joining $B$ and $W$ (Figure 2.23). Similar to $M$, the midpoint $S$ can be calculated from:

$$S = \frac{B + W}{2} \tag{2.105}$$

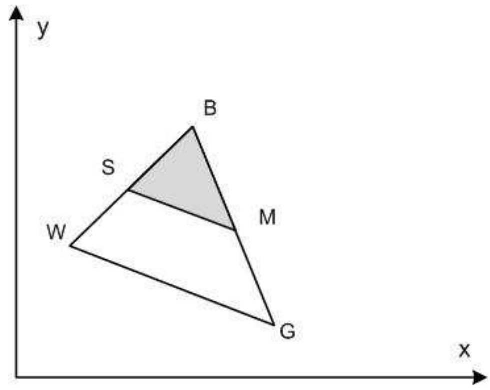


Figure 2.23: Shrinking toward B

The algorithm must terminate in a finite number of steps or iterations. The stop criterion can be one or all of the following:

- The simplex becomes sufficiently small, or the vertices are within a given tolerance $\varepsilon$
- The function does not change for successive iterations.
- The number of iterations or function evaluations exceeds a maximum value allowed.

For the two-dimensional case, the details are summarized in Algorithm 11, (Mathews and Fink, 2004).

**Example 2.5** *Use the Nelder-Mead algorithm to find the minimum of:*

$$f(x, y) = (x - 10)^2 + (y - 10)^2 \tag{2.106}$$

*We calculate the first steps and illustrate them graphically.*

**Algorithm 11** Nelder-Mead method

```
Define function f(x, y) and tolerance ε
Select 3 initial vertices V1, V2, V3
while stop criterion not fulfilled do
  Compute f(V1), f(V2), f(V3) and set the labels B, G, W
  Compute M = (B + G)/2 and R = 2M − W
  if f(R) < f(W) then
    Compute E=2R-M, f(E)
    if f(E) < f(R) then
      Replace W with E
    else
      Replace W with R
    end if
  else
    Compute C1 and C2 and choose C
    if f(C) < f(W) then
      Replace W with C
    else
      Compute S and f(S)
      Replace W with S
      Replace G with M
    end if
  end if
end while
```

**First iteration** *Start with three vertices conveniently chosen at the origin and on the axes (Figure 2.24):*

$$V_1 = (0,0), \quad V_2 = (2,0), \quad V_3 = (0,6) \tag{2.107}$$

*and calculate the function values $f(x,y)$ at these points:*

$$f(0,0) = 200, \quad f(2,0) = 164, \quad f(0,6) = 116 \tag{2.108}$$

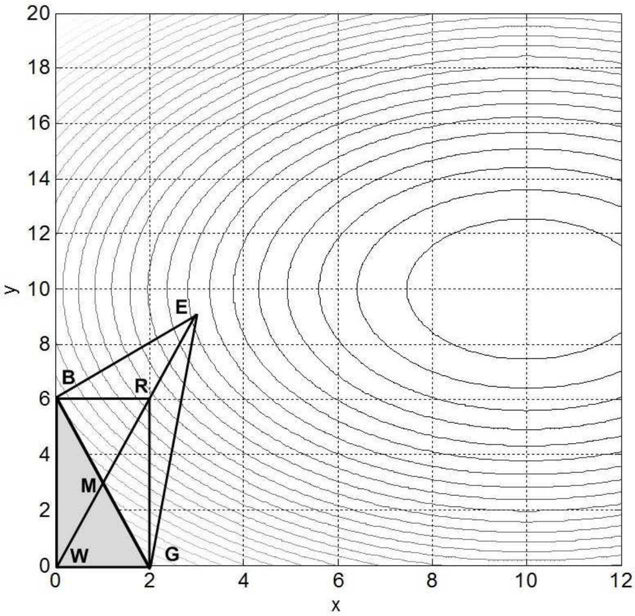


Figure 2.24: The first sequence of triangles

*Since $f(0,6) < f(2,0) < f(0,0)$, the vertices will be assigned the names:*

$$B = (0,6), \quad G = (2,0), \quad W = (0,0) \tag{2.109}$$

*The midpoint of BG is:*

$$M = \frac{B+G}{2} = (1,3) \tag{2.110}$$

*and the reflected point:*

$$R = 2M - W = (2, 6) \tag{2.111}$$

*The function value at* $R$*,* $f(R) = f(2,6) = 80 < f(G) < f(B)$ *and the extended point* $E$ *will be calculated from:*

