Newton-Raphson Method For Solving Nonlinear Equations Part III - Solving Multivariate Equations

Gary Schurman, MBE, CFA

March, 2016

Solutions to a system of nonlinear equations requires a different set of tools other than the standard matrix inversion method used to solve a system of linear equations. Nonlinear financial models that have multiple parameters can be solved using the Newton-Raphson method in its multi-dimensional (i.e vector) form rather than the simpler one-dimensional (i.e. scalar) form. In this white paper we will build the mathematics to solve the following hypothetical problem...

Our Hypothetical Problem

Imagine that we have the following system of three nonlinear equations with three unknowns...

$$f(x,y,z)=x^2+y^2+z^2=14 \text{ ...and... } g(x,y,z)=4\,x-3\,y^2+2\,z^3=46 \text{ ...and... } h(x,y,z)=3\,y^3-2\,z=18 \quad (1)$$

Question: What is the solution (or solutions) to this problem?

Newton-Raphson Method of Solving Nonlinear Multivariate Equations

The Newton-Raphson method, named after Isaac Newton and Joseph Raphson, is a method for finding successively better approximations for the roots of a real-valued function. The method converges quickly if the iterations begin near the desired root. To develop the mathematics for the multivariate case we will build on the methods developed in Parts I and II to this series (i.e. the univariate case).

Using our hypothetical problem above we will define the functions $f(x,y,z)$, $g(x,y,z)$ and $h(x,y,z)$ as follows...

$$f(x,y,z)=14 \text{ ...and... } g(x,y,z)=46 \text{ ...and... } h(x,y,z)=18 \quad (2)$$

We will define the iterated functions $f(\hat{x},\hat{y},\hat{z})$, $g(\hat{x},\hat{y},\hat{z})$ and $h(\hat{x},\hat{y},\hat{z})$ as follows...

$$f(\hat{x},\hat{y},\hat{z})=\hat{x}^2+\hat{y}^2+\hat{z}^2 \text{ ...and... } g(\hat{x},\hat{y},\hat{z})=4\,\hat{x}-3\,\hat{y}^2+2\,\hat{z}^3 \text{ ...and... } h(\hat{x},\hat{y},\hat{z})=3\,\hat{y}^3-2\,\hat{z} \quad (3)$$

The equations for the first order derivatives of equation $f(\hat{x},\hat{y},\hat{z})$ as defined in Equation (3) above are...

$$\frac{\delta}{\delta\hat{x}}\,f(\hat{x},\hat{y},\hat{z})=2\,\hat{x} \text{ ...and... } \frac{\delta}{\delta\hat{y}}\,f(\hat{x},\hat{y},\hat{z})=2\,\hat{y} \text{ ...and... } \frac{\delta}{\delta\hat{z}}\,f(\hat{x},\hat{y},\hat{z})=2\,\hat{z} \quad (4)$$

The equations for the first order derivatives of equation $g(\hat{x},\hat{y},\hat{z})$ as defined in Equation (3) above are...

$$\frac{\delta}{\delta\hat{x}}\,g(\hat{x},\hat{y},\hat{z})=4 \text{ ...and... } \frac{\delta}{\delta\hat{y}}\,g(\hat{x},\hat{y},\hat{z})=6\,\hat{y} \text{ ...and... } \frac{\delta}{\delta\hat{z}}\,g(\hat{x},\hat{y},\hat{z})=6\,\hat{z}^2 \quad (5)$$

The equations for the first order derivatives of equation $h(\hat{x},\hat{y},\hat{z})$ as defined in Equation (3) above are...

$$\frac{\delta}{\delta\hat{x}}\,h(\hat{x},\hat{y},\hat{z})=0 \text{ ...and... } \frac{\delta}{\delta\hat{y}}\,h(\hat{x},\hat{y},\hat{z})=9\,\hat{y}^2 \text{ ...and... } \frac{\delta}{\delta\hat{z}}\,h(\hat{x},\hat{y},\hat{z})=-2 \quad (6)$$

We can relate Equations (2) and (3) above via the following Taylor Series Expansion...

$$\begin{aligned}
f(x,y,z)&=f(\hat{x},\hat{y},\hat{z})+\frac{\delta}{\delta\hat{x}}\,f(x,y,z)\,(x-\hat{x})+\frac{\delta}{\delta\hat{y}}\,f(x,y,z)\,(y-\hat{y})+\frac{\delta}{\delta\hat{z}}\,f(x,y,z)\,(z-\hat{z})+e_f\\
g(x,y,z)&=g(\hat{x},\hat{y},\hat{z})+\frac{\delta}{\delta\hat{x}}\,g(x,y,z)\,(x-\hat{x})+\frac{\delta}{\delta\hat{y}}\,g(x,y,z)\,(y-\hat{y})+\frac{\delta}{\delta\hat{z}}\,g(x,y,z)\,(z-\hat{z})+e_g\\
h(x,y,z)&=h(\hat{x},\hat{y},\hat{z})+\frac{\delta}{\delta\hat{x}}\,h(x,y,z)\,(x-\hat{x})+\frac{\delta}{\delta\hat{y}}\,h(x,y,z)\,(y-\hat{y})+\frac{\delta}{\delta\hat{z}}\,h(x,y,z)\,(z-\hat{z})+e_h
\end{aligned} \quad (7)$$

