Newton's method in *n* dimensions for finding extrema

1. Apply Newton's method in *n* dimensions for finding the extrema of the function starting with the point $x = 0$ and $y = 0$:

$$f(x, y) = 3x^2 + 5y^2 + 4xy + 17x - 13y + 4$$

Answer: Let $\mathbf{u} = \begin{pmatrix} x \\ y \end{pmatrix}$ and let $\mathbf{u}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. The gradient and Hessian (the Jacobian of the gradient) of this function are:

$$\vec{\nabla} f(\mathbf{u}) = \begin{pmatrix} 6x + 4y + 17 \\ 4x + 10y - 13 \end{pmatrix} \text{ and } \mathbf{J}\left(\vec{\nabla} f\right)(\mathbf{u}) = \begin{pmatrix} 6 & 4 \\ 4 & 10 \end{pmatrix}.$$

Solving $\mathbf{J}\left(\vec{\nabla} f\right)(\mathbf{u}_0)\Delta(x, y) = -\vec{\nabla} f(\mathbf{u}_0)$ has us solve $\begin{pmatrix} 6 & 4 \\ 4 & 10 \end{pmatrix}\Delta\mathbf{u}_0 = \begin{pmatrix} -17 \\ 13 \end{pmatrix}$, which yields a solution $\Delta\mathbf{u}_0 = \begin{pmatrix} -\frac{111}{22} \\ \frac{73}{22} \end{pmatrix}$. Thus, our next approximation is $\mathbf{u}_1 \leftarrow \mathbf{u}_0 + \Delta\mathbf{u}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -\frac{111}{22} \\ \frac{73}{22} \end{pmatrix} = \begin{pmatrix} -\frac{111}{22} \\ \frac{73}{22} \end{pmatrix}$. Now, because the Hessian is a constant matrix, we only require one step, and we are done. This is the minimum.

2. Apply Newton's method in *n* dimensions for finding the extrema of the function starting with the point $x = 1$, $y = 1$ and $z = 1$:

$$f(x, y, z) = 4\cos(0.3xy) + 3\cos(0.2yz) + 3\cos(0.1xz)$$

Answer: Let $\mathbf{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ and let $\mathbf{u}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. The gradient and Hessian (the Jacobian of the gradient) of this function are:

$$\vec{\nabla} f(\mathbf{u}) = \begin{pmatrix} -1.2y\sin(0.3xy) - 0.3z\sin(0.1xz) \\ -0.6z\sin(0.2xz) - 1.2x\sin(0.3xy) \\ -0.3x\sin(0.1xz) - 0.6y\sin(0.2xz) \end{pmatrix}$$

and

$$\mathbf{J}(\vec{\nabla} f)(\mathbf{u}) = \begin{pmatrix} -0.36y^2\cos(0.3xy) - 0.03z^2\cos(0.1xz) & -1.2\sin(0.3xy) - 0.36xy\cos(0.3xy) & -0.3\sin(0.1xz) - 0.03xz\cos(0.1xz) \\ -1.2\sin(0.3xy) - 0.36xy\cos(0.3xy) & -0.36x^2\cos(0.3xy) - 0.12z^2\cos(0.1yz) & -0.6\sin(0.2yz) - 0.12yz\cos(0.2yz) \\ -0.3\sin(0.1xz) - 0.03xz\cos(0.1xz) & -0.6\sin(0.2yz) - 0.12yz\cos(0.2yz) & -0.03x^2\cos(0.1xz) - 0.12y^2\cos(0.2yz) \end{pmatrix}.$$

Solving $\mathbf{J}(\vec{\nabla} f)(\mathbf{u}_0)\Delta\mathbf{u}_0 = -\vec{\nabla} f(\mathbf{u}_0)$ means solving

$$\begin{pmatrix} -0.3737712610435590 & -0.6985453840788257 & -0.05980014995238921 \\ -0.6985453840788257 & -0.4615291254261672 & -0.2368095878179857 \\ -0.05980014995238921 & -0.2368095878179857 & -0.1474581142992898 \end{pmatrix} \Delta\mathbf{u}_0 = \begin{pmatrix} 0.3845742729876559 \\ 0.4738258464706442 \\ 0.1491516234710851 \end{pmatrix}$$

so $\mathbf{u}_1 \leftarrow \mathbf{u}_0 + \Delta\mathbf{u}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -0.3431921637687873 \\ -0.3388075915245113 \\ -0.3282002807826209 \end{pmatrix} = \begin{pmatrix} 0.6568078362312127 \\ 0.6611924084754887 \\ 0.6717997192173791 \end{pmatrix}$.

