

Nelder-Mead Method**Nelder-Mead Method**

A simplex method for finding a local minimum of a function of several variables has been devised by Nelder and Mead. For two variables, a simplex is a triangle, and the method is a pattern search that compares function values at the three vertices of a triangle. The worst vertex, where $f(x, y)$ is largest, is rejected and replaced with a new vertex. A new triangle is formed and the search is continued. The process generates a sequence of triangles (which might have different shapes), for which the function values at the vertices get smaller and smaller. The size of the triangles is reduced and the coordinates of the minimum point are found.

The algorithm is stated using the term *simplex* (a generalized triangle in N dimensions) and will find the minimum of a function of N variables. It is effective and computationally compact.

Initial Triangle BGW

Let $f(x, y)$ be the function that is to be minimized. To start, we are given three vertices of a triangle: $V_k = (x_k, y_k)$, $k = 1, 2, 3$. The function $f(x, y)$ is then evaluated at each of the three points: $z_k = f(x_k, y_k)$ for $k = 1, 2, 3$. The subscripts are then reordered so that $z_1 \leq z_2 \leq z_3$. We use the notation

$$(5) \quad \mathbf{B} = (x_1, y_1), \quad \mathbf{G} = (x_2, y_2), \quad \text{and} \quad \mathbf{W} = (x_3, y_3)$$

to help remember that \mathbf{B} is the best vertex, \mathbf{G} is good (next to best), and \mathbf{W} is the worst vertex.

Midpoint of the Good Side

The construction process uses the midpoint of the line segment joining \mathbf{B} and \mathbf{G} . It is found by averaging the coordinates:

$$(6) \quad \mathbf{M} = \frac{\mathbf{B} + \mathbf{G}}{2} = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right).$$

Reflection Using the Point \mathbf{R}

The function decreases as we move along the side of the triangle from \mathbf{W} to \mathbf{B} , and it decreases as we move along the side from \mathbf{W} to \mathbf{G} . Hence it is feasible that $f(x, y)$ takes on smaller values at points that lie away from \mathbf{W} on the opposite side of the line between \mathbf{B} and \mathbf{G} . We choose a test point \mathbf{R} that is obtained by “reflecting” the triangle through the side \overline{BG} . To determine \mathbf{R} , we first find the midpoint \mathbf{M} of the side \overline{BG} . Then draw the line segment from \mathbf{W} to \mathbf{M} and call its length d . This last segment is extended a distance d through \mathbf{M} to locate the point \mathbf{R} (see Figure 8.6). The vector formula for \mathbf{R} is

$$(7) \quad \mathbf{R} = \mathbf{M} + (\mathbf{M} - \mathbf{W}) = 2\mathbf{M} - \mathbf{W}.$$

Expansion Using the Point \mathbf{E}

If the function value at \mathbf{R} is smaller than the function value at \mathbf{W} , then we have moved in the correct direction toward the minimum. Perhaps the minimum is just a bit farther than the point \mathbf{R} . So we extend the line segment through \mathbf{M} and \mathbf{R} to the point \mathbf{E} . This forms an expanded triangle BGE . The point \mathbf{E} is found by moving an additional distance d along the line joining \mathbf{M} and \mathbf{R} (see Figure 8.7). If the function value at \mathbf{E} is less than the function value at \mathbf{R} , then we have found a better vertex than \mathbf{R} . The vector formula for \mathbf{E} is

$$(8) \quad \mathbf{E} = \mathbf{R} + (\mathbf{R} - \mathbf{M}) = 2\mathbf{R} - \mathbf{M}.$$

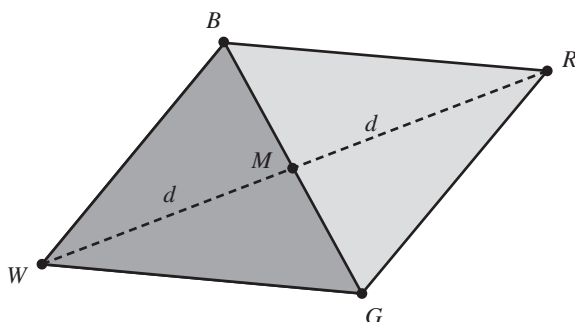


Figure 8.6 The triangle $\triangle BGW$ and midpoint M and reflected point R for the Nelder-Mead method.

