3.7 Iteration for Nonlinear Systems: Seidel and Newton's Methods

Iterative techniques will now be discussed that extend the methods of Chapter 2 and Section 3.6 to the case of systems of nonlinear functions. Consider the functions

$$\begin{aligned} f_1(x, y) &= x^2 - 2x - y + 0.5 \\ f_2(x, y) &= x^2 + 4y^2 - 4. \end{aligned} \tag{1}$$

We seek a method of solution for the system of nonlinear equations

$$f_1(x, y) = 0 \quad \text{and} \quad f_2(x, y) = 0. \tag{2}$$

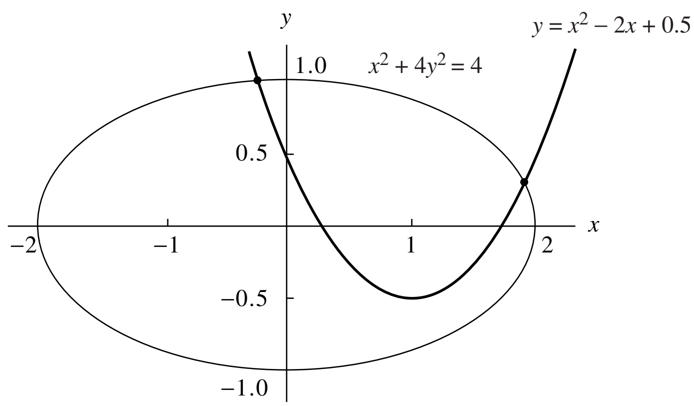

Figure 3.6 The graphs for the nonlinear system $y = x^2 - 2x + 0.5$ and $x^2 + 4y^2 = 4$.

The equations $f_1(x, y) = 0$ and $f_2(x, y) = 0$ implicitly define curves in the xy-plane. Hence a solution of the system (2) is a point (p, q) where the two curves cross (i.e., both $f_1(p, q) = 0$ and $f_2(p, q) = 0$). The curves for the system in (1) are well known:

$$\begin{aligned} x^2 - 2x + 0.5 &= 0 \quad \text{is the graph of a parabola,} \\ x^2 + 4y^2 - 4 &= 0 \quad \text{is the graph of an ellipse.} \end{aligned} \tag{3}$$

The graphs in Figure 3.6 show that there are two solution points and that they are in the vicinity of (−0.2, 1.0) and (1.9, 0.3).

The first technique is fixed-point iteration. A method must be devised for generating a sequence $\{(p_k, q_k)\}$ that converges to the solution (p, q). The first equation in (3) can be used to solve directly for x. However, a multiple of y can be added to each side of the second equation to get $x^2 + 4y^2 - 8y - 4 = -8y$. The choice of adding $-8y$ is crucial and will be explained later. We now have an equivalent system of equations:

$$\begin{aligned} x &= \frac{x^2 - y + 0.5}{2} \\ y &= \frac{-x^2 - 4y^2 + 8y + 4}{8}. \end{aligned} \tag{4}$$

These two equations are used to write the recursive formulas. Start with an initial point (p_0, q_0), and then compute the sequence $\{(p_{k+1}, q_{k+1})\}$ using

$$\begin{aligned} p_{k+1} = g_1(p_k, q_k) &= \frac{p_k^2 - q_k + 0.5}{2} \\ q_{k+1} = g_2(p_k, q_k) &= \frac{-p_k^2 - 4q_k^2 + 8q_k + 4}{8}. \end{aligned} \tag{5}$$

Table 3.5 Fixed-Point Iteration Using the Formulas in (5)

Case (i): Start with (0, 1)			Case (ii): Start with (2, 0)		
k	p_k	q_k	k	p_k	q_k
0	0.00	1.00	0	2.00	0.00
1	−0.25	1.00	1	2.25	0.00
2	−0.21875	0.9921875	2	2.78125	−0.1328125
3	−0.2221680	0.9939880	3	4.184082	−0.6085510
4	−0.2223147	0.9938121	4	9.307547	−2.4820360
5	−0.2221941	0.9938029	5	44.80623	−15.891091
6	−0.2222163	0.9938095	6	1,011.995	−392.60426
7	−0.2222147	0.9938083	7	512,263.2	−205,477.82
8	−0.2222145	0.9938084	This sequence is diverging.		
9	−0.2222146	0.9938084			

Case (i): If we use the starting value $(p_0, q_0) = (0, 1)$, then

$$p_1 = \frac{0^2 - 1 + 0.5}{2} = -0.25 \quad \text{and} \quad q_1 = \frac{-0^2 - 4(1)^2 + 8(1) + 4}{8} = 1.0.$$

Iteration will generate the sequence in case (i) of Table 3.5. In this case the sequence converges to the solution that lies near the starting value (0, 1).

