The two variable Newton method for solving equations

Let \( a(x,y) \) and \( b(x,y) \) be two differentiable functions of \( x \) and \( y \). In calculus we sometimes need to solve the equations

\[
(*) \quad a(x,y) = 0 \quad \text{and} \quad b(x,y) = 0
\]
simultaneously. Here are two examples.

**Example 1** Find the critical points of a function \( f(x,y) \).

We need to find simultaneous solutions to the equations

\[
\frac{\partial f}{\partial x}(x,y) = 0 \quad \text{and} \quad \frac{\partial f}{\partial y}(x,y) = 0.
\]
This is (*), where \( a(x,y) = \frac{\partial f}{\partial x}(x,y) \) and \( b(x,y) = \frac{\partial f}{\partial y}(x,y) \).

**Example 2** Find the maximum and minimum values of a function \( f(x,y) \) on a curve of the form \( g(x,y) = c \). The method of Lagrange multipliers says that we must solve the equation

\[
\nabla f(x,y) = \lambda \nabla g(x,y),
\]
which becomes the two equations

\[
(1) \quad \frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x} \quad \text{and} \quad (2) \quad \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}.
\]
Multiplying (1) by \( \frac{\partial g}{\partial y} \) and multiplying (2) by \( \frac{\partial g}{\partial x} \) we obtain

\[
\frac{\partial f}{\partial x} \frac{\partial g}{\partial y} = \frac{\partial f}{\partial y} \frac{\partial g}{\partial x}.
\]
Hence we must solve the equations (*), where \( a(x,y) = g(x,y) - c \) and \( b(x,y) = \frac{\partial f}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial x} \).

**Newton's method of solution**

In general there is no method for getting solutions to (*) that you can write down. However the Newton method gives a numerical procedure that solves the equation (*) to any desired degree of accuracy. The idea behind the Newton method is very simple. We describe it first and then illustrate the method with two examples.

**Step 1** By some method we find an approximate solution \( (x_o, y_o) \) to (*).

One way to find an approximate solution to (*) would be to graph the curves \( a(x,y) = 0 \) and \( b(x,y) = 0 \), and then look to see where these two curves intersect. A computer program with a zoom finder can be useful here.

**Step 2** We replace the linear equations (*) by the approximate equations

\[
(*)_o \quad a_o(x,y) = 0 \quad \text{and} \quad b_o(x,y) = 0
\]
where \( a_o(x,y) \) and \( b_o(x,y) \) are the linear approximations of \( a(x,y) \) and \( b(x,y) \) at \( (x_o, y_o) \). The equation \( (*)_o \) is a linear equation in two unknowns \( x \) and \( y \), and it can be easily solved by
hand. The solution \((x_1,y_1)\) to \((*)_o\) is not the true solution to the original equation \((*)\), but in general it is a better approximation than \((x_o,y_o)\).

Recall that the linear approximation of a function \(f(x,y)\) at a point \((x_o,y_o)\) is given by
\[
f_o(x,y) = A + B(x - x_o) + C(y - y_o)
\]
where \(A = f(x_o,y_o)\), \(B = \frac{\partial f}{\partial x}(x_o,y_o)\) and \(C = \frac{\partial f}{\partial y}(x_o,y_o)\), all real numbers.

**Step 3** We replace the linear equations \((*)\) by the approximate equations
\[
(*)_1 \quad a_1(x,y) = 0 \text{ and } b_1(x,y) = 0
\]
where \(a_1(x,y)\) and \(b_1(x,y)\) are the linear approximations of \(a(x,y)\) and \(b(x,y)\) at \((x_1,y_1)\). Let \((x_2,y_2)\) be the solution to \((*)_1\).

**Step 4** Continue as described above. In general we get a sequence of approximate solutions \((x_1,y_1), (x_2,y_2), ..., (x_k,y_k)\) that converge rapidly to the true solution if the beginning approximate solution \((x_o,y_o)\) is reasonably accurate. To obtain the next approximate solution \((x_{k+1},y_{k+1})\) we solve the equations
\[
(*)_k \quad a_k(x,y) = 0 \text{ and } b_k(x,y) = 0
\]
where \(a_k(x,y)\) and \(b_k(x,y)\) are the linear approximations of \(a(x,y)\) and \(b(x,y)\) at \((x_k,y_k)\).

