2

One-Dimensional Unconstrained Minimization

2.1 Introduction

Determination of the minimum of a real valued function of one variable, and the location of that minimum, plays an important role in nonlinear optimization. A one-dimensional minimization routine may be called several times in a multivariable problem. We show later that at the minimum point of a sufficiently smooth function, the slope is zero. If the slope and curvature information is available, minimum may be obtained by finding the location where the slope is zero and the curvature is positive. The need to determine the zero of a function occurs frequently in nonlinear optimization. Reliable and efficient ways of finding the minimum or a zero of a function are necessary for developing robust techniques for solving multivariable problems. We present the basic concepts involved in single variable minimization and zero finding.

2.2 Theory Related to Single Variable (Univariate) Minimization

We present the minimization ideas by considering a simple example. The first step is to determine the *objective function* that is to be optimized.

Example 2.1

Determine the objective function for building a minimum cost cylindrical refrigeration tank of volume 50 m^3, if the circular ends cost \$10 per m^2, the cylindrical wall costs \$6 per mm^2, and it costs \$80 per m^2 to refrigerate over the useful life.

Solution

Cylindrical tank of diameter x and volume V is shown in Fig. E2.1.

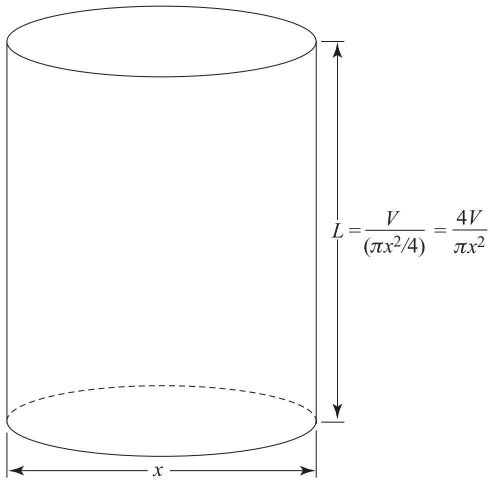

Figure E2.1. Refrigerated tank.

We denote the total cost as f. We have

$$\begin{aligned} f &= (10)(2)\left(\frac{\pi x^2}{4}\right) + (6)(\pi x L) + 80\left(2.\frac{\pi x^2}{4} + \pi x L\right) \\ &= 45\pi x^2 + 86\pi x L, \end{aligned}$$

Substituting

$$L = \frac{(50)(4)}{\pi x^2} = \frac{200}{\pi x^2},$$

we get

$$f = 45\pi x^2 + \frac{17200}{x}$$

The function $f(x)$ is the objective function to be minimized. One variable plot of the function over a sufficiently large range of x shows the distinct characteristics.

The one variable problem may be stated as

$$\text{minimize } f(x) \qquad \text{for all real } x \tag{2.1}$$

The point x^* is a *weak* local minimum if there exists a $\delta > 0$, such that $f(x^*) \leq f(x)$ for all x such that $|x - x^*| < \delta$, that is, $f(x^*) \leq f(x)$ for all x in a δ-neighborhood of x^*. The point x^* is a *strong* (or *strict*) local minimum if $f(x^*) \leq f(x)$ is replaced by $f(x^*) < f(x)$ in the preceding statement. Further, x^* is a *global minimum* if $f(x^*) < f(x)$ for all x. These cases are illustrated in Fig. 2.1. If a minimum does not exist, the function is not bounded below.

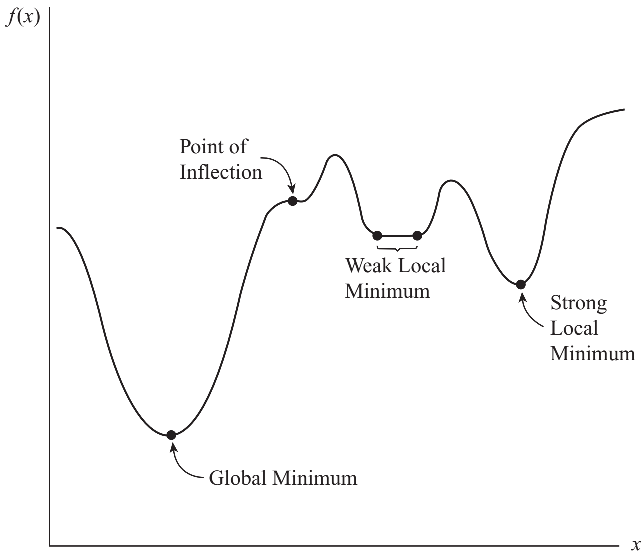

Figure 2.1. One-variable plot.