Case (ii): If we use the starting value $(p_0, q_0) = (2, 0)$, then

$$p_1 = \frac{2^2 - 0 + 0.5}{2} = 2.25 \quad \text{and} \quad q_1 = \frac{-2^2 - 4(0)^2 + 8(0) + 4}{8} = 0.0.$$

Iteration will generate the sequence in case (ii) of Table 3.5. In this case the sequence diverges away from the solution.

Iteration using formulas (5) cannot be used to find the second solution (1.900677, 0.3112186). To find this point, a different pair of iteration formulas are needed. Start with equation (3) and add $-2x$ to the first equation and $-11y$ to the second equation and get

$$x^2 - 4x - y + 0.5 = -2x \quad \text{and} \quad x^2 + 4y^2 - 11y - 4 = -11y.$$

These equations can then be used to obtain the iteration formulas

$$\begin{aligned} p_{k+1} &= g_1(p_k, q_k) = \frac{-p_k^2 + 4p_k + q_k - 0.5}{2} \\ q_{k+1} &= g_2(p_k, q_k) = \frac{-p_k^2 - 4q_k^2 + 11q_k + 4}{11}. \end{aligned} \tag{6}$$

Table 3.6 shows how to use (6) to find the second solution.

Table 3.6 Fixed-Point Iteration Using the Formulas in (6)

k	p_k	q_k
0	2.00	0.00
1	1.75	0.0
2	1.71875	0.0852273
3	1.753063	0.1776676
4	1.808345	0.2504410
8	1.903595	0.3160782
12	1.900924	0.3112267
16	1.900652	0.3111994
20	1.900677	0.3112196
24	1.900677	0.3112186

Theory

We want to determine why equations (6) were suitable for finding the solution near $(1.9, 0.3)$ and equations (5) were not. In Section 2.1 the size of the derivative at the fixed point was the necessary idea. When functions of several variables are used, the partial derivatives must be used. The generalization of "the derivative" for systems of functions of several variables is the Jacobian matrix. We will consider only a few introductory ideas regarding this topic. More details can be found in any textbook on advanced calculus.

Definition 3.8. Assume that $f_1(x, y)$ and $f_2(x, y)$ are functions of the independent variables x and y; then their ***Jacobian matrix*** $\boldsymbol{J}(x, y)$ is

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{bmatrix}. \tag{7}$$

Similarly, if $f_1(x, y, z)$, $f_2(x, y, z)$, and $f_3(x, y, z)$ are functions of the independent variables x, y, and z, then their 3×3 Jacobian matrix $\boldsymbol{J}(x, y, z)$ is defined as follows:

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_1}{\partial z} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} & \dfrac{\partial f_2}{\partial z} \\ \dfrac{\partial f_3}{\partial x} & \dfrac{\partial f_3}{\partial y} & \dfrac{\partial f_3}{\partial z} \end{bmatrix}. \tag{8}$$

▲

Example 3.30. Find the Jacobian matrix $\boldsymbol{J}(x, y, z)$ of order 3×3 at the point $(1, 3, 2)$ for the three functions

$$f_1(x, y, z) = x^3 - y^2 + y - z^4 + z^2$$
$$f_2(x, y, z) = xy + yz + xz$$
$$f_3(x, y, z) = \frac{y}{xz}.$$

The Jacobian matrix is

$$\boldsymbol{J}(x, y, z) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_1}{\partial z} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} & \dfrac{\partial f_2}{\partial z} \\ \dfrac{\partial f_3}{\partial x} & \dfrac{\partial f_3}{\partial y} & \dfrac{\partial f_3}{\partial z}. \end{bmatrix} = \begin{bmatrix} 3x^2 & -2y + 1 & -4z^3 + 2z \\ y + z & x + z & y + x \\ \dfrac{-y}{x^2 z} & \dfrac{1}{xz} & \dfrac{-y}{xz^2} \end{bmatrix}.$$

