An Introduction to Optimization

Second Edition

EDWIN K. P. CHONG
STANISLAW H. ŻAK

A Wiley-Interscience Publication
JOHN WILEY & SONS, INC.
New York / Chichester / Weinheim / Brisbane / Singapore / Toronto

Contents

6

Basics of Set-Constrained and Unconstrained Optimization

6.1 INTRODUCTION

In this chapter, we consider the optimization problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega. \end{array}$$

The function $f : \mathbb{R}^n \to \mathbb{R}$ that we wish to minimize is a real-valued function, and is called the *objective function*, or *cost function*. The vector $\boldsymbol{x}$ is an n-vector of independent variables, that is, $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n$. The variables $x_1, \ldots, x_n$ are often referred to as *decision variables*. The set Ω is a subset of $\mathbb{R}^n$, called the *constraint set* or *feasible set*.

The optimization problem above can be viewed as a decision problem that involves finding the "best" vector $\boldsymbol{x}$ of the decision variables over all possible vectors in Ω. By the "best" vector we mean the one that results in the smallest value of the objective function. This vector is called the *minimizer* of f over Ω. It is possible that there may be many minimizers. In this case, finding any of the minimizers will suffice.

There are also optimization problems that require maximization of the objective function. These problems, however, can be represented in the above form because maximizing f is equivalent to minimizing $-f$. Therefore, we can confine our attention to minimization problems without loss of generality.

The above problem is a general form of a *constrained* optimization problem, because the decision variables are constrained to be in the constraint set Ω. If $\Omega = \mathbb{R}^n$, then we refer to the problem as an *unconstrained* optimization problem. In this chapter, we discuss basic properties of the general optimization problem above,

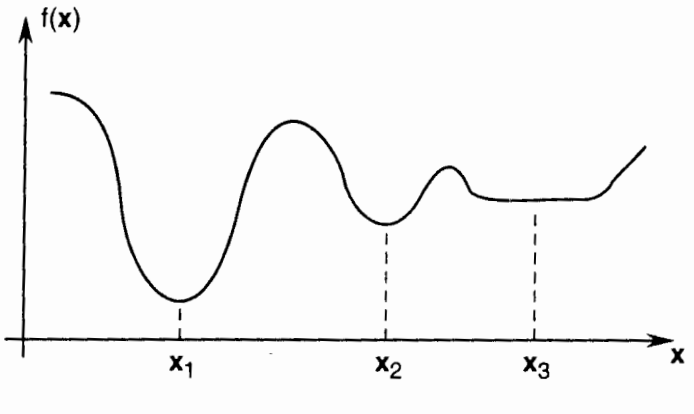

Figure 6.1 Examples of minimizers: $\boldsymbol{x}_1$: strict global minimizer; $\boldsymbol{x}_2$: strict local minimizer; $\boldsymbol{x}_3$: local (not strict) minimizer

which includes the unconstrained case. In the remaining chapters of this part, we deal with iterative algorithms for solving unconstrained optimization problems.

The constraint "$\boldsymbol{x} \in \Omega$" is called a *set constraint*. Often, the constraint set Ω takes the form $\Omega = \{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}) = 0,\ \boldsymbol{g}(\boldsymbol{x}) \leq 0\}$, where $\boldsymbol{h}$ and $\boldsymbol{g}$ are given functions. We refer to such constraints as *functional constraints*. The remainder of this chapter deals with general set constraints, including the special case where $\Omega = \mathbb{R}^n$. The case where $\Omega = \mathbb{R}^n$ is called the *unconstrained* case. In Parts III and IV, we consider constrained optimization problems with functional constraints.

In considering the general optimization problem above, we distinguish between two kinds of minimizers, as specified by the following definitions.

Definition 6.1 ***Local minimizer.*** Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is a real-valued function defined on some set $\Omega \subset \mathbb{R}^n$. A point $\boldsymbol{x}^* \in \Omega$ is a *local minimizer* of f over Ω if there exists $\epsilon > 0$ such that $f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$ for all $\boldsymbol{x} \in \Omega \setminus \{\boldsymbol{x}^*\}$ and $\|\boldsymbol{x} - \boldsymbol{x}^*\| < \epsilon$.

Global minimizer. A point $\boldsymbol{x}^* \in \Omega$ is a *global minimizer* of f over Ω if $f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$ for all $\boldsymbol{x} \in \Omega \setminus \{\boldsymbol{x}^*\}$. ∎

If, in the above definitions, we replace "$\geq$" with "$>$", then we have a *strict local minimizer* and a *strict global minimizer*, respectively.

In Figure 6.1, we graphically illustrate the above definitions for $n = 1$.

Given a real-valued function f, the notation $\arg\min f(\boldsymbol{x})$ denotes the argument that minimizes the function f (a point in the domain of f), assuming such a point is unique. For example, if $f : \mathbb{R} \to \mathbb{R}$ is given by $f(x) = (x + 1)^2 + 3$, then $\arg\min f(x) = -1$. If we write $\arg\min_{x \in \Omega}$, then we treat Ω as the domain of f. For example, for the function f above, $\arg\min_{x \geq 0} f(x) = 0$. In general, we can think of $\arg\min_{\boldsymbol{x} \in \Omega} f(\boldsymbol{x})$ as the global minimizer of f over Ω (assuming it exists and is unique).

Strictly speaking, an optimization problem is solved only when a global minimizer is found. However, global minimizers are, in general, difficult to find. Therefore, in practice, we often have to be satisfied with finding local minimizers.

6.2 CONDITIONS FOR LOCAL MINIMIZERS

In this section, we derive conditions for a point $\boldsymbol{x}^*$ to be a local minimizer. We use derivatives of a function $f : \mathbb{R}^n \to \mathbb{R}$. Recall that the first-order derivative of f, denoted Df, is

$$Df \triangleq \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right].$$

Note that the gradient ∇f is just the transpose of Df; that is, $\nabla f = (Df)^T$. The second derivative of $f : \mathbb{R}^n \to \mathbb{R}$ (also called the *Hessian* of f) is

$$\boldsymbol{F}(\boldsymbol{x}) \triangleq D^2 f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\boldsymbol{x}) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\boldsymbol{x}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\boldsymbol{x}) \end{bmatrix}.$$

Example 6.1 Let $f(x_1, x_2) = 5x_1 + 8x_2 + x_1x_2 - x_1^2 - 2x_2^2$. Then,

$$Df(\boldsymbol{x}) = (\nabla f(\boldsymbol{x}))^T = \left[\frac{\partial f}{\partial x_1}(\boldsymbol{x}), \frac{\partial f}{\partial x_2}(\boldsymbol{x})\right] = [5 + x_2 - 2x_1, 8 + x_1 - 4x_2],$$

and

$$\boldsymbol{F}(\boldsymbol{x}) = D^2 f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}) & \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}) \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -4 \end{bmatrix}.$$

■

Given an optimization problem with constraint set Ω, a minimizer may lie either in the interior or on the boundary of Ω. To study the case where it lies on the boundary, we need the notion of *feasible directions*.

