Computational Mathematics/Information Technology

Dr Oliver Kerr

2009–10

We continue to look at the problem of finding roots of systems of equations:

$$\mathbf{x} = (x_1, x_2, \ldots, x_N) \quad \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_N(\mathbf{x}))$$

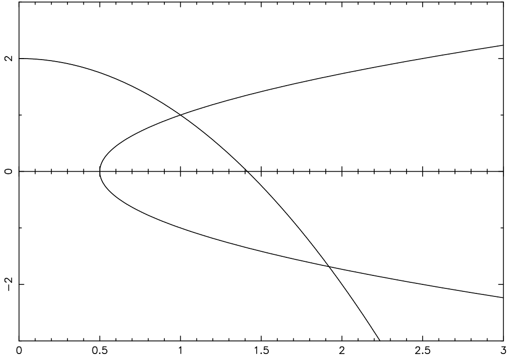
Or more briefly:

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}$$

Today we focus on nonlinear problems such as finding the roots of:

$$f_1(x_1, x_2) = 2 - x_1^2 - x_2 = 0 \quad \text{and} \quad f_2(x_1, x_2) = 2x_1 - x_2^2 - 1 = 0.$$

$$f_1(x_1, x_2) = 2 - x_1^2 - x_2 = 0 \quad \text{and} \quad f_2(x_1, x_2) = 2x_1 - x_2^2 - 1 = 0.$$



How do we set about this?

Non-linear simultaneous equations

We will look for a solution of the simultaneous equations

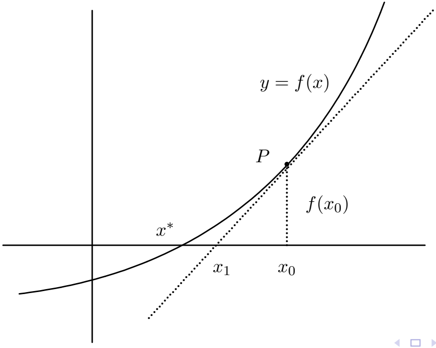
$$f(x, y) = 0 \quad \text{and} \quad g(x, y) = 0$$

We will develop the method for two equations in two unknowns, but it will becomes clear how the method can be extended to n equations in n unknowns.

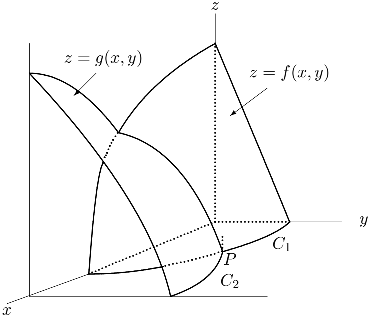
Note we introduce a third variable

$$z = f(x, y) = 0 \quad \text{and} \quad z = g(x, y) = 0$$

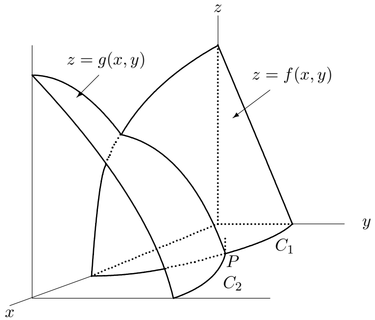
This will help us understand what is going on very much as we did for Newton's method:



If we plot $z = f(x, y)$ and $z = g(x, y)$ we get two surfaces:



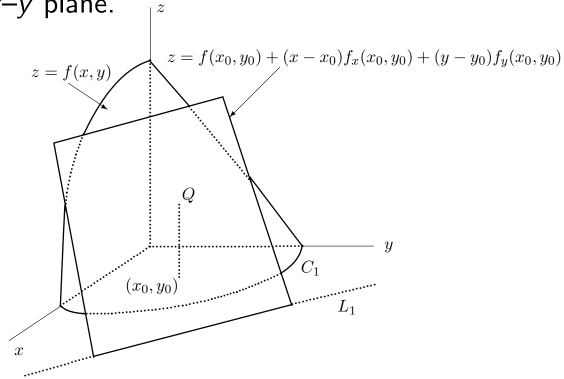
Since we require $z = 0$ we look to see where these two surfaces intersect on the x-y plane.