Thus the Jacobian evaluated at the point $(1, 3, 2)$ is the 3×3 matrix

$$\boldsymbol{J}(1, 3, 2) = \begin{bmatrix} 3 & -5 & -28 \\ 5 & 3 & 4 \\ -\frac{3}{2} & \frac{1}{2} & -\frac{3}{4} \end{bmatrix}. \qquad \blacksquare$$

Generalized Differential

For a function of several variables, the differential is used to show how changes of the independent variables affect the change in the dependent variables. Suppose that we have

$$(9) \qquad u = f_1(x, y, z), \qquad v = f_2(x, y, z), \qquad \text{and} \qquad w = f_3(x, y, z).$$

Suppose that the values of the functions in (9) are known at the point (x_0, y_0, z_0) and we wish to predict their value at a nearby point (x, y, z). Let du, dv, and dw denote differential changes in the dependent variables and dx, dy, and dz denote differential changes in the independent variables. These changes obey the relationships

$$(10) \qquad \begin{aligned} du &= \frac{\partial f_1}{\partial x}(x_0, y_0, z_0)\,dx + \frac{\partial f_1}{\partial y}(x_0, y_0, z_0)\,dy + \frac{\partial f_1}{\partial z}(x_0, y_0, z_0)\,dz, \\ dv &= \frac{\partial f_2}{\partial x}(x_0, y_0, z_0)\,dx + \frac{\partial f_2}{\partial y}(x_0, y_0, z_0)\,dy + \frac{\partial f_2}{\partial z}(x_0, y_0, z_0)\,dz, \\ dw &= \frac{\partial f_3}{\partial x}(x_0, y_0, z_0)\,dx + \frac{\partial f_3}{\partial y}(x_0, y_0, z_0)\,dy + \frac{\partial f_3}{\partial z}(x_0, y_0, z_0)\,dz. \end{aligned}$$

If vector notation is used, (10) can be compactly written by using the Jacobian matrix. The function changes are $d\boldsymbol{F}$ and the changes in the variables are denoted $d\boldsymbol{X}$.

$$(11) \qquad d\boldsymbol{F} = \begin{bmatrix} du \\ dv \\ dw \end{bmatrix} = \boldsymbol{J}(x_0, y_0, z_0) \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \boldsymbol{J}(x_0, y_0, z_0)\, d\boldsymbol{X}.$$

Example 3.31. Use the Jacobian matrix to find the differential changes (du, dv, dw) when the independent variables change from (1, 3, 2) to (1.02, 2.97, 2.01) for the system of functions

$$\begin{aligned} u &= f_1(x, y, z) = x^3 - y^2 + y - z^4 + z^2 \\ v &= f_2(x, y, z) = xy + yz + xz \\ w &= f_3(x, y, z) = \frac{y}{xz}. \end{aligned}$$

Use equation (11) with $\boldsymbol{J}(1, 3, 2)$ of Example 3.30 and the differential changes $(dx, dy, dz) = (0.02, -0.03, 0.01)$ to obtain

$$\begin{bmatrix} du \\ dv \\ dw \end{bmatrix} = \begin{bmatrix} 3 & -5 & -28 \\ 5 & 3 & 4 \\ -\frac{3}{2} & \frac{1}{2} & -\frac{3}{4} \end{bmatrix} \begin{bmatrix} 0.02 \\ -0.03 \\ 0.01 \end{bmatrix} = \begin{bmatrix} -0.07 \\ 0.05 \\ -0.0525 \end{bmatrix}.$$

Notice that the function values at (1.02, 2.97, 2.01) are close to the linear approximations obtained by adding the differentials $du = -0.07$, $dv = 0.05$, and $dw = -0.0525$ to the corresponding function values $f_1(1, 3, 2) = -17$, $f_2(1, 3, 2) = 11$, and $f_3(1, 3, 2) = 1.5$; that is,

$$\begin{aligned} f_1(1.02, 2.97, 2.01) &= -17.072 \approx -17.07 = f_1(1, 3, 2) + du \\ f_2(1.02, 2.97, 2.01) &= 11.0493 \approx 11.05 = f_2(1, 3, 2) + dv \\ f_3(1.02, 2.97, 2.01) &= 1.44864 \approx 1.4475 = f_3(1, 3, 2) + dw. \end{aligned} \qquad \blacksquare$$

Convergence near Fixed Points

The extensions of the definitions and theorems in Section 2.1 to the case of two and three dimensions are now given. The notation for N-dimensional functions has not been used. The reader can easily find these extensions in many books on numerical analysis.