Definition 6.2 ***Feasible direction.*** A vector $\boldsymbol{d} \in \mathbb{R}^n$, $\boldsymbol{d} \neq \boldsymbol{0}$, is a *feasible direction* at $\boldsymbol{x} \in \Omega$ if there exists $\alpha_0 > 0$ such that $\boldsymbol{x} + \alpha \boldsymbol{d} \in \Omega$ for all $\alpha \in [0, \alpha_0]$. ■

Figure 6.2 illustrates the notion of feasible directions.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real-valued function and let $\boldsymbol{d}$ be a feasible direction at $\boldsymbol{x} \in \Omega$. The *directional derivative of f in the direction $\boldsymbol{d}$*, denoted $\partial f / \partial \boldsymbol{d}$, is the real-valued function defined by

$$\frac{\partial f}{\partial \boldsymbol{d}}(\boldsymbol{x}) = \lim_{\alpha \to 0} \frac{f(\boldsymbol{x} + \alpha \boldsymbol{d}) - f(\boldsymbol{x})}{\alpha}.$$

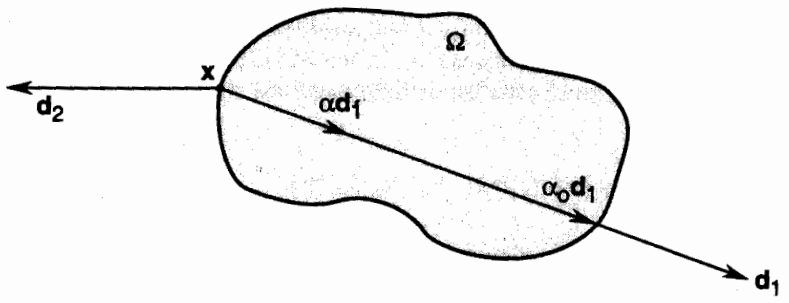

Figure 6.2 Two-dimensional illustration of feasible directions; $\boldsymbol{d}_1$ is a feasible direction, $\boldsymbol{d}_2$ is not a feasible direction

If $\|\boldsymbol{d}\| = 1$, then $\partial f/\partial \boldsymbol{d}$ is the rate of increase of f at $\boldsymbol{x}$ in the direction $\boldsymbol{d}$. To compute the above directional derivative, suppose that $\boldsymbol{x}$ and $\boldsymbol{d}$ are given. Then, $f(\boldsymbol{x} + \alpha \boldsymbol{d})$ is a function of α, and

$$\frac{\partial f}{\partial \boldsymbol{d}}(\boldsymbol{x}) = \frac{d}{d\alpha} f(\boldsymbol{x} + \alpha \boldsymbol{d})\bigg|_{\alpha=0}.$$

Applying the chain rule yields

$$\frac{\partial f}{\partial \boldsymbol{d}}(\boldsymbol{x}) = \frac{d}{d\alpha} f(\boldsymbol{x} + \alpha \boldsymbol{d})\bigg|_{\alpha=0} = \nabla f(\boldsymbol{x})^T \boldsymbol{d} = \langle \nabla f(\boldsymbol{x}), \boldsymbol{d} \rangle = \boldsymbol{d}^T \nabla f(\boldsymbol{x}).$$

In summary, if $\boldsymbol{d}$ is a unit vector, that is, $\|\boldsymbol{d}\| = 1$, then $\langle \nabla f(\boldsymbol{x}), \boldsymbol{d} \rangle$ is the rate of increase of f at the point $\boldsymbol{x}$ in the direction $\boldsymbol{d}$.

Example 6.2 Define $f : \mathbb{R}^3 \to \mathbb{R}$ by $f(\boldsymbol{x}) = x_1 x_2 x_3$, and let

$$\boldsymbol{d} = \left[\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}\right]^T.$$

The directional derivative of f in the direction $\boldsymbol{d}$ is

$$\frac{\partial f}{\partial \boldsymbol{d}}(\boldsymbol{x}) = \nabla f(\boldsymbol{x})^T \boldsymbol{d} = [x_2 x_3, x_1 x_3, x_1 x_2] \begin{bmatrix} 1/2 \\ 1/2 \\ 1/\sqrt{2} \end{bmatrix} = \frac{x_2 x_3 + x_1 x_3 + \sqrt{2} x_1 x_2}{2}.$$

Note that because $\|\boldsymbol{d}\| = 1$, the above is also the rate of increase of f at $\boldsymbol{x}$ in the direction $\boldsymbol{d}$. ∎

We are now ready to state and prove the following theorem.

Theorem 6.1 ***First-Order Necessary Condition (FONC).*** *Let Ω be a subset of $\mathbb{R}^n$ and $f \in \mathcal{C}^1$ a real-valued function on Ω. If $\boldsymbol{x}^*$ is a local minimizer of f over Ω, then for any feasible direction $\boldsymbol{d}$ at $\boldsymbol{x}^*$, we have*

$$\boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) \geq 0.$$

□

Proof. Define

$$\boldsymbol{x}(\alpha) = \boldsymbol{x}^* + \alpha \boldsymbol{d} \in \Omega.$$

Note that $\boldsymbol{x}(0) = \boldsymbol{x}^*$. Define the composite function

$$\phi(\alpha) = f(\boldsymbol{x}(\alpha)).$$

Then, by Taylor's theorem,

$$f(\boldsymbol{x}^* + \alpha \boldsymbol{d}) - f(\boldsymbol{x}^*) = \phi(\alpha) - \phi(0) = \phi'(0)\alpha + o(\alpha) = \alpha \boldsymbol{d}^T \nabla f(\boldsymbol{x}(0)) + o(\alpha),$$

where $\alpha \geq 0$ (recall the definition of $o(\alpha)$ ("little-oh of α") in Part I). Thus, if $\phi(\alpha) \geq \phi(0)$, that is, $f(\boldsymbol{x}^* + \alpha \boldsymbol{d}) \geq f(\boldsymbol{x}^*)$ for sufficiently small values of $\alpha > 0$ ($\boldsymbol{x}^*$ is a local minimizer), then we have to have $\boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) \geq 0$ (see Exercise 5.7). ∎

The above theorem is graphically illustrated in Figure 6.3.
An alternative way to express the FONC is:

$$\frac{\partial f}{\partial \boldsymbol{d}}(\boldsymbol{x}^*) \geq 0$$

for all feasible directions $\boldsymbol{d}$. In other words, if $\boldsymbol{x}^*$ is a local minimizer, then the rate of increase of f at $\boldsymbol{x}^*$ in any feasible direction $\boldsymbol{d}$ in Ω is nonnegative. Using directional derivatives, an alternative proof of Theorem 6.1 is as follows. Suppose that $\boldsymbol{x}^*$ is a local minimizer. Then, for any feasible direction $\boldsymbol{d}$, there exists $\bar{\alpha} > 0$ such that for all $\alpha \in (0, \bar{\alpha})$,

$$f(\boldsymbol{x}^*) \leq f(\boldsymbol{x}^* + \alpha \boldsymbol{d}).$$

Hence, for all $\alpha \in (0, \bar{\alpha})$, we have

$$\frac{f(\boldsymbol{x}^* + \alpha \boldsymbol{d}) - f(\boldsymbol{x}^*)}{\alpha} \geq 0.$$

Taking the limit as $\alpha \to 0$, we conclude that

$$\frac{\partial f}{\partial \boldsymbol{d}}(\boldsymbol{x}^*) \geq 0.$$

A special case of interest is when $\boldsymbol{x}^*$ is an interior point of Ω (see Section 4.4). In this case, any direction is feasible, and we have the following result.

Corollary 6.1 ***Interior case.*** *Let Ω be a subset of $\mathbb{R}^n$ and $f \in \mathcal{C}^1$ a real-valued function on Ω. If $\boldsymbol{x}^*$ is a local minimizer of f over Ω and if $\boldsymbol{x}^*$ is an interior point of Ω, then*

$$\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}.$$

□

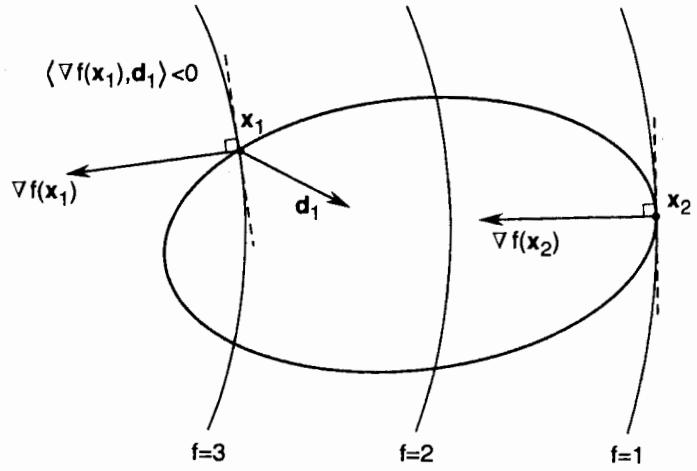

Figure 6.3 Illustration of the FONC for the constrained case; $\boldsymbol{x}_1$ does not satisfy the FONC, $\boldsymbol{x}_2$ satisfies the FONC

