point is found. This is because the search directions $\mathbf{S}_i$ might become dependent or almost dependent during numerical computation.

Convergence Criterion. The convergence criterion one would generally adopt in a method such as Powell's method is to stop the procedure whenever a minimization cycle produces a change in all variables less than one-tenth of the required accuracy. However, a more elaborate convergence criterion, which is more likely to prevent premature termination of the process, was given by Powell [6.7].

Example 6.7 Minimize $f(x_1,x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ from the starting point $\mathbf{X}_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$ using Powell's method.

SOLUTION

Cycle 1: Univariate Search

We minimize f along $\mathbf{S}_2 = \mathbf{S}_n = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$ from $\mathbf{X}_1$. To find the correct direction ($+\mathbf{S}_2$ or $-\mathbf{S}_2$) for decreasing the value of f, we take the probe length as $\varepsilon = 0.01$. As $f_1 = f(\mathbf{X}_1) = 0.0$, and

$$f^+ = f(\mathbf{X}_1 + \varepsilon\mathbf{S}_2) = f(0.0,0.01) = -0.0099 < f_1$$

f decreases along the direction $+\mathbf{S}_2$. To find the minimizing step length λ^* along $\mathbf{S}_2$, we minimize

$$f(\mathbf{X}_1 + \lambda\,\mathbf{S}_2) = f(0.0,\lambda) = \lambda^2 - \lambda$$

As $df/d\lambda = 0$ at $\lambda^* = \frac{1}{2}$, we have $\mathbf{X}_2 = \mathbf{X}_1 + \lambda^*\mathbf{S}_2 = \begin{Bmatrix} 0 \\ 0.5 \end{Bmatrix}$.

Next we minimize f along $\mathbf{S}_1 = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix}$ from $\mathbf{X}_2 = \begin{Bmatrix} 0.0 \\ 0.5 \end{Bmatrix}$. Since

$$f_2 = f(\mathbf{X}_2) = f(0.0,0.5) = -0.25$$

$$f^+ = f(\mathbf{X}_2 + \varepsilon\mathbf{S}_1) = f(0.01,0.50) = -0.2298 > f_2$$

$$f^- = f(\mathbf{X}_2 - \varepsilon\mathbf{S}_1) = f(-0.01,0.50) = -0.2698$$

f decreases along $-\mathbf{S}_1$. As $f(\mathbf{X}_2 - \lambda\mathbf{S}_1) = f(-\lambda,0.50) = 2\lambda^2 - 2\lambda - 0.25$, $df/d\lambda = 0$ at $\lambda^* = \frac{1}{2}$. Hence $\mathbf{X}_3 = \mathbf{X}_2 - \lambda^*\mathbf{S}_1 = \begin{Bmatrix} -0.5 \\ 0.5 \end{Bmatrix}$.

Now we minimize f along $\mathbf{S}_2 = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$ from $\mathbf{X}_3 = \begin{Bmatrix} -0.5 \\ 0.5 \end{Bmatrix}$. As $f_3 = f(\mathbf{X}_3) = -0.75$, $f^+ = f(\mathbf{X}_3 + \varepsilon\mathbf{S}_2) = f(-0.5,0.51) = -0.7599 < f_3$, f decreases along $+\mathbf{S}_2$ direction. Since

$$f(\mathbf{X}_3 + \lambda\mathbf{S}_2) = f(-0.5,\, 0.5 + \lambda) = \lambda^2 - \lambda - 0.75, \qquad \frac{df}{d\lambda} = 0 \text{ at } \lambda^* = \frac{1}{2}$$

This gives

$$\mathbf{X}_4 = \mathbf{X}_3 + \lambda^*\mathbf{S}_2 = \begin{Bmatrix} -0.5 \\ 1.0 \end{Bmatrix}$$

