbb{R}^n \to \mathbb{R}^n.$$

Let $F'(X)$, where the (i, j) entry is $\frac{\partial f_i}{\partial x_j}(X)$, be the $n \times n$ Jacobian matrix.

Newton's iteration is defined as

$$X^{(k+1)} = X^{(k)} + S^{(k)},$$

where $S^{(k)} \in \mathbb{R}^n$ is the solution of the linear system

$$F'(X^{(k)})S^{(k)} = -F(X^{(k)}).$$

EXAMPLE

Use the Newton-Raphson method with $\varepsilon_s = 0.0001$ to find the solution to the following nonlinear system of equations:

$$x_1^2 + x_1x_2 = 10 \qquad x_2 + 3x_1x_2^2 = 57$$

SOLUTION

In addition to requiring an initial guess, the Newton-Raphson method requires evaluating the derivatives of the functions $f_1 = x_1^2 + x_1x_2 - 10$ and $f_2 = x_2 + 3x_1x_2^2 - 57$. If $K_{ab} = \frac{\partial f_a}{\partial x_b}$, then it has the following form:

$$K = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 & x_1 \\ 3x_2^2 & 1 + 6x_1x_2 \end{pmatrix}$$

Assuming an initial guess of $x_1^{(0)} = 1.5$ and $x_2^{(0)} = 3.5$, then the vector f and the matrix K have components:

$$f = \begin{pmatrix} x_1^2 + x_1x_2 - 10 \\ x_2 + 3x_1x_2^2 - 57 \end{pmatrix} = \begin{pmatrix} -2.5 \\ 1.625 \end{pmatrix} \qquad K = \begin{pmatrix} 6.5 & 1.5 \\ 36.75 & 32.5 \end{pmatrix}$$

The components of the vector Δx can be computed as follows:

$$\Delta x = -K^{-1}f = \begin{pmatrix} 0.53603 \\ -0.65612 \end{pmatrix}$$

Therefore, the new estimates for $x_1^{(1)}$ and $x_2^{(1)}$ are:

$$x^{(1)} = x^{(0)} + \Delta x = \begin{pmatrix} 2.036029 \\ 2.843875 \end{pmatrix}$$

The approximate relative error is given by:

$$\varepsilon_r = \frac{\sqrt{(0.53603)^2 + (-0.65612)^2}}{\sqrt{(2.036029)^2 + (2.843875)^2}} = 0.2422$$

For the second iteration the vector f and the matrix K have components:

$$f = \begin{pmatrix} -0.06437 \\ -4.75621 \end{pmatrix} \qquad K = \begin{pmatrix} 6.9159 & 2.0360 \\ 24.2629 & 35.7413 \end{pmatrix}$$

The components of the vector Δx can be computed as follows:

$$\Delta x = -K^{-1}f = \begin{pmatrix} -0.03733 \\ 0.15841 \end{pmatrix}$$

Therefore, the new estimates for $x_1^{(2)}$ and $x_2^{(2)}$ are:

$$x^{(2)} = x^{(1)} + \Delta x = \begin{pmatrix} 1.998701 \\ 3.002289 \end{pmatrix}$$

The approximate relative error is given by:

$$\varepsilon_r = \frac{\sqrt{(-0.03733)^2 + (0.15841)^2}}{\sqrt{(1.998701)^2 + (3.002289)^2}} = 0.04512$$

For the third iteration the vector f and the matrix K have components:

$$f = \begin{pmatrix} -0.00452 \\ 0.04957 \end{pmatrix} \qquad K = \begin{pmatrix} 6.9997 & 1.9987 \\ 27.0412 & 37.0041 \end{pmatrix}$$

The components of the vector Δx can be computed as follows:

$$\Delta x = -K^{-1} f = \begin{pmatrix} 0.00130 \\ -0.00229 \end{pmatrix}$$

Therefore, the new estimates for $x_1^{(3)}$ and $x_2^{(3)}$ are:

$$x^{(3)} = x^{(2)} + \Delta x = \begin{pmatrix} 2.00000 \\ 2.999999 \end{pmatrix}$$

The approximate relative error is given by:

$$\varepsilon_r = \frac{\sqrt{(0.00130)^2 + (-0.00229)^2}}{\sqrt{(2.00000)^2 + (2.999999)^2}} = 0.00073$$

Finally, for the fourth iteration the vector f and the matrix K have components:

$$f = \begin{pmatrix} 0.00000 \\ -0.00002 \end{pmatrix} \qquad K = \begin{pmatrix} 7 & 2 \\ 27 & 37 \end{pmatrix}$$

The components of the vector Δx can be computed as follows:

$$\Delta x = -K^{-1} f = \begin{pmatrix} 0.00000 \\ 0.00000 \end{pmatrix}$$

Therefore, the new estimates for $x_1^{(4)}$ and $x_2^{(4)}$ are:

$$x^{(4)} = x^{(3)} + \Delta x = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

The approximate relative error is given by:

$$\varepsilon_r = \frac{\sqrt{(0.00000)^2 + (0.00000)^2}}{\sqrt{(2)^2 + (3)^2}} = 0.00000$$

Therefore, convergence is achieved after 4 iterations which is much faster than the 9 iterations in the fixed-point iteration method. The following is the Microsoft Excel table showing the values generated in every iteration:

Iteration	x_in		K matrix		f	Δ x		x_out		er
1	x_1 in	1.5	6.5	1.5	-2.50000	Δ x_1	0.53603	x_1 out	2.036029	0.24224
	x_2 in	3.5	36.75	32.5	1.62500	Δ x_2	-0.65612	x_2 out	2.843875	
2	x_1 in	2.03603	6.9159	2.0360	-0.06437	Δ x_1	-0.03733	x_1 out	1.998701	0.04512
	x_2 in	2.84388	24.2629	35.7413	-4.75621	Δ x_2	0.15841	x_2 out	3.002289	
3	x_1 in	1.99870	6.9997	1.9987	-0.00452	Δ x_1	0.00130	x_1 out	2.000000	0.00073
	x_2 in	3.00229	27.0412	37.0041	0.04957	Δ x_2	-0.00229	x_2 out	2.999999	
4	x_1 in	2.000000	7.0000	2.0000	0.00000	Δ x_1	0.00000	x_1 out	2.000000	0.0000002
	x_2 in	2.999999	27.0000	37.0000	-0.00002	Δ x_2	0.00000	x_2 out	3.000000	

Details of the Generalization

$$\begin{aligned} f_1(x_1, x_2, \cdots, x_n) &= 0 \\ f_2(x_1, x_2, \cdots, x_n) &= 0 \\ &\vdots \\ f_n(x_1, x_2, \cdots, x_n) &= 0 \end{aligned}$$

If the components of one iteration $x^{(i)} \in \mathbb{R}^n$ are known as: $x_1^{(i)}, x_2^{(i)}, \cdots, x_n^{(i)}$, then, the Taylor expansion of the first equation around these components is given by:

$$\begin{aligned} f_1\left(x_1^{(i+1)}, x_2^{(i+1)}, \cdots, x_n^{(i+1)}\right) \approx & f_1\left(x_1^{(i)}, x_2^{(i)}, \cdots, x_n^{(i)}\right) + \\ & \left.\frac{\partial f_1}{\partial x_1}\right|_{x^{(i)}} \left(x_1^{(i+1)} - x_1^{(i)}\right) + \left.\frac{\partial f_1}{\partial x_2}\right|_{x^{(i)}} \left(x_2^{(i+1)} - x_2^{(i)}\right) + \cdots + \\ & \left.\frac{\partial f_1}{\partial x_n}\right|_{x^{(i)}} \left(x_n^{(i+1)} - x_n^{(i)}\right) \end{aligned}$$

Applying the Taylor expansion in the same manner for $f_2, \cdots, f_n$, we obtained the following system of linear equations with the unknowns being the components of the vector $x^{(i+1)}$:

$$\begin{pmatrix} f_1\left(x^{(i+1)}\right) \\ f_2\left(x^{(i+1)}\right) \\ \vdots \\ f_n\left(x^{(i+1)}\right) \end{pmatrix} = \begin{pmatrix} f_1\left(x^{(i)}\right) \\ f_2\left(x^{(i)}\right) \\ \vdots \\ f_n\left(x^{(i)}\right) \end{pmatrix} + \begin{pmatrix} \frac{\partial f_1}{\partial x_1}\big|_{x^{(i)}} & \frac{\partial f_1}{\partial x_2}\big|_{x^{(i)}} & \cdots & \frac{\partial f_1}{\partial x_n}\big|_{x^{(i)}} \\ \frac{\partial f_2}{\partial x_1}\big|_{x^{(i)}} & \frac{\partial f_2}{\partial x_2}\big|_{x^{(i)}} & \cdots & \frac{\partial f_2}{\partial x_n}\big|_{x^{(i)}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_n}{\partial x_1}\big|_{x^{(i)}} & \frac{\partial f_n}{\partial x_2}\big|_{x^{(i)}} & \cdots & \frac{\partial f_n}{\partial x_n}\big|_{x^{(i)}} \end{pmatrix} \begin{pmatrix} \left(x_1^{(i+1)} - x_1^{(i)}\right) \\ \left(x_2^{(i+1)} - x_2^{(i)}\right) \\ \vdots \\ \left(x_n^{(i+1)} - x_n^{(i)}\right) \end{pmatrix}$$

By setting the left hand side to zero (which is the desired value for the functions $f_1, f_2, \cdots, f_n$, then, the system can be written as:

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}\big|_{x^{(i)}} & \frac{\partial f_1}{\partial x_2}\big|_{x^{(i)}} & \cdots & \frac{\partial f_1}{\partial x_n}\big|_{x^{(i)}} \\ \frac{\partial f_2}{\partial x_1}\big|_{x^{(i)}} & \frac{\partial f_2}{\partial x_2}\big|_{x^{(i)}} & \cdots & \frac{\partial f_2}{\partial x_n}\big|_{x^{(i)}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_n}{\partial x_1}\big|_{x^{(i)}} & \frac{\partial f_n}{\partial x_2}\big|_{x^{(i)}} & \cdots & \frac{\partial f_n}{\partial x_n}\big|_{x^{(i)}} \end{pmatrix} \begin{pmatrix} \left(x_1^{(i+1)} - x_1^{(i)}\right) \\ \left(x_2^{(i+1)} - x_2^{(i)}\right) \\ \vdots \\ \left(x_n^{(i+1)} - x_n^{(i)}\right) \end{pmatrix} = - \begin{pmatrix} f_1\left(x^{(i)}\right) \\ f_2\left(x^{(i)}\right) \\ \vdots \\ f_n\left(x^{(i)}\right) \end{pmatrix}$$

Setting $K_{ab} = \frac{\partial f_a}{\partial x_b}\big|_{x^{(i)}}$, the above equation can be written in matrix form as follows:

$$K\Delta x = -f$$

where K is an $n \times n$ matrix, $-f$ is a vector of n components and Δx is an n-dimensional vector with the components $\left(x_1^{(i+1)} - x_1^{(i)}\right), \left(x_2^{(i+1)} - x_2^{(i)}\right), \cdots, \left(x_n^{(i+1)} - x_n^{(i)}\right)$. If K is invertible, then, the above system can be solved as follows:

$$\Delta x = -K^{-1} f \Rightarrow x^{(i+1)} = x^{(i)} + \Delta x$$