Proof. Suppose that f has a local minimizer $\boldsymbol{x}^*$ that is an interior point of Ω. Because $\boldsymbol{x}^*$ is an interior point of Ω, the set of feasible directions at $\boldsymbol{x}^*$ is the whole of $\mathbb{R}^n$. Thus, for any $\boldsymbol{d} \in \mathbb{R}^n$, $\boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) \geq 0$ and $-\boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) \geq 0$. Hence, $\boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) = 0$ for all $\boldsymbol{d} \in \mathbb{R}^n$, which implies that $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$. ∎

Example 6.3 Consider the problem

$$\begin{aligned} \text{minimize} \quad & x_1^2 + 0.5x_2^2 + 3x_2 + 4.5 \\ \text{subject to} \quad & x_1, x_2 \geq 0. \end{aligned}$$

Questions:

a. Is the first-order necessary condition (FONC) for a local minimizer satisfied at $\boldsymbol{x} = [1, 3]^T$?

b. Is the FONC for a local minimizer satisfied at $\boldsymbol{x} = [0, 3]^T$?

c. Is the FONC for a local minimizer satisfied at $\boldsymbol{x} = [1, 0]^T$?

d. Is the FONC for a local minimizer satisfied at $\boldsymbol{x} = [0, 0]^T$?

Answers: First, let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(\boldsymbol{x}) = x_1^2 + 0.5x_2^2 + 3x_2 + 4.5$, where $\boldsymbol{x} = [x_1, x_2]^T$. A plot of the level sets of f is shown in Figure 6.4.

a. At $\boldsymbol{x} = [1, 3]^T$, we have $\nabla f(\boldsymbol{x}) = [2x_1, x_2 + 3]^T = [2, 6]^T$. The point $\boldsymbol{x} = [1, 3]^T$ is an interior point of $\Omega = \{\boldsymbol{x} : x_1 \geq 0, x_2 \geq 0\}$. Hence, the FONC requires $\nabla f(\boldsymbol{x}) = \boldsymbol{0}$. The point $\boldsymbol{x} = [1, 3]^T$ does not satisfy the FONC for a local minimizer.

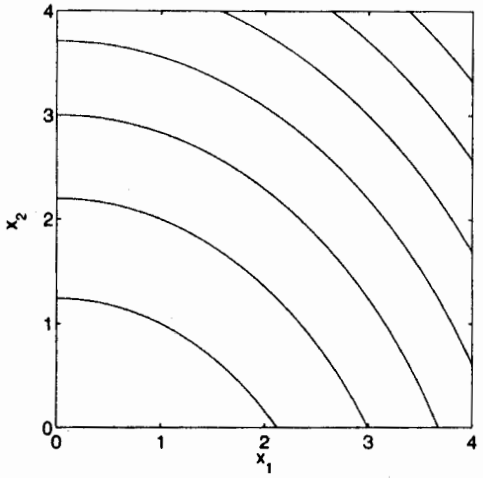

Figure 6.4 Level sets of the function in Example 6.3

b. At $\boldsymbol{x} = [0,3]^T$, we have $\nabla f(\boldsymbol{x}) = [0,6]^T$, and hence $\boldsymbol{d}^T \nabla f(\boldsymbol{x}) = 6d_2$, where $\boldsymbol{d} = [d_1, d_2]^T$. For $\boldsymbol{d}$ to be feasible at $\boldsymbol{x}$, we need $d_1 \geq 0$, and d_2 can take an arbitrary value in $\mathbb{R}$. The point $\boldsymbol{x} = [0,3]^T$ does not satisfy the FONC for a minimizer because d_2 is allowed to be less than zero. For example, $\boldsymbol{d} = [1,-1]^T$ is a feasible direction, but $\boldsymbol{d}^T \nabla f(\boldsymbol{x}) = -6 < 0$.

c. At $\boldsymbol{x} = [1,0]^T$, we have $\nabla f(\boldsymbol{x}) = [2,3]^T$, and hence $\boldsymbol{d}^T \nabla f(\boldsymbol{x}) = 2d_1 + 3d_2$. For $\boldsymbol{d}$ to be feasible, we need $d_2 \geq 0$, and d_1 can take an arbitrary value in $\mathbb{R}$. For example, $\boldsymbol{d} = [-5,1]^T$ is a feasible direction. But $\boldsymbol{d}^T \nabla f(\boldsymbol{x}) = -7 < 0$. Thus, $\boldsymbol{x} = [1,0]^T$ does not satisfy the FONC for a local minimizer.

d. At $\boldsymbol{x} = [0,0]^T$, we have $\nabla f(\boldsymbol{x}) = [0,3]^T$, and hence $\boldsymbol{d}^T \nabla f(\boldsymbol{x}) = 3d_2$. For $\boldsymbol{d}$ to be feasible, we need $d_2 \geq 0$ and $d_1 \geq 0$. Hence, $\boldsymbol{x} = [0,0]^T$ satisfies the FONC for a local minimizer.

∎

Example 6.4 Figure 6.5 shows a simplified model of a cellular wireless system (the distances shown have been scaled down to make the calculations simpler). A mobile user (also called a "mobile") is located at position x (see Figure 6.5).

There are two basestation antennas, one for the primary basestation and another for the neighboring basestation. Both antennas are transmitting signals to the mobile user, at equal power. However, the power of the received signal as measured by the mobile is the reciprocal of the squared distance from the associated antenna (primary or neighboring basestation). We are interested in finding the position of the mobile that maximizes the *signal-to-interference ratio*, which is the ratio of the received signal power from the primary basestation to the received signal power from the neighboring basestation.

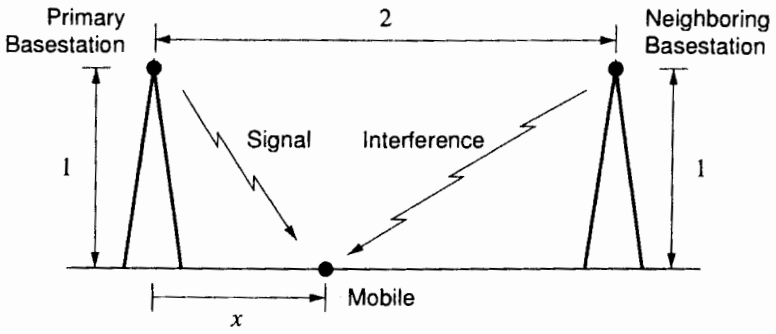

Figure 6.5 Simplified cellular wireless system in Example 6.4

We use the FONC to solve this problem. The squared distance from the mobile to the primary antenna is $1 + x^2$, while the squared distance from the mobile to the neighboring antenna is $1 + (2 - x)^2$. Therefore, the signal-to-interference ratio is

$$f(x) = \frac{1 + x^2}{1 + (2 - x)^2}.$$

We have

$$\begin{aligned} f'(x) &= \frac{-2x(1 + (2 - x)^2) - 2(2 - x)(1 + x^2)}{1 + (2 - x)^2} \\ &= \frac{4(x^2 - 2x - 1)}{1 + (2 - x)^2}. \end{aligned}$$

By the FONC, at the optimal position x^*, we have $f'(x^*) = 0$. Hence, either $x^* = 1 - \sqrt{2}$ or $x^* = 1 + \sqrt{2}$. Evaluating the objective function at these two candidate points, it easy to see that $x^* = 1 - \sqrt{2}$ is the optimal position. ∎

We now derive a second-order necessary condition that is satisfied by a local minimizer.