Cycle 2: Pattern Search

Now we generate the first pattern direction as

$$\mathbf{S}_p^{(1)} = \mathbf{X}_4 - \mathbf{X}_2 = \begin{Bmatrix} -\frac{1}{2} \\ 1 \end{Bmatrix} - \begin{Bmatrix} 0 \\ \frac{1}{2} \end{Bmatrix} = \begin{Bmatrix} -0.5 \\ 0.5 \end{Bmatrix}$$

and minimize f along $\mathbf{S}_p^{(1)}$ from $\mathbf{X}_4$. Since

$$f_4 = f(\mathbf{X}_4) = -1.0$$

$$f^+ = f(\mathbf{X}_4 + \varepsilon\mathbf{S}_p^{(1)}) = f(-0.5 - 0.005,\, 1 + 0.005)$$

$$= f(-0.505,1.005) = -1.004975$$

f decreases in the positive direction of $\mathbf{S}_p^{(1)}$. As

$$f(\mathbf{X}_4 + \lambda\mathbf{S}_p^{(1)}) = f(-0.5 - 0.5\lambda,\, 1.0 + 0.5\lambda)$$

$$= 0.25\lambda^2 - 0.50\lambda - 1.00,$$

$\dfrac{df}{d\lambda} = 0$ at $\lambda^* = 1.0$ and hence

$$\mathbf{X}_5 = \mathbf{X}_4 + \lambda^*\mathbf{S}_p^{(1)} = \begin{Bmatrix} -\frac{1}{2} \\ 1 \end{Bmatrix} + 1.0\begin{Bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{Bmatrix} = \begin{Bmatrix} -1.0 \\ 1.5 \end{Bmatrix}$$

The point $\mathbf{X}_5$ can be identified to be the optimum point.

If we do not recognize $\mathbf{X}_5$ as the optimum point at this stage, we proceed to minimize f along the direction $\mathbf{S}_2 = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$ from $\mathbf{X}_5$. Then we would obtain

$$f_5 = f(\mathbf{X}_5) = -1.25, \quad f^+ = f(\mathbf{X}_5 + \varepsilon\mathbf{S}_2) > f_5,$$
$$\text{and} \quad f^- = f(\mathbf{X}_5 - \varepsilon\mathbf{S}_2) > f_5$$

This shows that f cannot be minimized along $\mathbf{S}_2$, and hence $\mathbf{X}_5$ will be the optimum point. In this example the convergence has been achieved in the second cycle itself. This is to be expected in this case, as f is a quadratic function, and the method is a quadratically convergent method.

6.8 ROSENBROCK'S METHOD OF ROTATING COORDINATES

The method of rotating coordinates, given by Rosenbrock [6.8], can be considered as a further development of the Hooke and Jeeves method. In this method the coordinate system is rotated in each stage of minimization in such a manner that the first axis is oriented toward the locally estimated direction of the valley and all the other axes are made mutually orthogonal and normal to the first one. The details can be found in Refs. [6.8] and [6.9].

6.9 SIMPLEX METHOD

Definition: Simplex. The geometric figure formed by a set of $n + 1$ points in an n-dimensional space is called a *simplex*. When the points are equidistant, the simplex is said to be *regular*. Thus in two dimensions, the simplex is a triangle, and in three dimensions, it is a tetrahedron.

The basic idea in the simplex method† is to compare the values of the objective function at the $n + 1$ vertices of a general simplex and move the simplex gradually toward the optimum point during the iterative process. The following equations can be used to generate the vertices of a regular simplex (equilateral triangle in two-dimensional space) of size a in the n-dimensional space [6.10]:

$$\mathbf{X}_i = \mathbf{X}_0 + p\mathbf{u}_i + \sum_{j=1, j \neq i}^{n} q\mathbf{u}_j, \quad i = 1,2,\ldots,n \tag{6.46}$$

where

$$p = \frac{a}{n\sqrt{2}}(\sqrt{n+1} + n - 1) \quad \text{and} \quad q = \frac{a}{n\sqrt{2}}(\sqrt{n+1} - 1) \tag{6.47}$$

†This simplex method should not be confused with the simplex method of linear programming.