Minimizing Functions of Several Variables

Let \( f(x_1, x_2) \) be a function of two variables whose partial derivatives \( g_i = \frac{\partial f}{\partial x_i} \) can be computed or approximated easily. One type of algorithm for finding a minimum of \( f \) attempts to optimize the process for quadratic functions, assuming that near a minimum the function is nearly quadratic.

Given a starting point \( x_0 \), one takes a straight line path \( L : \{ x = x_0 + tv_0 \} \), where \( v_0 \) is a vector pointing in a suitable direction, and looks for a point \( x_1 = x_0 + t v \) on \( L \) such that \( f(x_1) < f(x_0) \). For example, one might determine \( x_1 \) by minimizing the function \( \phi(t) = f(x_0 + t v_0) \) of one variable. Then this process is repeated.

The method of steepest descent takes for \( v_0 \) the negative gradient \( v_0 = -g_0 = - (\nabla f)_0 \).

Project, Part I

a. Implement the steepest descent method. You may set it up so that you can use eyeball to determine a suitable value of \( t \) at each stage, if you wish.

b. Try your method on various quadratic functions. Make contour plots of the functions you use.

c. Using the function
   \[ 2x + y + x^2 - y^2 + (x - y^2)^2, \]
   run your program starting at the point \((-1, -1.3)\).

d. Using
   \[ (1 - x)^2 + 100(y - x^2)^2, \]
   run your program starting at \((0.676, 0.443)\).

e. Comment on the limitations of the steepest descent method.

Project, Part II

To overcome the limitations of the steepest descent method, some algorithms choose a direction other than the negative gradient. The conjugate gradients method proceeds this way: As a first step, steepest descent is used to obtain the point \( x_1 \). Then if \( g_i \) is the gradient at \( x_i \), the second search direction is taken to be

\[ v_1 = -g_1 - \frac{(g_1 \cdot g_1)}{(g_0 \cdot g_0)} g_0. \]

a. Implement the conjugate gradients method.

b. Try the method on the same functions as before.

c. Explain the phenomena you observe analytically.

Project, Part III (optional)

Try the conjugate gradient method on functions of three variables.