Theorem 6.2 ***Second-Order Necessary Condition (SONC).*** *Let $\Omega \subset \mathbb{R}^n$, $f \in C^2$ a function on Ω, $\boldsymbol{x}^*$ a local minimizer of f over Ω, and $\boldsymbol{d}$ a feasible direction at $\boldsymbol{x}^*$. If $\boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) = 0$, then*

$$\boldsymbol{d}^T \boldsymbol{F}(\boldsymbol{x}^*) \boldsymbol{d} \geq 0,$$

where $\boldsymbol{F}$ is the Hessian of f. □

Proof. We prove the result by contradiction. Suppose that there is a feasible direction $\boldsymbol{d}$ at $\boldsymbol{x}^*$ such that $\boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) = 0$ and $\boldsymbol{d}^T \boldsymbol{F}(\boldsymbol{x}^*) \boldsymbol{d} < 0$. Let $\boldsymbol{x}(\alpha) = \boldsymbol{x}^* + \alpha \boldsymbol{d}$ and define the composite function $\phi(\alpha) = f(\boldsymbol{x}^* + \alpha \boldsymbol{d}) = f(\boldsymbol{x}(\alpha))$. Then, by Taylor's theorem

$$\phi(\alpha) = \phi(0) + \phi''(0) \frac{\alpha^2}{2} + o(\alpha^2),$$

where by assumption $\phi'(0) = \boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) = 0$, and $\phi''(0) = \boldsymbol{d}^T \boldsymbol{F}(\boldsymbol{x}^*)\boldsymbol{d} < 0$. For sufficiently small α,

$$\phi(\alpha) - \phi(0) = \phi''(0)\frac{\alpha^2}{2} + o(\alpha^2) < 0,$$

that is,

$$f(\boldsymbol{x}^* + \alpha \boldsymbol{d}) < f(\boldsymbol{x}^*),$$

which contradicts the assumption that $\boldsymbol{x}^*$ is a local minimizer. Thus,

$$\phi''(0) = \boldsymbol{d}^T \boldsymbol{F}(\boldsymbol{x}^*)\boldsymbol{d} \geq 0.$$

∎

Corollary 6.2 ***Interior Case.*** *Let $\boldsymbol{x}^*$ be an interior point of $\Omega \subset \mathbb{R}^n$. If $\boldsymbol{x}^*$ is a local minimizer of $f : \Omega \to \mathbb{R}$, $f \in C^2$, then*

$$\nabla f(\boldsymbol{x}^*) = \boldsymbol{0},$$

and $\boldsymbol{F}(\boldsymbol{x}^*)$ is positive semidefinite ($\boldsymbol{F}(\boldsymbol{x}^*) \geq 0$); that is, for all $\boldsymbol{d} \in \mathbb{R}^n$,

$$\boldsymbol{d}^T \boldsymbol{F}(\boldsymbol{x}^*)\boldsymbol{d} \geq 0.$$

□

Proof. If $\boldsymbol{x}^*$ is an interior point then all directions are feasible. The result then follows from Corollary 6.1 and Theorem 6.2. ∎

In the examples below, we show that the necessary conditions are *not* sufficient.

Example 6.5 Consider a function of one variable $f(x) = x^3$, $f : \mathbb{R} \to \mathbb{R}$. Because $f'(0) = 0$, and $f''(0) = 0$, the point $x = 0$ satisfies both the FONC and SONC. However, $x = 0$ is not a minimizer (see Figure 6.6). ∎

Example 6.6 Consider a function $f : \mathbb{R}^2 \to \mathbb{R}$, where $f(\boldsymbol{x}) = x_1^2 - x_2^2$. The FONC requires that $\nabla f(\boldsymbol{x}) = [2x_1, -2x_2]^T = \boldsymbol{0}$. Thus, $\boldsymbol{x} = [0, 0]^T$ satisfies the FONC. The Hessian matrix of f is

$$\boldsymbol{F}(\boldsymbol{x}) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}.$$

The Hessian matrix is indefinite; that is, for some $\boldsymbol{d}_1 \in \mathbb{R}^2$ we have $\boldsymbol{d}_1^T \boldsymbol{F} \boldsymbol{d}_1 > 0$, e.g., $\boldsymbol{d}_1 = [1, 0]^T$; and, for some $\boldsymbol{d}_2$, we have $\boldsymbol{d}_2^T \boldsymbol{F} \boldsymbol{d}_2 < 0$, e.g., $\boldsymbol{d}_2 = [0, 1]^T$. Thus, $\boldsymbol{x} = [0, 0]^T$ does not satisfy the SONC, and hence it is not a minimizer. The graph of $f(\boldsymbol{x}) = x_1^2 - x_2^2$ is shown in Figure 6.7. ∎

We now derive sufficient conditions that imply that $\boldsymbol{x}^*$ is a local minimizer.

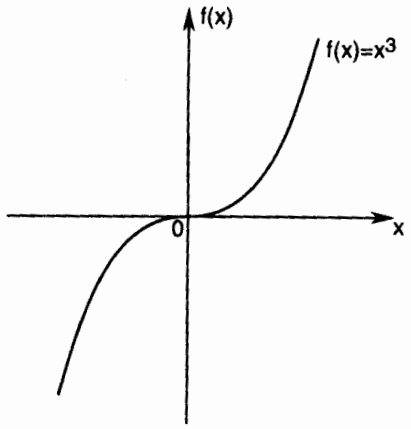

Figure 6.6 The point 0 satisfies the FONC and SONC, but is not a minimizer

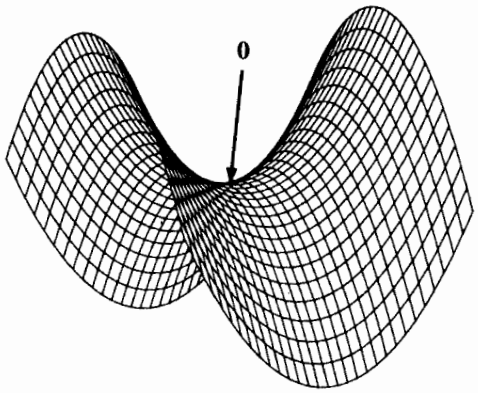

Figure 6.7 Graph of $f(\boldsymbol{x}) = x_1^2 - x_2^2$. The point 0 satisfies the FONC but not SONC; this point is not a minimizer

Theorem 6.3 ***Second-Order Sufficient Condition (SOSC), Interior Case.*** *Let $f \in C^2$ be defined on a region in which $\boldsymbol{x}^*$ is an interior point. Suppose that*

1. $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$; *and*
2. $\boldsymbol{F}(\boldsymbol{x}^*) > 0$.

Then, $\boldsymbol{x}^*$ is a strict local minimizer of f.

Proof. Because $f \in C^2$, we have $\boldsymbol{F}(\boldsymbol{x}^*) = \boldsymbol{F}^T(\boldsymbol{x}^*)$. Using assumption 2 and Rayleigh's inequality it follows that if $\boldsymbol{d} \neq \boldsymbol{0}$, then $0 < \lambda_{\min}(\boldsymbol{F}(\boldsymbol{x}^*))\|\boldsymbol{d}\|^2 \leq \boldsymbol{d}^T\boldsymbol{F}(\boldsymbol{x}^*)\boldsymbol{d}$. By Taylor's theorem and assumption 1,

$$f(\boldsymbol{x}^* + \boldsymbol{d}) - f(\boldsymbol{x}^*) = \frac{1}{2}\boldsymbol{d}^T\boldsymbol{F}(\boldsymbol{x}^*)\boldsymbol{d} + o(\|\boldsymbol{d}\|^2) \geq \frac{\lambda_{\min}(\boldsymbol{F}(\boldsymbol{x}^*))}{2}\|\boldsymbol{d}\|^2 + o(\|\boldsymbol{d}\|^2)$$

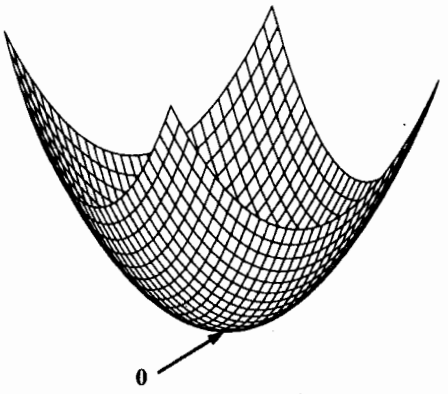

Figure 6.8 Graph of $f(\boldsymbol{x}) = x_1^2 + x_2^2$

Hence, for all $\boldsymbol{d}$ such that $\|\boldsymbol{d}\|$ is sufficiently small,

$$f(\boldsymbol{x}^* + \boldsymbol{d}) > f(\boldsymbol{x}^*),$$

and the proof is completed. ∎

Example 6.7 Let $f(\boldsymbol{x}) = x_1^2 + x_2^2$. We have $\nabla f(\boldsymbol{x}) = [2x_1, 2x_2]^T = \boldsymbol{0}$ if and only if $\boldsymbol{x} = [0,0]^T$. For all $\boldsymbol{x} \in \mathbb{R}^2$, we have

$$\boldsymbol{F}(\boldsymbol{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} > 0.$$