Definition 3.9. A ***fixed point*** for the system of two equations

$$(12) \qquad x = g_1(x, y) \quad \text{and} \quad y = g_2(x, y)$$

is a point (p, q) such that $p = g_1(p, q)$ and $q = g_2(p, q)$. Similarly, in three dimensions a fixed point for the system

$$(13) \qquad x = g_1(x, y, z), \quad y = g_2(x, y, z), \quad \text{and} \quad z = g_3(x, y, z)$$

is a point (p, q, r) such that $p = g_1(p, q, r)$, $q = g_2(p, q, r)$, and $r = g_3(p, q, r)$. ▲

Definition 3.10. For the functions (12), ***fixed-point iteration*** is

$$p_{k+1} = g_1(p_k, q_k) \quad \text{and} \quad q_{k+1} = g_2(p_k, q_k) \tag{14}$$

for $k = 0, 1, \ldots$. Similarly, for the functions (13), ***fixed-point iteration*** is

$$\begin{aligned} p_{k+1} &= g_1(p_k, q_k, r_k) \\ q_{k+1} &= g_2(p_k, q_k, r_k) \\ r_{k+1} &= g_3(p_k, q_k, r_k) \end{aligned} \tag{15}$$

for $k = 0, 1, \ldots$. ▲

Theorem 3.17 (Fixed-Point Iteration). Assume that the functions in (12) and (13) and their first partial derivatives are continuous on a region that contains the fixed point (p, q) or (p, q, r), respectively. If the starting point is chosen sufficiently close to the fixed point, then one of the following cases applies.

Case (i): Two dimensions. If (p_0, q_0) is sufficiently close to (p, q) and if

$$\begin{aligned} \left|\frac{\partial g_1}{\partial x}(p, q)\right| + \left|\frac{\partial g_1}{\partial y}(p, q)\right| &< 1, \\ \left|\frac{\partial g_2}{\partial x}(p, q)\right| + \left|\frac{\partial g_2}{\partial y}(p, q)\right| &< 1, \end{aligned} \tag{16}$$

then the iteration in (14) converges to the fixed point (p, q).

Case (ii): Three dimensions. If (p_0, q_0, r_0) is sufficiently close to (p, q, r) and if

$$\begin{aligned} \left|\frac{\partial g_1}{\partial x}(p, q, r)\right| + \left|\frac{\partial g_1}{\partial y}(p, q, r)\right| + \left|\frac{\partial g_1}{\partial z}(p, q, r)\right| &< 1, \\ \left|\frac{\partial g_2}{\partial x}(p, q, r)\right| + \left|\frac{\partial g_2}{\partial y}(p, q, r)\right| + \left|\frac{\partial g_2}{\partial z}(p, q, r)\right| &< 1, \\ \left|\frac{\partial g_3}{\partial x}(p, q, r)\right| + \left|\frac{\partial g_3}{\partial y}(p, q, r)\right| + \left|\frac{\partial g_3}{\partial z}(p, q, r)\right| &< 1, \end{aligned} \tag{17}$$

then the iteration in (15) converges to the fixed point (p, q, r).

If conditions (16) or (17) are not met, the iteration might diverge. This will usually be the case if the sum of the magnitudes of the partial derivatives is much larger than 1. Theorem 3.17 can be used to show why the iteration (5) converged to the fixed point near $(-0.2, 1.0)$. The partial derivatives are

$$\begin{aligned} \frac{\partial}{\partial x} g_1(x, y) &= x, & \frac{\partial}{\partial y} g_1(x, y) &= -\frac{1}{2}, \\ \frac{\partial}{\partial x} g_2(x, y) &= -\frac{x}{4}, & \frac{\partial}{\partial y} g_2(x, y) &= -y + 1. \end{aligned}$$

Indeed, for all (x, y) satisfying $-0.5 < x < 0.5$ and $0.5 < y < 1.5$, the partial derivatives satisfy

$$\left|\frac{\partial}{\partial x} g_1(x, y)\right| + \left|\frac{\partial}{\partial y} g_1(x, y)\right| = |x| + |-0.5| < 1,$$
$$\left|\frac{\partial}{\partial x} g_2(x, y)\right| + \left|\frac{\partial}{\partial y} g_2(x, y)\right| = \frac{|-x|}{4} + |-y+1| < 0.625 < 1.$$