$$E = 2R - M = (3, 9) \tag{2.112}$$

*The new triangle is* $\Delta BGE$*.*

**Second iteration** *Since the function value at* $E$ *is* $f(3,9) = 50 < f(G) < f(B)$*, the vertices labels for the current stage are (Figure 2.25):*

$$B = (3, 9), \quad G = (0, 6), \quad W = (2, 0) \tag{2.113}$$

*The new midpoint of* $BG$ *and the reflected point* $R$ *are:*

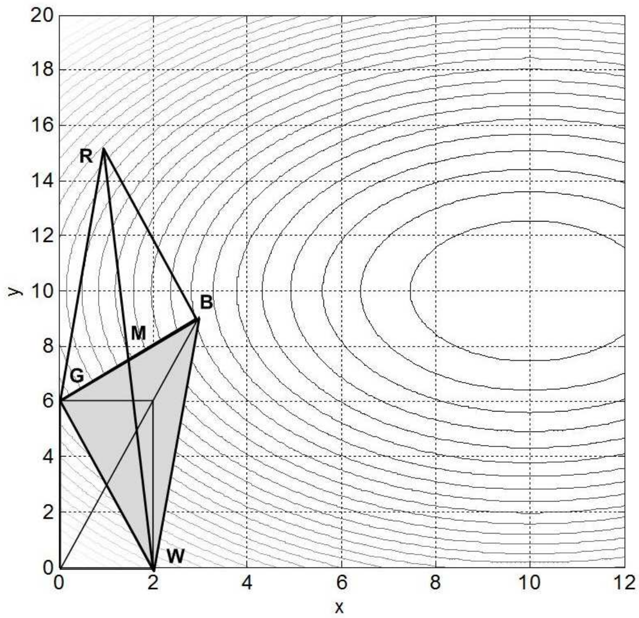


Figure 2.25: The second sequence of triangles

$$M = \frac{B+G}{2} = (1.5, 7.5), \quad R = 2M - W = (1, 15) \tag{2.114}$$

*The value of $f$ at the reflected point is: $f(1,15) = 106$. It is less than $f(G) = f(0,6) = 116$ but greater than $f(B) = f(3,9) = 50$. Therefore, the extended point will not be calculated and the new triangle is $\Delta RGB$.*

**Third iteration** *Because $f(3,9) < f(1,15) < f(0,6)$, the new labels are (Figure 2.26):*

$$B = (3,9), \quad G = (1,15), \quad W = (0,6) \tag{2.115}$$

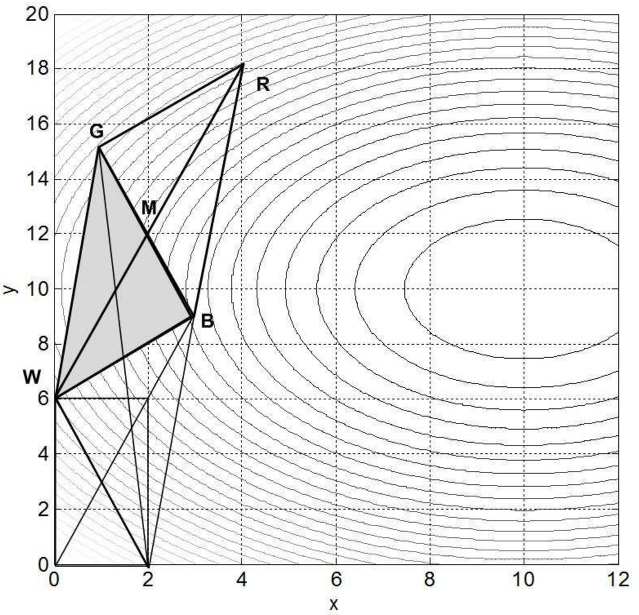


Figure 2.26: The third sequence of triangles

*A new midpoint and a reflected point are calculated:*

$$M = \frac{B+G}{2} = (2, 12), \quad R = 2M - W = (4, 18) \tag{2.116}$$

*and the function takes at $R$ the value $f(R) = f(4,18) = 100$. This is again less than the value at $G$ but it is not less than the value at $B$, thus the point $R$ will be a vertex of the next triangle $\Delta RGB$.*