The point $\boldsymbol{x} = [0,0]^T$ satisfies the FONC, SONC, and SOSC. It is a strict local minimizer. Actually $\boldsymbol{x} = [0,0]^T$ is a strict global minimizer. Figure 6.8 shows the graph of $f(\boldsymbol{x}) = x_1^2 + x_2^2$. ∎

In this chapter, we presented a theoretical basis for the solution of nonlinear unconstrained problems. In the following chapters, we are concerned with iterative methods of solving such problems. Such methods are of great importance in practice. Indeed, suppose that one is confronted with a highly nonlinear function of 20 variables. Then, the FONC requires the solution of 20 nonlinear simultaneous equations for 20 variables. These equations, being nonlinear, will normally have multiple solutions. In addition, we would have to compute 210 second derivatives (provided $f \in \mathcal{C}^2$) to use the SONC or SOSC. We begin our discussion of iterative methods in the next chapter with search methods for functions of one variable.

EXERCISES

6.1 Consider the problem

$$\text{minimize} \qquad f(\boldsymbol{x})$$

From the above, we see that

$$r_1 \cdots r_N = \frac{1}{F_{N+1}}$$

if and only if

$$r_1 = \frac{F_N}{F_{N+1}}.$$

The above is simply the value of r_1 for the Fibonacci search method. Note that fixing r_1 uniquely determines $r_2, \ldots, r_N$. ∎

For further discussion on the Fibonacci search method and its variants, see [96].

7.3 NEWTON'S METHOD

Suppose again that we are confronted with the problem of minimizing a function f of a single real variable x. We assume now that at each measurement point $x^{(k)}$ we can calculate $f(x^{(k)})$, $f'(x^{(k)})$, and $f''(x^{(k)})$. We can fit a quadratic function through $x^{(k)}$ that matches its first and second derivatives with that of the function f. This quadratic has the form

$$q(x) = f(x^{(k)}) + f'(x^{(k)})(x - x^{(k)}) + \frac{1}{2}f''(x^{(k)})(x - x^{(k)})^2.$$

Note that $q(x^{(k)}) = f(x^{(k)})$, $q'(x^{(k)}) = f'(x^{(k)})$, and $q''(x^{(k)}) = f''(x^{(k)})$. Then, instead of minimizing f, we minimize its approximation q. The first-order necessary condition for a minimizer of q yields

$$0 = q'(x) = f'(x^{(k)}) + f''(x^{(k)})(x - x^{(k)}).$$

Setting $x = x^{(k+1)}$, we obtain

$$x^{(k+1)} = x^{(k)} - \frac{f'(x^{(k)})}{f''(x^{(k)})}.$$

Example 7.3 Using Newton's method, find the minimizer of

$$f(x) = \frac{1}{2}x^2 - \sin x.$$

The initial value is $x^{(0)} = 0.5$. The required accuracy is $\epsilon = 10^{-5}$, in the sense that we stop when $|x^{(k+1)} - x^{(k)}| < \epsilon$.

We compute

$$f'(x) = x - \cos x, \qquad f''(x) = 1 + \sin x.$$

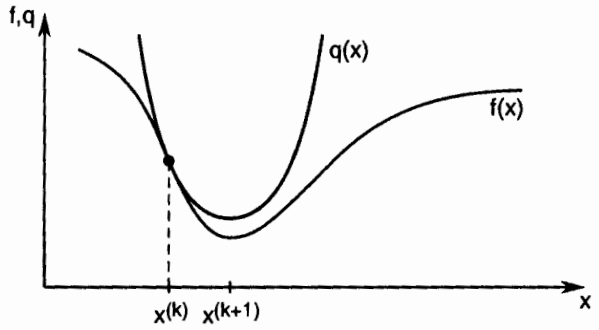

Figure 7.6 Newton's algorithm with $f''(x) > 0$

Hence,

$$\begin{aligned} x^{(1)} &= 0.5 - \left[\frac{0.5 - \cos 0.5}{1 + \sin 0.5}\right] \\ &= 0.5 - \left[\frac{-0.3775}{1.479}\right] \\ &= 0.7552. \end{aligned}$$

Proceeding in a similar manner, we obtain

$$\begin{aligned} x^{(2)} &= x^{(1)} - \frac{f'(x^{(1)})}{f''(x^{(1)})} = x^{(1)} - \frac{0.02710}{1.685} = 0.7391, \\ x^{(3)} &= x^{(2)} - \frac{f'(x^{(2)})}{f''(x^{(2)})} = x^{(2)} - \frac{9.461 \times 10^{-5}}{1.673} = 0.7390, \\ x^{(4)} &= x^{(3)} - \frac{f'(x^{(3)})}{f''(x^{(3)})} = x^{(3)} - \frac{1.17 \times 10^{-9}}{1.673} = 0.7390. \end{aligned}$$

Note that $|x^{(4)} - x^{(3)}| < \epsilon = 10^{-5}$. Furthermore, $f'(x^{(4)}) = -8.6 \times 10^{-6} \approx 0$. Observe that $f''(x^{(4)}) = 1.673 > 0$, so we can assume that $x^* \approx x^{(4)}$ is a strict minimizer. ∎

Newton's method works well if $f''(x) > 0$ everywhere (see Figure 7.6). However, if $f''(x) < 0$ for some x, Newton's method may fail to converge to the minimizer (see Figure 7.7).

Newton's method can also be viewed as a way to drive the first derivative of f to zero. Indeed, if we set $g(x) = f'(x)$, then we obtain a formula for iterative solution of the equation $g(x) = 0$:

$$x^{(k+1)} = x^{(k)} - \frac{g(x^{(k)})}{g'(x^{(k)})}.$$

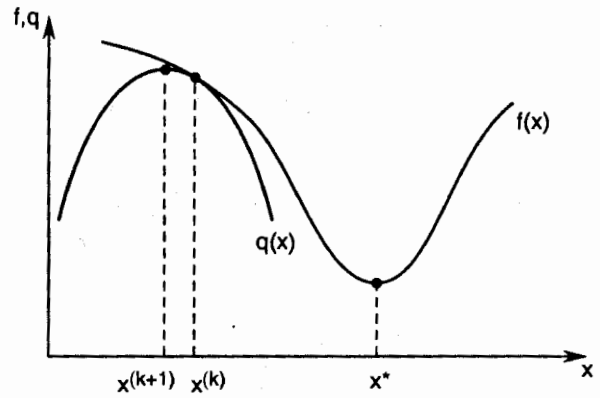

Figure 7.7 Newton's algorithm with $f''(x) < 0$

Example 7.4 We apply Newton's method to improve a first approximation, $x^{(0)} = 12$, to the root of the equation

$$g(x) = x^3 - 12.2x^2 + 7.45x + 42 = 0.$$

We have $g'(x) = 3x^2 - 24.4x + 7.45$.

Performing two iterations yields

$$\begin{aligned} x^{(1)} &= 12 - \frac{102.6}{146.65} = 11.33 \\ x^{(2)} &= 11.33 - \frac{14.73}{116.11} = 11.21. \end{aligned}$$

∎

Newton's method for solving equations of the form $g(x) = 0$ is also referred to as *Newton's method of tangents*. This name is easily justified if we look at a geometric interpretation of the method when applied to the solution of the equation $g(x) = 0$ (see Figure 7.8).

