SYSTEMS OF NONLINEAR EQUATIONS

Widely used in the mathematical modeling of real world phenomena.
We introduce some numerical methods for their solution.

For better intuition, we examine systems of two nonlinear equations and numerical methods for their solution. We then generalize to systems of an arbitrary order.

The Problem: Consider solving a system of two nonlinear equations

$$\begin{aligned} f(x,y) &= 0 \\ g(x,y) &= 0 \end{aligned} \tag{1}$$

Example: Consider solving the system

$$\begin{aligned} f(x,y) &\equiv x^2 + 4y^2 - 9 &= 0 \\ g(x,y) &\equiv 18y - 14x^2 + 45 &= 0 \end{aligned} \tag{2}$$

A graph of $z = f(x,y)$ is given in Figure 1, along with the curve for $f(x,y) = 0$.

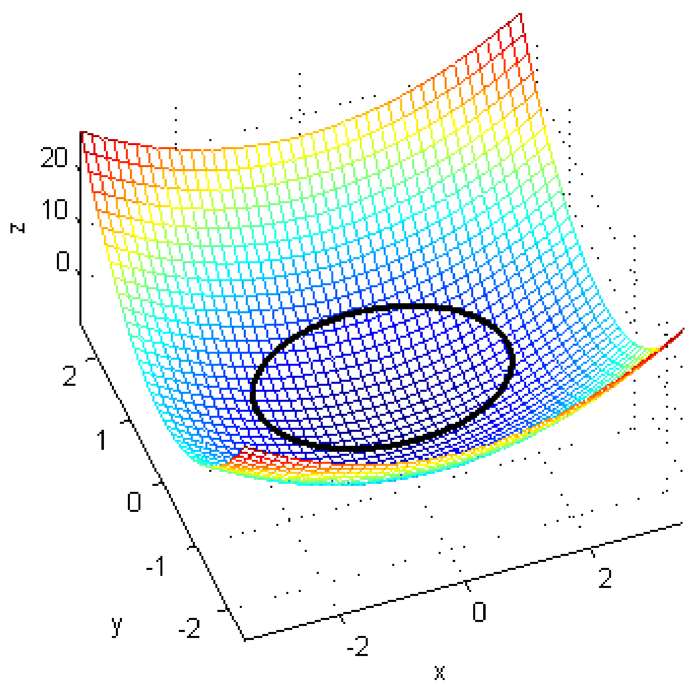

Figure 1: Graph of $z = x^2 + 4y^2 - 9$, along with $z = 0$ on that surface

To visualize the points (x, y) that satisfy simultaneously both equations in (1), look at the intersection of the zero curves of the functions f and g. Figure 2 illustrates this. The zero curves intersect at four points, each of which corresponds to a solution of the system (2). For example, there is a solution near the point $(1, -1)$.

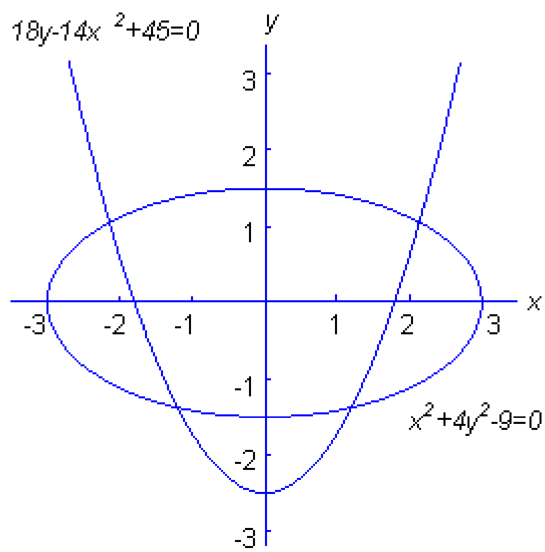

Figure 2: The graphs of $f(x, y) = 0$ and $g(x, y) = 0$

TANGENT PLANE APPROXIMATION

The graph of $z = f(x, y)$ is a surface in xyz-space.