Therefore, the partial derivative conditions in (16) are met and Theorem 3.17 implies that fixed-point iteration will converge to $(p, q) \approx (-0.2222146, 0.9938084)$. Notice that near the other fixed point $(1.90068, 0.31122)$ the partial derivatives do not meet the conditions in (16); hence convergence is not guaranteed. That is,

$$\left|\frac{\partial}{\partial x} g_1(1.90068, 0.31122)\right| + \left|\frac{\partial}{\partial y} g_1(1.90068, 0.31122)\right| = 2.40068 > 1,$$
$$\left|\frac{\partial}{\partial x} g_2(1.90068, 0.31122)\right| + \left|\frac{\partial}{\partial y} g_2(1.90068, 0.31122)\right| = 1.16395 > 1.$$

Seidel Iteration

An improvement, analogous to the Gauss-Seidel method for linear systems, of fixed-point iteration can be made. Suppose that p_{k+1} is used in the calculation of q_{k+1} (in three dimensions both p_{k+1} and q_{k+1} are used to compute r_{k+1}). When these modifications are incorporated in formulas (14) and (15), the method is called ***Seidel iteration***:

$$p_{k+1} = g_1(p_k, q_k) \quad \text{and} \quad q_{k+1} = g_2(p_{k+1}, q_k) \tag{18}$$

and

$$\begin{aligned} p_{k+1} &= g_1(p_k, q_k, r_k) \\ q_{k+1} &= g_2(p_{k+1}, q_k, r_k) \\ r_{k+1} &= g_3(p_{k+1}, q_{k+1}, r_k). \end{aligned} \tag{19}$$

Program 3.6 will implement Seidel iteration for nonlinear systems. Implementation of fixed-point iteration is left for the reader.

Newton's Method for Nonlinear Systems

We now outline the derivation of Newton's method in two dimensions. Newton's method can easily be extended to higher dimensions.

Consider the system

$$\begin{aligned} u &= f_1(x, y) \\ v &= f_2(x, y), \end{aligned} \tag{20}$$

which can be considered a transformation from the xy-plane to the uv-plane. We are interested in the behavior of this transformation near the point (x_0, y_0) whose image is the point (u_0, v_0). If the two functions have continuous partial derivatives, then the differential can be used to write a system of linear approximations that is valid near the point (x_0, y_0):

$$
\begin{aligned}
u - u_0 &\approx \frac{\partial}{\partial x} f_1(x_0, y_0)(x - x_0) + \frac{\partial}{\partial y} f_1(x_0, y_0)(y - y_0), \\
v - v_0 &\approx \frac{\partial}{\partial x} f_2(x_0, y_0)(x - x_0) + \frac{\partial}{\partial y} f_2(x_0, y_0)(y - y_0).
\end{aligned} \tag{21}
$$

The system (21) is a local linear transformation that relates small changes in the independent variables to small changes in the dependent variable. When the Jacobian matrix $\boldsymbol{J}(x_0, y_0)$ is used, this relationship is easier to visualize:

$$
\begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x} f_1(x_0, y_0) & \frac{\partial}{\partial y} f_1(x_0, y_0) \\ \frac{\partial}{\partial x} f_2(x_0, y_0) & \frac{\partial}{\partial y} f_2(x_0, y_0) \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}. \tag{22}
$$

If the system in (20) is written as a vector function $\boldsymbol{V} = \boldsymbol{F}(\boldsymbol{X})$, the Jacobian $\boldsymbol{J}(x, y)$ is the two-dimensional analog of the derivative, because (22) can be written as

$$
\Delta \boldsymbol{F} \approx \boldsymbol{J}(x_0, y_0)\, \Delta \boldsymbol{X}. \tag{23}
$$

We now use (23) to derive Newton's method in two dimensions.