If we draw a tangent to $g(x)$ at the given point $x^{(k)}$, then the tangent line intersects the x-axis at the point $x^{(k+1)}$, which we expect to be closer to the root x^* of $g(x) = 0$. Note that the slope of $g(x)$ at $x^{(k)}$ is

$$g'(x^{(k)}) = \frac{g(x^{(k)})}{x^{(k)} - x^{(k+1)}}.$$

Hence,

$$x^{(k+1)} = x^{(k)} - \frac{g(x^{(k)})}{g'(x^{(k)})}.$$

Newton's method of tangents may fail if the first approximation to the root is such that the ratio $g(x^{(0)})/g'(x^{(0)})$ is not small enough (see Figure 7.9). Thus, an initial approximation to the root is very important.

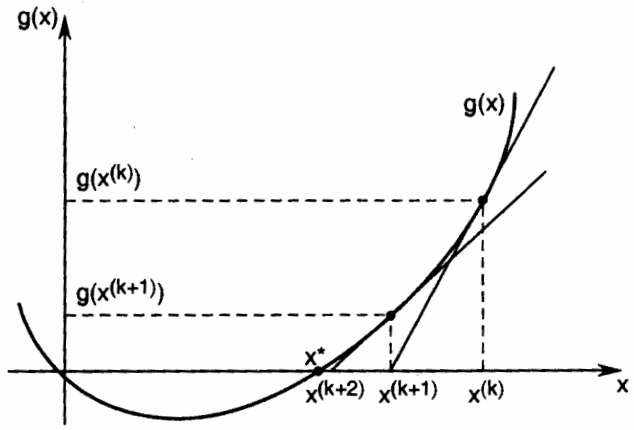

Figure 7.8 Newton's method of tangents

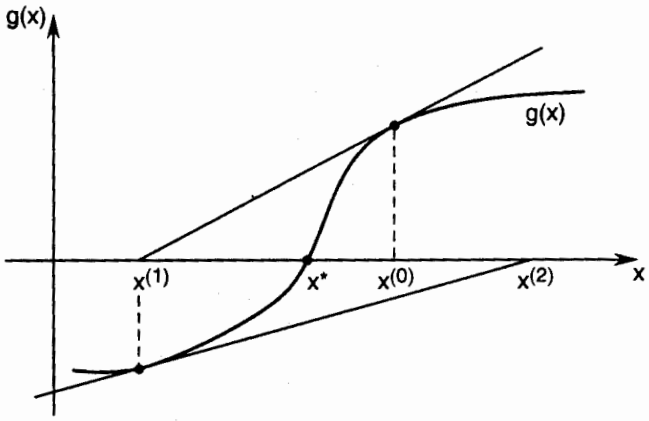

Figure 7.9 Example where Newton's method of tangents fails to converge to the root x^* of $g(x) = 0$

7.4 SECANT METHOD

Newton's method for minimizing f uses second derivatives of f:

$$x^{(k+1)} = x^{(k)} - \frac{f'(x^{(k)})}{f''(x^{(k)})}.$$

If the second derivative is not available, we may attempt to approximate it using first derivative information. In particular, we may approximate $f''(x^{(k)})$ above with

$$\frac{f'(x^{(k)}) - f'(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}.$$

Using the above approximation of the second derivative, we obtain the algorithm

$$x^{(k+1)} = x^{(k)} - \frac{x^{(k)} - x^{(k-1)}}{f'(x^{(k)}) - f'(x^{(k-1)})} f'(x^{(k)}).$$

The above algorithm is called the *secant method*. Note that the algorithm requires two initial points to start it, which we denote $x^{(-1)}$ and $x^{(0)}$. The secant algorithm can be represented in the following equivalent form:

$$x^{(k+1)} = \frac{f'(x^{(k)})x^{(k-1)} - f'(x^{(k-1)})x^{(k)}}{f'(x^{(k)}) - f'(x^{(k-1)})}.$$

Observe that, like Newton's method, the secant method does not directly involve values of $f(x^{(k)})$. Instead, it tries to drive the derivative f' to zero. In fact, as we did for Newton's method, we can interpret the secant method as an algorithm for solving equations of the form $g(x) = 0$. Specifically, the secant algorithm for finding a root of the equation $g(x) = 0$ takes the form

$$x^{(k+1)} = x^{(k)} - \frac{x^{(k)} - x^{(k-1)}}{g(x^{(k)}) - g(x^{(k-1)})} g(x^{(k)}),$$

or, equivalently,

$$x^{(k+1)} = \frac{g(x^{(k)})x^{(k-1)} - g(x^{(k-1)})x^{(k)}}{g(x^{(k)}) - g(x^{(k-1)})}.$$

The secant method for root finding is illustrated in Figure 7.10 (compare this with Figure 7.8). Unlike Newton's method, which uses the slope of g to determine the next point, the secant method uses the "secant" between the $(k-1)$st and kth points to determine the $(k+1)$st point.

Example 7.5 We apply the secant method to find the root of the equation

$$g(x) = x^3 - 12.2x^2 + 7.45x + 42 = 0.$$

We perform two iterations, with starting points $x^{(-1)} = 13$ and $x^{(0)} = 12$. We obtain

$$\begin{aligned} x^{(1)} &= 11.40 \\ x^{(2)} &= 11.25. \end{aligned}$$

■

Example 7.6 Suppose the voltage across a resistor in a circuit decays according to the model $V(t) = e^{-Rt}$, where $V(t)$ is the voltage at time t, and R is the resistance value.

Given measurements $V_1, \ldots, V_n$ of the voltage at times $t_1, \ldots, t_n$, respectively, we wish to find the best estimate of R. By the "best estimate" we mean the value

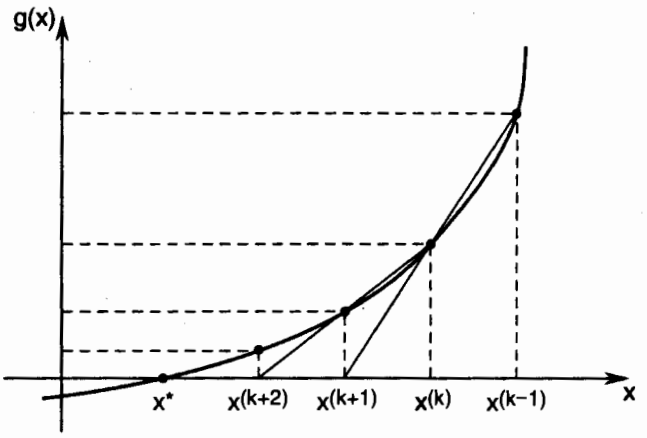

Figure 7.10 Secant method for root finding

of R that minimizes the total squared error between the measured voltages and the voltages predicted by the model.

We derive an algorithm to find the best estimate of R using the secant method. The objective function is:

$$f(R) = \sum_{i=1}^{n} (V_i - e^{-Rt_i})^2.$$

Hence, we have

$$f'(R) = 2 \sum_{i=1}^{n} (V_i - e^{-Rt_i}) e^{-Rt_i} t_i.$$

The secant algorithm for the problem is:

$$R_{k+1} = R_k - \left(\frac{R_k - R_{k-1}}{\sum_{i=1}^{n} (V_i - e^{-R_k t_i}) e^{-R_k t_i} t_i - (V_i - e^{-R_{k-1} t_i}) e^{-R_{k-1} t_i} t_i} \right) \times \sum_{i=1}^{n} (V_i - e^{-R_k t_i}) e^{-R_k t_i} t_i.$$

For further reading on the secant method, see [20].

7.5 REMARKS ON LINE SEARCH METHODS

One-dimensional search methods play an important role in multidimensional optimization problems. In particular, iterative algorithms for solving such optimization

9

Newton's Method

9.1 INTRODUCTION

Recall that the method of steepest descent uses only first derivatives (gradients) in selecting a suitable search direction. This strategy is not always the most effective. If higher derivatives are used, the resulting iterative algorithm may perform better than the steepest descent method. Newton's method (sometimes called the Newton-Raphson method) uses first and second derivatives and indeed does perform better than the steepest descent method if the initial point is close to the minimizer. The idea behind this method is as follows. Given a starting point, we construct a quadratic approximation to the objective function that matches the first and second derivative values at that point. We then minimize the approximate (quadratic) function instead of the original objective function. We use the minimizer of the approximate function as the starting point in the next step and repeat the procedure iteratively. If the objective function is quadratic, then the approximation is exact, and the method yields the true minimizer in one step. If, on the other hand, the objective function is not quadratic, then the approximation will provide only an estimate of the position of the true minimizer. Figure 9.1 illustrates the above idea.