For $(x, y) \approx (x_0, y_0)$, we approximate $f(x, y)$ by a plane that is tangent to the surface at the point $(x_0, y_0, f(x_0, y_0))$. The equation of this plane is $z = p(x, y)$ with

$$\begin{aligned} p(x,y) \ \equiv \ & f(x_0, y_0) \\ & +(x - x_0) f_x(x_0, y_0) + (y - y_0) f_y(x_0, y_0), \end{aligned} \tag{3}$$

$$f_x(x, y) = \frac{\partial f(x, y)}{\partial x}, \qquad f_y(x, y) = \frac{\partial f(x, y)}{\partial y},$$

the partial derivatives of f with respect to x and y, respectively.

If $(x, y) \approx (x_0, y_0)$, then

$$f(x, y) \approx p(x, y)$$

For functions of two variables, $p(x, y)$ is the linear Taylor polynomial approximation to $f(x, y)$.

NEWTON'S METHOD

Let $\alpha = (\xi, \eta)$ denote a solution of the system

$$
\begin{aligned}
f(x, y) &= 0 \\
g(x, y) &= 0
\end{aligned}
$$

Let $(x_0, y_0) \approx (\xi, \eta)$ be an initial guess at the solution.

Approximate the surface $z = f(x, y)$ with the tangent plane at $(x_0, y_0, f(x_0, y_0))$.

If $f(x_0, y_0)$ is sufficiently close to zero, then the zero curve of $p(x, y)$ will be an approximation of the zero curve of $f(x, y)$ for those points (x, y) near (x_0, y_0).

Because the graph of $z = p(x, y)$ is a plane, its zero curve is simply a straight line.

Example. Consider $f(x, y) \equiv x^2 + 4y^2 - 9$. Then

$$f_x(x, y) = 2x, \qquad f_y(x, y) = 8y$$

At $(x_0, y_0) = (1, -1)$,

$$f(x_0, y_0) = -4, \quad f_x(x_0, y_0) = 2, \quad f_y(x_0, y_0) = -8$$

At $(1, -1, -4)$ the tangent plane to the surface $z = f(x, y)$ has the equation

$$z = p(x, y) \equiv -4 + 2(x - 1) - 8(y + 1)$$

The graphs of the zero curves of $f(x, y)$ and $p(x, y)$ for (x, y) near (x_0, y_0) are given in Figure 3.

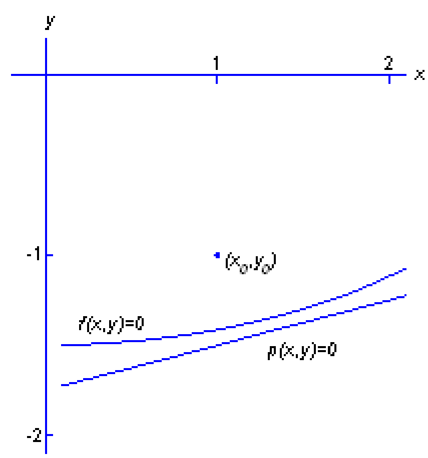

Figure 3: $f(x, y) = 0$ and $p(x, y) = 0$

The tangent plane to the surface $z = g(x, y)$ at $(x_0, y_0, g(x_0, y_0))$ has the equation $z = q(x, y)$ with

$$q(x, y) \equiv g(x_0, y_0) + (x - x_0) g_x(x_0, y_0) + (y - y_0) g_y(x_0, y_0)$$

Recall that the solution $\alpha = (\xi, \eta)$ to

$$\begin{aligned} f(x, y) &= 0 \\ g(x, y) &= 0 \end{aligned}$$

is the intersection of the zero curves of $z = f(x, y)$ and $z = g(x, y)$.

Approximate these zero curves by those of the tangent planes $z = p(x, y)$ and $z = q(x, y)$. The intersection of these latter zero curves gives an approximate solution to the above nonlinear system.

Denote the solution to

$$\begin{aligned} p(x, y) &= 0 \\ q(x, y) &= 0 \end{aligned}$$

by (x_1, y_1).

Example. Return to the equations

$$\begin{aligned} f(x,y) &\equiv x^2 + 4y^2 - 9 &= 0 \\ g(x,y) &\equiv 18y - 14x^2 + 45 &= 0 \end{aligned}$$

with $(x_0, y_0) = (1, -1)$.

The use of the zero curves of the tangent plane approximations is illustrated in Figure 4.

