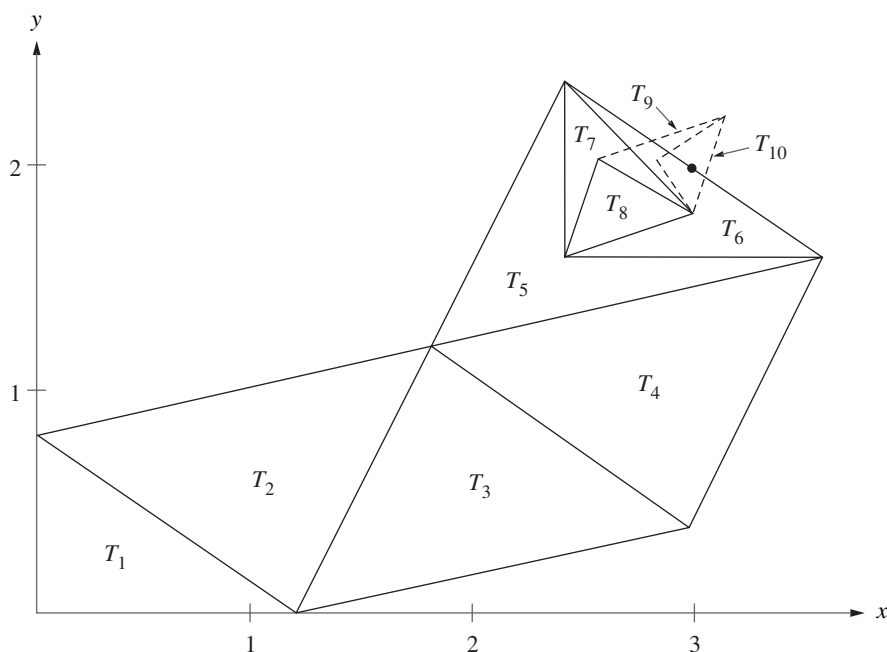


**Powell's Method****Powell's Method**

Let  $\mathbf{X}_0$  be an initial guess at the location of the minimum of the function  $z = f(x_1, x_2, \dots, x_N)$ . Assume that the partial derivatives of the function are not available. An intuitively appealing approach to approximating a minimum of the function  $f$  is to generate the next approximation  $\mathbf{X}_1$  by proceeding successively to a minimum of  $f$  along each of the  $N$  standard base vectors. The process generates the sequence of



**Figure 8.10** The sequence of triangles  $\{T_k\}$  converging to the point  $(3, 2)$  for the Nelder-Mead method.

points  $\mathbf{X}_0 = \mathbf{P}_0, \mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_N = \mathbf{X}_1$ . Along each standard base vector the function  $f$  is a function of one variable. Thus the minimization of  $f$  requires the application of either the golden ratio or Fibonacci searches (Section 8.1) on an interval over which the function is unimodal. The iteration is then repeated to generate a sequence of points  $\{\mathbf{X}_k\}_{k=0}^{\infty}$ . Unfortunately, the method is, in general, inefficient due to the geometry of multivariable functions. But the step from the point  $\mathbf{X}_0$  to the point  $\mathbf{X}_1$  is the first step of **Powell's method**.

The essence of Powell's method is to add two steps to the process described in the preceding paragraph. The vector  $\mathbf{P}_N - \mathbf{P}_0$  represents, in some sense, the *average direction* moved during each iteration. Thus the point  $\mathbf{X}_1$  is determined to be the point at which the minimum of the function  $f$  occurs along the vector  $\mathbf{P}_N - \mathbf{P}_0$ . As before,  $f$  is a function of one variable along this vector and the minimization requires an application of the golden ratio or Fibonacci searches. Finally, since the vector  $\mathbf{P}_N - \mathbf{P}_0$  was such a *good* direction, it replaces one of the direction vectors for the next iteration. The iteration is then repeated using the new set of direction vectors to generate a sequence of points  $\{\mathbf{X}_k\}_{k=0}^{\infty}$ . The process is outlined below.