We can obtain a quadratic approximation to the given twice continuously differentiable objection function $f : \mathbb{R}^n \to \mathbb{R}$ using the Taylor series expansion of f about the current point $\boldsymbol{x}^{(k)}$, neglecting terms of order three and higher. We obtain

$$f(\boldsymbol{x}) \approx f(\boldsymbol{x}^{(k)}) + (\boldsymbol{x} - \boldsymbol{x}^{(k)})^T \boldsymbol{g}^{(k)} + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^{(k)})^T \boldsymbol{F}(\boldsymbol{x}^{(k)})(\boldsymbol{x} - \boldsymbol{x}^{(k)}) \triangleq q(\boldsymbol{x}),$$

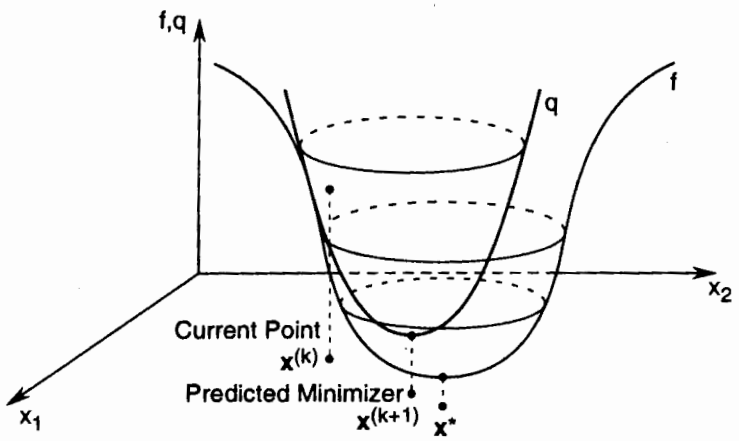

Figure 9.1 Quadratic approximation to the objective function using first and second derivatives

where, for simplicity, we use the notation $g^{(k)} = \nabla f(x^{(k)})$. Applying the FONC to q yields

$$0 = \nabla q(x) = g^{(k)} + F(x^{(k)})(x - x^{(k)}).$$

If $F(x^{(k)}) > 0$, then q achieves a minimum at

$$x^{(k+1)} = x^{(k)} - F(x^{(k)})^{-1} g^{(k)}.$$

This recursive formula represents Newton's method.

Example 9.1 Use Newton's method to minimize the Powell function:

$$f(x_1, x_2, x_3, x_4) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4.$$

Use as the starting point $x^{(0)} = [3, -1, 0, 1]^T$. Perform three iterations.

Note that $f(x^{(0)}) = 215$. We have

$$\nabla f(x) = \begin{bmatrix} 2(x_1 + 10x_2) + 40(x_1 - x_4)^3 \\ 20(x_1 + 10x_2) + 4(x_2 - 2x_3)^3 \\ 10(x_3 - x_4) - 8(x_2 - 2x_3)^3 \\ -10(x_3 - x_4) - 40(x_1 - x_4)^3 \end{bmatrix},$$

and $F(x)$ is given by

$$\begin{bmatrix} 2 + 120(x_1 - x_4)^2 & \cdot 20 & 0 & -120(x_1 - x_4)^2 \\ 20 & 200 + 12(x_2 - 2x_3)^2 & -24(x_2 - 2x_3)^2 & 0 \\ 0 & -24(x_2 - 2x_3)^2 & 10 + 48(x_2 - 2x_3)^2 & -10 \\ -120(x_1 - x_4)^2 & 0 & -10 & 10 + 120(x_1 - x_4)^2 \end{bmatrix}$$

Iteration 1.

$$
\begin{aligned}
g^{(0)} &= [306, -144, -2, -310]^T, \\
F(x^{(0)}) &= \begin{bmatrix} 482 & 20 & 0 & -480 \\ 20 & 212 & -24 & 0 \\ 0 & -24 & 58 & -10 \\ -480 & 0 & -10 & 490 \end{bmatrix}, \\
F(x^{(0)})^{-1} &= \begin{bmatrix} .1126 & -.0089 & .0154 & .1106 \\ -.0089 & .0057 & .0008 & -.0087 \\ .0154 & .0008 & .0203 & .0155 \\ .1106 & -.0087 & .0155 & .1107 \end{bmatrix}, \\
F(x^{(0)})^{-1} g^{(0)} &= [1.4127, -0.8413, -0.2540, 0.7460]^T.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
x^{(1)} &= x^{(0)} - F(x^{(0)})^{-1} g^{(0)} = [1.5873, -0.1587, 0.2540, 0.2540]^T, \\
f(x^{(1)}) &= 31.8.
\end{aligned}
$$

Iteration 2.

$$
\begin{aligned}
g^{(1)} &= [94.81, -1.179, 2.371, -94.81]^T, \\
F(x^{(1)}) &= \begin{bmatrix} 215.3 & 20 & 0 & -213.3 \\ 20 & 205.3 & -10.67 & 0 \\ 0 & -10.67 & 31.34 & -10 \\ -213.3 & 0 & -10 & 223.3 \end{bmatrix}, \\
F(x^{(1)})^{-1} g^{(1)} &= [0.5291, -0.0529, 0.0846, 0.0846]^T.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
x^{(2)} &= x^{(1)} - F(x^{(1)})^{-1} g^{(1)} = [1.0582, -0.1058, 0.1694, 0.1694]^T, \\
f(x^{(2)}) &= 6.28.
\end{aligned}
$$

Iteration 3.

$$
\begin{aligned}
g^{(2)} &= [28.09, -0.3475, 0.7031, -28.08]^T, \\
F(x^{(2)}) &= \begin{bmatrix} 96.80 & 20 & 0 & -94.80 \\ 20 & 202.4 & -4.744 & 0 \\ 0 & -4.744 & 19.49 & -10 \\ -94.80 & 0 & -10 & 104.80 \end{bmatrix}, \\
x^{(3)} &= [0.7037, -0.0704, 0.1121, 0.1111]^T, \\
f(x^{(3)}) &= 1.24.
\end{aligned}
$$

■

Observe that the kth iteration of Newton's method can be written in two steps as

1. Solve $\boldsymbol{F}(\boldsymbol{x}^{(k)})\boldsymbol{d}^{(k)} = -\boldsymbol{g}^{(k)}$ for $\boldsymbol{d}^{(k)}$;
2. Set $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \boldsymbol{d}^{(k)}$.

Step 1 requires the solution of an $n \times n$ system of linear equations. Thus, an efficient method for solving systems of linear equations is essential when using Newton's method.

As in the one-variable case, Newton's method can also be viewed as a technique for iteratively solving the equation

$$\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0},$$

where $\boldsymbol{x} \in \mathbb{R}^n$, and $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^n$. In this case, $\boldsymbol{F}(\boldsymbol{x})$ is the Jacobian matrix of $\boldsymbol{g}$ at $\boldsymbol{x}$, that is, $\boldsymbol{F}(\boldsymbol{x})$ is the $n \times n$ matrix whose (i,j) entry is $(\partial g_i/\partial x_j)(\boldsymbol{x})$, $i,j = 1,2,\ldots,n$.

9.2 ANALYSIS OF NEWTON'S METHOD

As in the one-variable case, there is no guarantee that Newton's algorithm heads in the direction of decreasing values of the objective function if $\boldsymbol{F}(\boldsymbol{x}^{(k)})$ is not positive definite (recall Figure 7.7 illustrating Newton's method for functions of one variable when $f'' < 0$). Moreover, even if $\boldsymbol{F}(\boldsymbol{x}^{(k)}) > 0$, Newton's method may not be a descent method; that is, it is possible that $f(\boldsymbol{x}^{(k+1)}) \geq f(\boldsymbol{x}^{(k)})$. For example, this may occur if our starting point $\boldsymbol{x}^{(0)}$ is far away from the solution. See the end of this section for a possible remedy to this problem. Despite the above drawbacks, Newton's method has superior convergence properties when the starting point is near the solution, as we shall see in the remainder of this section.