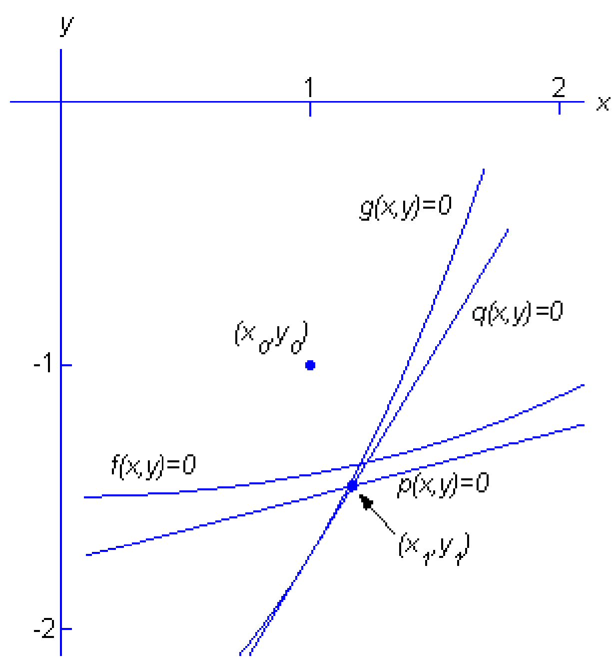

Figure 4: $f = g = 0$ and $p = q = 0$

Calculating (x_1, y_1). To find the intersection of the zero curves of the tangent planes, we must solve the linear system

$$f(x_0, y_0) + (x - x_0) f_x(x_0, y_0) + (y - y_0) f_y(x_0, y_0) = 0$$
$$g(x_0, y_0) + (x - x_0) g_x(x_0, y_0) + (y - y_0) g_y(x_0, y_0) = 0$$

The solution is denoted by (x_1, y_1). In matrix form,

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

It is actually computed as follows. Define δ_x and δ_y to be the solution of the linear system

$$\begin{bmatrix} f_x(x_0, y_0) & f_y(x_0, y_0) \\ g_x(x_0, y_0) & g_y(x_0, y_0) \end{bmatrix} \begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix} = - \begin{bmatrix} f(x_0, y_0) \\ g(x_0, y_0) \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix}$$

Usually the point (x_1, y_1) is closer to the solution α than is the original point (x_0, y_0).

Continue this process, using (x_1, y_1) as a new initial guess. Obtain an improved estimate (x_2, y_2).

This iteration process is continued until a solution with sufficient accuracy is obtained.

The general iteration is given by

$$\left[\begin{array}{cc} f_x(x_k, y_k) & f_y(x_k, y_k) \\ g_x(x_k, y_k) & g_y(x_k, y_k) \end{array}\right] \left[\begin{array}{c} \delta_{x,k} \\ \delta_{y,k} \end{array}\right] = -\left[\begin{array}{c} f(x_k, y_k) \\ g(x_k, y_k) \end{array}\right]$$

$$\left[\begin{array}{c} x_{k+1} \\ y_{k+1} \end{array}\right] = \left[\begin{array}{c} x_k \\ y_k \end{array}\right] + \left[\begin{array}{c} \delta_{x,k} \\ \delta_{y,k} \end{array}\right], \quad k = 0, 1, \ldots \tag{4}$$

This is *Newton's method* for solving

$$\begin{aligned} f(x, y) &= 0 \\ g(x, y) &= 0 \end{aligned}$$

Many numerical methods for solving nonlinear systems are variations on Newton's method.

Example. Consider again the system

$$\begin{aligned} f(x,y) &\equiv x^2+4y^2-9 &&= 0 \\ g(x,y) &\equiv 18y-14x^2+45 &&= 0 \end{aligned}$$

Newton's method (4) becomes

$$\begin{bmatrix} 2x_k & 8y_k \\ -28x_k & 18 \end{bmatrix} \begin{bmatrix} \delta_{x,k} \\ \delta_{y,k} \end{bmatrix} = - \begin{bmatrix} x_k^2+4y_k^2-9 \\ 18y_k-14x_k^2+45 \end{bmatrix}$$

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \end{bmatrix} + \begin{bmatrix} \delta_{x,k} \\ \delta_{y,k} \end{bmatrix} \tag{5}$$

Choose $(x_0,y_0)=(1,-1)$. The resulting Newton iterates are given in Table 1, along with

$$\text{Error } = \|\alpha-(x_k,y_k)\| \equiv \max\{|\xi-x_k|\,,|\eta-y_k|\}$$

Table 1: Newton iterates for the system (5)

k	x_k	y_k
0	1.0	-1.0
1	1.170212765957447	-1.457446808510638
2	1.202158829506705	-1.376760321923060
3	1.203165807091535	-1.374083486949713
4	1.203166963346410	-1.374080534243534
5	1.203166963347774	-1.374080534239942

k	*Error*
0	$3.74E-1$
1	$8.34E-2$
2	$2.68E-3$
3	$2.95E-6$
4	$3.59E-12$
5	$2.22E-16$

The final iterate is accurate to the precision of the computer arithmetic.