Let  $\mathbf{X}_0$  be an initial guess at the location of the minimum of the function  $z = f(x_1, x_2, \dots, x_N)$ ,  $\{\mathbf{E}_k = [0 \ 0 \ \dots \ 0 \ 1_k \ 0 \ \dots \ 0] : k = 1, 2, \dots, N\}$  be the set of stan-

**Table 8.6** Function Values at Various Triangles for Example 8.6

$k$	Best point	Good point	Worst point
1	$f(1.2, 0.0) = -3.36$	$f(0.0, 0.8) = -0.16$	$f(0.0, 0.0) = 0.00$
2	$f(1.8, 1.2) = -5.88$	$f(1.2, 0.0) = -3.36$	$f(0.0, 0.8) = -0.16$
3	$f(1.8, 1.2) = -5.88$	$f(3.0, 0.4) = -4.44$	$f(1.2, 0.0) = -3.36$
4	$f(3.6, 1.6) = -6.24$	$f(1.8, 1.2) = -5.88$	$f(3.0, 0.4) = -4.44$
5	$f(3.6, 1.6) = -6.24$	$f(2.4, 2.4) = -6.24$	$f(1.8, 1.2) = -5.88$
6	$f(2.4, 1.6) = -6.72$	$f(3.6, 1.6) = -6.24$	$f(2.4, 2.4) = -6.24$
7	$f(3.0, 1.8) = -6.96$	$f(2.4, 1.6) = -6.72$	$f(2.4, 2.4) = -6.24$
8	$f(3.0, 1.8) = -6.96$	$f(2.55, 2.05) = -6.7725$	$f(2.4, 1.6) = -6.72$
9	$f(3.0, 1.8) = -6.96$	$f(3.15, 2.25) = -6.9525$	$f(2.55, 2.05) = -6.7725$
10	$f(3.0, 1.8) = -6.96$	$f(2.8125, 2.0375) = -6.95640625$	$f(3.15, 2.25) = -6.9525$

dard base vectors,

$$(9) \quad \mathbf{U} = [\mathbf{U}'_1 \mathbf{U}'_2 \cdots \mathbf{U}'_N] = [\mathbf{E}'_1 \mathbf{E}'_2 \cdots \mathbf{E}'_N],$$

and  $i = 0$ .

- (i) Set  $\mathbf{P}_0 = \mathbf{X}_i$ .
- (ii) For  $k = 1, 2, \dots, N$  find the value of  $\gamma_k$  that minimizes  $f(\mathbf{P}_{k-1} + \gamma_k \mathbf{U}_k)$  and set  $\mathbf{P}_k = \mathbf{P}_{k-1} + \gamma_k \mathbf{U}_k$ .
- (iii) Set  $i = i + 1$ .
- (iv) Set  $\mathbf{U}_j = \mathbf{U}_{j+1}$  for  $j = 1, 2, \dots, N - 1$ . Set  $\mathbf{U}_N = \mathbf{P}_N - \mathbf{P}_0$ .
- (v) Find the value of  $\gamma$  that minimizes  $f(\mathbf{P}_0 + \gamma \mathbf{U}_N)$ . Set  $\mathbf{X}_i = \mathbf{P}_0 + \gamma \mathbf{U}_N$ .
- (vi) Repeat steps (i) through (v).

**Example 8.7.** Use the process described in the preceding paragraph to find  $\mathbf{X}_1$  and  $\mathbf{X}_2$  for the function  $f(x, y) = \cos(x) + \sin(y)$ . Use the initial point  $\mathbf{X}_0 = (5.5, 2)$ .

Let  $\mathbf{U} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$  and  $\mathbf{P}_0 = \mathbf{X}_0 = (5.5, 2)$ . When  $i = 1$  the function

$$\begin{aligned}
 f(\mathbf{P}_0 + \gamma_1 \mathbf{U}_1) &= f((5.5, 2) + \gamma_1(1, 0)) \\
 &= f(5.5 + \gamma_1, 2) \\
 &= \cos(5.5 + \gamma_1) + \sin(2)
 \end{aligned}$$

has a minimum at  $\gamma_1 = -2.3584042$ . Thus  $\mathbf{P}_1 = (3.1415958, 2)$ . When  $i = 2$  the function

$$\begin{aligned} f(\mathbf{P}_1 + \gamma_2 \mathbf{U}_2) &= f((3.1415958, 2) + \gamma_2(0, 1)) \\ &= f(3.1415958, 2 + \gamma_2) \\ &= \cos(3.1415958) + \sin(2 + \gamma_2) \end{aligned}$$