**A matrix formula for the approximate solutions \((x_k,y_k)\)**

In order to solve the approximate equations \((*)_k\) above rapidly with a computer it is useful to have a formula that one can program into the computer. First we define the

Jacobian matrix \(J(x,y) = \begin{bmatrix} \frac{\partial a}{\partial x} & \frac{\partial a}{\partial y} \\ \frac{\partial b}{\partial x} & \frac{\partial b}{\partial y} \end{bmatrix}\). If we write \((x_k,y_k)\) and \((x_{k+1},y_{k+1})\) as column vectors

\[
\begin{bmatrix} x_k \\ y_k \end{bmatrix} \text{ and } \begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix}
\]

then the equation \((*)_k\) can be written in matrix form as

\[
\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a(x_k) \\ b(y_k) \end{bmatrix} + J(x_k,y_k) \begin{bmatrix} x-x_k \\ y-y_k \end{bmatrix} + J(x_k,y_k) \begin{bmatrix} x_k \\ y_k \end{bmatrix} - J(x_k,y_k)^{-1} \begin{bmatrix} a(x_k) \\ b(y_k) \end{bmatrix}
\]

The solution \(\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix}\) to \((*)_k\) in matrix form now becomes

\[
(*)_k \quad \begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \end{bmatrix} - J(x_k,y_k)^{-1} \begin{bmatrix} a(x_k) \\ b(y_k) \end{bmatrix}
\]

where \(J(x_k,y_k)^{-1}\) denotes the inverse of the matrix \(J(x_k,y_k)\).
Example 1  Find an intersection point of the circle \( x^2 + y^2 = 1 \) and the parabola \( y = x^2 \).

We need to solve the equations

\[
(*) \quad \begin{align*}
    a(x, y) &= 0 \quad \text{where } a(x, y) = y - x^2 \\
    b(x, y) &= 0 \quad \text{where } b(x, y) = x^2 + y^2 - 1
\end{align*}
\]

We pick \((x_0, y_0) = (1, 1)\) as an approximate solution. If \(a_o(x, y)\) and \(b_o(x, y)\) denote the linear approximations at \((x_0, y_0)\), then we obtain \(a_o(x, y) = -2x + y + 1\) and \(b_o(x, y) = 2x + 2y - 3\). Hence the equations \(a_o(x, y) = 0\) and \(b_o(x, y) = 0\) have the solution \((x_1, y_1) = (5/6, 2/3)\).

If \(a_1(x, y)\) and \(b_1(x, y)\) denote the linear approximations at \((x_1, y_1)\), then we obtain \(a_1(x, y) = -(5/3)x + y + (25/36)\) and \(b_1(x, y) = (5/3)x + (4/3)y - (77/36)\). Hence the equations \(a_1(x, y) = 0\) and \(b_1(x, y) = 0\) have the solution \((x_2, y_2) = (331/420, 13/21)\).

If we continue this process we obtain the following approximate solutions

\[
\begin{align*}
    (x_1, y_1) &= (5/6, 2/3) \approx (0.8333, 0.6667) \\
    (x_2, y_2) &= (331/420, 13/21) \approx (0.7881, 0.6190) \\
    (x_3, y_3) &\approx (0.7861, 0.6180). \quad \text{In two more steps we obtain 15 digit accuracy} \\
    (x_5, y_5) &\approx (0.78615137757423, 0.618033988749894)
\end{align*}
\]

Example 2  Find a critical point of the function \(f(x, y) = 2x^2y^2 + x^2y - 2x - y^2\)

We calculate

\[
\begin{align*}
    a(x, y) &= \frac{\partial f}{\partial x}(x, y) = 4xy^2 + 2xy - 2 \\
    b(x, y) &= \frac{\partial f}{\partial y}(x, y) = 4x^2y + x^2 - 2y
\end{align*}
\]

The critical points of \(f(x, y)\) are the solutions to the equations \(a(x, y) = 0\) and \(b(x, y) = 0\).

We start with the approximate solution \((x_o, y_o) = (1, -1)\) and we obtain the equations

\[
\begin{align*}
    0 &= a_o(x, y) = 0 + 2(x-1) - 6(y+1) = 2x - 6y - 8 \\
    0 &= b_o(x, y) = -1 - 6(x-1) + 2(y+1) = -6x + 2y + 7
\end{align*}
\]

The solution to these equations is the next approximate solution

\[
(x_1, y_1) = (13/16, -17/16) = (0.8125, -1.0625) \quad \text{Continuing we obtain}
\]

\[
(x_2, y_2) \approx (.8069, -1.0759) \quad \text{In two more steps we obtain 15 digit accuracy}
\]

\[
(x_4, y_4) \approx (0.807068419897086, -1.075848752008690)
\]