**Fourth iteration** *Because $f(3,9) < f(4,18) < f(1,15)$, the new labels are (Fig-*

*ure 2.27):*

$$B = (3, 9), \quad G = (4, 18), \quad W = (1, 15) \tag{2.117}$$

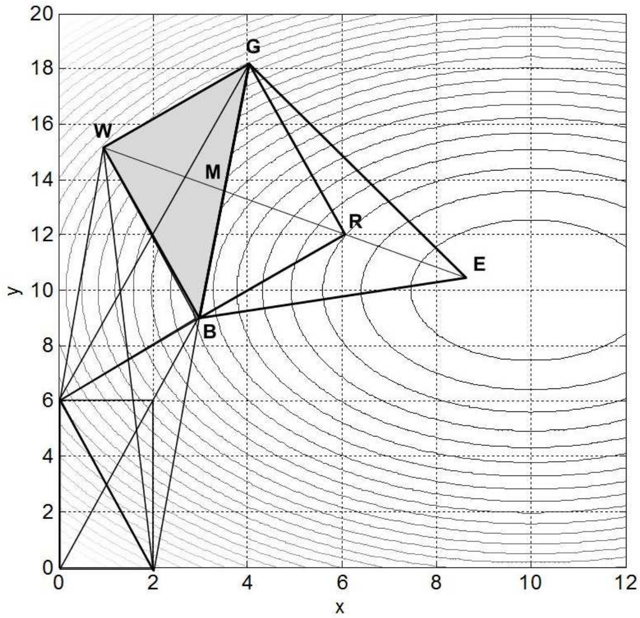


Figure 2.27: The fourth sequence of triangles

*The midpoint of $BG$ and $R$ result as:*

$$M = \frac{B+G}{2} = (3.5, 13.5), \quad R = 2M - W = (6, 12) \tag{2.118}$$

*Because $f(R) = f(6, 12) = 20 < f(B)$, we shall build the extended point $E$:*

$$E = 2R - M = (8.5, 10.5) \tag{2.119}$$

*The function takes here the value $f(E) = f(8.5, 10.5) = 2.5$, thus the new triangle is $\Delta GBE$.*

*The calculations continue until the function values at the vertices have almost equal values (with some allowable tolerance $\varepsilon$). It will still take a large number of steps until the procedure will reach the minimum point $(10, 10)$ where the function* *value is* $0$*, but the steps performed above show that the function is decreasing at each iteration.*

#### 2.2.6.2 Rosenbrock method

The main drawback of the Nelder-Mead method is that general convergence properties have not been proven. An alternative for a derivative-free algorithm is the method of Rosenbrock whose global convergence to a local optima is assured.

This zero-order method solves the unconstrained optimization problem of minimizing a nonlinear function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, (Rosenbrock, 1960). Starting at an initial point $\mathbf{x}_0$, the method explores locally along $n$ orthogonal directions seeking for a better solution. The first set of directions is usually composed of the base vectors of an $n$-dimensional coordinate system. During the iterative process, this set is recalculated such that the first vector to point into the direction of the maximum function decrease.

**Example 2.6** *Consider the orthogonal coordinate system* $x_1 - x_2$ *shown in Figure 2.28. A set of normalized vectors in the direction of the axes is* $(\mathbf{d}_1, \mathbf{d}_2)$ *given*

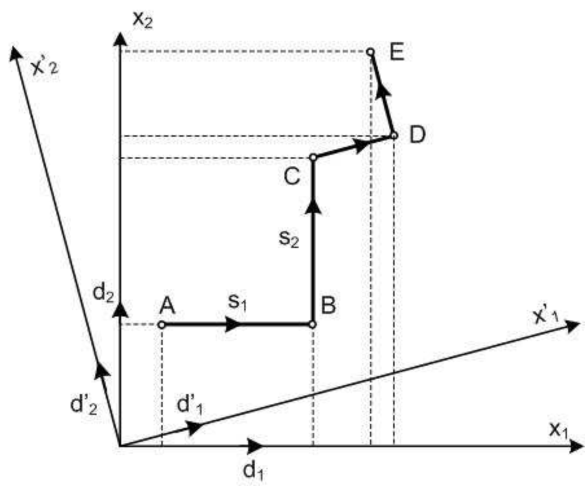


Figure 2.28: Rotation of a coordinate system

*by:*

$$\mathbf{d}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{d}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{2.120}$$