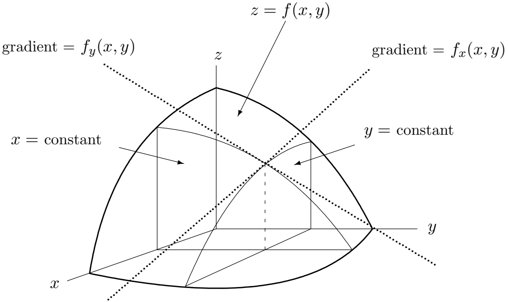


The surfaces will intersect on the x–y plane in two curves, C_1 and C_2. We are looking for the intersections of these two curves, this will be at the point P.

Our initial guess at x and y will, in general, not be on either C_1 or C_2.

Previously for Newton's method we used a *tangent* to the curve to estimate where it crossed the x-axis.
Here we will use a *tangent surface* to estimate where the surface crosses the x–y plane.





We start with a point (x_0, y_0) and construct the line which is a tangent to the surface in the plane $y = y_0$. **In this plane y is constant**

If $z = f(x)$ the tangent line through $x = x_0$ and $z = f(x_0)$ is given by

$$z = f(x_0) + (x - x_0)f'(x_0)$$

If $z = f(z)$ the tangent line through $x = x_0$ and $z = f(x_0, y_0)$ in the plane $y = y_0$ is given by

$$z = f(x_0, y_0) + (x - x_0)f_x(x_0, y_0), \qquad y = y_0$$

where $f_x(x_0, y_0)$ is the derivative of $f(x, y)$ with respect x keeping y constant. This is called **partial differentiation**.

Partial differentiation

To find the **partial derivative of** $f(x, y)$ **with respect to** x we do normal differentiation but treating y as a constant. The derivative is written as

$$f_x(x, y) \quad \text{or} \quad \frac{\partial f}{\partial x}$$

Example: If $f(x, y) = 2 - x^2 y + 4x + 5y$, find $f_x(x, y)$.

Partial differentiation

To find the **partial derivative of** $f(x, y)$ **with respect to** x we do normal differentiation but treating y as a constant. The derivative is written as

$$f_x(x, y) \quad \text{or} \quad \frac{\partial f}{\partial x}$$

Example: If $f(x, y) = 2 - x^2 y + 4x + 5y$, find $f_x(x, y)$.

$$f_x(x, y) = -2xy + 4$$

In a similar way we can find the **partial derivative of $f(x, y)$ with respect to y**. We do normal differentiation with y as our independent variable, but treating x as a constant. The derivative is written as

$$f_y(x, y) \quad \text{or} \quad \frac{\partial f}{\partial y}$$

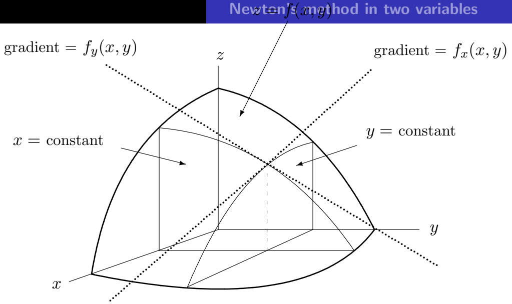
Example: If $f(x, y) = 2 - x^2y + 4x + 5y$, find $f_y(x, y)$.

In a similar way we can find the **partial derivative of $f(x, y)$ with respect to y**. We do normal differentiation with y as our independent variable, but treating x as a constant. The derivative is written as

$$f_y(x, y) \quad \text{or} \quad \frac{\partial f}{\partial y}$$

Example: If $f(x, y) = 2 - x^2y + 4x + 5y$, find $f_y(x, y)$.