Optimality Conditions

If the function f in Eq. (2.1) has continuous second-order derivatives everywhere (i.e., C^2 continuous), the necessary conditions for x^* to be a local minimum are given by

$$f'(x^*) = 0 \tag{2.2}$$

$$f''(x^*) \geq 0 \tag{2.3}$$

where $f'(x^*)$ represents the derivative of f evaluated at $x = x^*$. In deriving these conditions, we make use of Taylor's expansion in the neighborhood of x^*. For a small real number $h > 0$, we use the first-order expansion at $x^* + h$ and $x^* - h$,

$$f(x^* + h) = f(x^*) + hf'(x^*) + O(h^2) \tag{2.4a}$$

$$f(x^* - h) = f(x^*) - hf'(x^*) + O(h^2) \tag{2.4b}$$

where the term $O(h^2)$ can be understood from the definition: $O(h^n)/h^r \to 0$ as $h \to 0$ for $0 \leq r < n$. For sufficiently small h, the remainder term can be dropped in comparison to the linear term. Thus, for $f(x^*)$ to be minimum, we need

$$f(x^* + h) - f(x^*) \approx \quad hf'(x^*) \geq 0$$
$$f(x^* - h) - f(x^*) \approx -hf'(x^*) \geq 0$$

Since h is not zero, these inequalities are satisfied when $f'(x^*) = 0$, which is Eq. (2.2). Condition given in Eq. (2.3) can be obtained using the second-order expansion

$$f(x^* + h) = f(x^*) + hf'(x^*) + \frac{h^2}{2} f''(x^*) + O(h^3)$$

At the minimum, $f'(x^*)$ is zero. Also the term $\frac{h^2}{2} f''(x^*)$dominates the remainder term $O(h^3)$. Thus,

$$f(x^* + h) - f(x^*) \approx \frac{h^2}{2} f''(x^*) \geq 0$$

Since h^2 is always positive, we need $f''(x^*) \geq 0$.

Points of inflection or flat regions shown in Fig. 2.1 satisfy the necessary conditions.

The *sufficient conditions* for x^* to be a strong local minimum are

$$f'(x^*) = 0 \tag{2.5}$$

$$f''(x^*) > 0 \tag{2.6}$$

These follow from the definition of a strong local minimum.

Example 2.2

Find the dimensions of the minimum cost refrigeration tank of Example 2.1 using optimality conditions.

Solution

We derived the objective function

$$\min\ f(x) = 45\pi x^2 + \frac{17200}{x}$$

$$f'(x) = 90\pi x - \frac{17200}{x^2} = 0$$

$$x^3 = \frac{17200}{90\pi} = 60.833$$

Diameter $\quad x = 3.93\,\text{m}$

Length $\quad L = \dfrac{200}{\pi x^2} = 4.12\,\text{m}$

Cost $\quad f = 45\pi x^2 + \dfrac{17200}{x} = \$6560$

Also $\quad f''(x) = 90\pi + \dfrac{(3)(17200)}{x^3}$

which is strictly positive at $x = 3.93$ m. Thus, the solution is a strict or strong minimum.

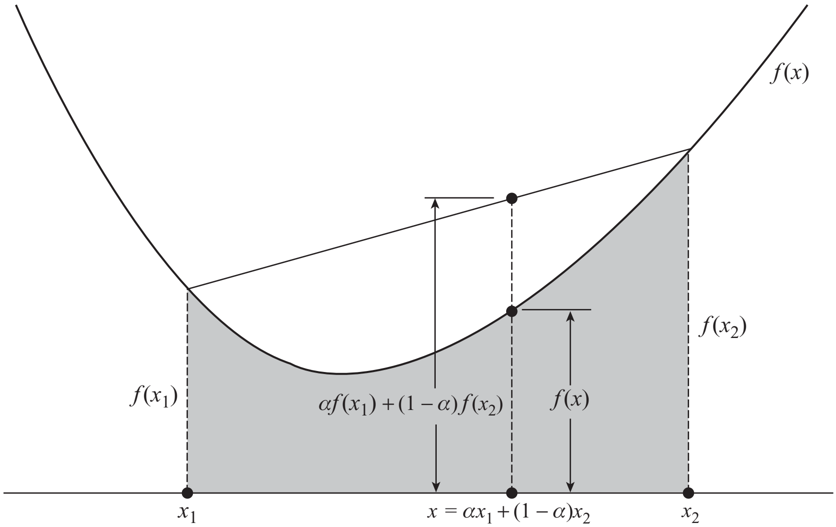

Figure 2.2. Convex function.