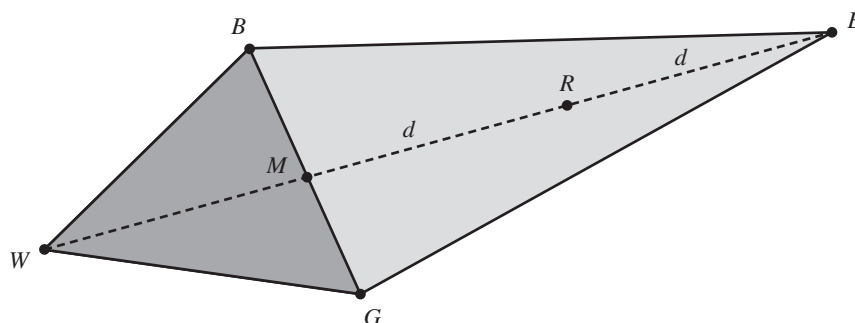


Figure 8.7 The triangle $\triangle BGW$ and point R and extended point E .

Contraction Using the Point C

If the function values at R and W are the same, another point must be tested. Perhaps the function is smaller at M , but we cannot replace W with M because we must have a triangle. Consider the two midpoints C_1 and C_2 of the line segments \overline{WM} and \overline{MR} , respectively (see Figure 8.8). The point with the smaller function value is called C , and the new triangle is BGC . Note. The choice between C_1 and C_2 might seem inappropriate for the two-dimensional case, but it is important in higher dimensions.

Shrink toward B

If the function value at C is not less than the value at W , the points G and W must be shrunk toward B (see Figure 8.9). The point G is replaced with M , and W is replaced with S , which is the midpoint of the line segment joining B with W .

Logical Decisions for Each Step

A computationally efficient algorithm should perform function evaluations only if needed. In each step, a new vertex is found, which replaces W . As soon as it is

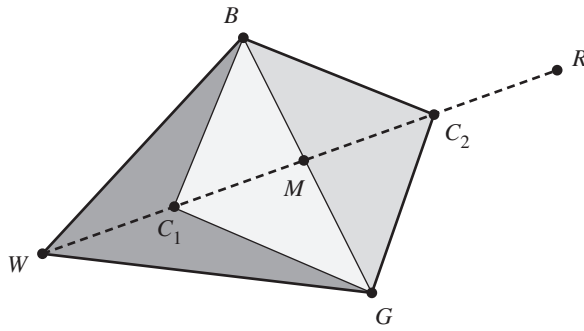


Figure 8.8 The contraction point C_1 or C_2 for Nelder-Mead method.

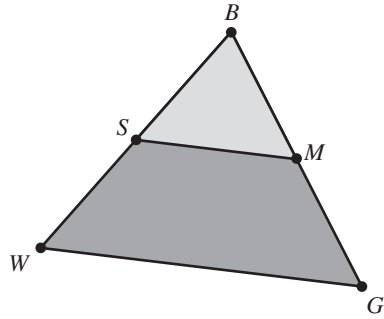


Figure 8.9 Shrinking the triangle toward B .

found, further investigation is not needed, and the iteration step is completed. The logical details for two-dimensional cases are explained in Table 8.5.

Example 8.6. Use the Nelder-Mead algorithm to find the minimum of $f(x, y) = x^2 - 4x + y^2 - y - xy$. Start with the three vertices

$$V_1 = (0, 0), \quad V_2 = (1.2, 0.0), \quad V_3 = (0.0, 0.8).$$

The function $f(x, y)$ takes on the values

$$f(0, 0) = 0.0, \quad f(1.2, 0.0) = -3.36, \quad f(0.0, 0.8) = -0.16.$$

The function values must be compared to determine B , G , and W ;

$$B = (1.2, 0.0), \quad G = (0.0, 0.8), \quad W = (0, 0).$$

The vertex $W = (0, 0)$ will be replaced. The points M and R are

$$M = \frac{B + G}{2} = (0.6, 0.4) \quad \text{and} \quad R = 2M - W = (1.2, 0.8).$$

The function value $f(R) = f(1.2, 0.8) = -4.48$ is less than $f(G)$, so the situation is case (i). Since $f(R) \leq f(B)$, we have moved in the right direction, and the vertex E must