$$f_y(x, y) = -x^2 + 5$$



The tangent line through $x = x_0$, $y = y_0$ and $z = f(x_0, y_0)$ in the plane $y = y_0$ is given by

$$z = f(x_0, y_0) + (x - x_0) f_x(x_0, y_0), \qquad y = y_0$$

The tangent line through $x = x_0$, $y = y_0$ and $z = f(x_0, y_0)$ in the plane $x = y_0$ is given by

$$z = f(x_0, y_0) + (y - y_0) f_y(x_0, y_0), \qquad x = x_0$$

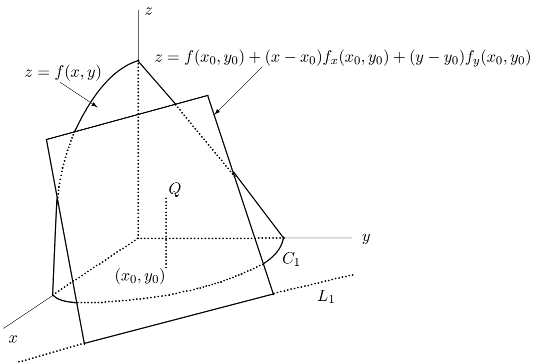
We have the two tangent lines

$$z = f(x_0, y_0) + (x - x_0)f_x(x_0, y_0), \qquad y = y_0$$

$$z = f(x_0, y_0) + (y - y_0)f_y(x_0, y_0), \qquad x = x_0$$

The tangent plane passes through $x = x_0$, $y = y_0$ and $z = f(x_0, y_0)$ and will include these two lines. This plane is given by

$$z = f(x_0, y_0) + (x - x_0)f_x(x_0, y_0) + (y - y_0)f_y(x_0, y_0).$$



The tangent plane will cut the x–y plane along the line L_1 given by

$$0 = f(x_0, y_0) + (x - x_0) f_x(x_0, y_0) + (y - y_0) f_y(x_0, y_0).$$

We will use this as an approximation to the curve C_1.

When x_0 and y_0 are close to C_1 we expect L_1 to be a good local approximation to C_1.

Similarly, if C_2 is the line in the x-y plane given by $g(x, y) = 0$ we expect to be able to approximate this by the line L_2, which is given by

$$0 = g(x_0, y_0) + (x - x_0)g_x(x_0, y_0) + (y - y_0)g_y(x_0, y_0).$$

We want to find where L_1 and L_2 cross to get our next guess at the root.

Newton's method in two variables

To Find where L_1 and L_2 cross we proceed as follows:

- We need to solve

$$\begin{aligned} 0 &= f(x_0, y_0) + (x - x_0) f_x(x_0, y_0) + (y - y_0) f_y(x_0, y_0) \\ 0 &= g(x_0, y_0) + (x - x_0) g_x(x_0, y_0) + (y - y_0) g_y(x_0, y_0) \end{aligned}$$

- Write in matrix notation:

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}_{(x_0,y_0)} + \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}_{(x_0,y_0)} \begin{pmatrix} x_1 - x_0 \\ y_1 - y_0 \end{pmatrix}$$

Notation: we have placed (x_0, y_0) outside the brackets to indicates that the functions inside the brackets are evaluated at (x_0, y_0).

- Rearrange:

$$-\begin{pmatrix} f \\ g \end{pmatrix}_{(x_0,y_0)} = \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}_{(x_0,y_0)} \begin{pmatrix} x_1 - x_0 \\ y_1 - y_0 \end{pmatrix}$$

- Use the inverse of the matrix:

$$-\begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}^{-1}_{(x_0,y_0)} \begin{pmatrix} f \\ g \end{pmatrix}_{(x_0,y_0)} = \begin{pmatrix} x_1 - x_0 \\ y_1 - y_0 \end{pmatrix}$$

- Split the vector on the right:

$$-\begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}^{-1}_{(x_0,y_0)} \begin{pmatrix} f \\ g \end{pmatrix}_{(x_0,y_0)} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

- Rearrange:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}^{-1}_{(x_0,y_0)} \begin{pmatrix} f \\ g \end{pmatrix}_{(x_0,y_0)}$$

We expect (hope?) that $x = x_1$ and $y = y_1$ will be a better approximation to the root of $f(x, y) = g(x, y) = 0$.