has a minimum at  $\gamma_2 = 2.7123803$ . Thus  $\mathbf{P}_2 = (3.1415958, 4.7123803)$ . Set  $\mathbf{U}'_2 = (\mathbf{P}_2 - \mathbf{P}_0)'$  and

$$\mathbf{U} = \begin{bmatrix} 0 & -2.3584042 \\ 1 & 2.7123803 \end{bmatrix}.$$

The function

$$\begin{aligned} f(\mathbf{P}_0 + \gamma \mathbf{U}_2) &= f((5.5, 2) + \gamma(-2.3584042, 2.7123803)) \\ &= f(5.5 - 2.3584042\gamma, 2 + 2.7123803\gamma) \\ &= \cos(5.5 - 2.3584042\gamma) + \sin(2 + 2.7123803\gamma) \end{aligned}$$

has a minimum at  $\gamma = 0.9816697$ . Thus  $\mathbf{X}_1 = (3.1848261, 4.6626615)$ .

Set  $\mathbf{P}_0 = \mathbf{X}_1$ . When  $i = 1$  the function

$$\begin{aligned} f(\mathbf{P}_0 + \gamma_1 \mathbf{U}_1) &= f((3.1848261, 4.6626615) + \gamma_1(0, 1)) \\ &= f(3.1848261, 4.6626615 + \gamma_1) \\ &= \cos(3.1848261) + \sin(4.6626615 + \gamma_1) \end{aligned}$$

has a minimum at  $\gamma_1 = 0.0497117$ . Thus  $\mathbf{P}_1 = (3.1848261, 4.7123732)$ . When  $i = 2$  the function

$$\begin{aligned} f(\mathbf{P}_1 + \gamma_2 \mathbf{U}_2) &= f((3.1848261, 4.7123732) + \gamma_2(-2.3584042, 2.7123809)) \\ &= f(3.1848261 - 2.3584042\gamma_2, 4.7123732 + 2.7123809\gamma_2) \\ &= \cos(3.1848261 - 2.3584042\gamma_2) + \sin(4.7123732 + 2.7123809\gamma_2) \end{aligned}$$

has a minimum at  $\gamma_2 = 0.0078820$ . Thus  $\mathbf{P}_2 = (3.1662373, 4.7337521)$ . Set  $\mathbf{U}'_2 = (\mathbf{P}_2 - \mathbf{P}_0)'$  and

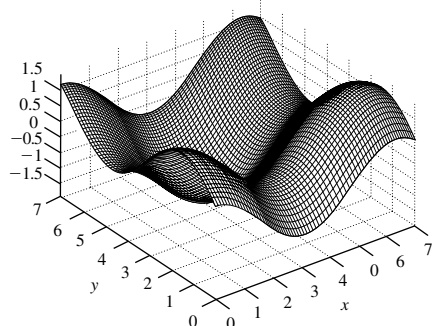
$$\mathbf{U} = \begin{bmatrix} -2.3584042 & -0.0185889 \\ 2.7123803 & 0.0710906 \end{bmatrix}.$$

The function

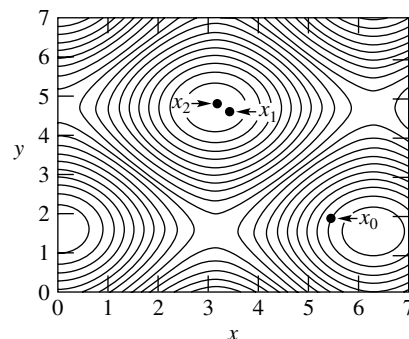
$$\begin{aligned} f(\mathbf{P}_0 + \gamma \mathbf{U}_2) &= f((3.1848261, 4.6626615) + \gamma(-0.0185889, 0.0710906)) \\ &= f(3.1848261 - 0.0185889\gamma, 4.6626615 + 0.0710906\gamma) \\ &= \cos(3.1848261 - 0.0185889\gamma) + \sin(4.6626615 + 0.0710906\gamma) \end{aligned}$$

has a minimum at  $\gamma = 0.8035684$ . Thus  $\mathbf{X}_2 = (3.1698887, 4.7197876)$ .