Next, solving $\mathbf{J}(\vec{\nabla} f)(\mathbf{u}_1)\Delta\mathbf{u}_1 = -\vec{\nabla} f(\mathbf{u}_1)$ means solving

$$\begin{pmatrix} -0.1695756181637984 & -0.3109121269152401 & -0.02645742007379728 \\ -0.3109121269152401 & -0.2079308025887898 & -0.1063250473793330 \\ -0.02645742007379728 & -0.1063250473793330 & -0.06518346840058194 \end{pmatrix} \Delta\mathbf{u}_1 = \begin{pmatrix} 0.1119682305523830 \\ 0.1381563901525302 \\ 0.04388851921981684 \end{pmatrix}$$

so

$$\mathbf{u}_2 \leftarrow \mathbf{u}_1 + \Delta\mathbf{u}_1 = \begin{pmatrix} 0.6568078362312127 \\ 0.6611924084754887 \\ 0.6717997192173791 \end{pmatrix} + \begin{pmatrix} -0.2201301573544340 \\ -0.2210578130761832 \\ -0.2233765874101995 \end{pmatrix} = \begin{pmatrix} 0.4366776788767787 \\ 0.4401345953993055 \\ 0.4484231318071796 \end{pmatrix}.$$

Next, solving $\mathbf{J}\left(\vec{\nabla} f\right)\left(\mathbf{u}_2\right)\Delta\mathbf{u}_2 = -\vec{\nabla} f\left(\mathbf{u}_2\right)$ means solving

$$\begin{pmatrix} -0.07565409558352892 & -0.1382284914187828 & -0.01174748070569741 \\ -0.1382284914187828 & -0.09264458231717509 & -0.04734336895344803 \\ -0.01174748070569741 & -0.04734336895344803 & -0.02894763248065672 \end{pmatrix} \Delta\mathbf{u}_2 = \begin{pmatrix} 0.03307052837521824 \\ 0.04081507280969326 \\ 0.01298652913100026 \end{pmatrix}$$

so

$$\mathbf{u}_3 \leftarrow \mathbf{u}_2 + \Delta\mathbf{u}_2 = \begin{pmatrix} 0.4366776788767787 \\ 0.4401345953993055 \\ 0.4484231318071796 \end{pmatrix} + \begin{pmatrix} -0.1457140759901665 \\ -0.1467971142457367 \\ -0.1494037861856176 \end{pmatrix} = \begin{pmatrix} 0.2909636028866122 \\ 0.2933374811535688 \\ 0.2990193456215620 \end{pmatrix}$$

3. What is the value of the function at $\mathbf{u}_3$? What happened?

Answer: Evaluating the function at this point yields that $f(\mathbf{u}_3) = 9.998113670725047$. We appear to have converged to a local maximum, not a local minimum.

4. Apply Newton's method again, but this time starting, as an initial point, the point we found using the Hooke-Jeeves method.

Answer: If we let $\mathbf{u}_0 = \begin{pmatrix} 4.75 \\ 2.25 \\ 6.75 \end{pmatrix}$ where $f(\mathbf{u}_0) = -9.969134842562037$, our sequence of iterations now proceed as follows:

1. $\mathbf{u}_1 = \begin{pmatrix} 4.565461888759592 \\ 2.291460996599569 \\ 6.872170933763010 \end{pmatrix}$ where $f(\mathbf{u}_1) = -9.999862067246340$
2. $\mathbf{u}_2 = \begin{pmatrix} 4.576414691575024 \\ 2.288237615902786 \\ 6.864720091357170 \end{pmatrix}$ where $f(\mathbf{u}_2) = -9.999999998003700$
3. $\mathbf{u}_3 = \begin{pmatrix} 4.576456163707606 \\ 2.288228082367500 \\ 6.864684246790537 \end{pmatrix}$ where $f(\mathbf{u}_3) = -10$

5. If the correct solution to seventeen digits of precision is $\mathbf{u} = \begin{pmatrix} 4.5764561643188450 \\ 2.2882280821594225 \\ 6.8646842464782675 \end{pmatrix}$, does the above show quadratic convergence of the approximations, and super-quadratic ($\omega(h^2)$) convergence of the value of the function?

Answer: Yes, for if we look at the errors if the approximations of the **u** values, we have

0.2114978145485270, 0.01368854694076882, 0.00005563933937551418, 0.0000000007172322651031534

where each subsequent approximation appears to have the previous error approximately squared, while the errors in the value of the function decrease by

0.030865157437963, 0.000137932753660, 0.000000001996300, 0

and the errors seem to be greater than the previous error squared.