Convexity ideas can be used in defining optimality. A set S is called a *convex set* if for any two points in the set, every point on the line joining the two points is in the set. Alternatively, the set S is *convex* if for every pair of points x_1 and x_2 in S, and every α such that $0 \le \alpha \le 1$, the point $\alpha\, x_1 + (1-\alpha)\, x_2$ is in S. As an example, the set of all real numbers R^1 is a convex set. Any closed interval of R^1 is also a convex set.

A function $f(x)$ is called a *convex function* defined on the convex set S if for every pair of points x_1 and x_2 in S, and $0 \le \alpha \le 1$, the following condition is satisfied

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2) \tag{2.7}$$

Geometrically, this means that the graph of the function between the two points lies below the line segment joining the two points on the graph as shown in Fig. 2.2. We observe that the function defined in Example 2.1 is convex for $x > 0$. The function f is *strictly convex* if $f(\alpha x_1 + (1-\alpha)x_2) < \alpha f(x_1) + (1-\alpha) f(x_2)$ for $0 < \alpha < 1$. A function f is said to be *concave* if $-f$ is *convex*.

Example 2.3

Prove that $f = |x|, x \in R^1$, is a convex function.

Using the triangular inequality $|x + y| \le |x| + |y|$, we have, for any two real numbers x_1 and x_2 and $0 < \alpha < 1$,

$$\begin{aligned} f(\alpha x_1 + (1-\alpha)x_2) = \left|\alpha x_1 + (1-\alpha)\, x_2\right| &\le \alpha\, |x_1| + (1-\alpha)\, |x_2| \\ &= \alpha f(x_1) + (1-\alpha) f(x_2) \end{aligned}$$

which is the inequality defining a convex function in (2.7).

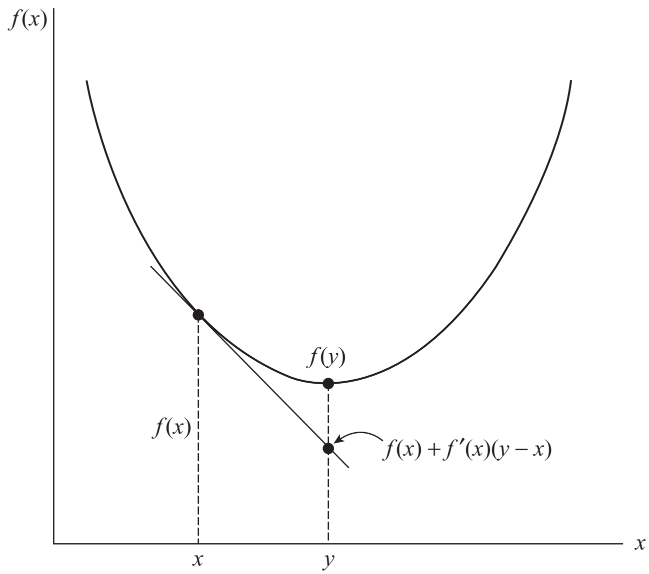

Figure 2.3. Convex function property.

Properties of Convex Functions

1. If f has continuous first derivatives then f is convex over a convex set S if and only if for every x and y in S,

$$f(y) \geq f(x) + f'(x)(y - x) \tag{2.8}$$

   This means that the graph of the function lies above the tangent line drawn at point as shown in Fig. 2.3.
2. If f has continuous second derivatives, then f is convex over a convex set S if and only if for every x in S,

$$f''(x) \geq 0 \tag{2.9}$$

3. If $f(x^*)$ is a local minimum for a convex function f on a convex set S, then it is also a global minimum.
4. If f has continuous first derivatives on a convex set S and for a point x^* in S, $f'(x^*)(y - x^*) \geq 0$ for every y in S, then x^* is a global minimum point of f over S. This follows from property (2.8) and the definition of global minimum.

Thus, we have used convexity ideas in characterizing the global minimum of a function. Further, the sufficiency condition (2.6) is equivalent to f being locally strictly convex.