We use the iterative scheme

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} - \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}^{-1}_{(x_{n-1},y_{n-1})} \begin{pmatrix} f \\ g \end{pmatrix}_{(x_{n-1},y_{n-1})} \quad n = 1, 2 \ldots$$

starting with some point (x_0, y_0) close to a solution.

We expect (hope?) that $x = x_1$ and $y = y_1$ will be a better approximation to the root of $f(x,y) = g(x,y) = 0$.

We use the iterative scheme

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} - \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}^{-1}_{(x_{n-1},y_{n-1})} \begin{pmatrix} f \\ g \end{pmatrix}_{(x_{n-1},y_{n-1})} \qquad n = 1, 2 \ldots$$

starting with some point (x_0, y_0) close to a solution.

Note equivalence to Newton's method in one variable:

$$x_n = x_{n-1} - (f'(x_{n-1}))^{-1} f(x_{n-1}) = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$$

The extension to N equations in N unknowns is reasonably obvious. The matrix of partial derivatives will be $N \times N$ and the column vectors will have N entries.

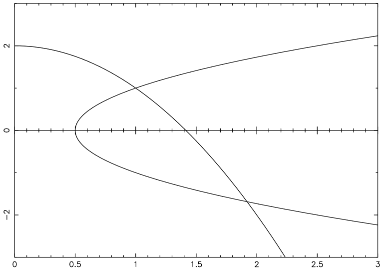
For example, for $N = 3$, to solve $f(x, y, z) = g(x, y, z) = h(x, y, z) = 0$ near the point (x_0, y_0, z_0) the scheme becomes:

$$\begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} = \begin{pmatrix} x_{n-1} \\ y_{n-1} \\ z_{n-1} \end{pmatrix} - \begin{pmatrix} f_x & f_y & f_z \\ g_x & g_y & g_z \\ h_x & h_y & h_z \end{pmatrix}^{-1} \begin{pmatrix} f \\ g \\ h \end{pmatrix} \qquad n = 1, 2 \ldots$$

where the functions and derivatives on the right-hand side of the equation are evaluated at $(x_{n-1}, y_{n-1}, z_{n-1})$.

Example: Using recent notation, we want to find the root of

$$f(x, y) = 2 - x^2 - y = 0 \quad \text{and} \quad g(x, y) = 2x - y^2 - 1 = 0.$$



We want to obtain the solution close to the point $x = 0.5$, $y = 0.5$ (not a particularly good guess).

To calculate the matrix we have to obtain four partial derivatives:

- $f(x,y) = 2 - x^2 - y \quad \Rightarrow \quad f_x(x,y) = -2x$ and $f_y(x,y) = -1$
- $g(x,y) = 2x - y^2 - 1 \quad \Rightarrow \quad g_x(x,y) = 2$ and $g_y(x,y) = -2y$

Thus the first step of the scheme is given by:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}^{-1}_{(x_0,y_0)} \begin{pmatrix} f \\ g \end{pmatrix}_{(x_0,y_0)}$$

$$\Rightarrow \quad \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} -2x_0 & -1 \\ 2 & -2y_0 \end{pmatrix}^{-1} \begin{pmatrix} 2 - x_0^2 - y_0 \\ 2x_0 - y_0^2 - 1 \end{pmatrix}$$

Substituting in our initial guess $x_0 = 0.5$ and $y_0 = 0,5$:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} - \begin{pmatrix} -1 & -1 \\ 2 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 1.25 \\ -0.25 \end{pmatrix}$$

$$= \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} - \begin{pmatrix} -1/3 & 1/3 \\ -2/3 & -1/3 \end{pmatrix} \begin{pmatrix} 1.25 \\ -0.25 \end{pmatrix}$$

$$= \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} - \begin{pmatrix} -1/2 \\ -3/4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1.25 \end{pmatrix}$$

Continuing this scheme to iterate gives the following results:

$n =$	0	1	2	3	4
$x_n =$	0.5	1.00	0.99	1.00	1.00
$y_n =$	0.5	1.25	1.00	1.00	1.00

We see that the scheme converges rapidly to the root $x = 1$, $y = 1$.(You can check this is indeed the root.)