Consider the system (20) with u and v set equal to zero:

$$
\begin{aligned}
0 &= f_1(x, y) \\
0 &= f_2(x, y).
\end{aligned} \tag{24}
$$

Suppose that (p, q) is a solution of (24); that is,

$$
\begin{aligned}
0 &= f_1(p, q) \\
0 &= f_2(p, q).
\end{aligned} \tag{25}
$$

To develop Newton's method for solving (24), we need to consider small changes in the functions near the point (p_0, q_0):

$$
\begin{aligned}
\Delta u &= u - u_0, & \Delta p &= x - p_0. \\
\Delta v &= v - v_0, & \Delta q &= y - q_0.
\end{aligned} \tag{26}
$$

Set $(x, y) = (p, q)$ in (20) and use (25) to see that $(u, v) = (0, 0)$. Hence the changes in the dependent variables are

$$
\begin{aligned}
u - u_0 &= f_1(p, q) - f_1(p_0, q_0) = 0 - f_1(p_0, q_0) \\
v - v_0 &= f_2(p, q) - f_2(p_0, q_0) = 0 - f_2(p_0, q_0).
\end{aligned} \tag{27}
$$

Use the result of (27) in (22) to get the linear transformation

$$\begin{bmatrix} \dfrac{\partial}{\partial x} f_1(p_0, q_0) & \dfrac{\partial}{\partial y} f_1(p_0, q_0) \\ \dfrac{\partial}{\partial x} f_2(p_0, q_0) & \dfrac{\partial}{\partial y} f_2(p_0, q_0) \end{bmatrix} \begin{bmatrix} \Delta p \\ \Delta q \end{bmatrix} \approx - \begin{bmatrix} f_1(p_0, q_0) \\ f_2(p_0, q_0) \end{bmatrix}. \tag{28}$$

If the Jacobian $\boldsymbol{J}(p_0, q_0)$ in (28) is nonsingular, we can solve for $\Delta \boldsymbol{P} = \begin{bmatrix} \Delta p & \Delta q \end{bmatrix}' = \begin{bmatrix} p & q \end{bmatrix}' - \begin{bmatrix} p_0 & q_0 \end{bmatrix}'$ as follows:

$$\Delta \boldsymbol{P} \approx -\boldsymbol{J}(p_0, q_0)^{-1} \boldsymbol{F}(p_0, q_0). \tag{29}$$

Then the next approximation $\boldsymbol{P}_1$ to the solution $\boldsymbol{P} = \begin{bmatrix} p & q \end{bmatrix}'$ is

$$\boldsymbol{P}_1 = \boldsymbol{P}_0 + \Delta \boldsymbol{P} = \boldsymbol{P}_0 - \boldsymbol{J}(p_0, q_0)^{-1} \boldsymbol{F}(p_0, q_0). \tag{30}$$

Notice that (30) is the generalization of Newton's method for the one-variable case; that is, $p_1 = p_0 - f(p_0)/f'(p_0)$.

Outline of Newton's Method

Suppose that $\boldsymbol{P}_k$ has been obtained.

Step 1. Evaluate the function

$$\boldsymbol{F}(\boldsymbol{P}_k) = \begin{bmatrix} f_1(p_k, q_k) \\ f_2(p_k, q_k) \end{bmatrix}.$$

Step 2. Evaluate the Jacobian

$$\boldsymbol{J}(\boldsymbol{P}_k) = \begin{bmatrix} \dfrac{\partial}{\partial x} f_1(p_k, q_k) & \dfrac{\partial}{\partial y} f_1(p_k, q_k) \\ \dfrac{\partial}{\partial x} f_2(p_k, q_k) & \dfrac{\partial}{\partial y} f_2(p_k, q_k) \end{bmatrix}.$$

Step 3. Solve the linear system

$$\boldsymbol{J}(\boldsymbol{P}_k)\Delta \boldsymbol{P} = -\boldsymbol{F}(\boldsymbol{P}_k) \qquad \text{for } \Delta \boldsymbol{P}.$$

Step 4. Compute the next point:

$$\boldsymbol{P}_{k+1} = \boldsymbol{P}_k + \Delta \boldsymbol{P}.$$

Now, repeat the process.

Example 3.32. Consider the nonlinear system

$$
\begin{aligned}
0 &= x^2 - 2x - y + 0.5 \\
0 &= x^2 + 4y^2 - 4.
\end{aligned}
$$

Use Newton's method with the starting value $(p_0, q_0) = (2.00, 0.25)$ and compute (p_1, q_1), (p_2, q_2), and (p_3, q_3).