The function  $f(x, y) = \cos(x) + \sin(y)$  has a relative minimum at the point  $\mathbf{P} = (\pi, 3\pi/2)$ . The graph of  $f$  is shown in Figure 8.11. Figure 8.12 shows a contour plot of the function  $f$  and the relative positions of the points  $\mathbf{X}_0$ ,  $\mathbf{X}_1$ , and  $\mathbf{X}_2$ . ■



**Figure 8.11** The graph of  $f(x, y) = \cos(x) + \sin(y)$ .



**Figure 8.12** The contour graph of  $f(x, y) = \cos(x) + \sin(y)$ .

In step (iv) of the previous process the first vector  $\mathbf{U}_1$  was discarded and the average direction vector  $\mathbf{P}_N - \mathbf{P}_0$  was added to the list of direction vectors. In fact, it would be better to discard the vector  $\mathbf{U}_r$  along which the greatest decrease in  $f$  occurred. It seems reasonable that the vector  $\mathbf{U}_r$  is a large component of the average direction vector  $\mathbf{U}_N = \mathbf{P}_N - \mathbf{P}_0$ . Thus, as the number of iterations increase, the set of direction vectors will tend to become *linearly dependent*. When the set becomes linearly dependent one or more of the directions will be lost and it is likely that the set of points  $\{\mathbf{X}\}_{k=0}^{\infty}$  will not converge to the point at which the local minimum occurs. Furthermore, in step (iv) it was assumed that the average direction vector represented a good direction in which to continue the search. But that may not be the case.

### Outline of Powell's Method

- (i) Set  $\mathbf{P}_0 = \mathbf{X}_i$ .
- (ii) For  $k = 1, 2, \dots, N$  find the value of  $\gamma_k$  that minimizes  $f(\mathbf{P}_{k-1} + \gamma_k \mathbf{U}_k)$  and set  $\mathbf{P}_k = \mathbf{P}_{k-1} + \gamma_k \mathbf{U}_k$ .
- (iii) Set  $r$  and  $\mathbf{U}_r$  equal to the maximum decrease in  $f$  and the direction of the maximum decrease, respectively, over all the direction vectors in step (ii).
- (iv) Set  $i = i + 1$ .
- (v) If  $f(2\mathbf{P}_N - \mathbf{P}_0) \geq f(\mathbf{P}_0)$  or

$$2(f(\mathbf{P}_0) - 2f(\mathbf{P}_N) + f(2\mathbf{P}_N - \mathbf{P}_0))(f(\mathbf{P}_0) - f(\mathbf{P}_N) - r)^2 \geq r(f(\mathbf{P}_0) - f(2\mathbf{P}_N - \mathbf{P}_0))^2,$$

then set  $\mathbf{X}_i = \mathbf{P}_N$  and return to step (i). Otherwise, go to step (vi).

- (vi) Set  $\mathbf{U}_r = \mathbf{P}_N - \mathbf{P}_0$ .
- (vii) Find the value of  $\gamma$  that minimizes  $f(\mathbf{P}_0 + \gamma\mathbf{U}_r)$ . Set  $\mathbf{X}_i = \mathbf{P}_0 + \gamma\mathbf{U}_r$ .
- (viii) Repeat steps (i) through (vii).

If the conditions in step (v) are satisfied, then the set of direction vectors is left unchanged. The first inequality in step (v) indicates that there is no further decrease in the value of  $f$  in the average direction  $\mathbf{P}_N - \mathbf{P}_0$ . The second inequality indicates that the decrease in the function  $f$  in the direction of greatest decrease  $\mathbf{U}_r$  was not a major part of the total decrease in  $f$  in step (ii). If the conditions in step (v) are not satisfied, then the direction of greatest decrease  $\mathbf{U}_r$  is replaced with the average direction from step (ii);  $\mathbf{P}_N - \mathbf{P}_0$ . In step (vii) the function is minimized in this direction. Stopping criteria based on the magnitudes  $\|\mathbf{X}_i - \mathbf{X}_{i-1}\|$  or  $\|f(\mathbf{X}_i)\|$  are typically found in steps (v) and (vii).

**Numerical Methods Using Matlab, 4<sup>th</sup> 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/>

# NUMERICAL METHODS USING MATLAB

FOURTH EDITION



JOHN H. MATHEWS • KURTIS D. FINK