Using the equations above we will define the following matrix and vectors...

$$\vec{\mathbf{v}} = \begin{bmatrix} x - \hat{x} \\ y - \hat{y} \\ z - \hat{z} \end{bmatrix} \;,\; \mathbf{A} = \begin{bmatrix} \frac{\delta}{\delta \hat{x}} f(x,y,z) & \frac{\delta}{\delta \hat{x}} g(x,y,z) & \frac{\delta}{\delta \hat{x}} h(x,y,z) \\ \frac{\delta}{\delta \hat{y}} f(x,y,z) & \frac{\delta}{\delta \hat{y}} g(x,y,z) & \frac{\delta}{\delta \hat{y}} h(x,y,z) \\ \frac{\delta}{\delta \hat{z}} f(x,y,z) & \frac{\delta}{\delta \hat{z}} g(x,y,z) & \frac{\delta}{\delta \hat{z}} h(x,y,z) \end{bmatrix} \;,\; \vec{\mathbf{w}} = \begin{bmatrix} f(x,y,z) - f(\hat{x},\hat{y},\hat{z}) \\ g(x,y,z) - g(\hat{x},\hat{y},\hat{z}) \\ h(x,y,z) - h(\hat{x},\hat{y},\hat{z}) \end{bmatrix} \tag{8}$$

Using the matrix and vector definitions in Equation (8) above we can write the system of nonlinear equations in Equation (7) above as...

$$\vec{\mathbf{v}}^T \mathbf{A} = \vec{\mathbf{w}}^T \tag{9}$$

If we multiply Equation (9) by the inverse of matrix A then that equation becomes...

$$\vec{\mathbf{v}}^T \mathbf{A} \mathbf{A}^{-1} = \vec{\mathbf{w}}^T \mathbf{A}^{-1} \text{ ...such that... } \vec{\mathbf{v}}^T = \vec{\mathbf{w}}^T \mathbf{A}^{-1} \tag{10}$$

To solve the system of nonlinear equations we iterate Equation (10) above until $\vec{\mathbf{v}} = \vec{\mathbf{0}}$. The steps to solve the system of multivariate nonlinear equations via the Newton-Raphson method are as follows...

Step	Action
1	Determine a guess value for $\vec{\mathbf{v}}$
2	Use Equation (10) above to calculate the new value of $\vec{\mathbf{v}}$
3	If new $\vec{\mathbf{v}} = \vec{\mathbf{0}}$ then exit this routine
4	Set new $\vec{\mathbf{v}} =$ old $\vec{\mathbf{v}} + \vec{\mathbf{v}}$ from Step 2
5	Go to Step 2

The Solution to Our Hypothetical Problem

The solution the our hypothetical problem from above is...

Iteration	$\hat{x}$	$\hat{y}$	$\hat{z}$	$f(\hat{x},\hat{y},\hat{z})$	$g(\hat{x},\hat{y},\hat{z})$	$h(\hat{x},\hat{y},\hat{z})$	$x-\hat{x}$	$y-\hat{y}$	$z-\hat{z}$
1	5.00	5.00	5.00	75.00	195.00	365.00	-3.33	-1.55	- 1.22
2	1.67	3.45	3.78	28.99	79.47	115.30	-1.35	-0.92	-0.55
3	0.32	2.53	3.24	16.96	49.95	41.93	1.10	-0.42	-0.24
4	1.42	2.10	3.00	15.45	46.51	21.87	-0.35	-0.10	-0.01
5	1.07	2.00	3.00	14.13	45.97	18.18	-0.07	0.00	0.00
6	1.00	2.00	3.00	14.00	46.00	18.00	0.00	0.00	0.00

The solution to our problem per the table above is the following vector...

$$\vec{\mathbf{v}} = \begin{bmatrix} 1.00 \\ 2.00 \\ 3.00 \end{bmatrix} \tag{11}$$

Note that our initial guess vector is very important. If the vector is all zeros then matrix A cannot be inverted such that a solution cannot be found. If the vector is all ones then the solution will not converge such that a solution cannot be found. Note that if the vector is all negative ones then the solution is...

$$\vec{\mathbf{v}} = \begin{bmatrix} -0.54 \\ 2.01 \\ 3.11 \end{bmatrix} \tag{12}$$