The function vector and Jacobian matrix are

$$\boldsymbol{F}(x, y) = \begin{bmatrix} x^2 - 2x - y + 0.5 \\ x^2 + 4y^2 - 4 \end{bmatrix}, \quad \boldsymbol{J}(x, y) = \begin{bmatrix} 2x - 2 & -1 \\ 2x & 8y \end{bmatrix}.$$

At the point (2.00, 0.25) they take on the values

$$\boldsymbol{F}(2.00, 0.25) = \begin{bmatrix} 0.25 \\ 0.25 \end{bmatrix}, \quad \boldsymbol{J}(2.00, 0.25) = \begin{bmatrix} 2.0 & -1.0 \\ 4.0 & 2.0 \end{bmatrix}.$$

The differentials Δp and Δq are solutions of the linear system

$$\begin{bmatrix} 2.0 & -1.0 \\ 4.0 & 2.0 \end{bmatrix} \begin{bmatrix} \Delta p \\ \Delta q \end{bmatrix} = - \begin{bmatrix} 0.25 \\ 0.25 \end{bmatrix}.$$

A straightforward calculation reveals that

$$\Delta \boldsymbol{P} = \begin{bmatrix} \Delta p \\ \Delta q \end{bmatrix} = \begin{bmatrix} -0.09375 \\ 0.0625 \end{bmatrix}.$$

The next point in the iteration is

$$\boldsymbol{P}_1 = \boldsymbol{P}_0 + \Delta \boldsymbol{P} = \begin{bmatrix} 2.00 \\ 0.25 \end{bmatrix} + \begin{bmatrix} -0.09375 \\ 0.0625 \end{bmatrix} = \begin{bmatrix} 1.90625 \\ 0.3125 \end{bmatrix}.$$

Similarly, the next two points are

$$\boldsymbol{P}_2 = \begin{bmatrix} 1.900691 \\ 0.311213 \end{bmatrix} \quad \text{and} \quad \boldsymbol{P}_3 = \begin{bmatrix} 1.900677 \\ 0.311219 \end{bmatrix}.$$

The coordinates of $\boldsymbol{P}_3$ are accurate to six decimal places. Calculations for finding $\boldsymbol{P}_2$ and $\boldsymbol{P}_3$ are summarized in Table 3.7. ∎

Table 3.7 Function Values, Jacobian Matrices, and Differentials Required for Each Iteration in Newton's Solution to Example 3.32

$\mathbf{P}_k$	Solution of the linear system $J(\mathbf{P}_k)\Delta\mathbf{P} = -\mathbf{F}(\mathbf{P}_k)$	$\mathbf{P}_k + \Delta\mathbf{P}$
$\begin{bmatrix} 2.00 \\ 0.25 \end{bmatrix}$	$\begin{bmatrix} 2.0 & -1.0 \\ 4.0 & 2.0 \end{bmatrix} \begin{bmatrix} -0.09375 \\ 0.0625 \end{bmatrix} = - \begin{bmatrix} 0.25 \\ 0.25 \end{bmatrix}$	$\begin{bmatrix} 1.90625 \\ 0.3125 \end{bmatrix}$
$\begin{bmatrix} 1.90625 \\ 0.3125 \end{bmatrix}$	$\begin{bmatrix} 1.8125 & -1.0 \\ 3.8125 & 2.5 \end{bmatrix} \begin{bmatrix} -0.005559 \\ -0.001287 \end{bmatrix} = - \begin{bmatrix} 0.008789 \\ 0.024414 \end{bmatrix}$	$\begin{bmatrix} 1.900691 \\ 0.311213 \end{bmatrix}$
$\begin{bmatrix} 1.900691 \\ 0.311213 \end{bmatrix}$	$\begin{bmatrix} 1.801381 & -1.000000 \\ 3.801381 & 2.489700 \end{bmatrix} \begin{bmatrix} -0.000014 \\ 0.000006 \end{bmatrix} = - \begin{bmatrix} 0.000031 \\ 0.000038 \end{bmatrix}$	$\begin{bmatrix} 1.900677 \\ 0.311219 \end{bmatrix}$

Implementation of Newton's method can require the determination of several partial derivatives. It is permissible to use numerical approximations for the values of these partial derivatives, but care must be taken to determine the proper step size. In higher dimensions it is necessary to use the methods for solving linear systems introduced earlier in this chapter to solve for $\Delta\mathbf{P}$.

Numerical Methods Using Matlab, 4th Edition, 2004

John H. Mathews and Kurtis K. Fink

ISBN: 0-13-065248-2

Prentice-Hall Inc.

Upper Saddle River, New Jersey, USA

http://vig.prenhall.com/