Table 8.5 Logical Decisions for the Nelder-Mead Algorithm

IF $f(R) < f(G)$, THEN Perform Case (i) {either reflect or extend}	
ELSE Perform Case (ii) {either contract or shrink}	
BEGIN {Case (i).}	BEGIN {Case (ii).}
IF $f(B) < f(R)$ THEN	IF $f(R) < f(W)$ THEN
replace W with R	replace W with R
ELSE	Compute $C = (W + M)/2$
Compute E and $f(E)$	or $C = (M + R)/2$ and $f(C)$
IF $f(E) < f(B)$ THEN	IF $f(C) < f(W)$ THEN
replace W with E	replace W with C
ELSE	ELSE
replace W with R	Compute S and $f(S)$
ENDIF	replace W with S
ENDIF	replace G with M
END {Case (i).}	ENDIF
	END {Case (ii).}

be constructed:

$$E = 2R - M = 2(1.2, 0.8) - (0.6, 0.4) = (1.8, 1.2).$$

The function value $f(E) = f(1.8, 1.2) = -5.88$ is less than $f(B)$, and the new triangle has vertices

$$V_1 = (1.8, 1.2), \quad V_2 = (1.2, 0.0), \quad V_3 = (0.0, 0.8).$$

The process continues and generates a sequence of triangles that converges down on the solution point $(3, 2)$ (see Figure 8.10). Table 8.6 gives the function values at vertices of the triangle for several steps in the iteration. A computer implementation of the algorithm continued until the thirty-third step, where the best vertex was $B = (2.99996456, 1.99983839)$ and $f(B) = -6.99999998$. These values are approximations to $f(3, 2) = -7$ found in Example 8.5. The reason that the iteration quit before $(3, 2)$ was obtained is that the function is flat near the minimum. The function values $f(B)$, $f(G)$, and $f(W)$ were checked (see Table 8.6) and found to be the same (this is an example of round-off error), and the algorithm was terminated. ■

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

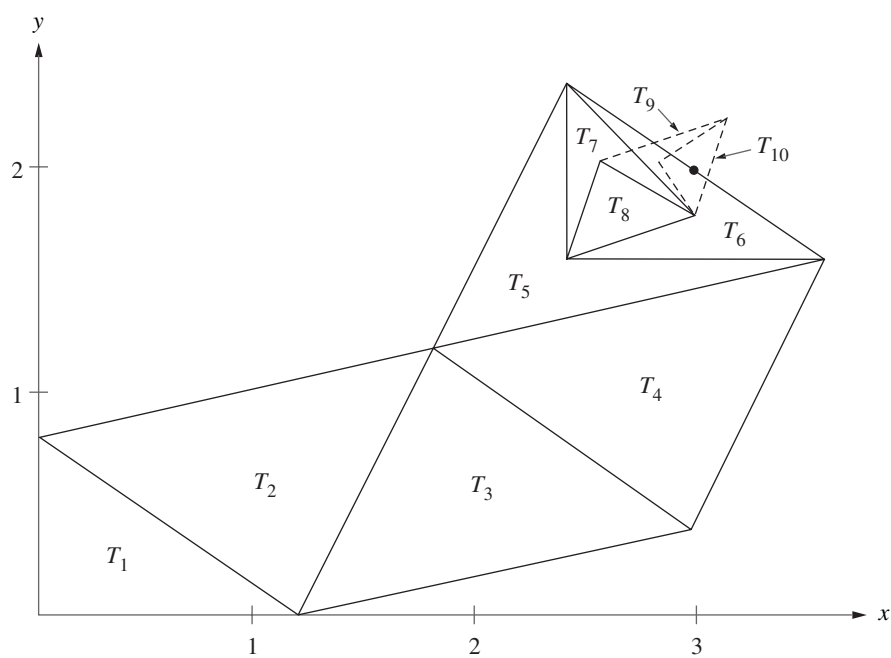


Figure 8.10 The sequence of triangles $\{T_k\}$ converging to the point $(3, 2)$ for the Nelder-Mead method.

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	$f(3.0, 1.8) = -6.96$	$f(2.8125, 2.0375) = -6.95640625$	$f(3.15, 2.25) = -6.9525$

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/>

NUMERICAL METHODS USING MATLAB

FOURTH EDITION



JOHN H. MATHEWS • KURTIS D. FINK