AN ALTERNATIVE NOTATION. We introduce a more general notation for the preceding work. The problem to be solved is

$$\begin{aligned} F_1(x_1, x_2) &= 0 \\ F_2(x_1, x_2) &= 0 \end{aligned} \tag{6}$$

Introduce

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad F(x) = \begin{bmatrix} F_1(x_1, x_2) \\ F_2(x_1, x_2) \end{bmatrix}$$

$$F'(x) = \begin{bmatrix} \dfrac{\partial F_1}{\partial x_1} & \dfrac{\partial F_1}{\partial x_2} \\ \dfrac{\partial F_2}{\partial x_1} & \dfrac{\partial F_2}{\partial x_2} \end{bmatrix}$$

$F'(x)$ is called the *Frechet derivative* of $F(x)$. It is a generalization to higher dimensions of the ordinary derivative of a function of one variable.

The system (6) can now be written as

$$F(x) = 0 \tag{7}$$

A solution of this equation will be denoted by α.

Newton's method becomes

$$\begin{aligned} F'(x^{(k)})\delta^{(k)} &= -F(x^{(k)}) \\ x^{(k+1)} &= x^{(k)} + \delta^{(k)} \end{aligned} \tag{8}$$

for $k = 0, 1, \ldots$.

A shorter and mathematically equivalent form:

$$x^{(k+1)} = x^{(k)} - \left[F'(x^{(k)})\right]^{-1} F(x^{(k)}) \tag{9}$$

for $k = 0, 1, \ldots$.

This last formula is often used in discussing and analyzing Newton's method for nonlinear systems. But (8) is used for practical computations, since it is usually less expensive to solve a linear system than to find the inverse of the coefficient matrix.

Note the analogy of (9) with Newton's method for a single equation.

THE GENERAL NEWTON METHOD

Consider the system of n nonlinear equations

$$\begin{array}{c} F_1(x_1, \ldots, x_n) = 0 \\ \vdots \\ F_n(x_1, \ldots, x_n) = 0 \end{array} \tag{10}$$

Define

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad F(x) = \begin{bmatrix} F_1(x_1, \ldots, x_n) \\ \vdots \\ F_n(x_1, \ldots, x_n) \end{bmatrix}$$

$$F'(x) = \begin{bmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial F_n}{\partial x_1} & \cdots & \dfrac{\partial F_n}{\partial x_n} \end{bmatrix}$$

The nonlinear system (10) can be written as

$$F(x) = 0$$

Its solution is denoted by $\alpha \in \mathbb{R}^n$.

Newton's method is

$$\begin{aligned} F'(x^{(k)})\delta^{(k)} &= -F(x^{(k)}) \\ x^{(k+1)} &= x^{(k)} + \delta^{(k)}, \qquad k = 0, 1, \ldots \end{aligned} \tag{11}$$

Alternatively, as before,

$$x^{(k+1)} = x^{(k)} - \left[F'(x^{(k)})\right]^{-1} F(x^{(k)}), \quad k = 0, 1, \ldots \tag{12}$$

This formula is often used in theoretical discussions of Newton's method for nonlinear systems. But (11) is used for practical computations, since it is usually less expensive to solve a linear system than to find the inverse of the coefficient matrix.

CONVERGENCE. Under suitable hypotheses, it can be shown that there is a $c > 0$ for which

$$\left\|\alpha - x^{(k+1)}\right\| \leq c \left\|\alpha - x^{(k)}\right\|^2, \quad k = 0, 1, \ldots$$

$$\left\|\alpha - x^{(k)}\right\| \equiv \max_{1 \leq i \leq n} \left|\alpha_i - x_i^{(k)}\right|$$

Newton's method is *quadratically convergent*.