The convergence analysis of Newton's method when f is a quadratic function is straightforward. In fact, Newton's method reaches the point $\boldsymbol{x}^*$ such that $\nabla f(\boldsymbol{x}^*) = 0$ in just one step starting from any initial point $\boldsymbol{x}^{(0)}$. To see this, suppose that $\boldsymbol{Q} = \boldsymbol{Q}^T$ is invertible, and

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{b}.$$

Then,

$$\boldsymbol{g}(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) = \boldsymbol{Q}\boldsymbol{x} - \boldsymbol{b},$$

and

$$\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{Q}.$$

Hence, given any initial point $\boldsymbol{x}^{(0)}$, by Newton's algorithm

$$\begin{aligned} \boldsymbol{x}^{(1)} &= \boldsymbol{x}^{(0)} - \boldsymbol{F}(\boldsymbol{x}^{(0)})^{-1}\boldsymbol{g}^{(0)} \\ &= \boldsymbol{x}^{(0)} - \boldsymbol{Q}^{-1}[\boldsymbol{Q}\boldsymbol{x}^{(0)} - \boldsymbol{b}] \\ &= \boldsymbol{Q}^{-1}\boldsymbol{b} \\ &= \boldsymbol{x}^*. \end{aligned}$$

because $g^{(k)} = Qx^{(k)} - b$ and $Qx^* = b$. Thus,

$$\beta_k = -\frac{d^{(k)T} g^{(k)}}{d^{(k)T} Q d^{(k)}} = \alpha_k$$

and $x^* = x^{(n)}$, which completes the proof. ∎

Example 10.2 Find the minimizer of

$$f(x_1, x_2) = \frac{1}{2} x^T \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} x - x^T \begin{bmatrix} -1 \\ 1 \end{bmatrix}, x \in \mathbb{R}^2,$$

using the conjugate direction method with the initial point $x^{(0)} = [0, 0]^T$, and Q-conjugate directions $d^{(0)} = [1, 0]^T$ and $d^{(1)} = \left[-\frac{3}{8}, \frac{3}{4}\right]^T$.

We have

$$g^{(0)} = -b = [1, -1]^T,$$

and hence

$$\alpha_0 = -\frac{g^{(0)T} d^{(0)}}{d^{(0)T} Q d^{(0)}} = -\frac{[1, -1] \begin{bmatrix} 1 \\ 0 \end{bmatrix}}{[1, 0] \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}} = -\frac{1}{4}.$$

Thus,

$$x^{(1)} = x^{(0)} + \alpha_0 d^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \frac{1}{4} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{4} \\ 0 \end{bmatrix}.$$

To find $x^{(2)}$, we compute

$$g^{(1)} = Qx^{(1)} - b = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} -\frac{1}{4} \\ 0 \end{bmatrix} - \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{3}{2} \end{bmatrix},$$

and

$$\alpha_1 = -\frac{g^{(1)T} d^{(1)}}{d^{(1)T} Q d^{(1)}} = -\frac{[0, -\frac{3}{2}] \begin{bmatrix} -\frac{3}{8} \\ \frac{3}{4} \end{bmatrix}}{[-\frac{3}{8}, \frac{3}{4}] \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} -\frac{3}{8} \\ \frac{3}{4} \end{bmatrix}} = 2.$$

Therefore,

$$x^{(2)} = x^{(1)} + \alpha_1 d^{(1)} = \begin{bmatrix} -\frac{1}{4} \\ 0 \end{bmatrix} + 2 \begin{bmatrix} -\frac{3}{8} \\ \frac{3}{4} \end{bmatrix} = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix}.$$

Because f is a quadratic function in two variables, $x^{(2)} = x^*$. ∎

For a quadratic function of n variables, the conjugate direction method reaches the solution after n steps. As we shall see below, the method also possesses a certain

If $j < k$, then $\boldsymbol{d}^{(k)T}\boldsymbol{Q}\boldsymbol{d}^{(j)} = 0$, by virtue of the induction hypothesis. Hence, we have

$$\boldsymbol{d}^{(k+1)T}\boldsymbol{Q}\boldsymbol{d}^{(j)} = -\boldsymbol{g}^{(k+1)T}\boldsymbol{Q}\boldsymbol{d}^{(j)}.$$

But $\boldsymbol{g}^{(j+1)} = \boldsymbol{g}^{(j)} + \alpha_j \boldsymbol{Q}\boldsymbol{d}^{(j)}$. Because $\boldsymbol{g}^{(k+1)T}\boldsymbol{g}^{(i)} = 0, i = 0, \ldots, k$,

$$\boldsymbol{d}^{(k+1)T}\boldsymbol{Q}\boldsymbol{d}^{(j)} = -\boldsymbol{g}^{(k+1)T}\frac{(\boldsymbol{g}^{(j+1)} - \boldsymbol{g}^{(j)})}{\alpha_j} = 0.$$

Thus,

$$\boldsymbol{d}^{(k+1)T}\boldsymbol{Q}\boldsymbol{d}^{(j)} = 0, \quad j = 0, \ldots, k-1.$$

It remains to be shown that $\boldsymbol{d}^{(k+1)T}\boldsymbol{Q}\boldsymbol{d}^{(k)} = 0$. We have

$$\boldsymbol{d}^{(k+1)T}\boldsymbol{Q}\boldsymbol{d}^{(k)} = (-\boldsymbol{g}^{(k+1)} + \beta_k \boldsymbol{d}^{(k)})^T \boldsymbol{Q}\boldsymbol{d}^{(k)}.$$

Using the expression for β_k, we get $\boldsymbol{d}^{(k+1)T}\boldsymbol{Q}\boldsymbol{d}^{(k)} = 0$, which completes the proof. ∎

Example 10.3 Consider the quadratic function

$$f(x_1, x_2, x_3) = \frac{3}{2}x_1^2 + 2x_2^2 + \frac{3}{2}x_3^2 + x_1x_3 + 2x_2x_3 - 3x_1 - x_3.$$

We find the minimizer using the conjugate gradient algorithm, using the starting point $\boldsymbol{x}^{(0)} = [0, 0, 0]^T$.

We can represent f as

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{b},$$

where

$$\boldsymbol{Q} = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}.$$

We have

$$\boldsymbol{g}(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) = \boldsymbol{Q}\boldsymbol{x} - \boldsymbol{b} = [3x_1 + x_3 - 3, 4x_2 + 2x_3, x_1 + 2x_2 + 3x_3 - 1]^T.$$

Hence,

$$\begin{aligned} \boldsymbol{g}^{(0)} &= [-3, 0, -1]^T, \\ \boldsymbol{d}^{(0)} &= -\boldsymbol{g}^{(0)}, \\ \alpha_0 &= -\frac{\boldsymbol{g}^{(0)T}\boldsymbol{d}^{(0)}}{\boldsymbol{d}^{(0)T}\boldsymbol{Q}\boldsymbol{d}^{(0)}} = \frac{10}{36} = 0.2778, \end{aligned}$$

and

$$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \alpha_0 \boldsymbol{d}^{(0)} = [0.8333, 0, 0.2778]^T.$$

The next stage yields

$$\begin{aligned} \boldsymbol{g}^{(1)} &= \nabla f(\boldsymbol{x}^{(1)}) = [-0.2222, 0.5556, 0.6667]^T, \\ \beta_0 &= \frac{\boldsymbol{g}^{(1)T}\boldsymbol{Q}\boldsymbol{d}^{(0)}}{\boldsymbol{d}^{(0)T}\boldsymbol{Q}\boldsymbol{d}^{(0)}} = 0.08025. \end{aligned}$$

We can now compute

$$\boldsymbol{d}^{(1)} = -\boldsymbol{g}^{(1)} + \beta_0 \boldsymbol{d}^{(0)} = [0.4630, -0.5556, -0.5864]^T.$$

Hence,

$$\alpha_1 = -\frac{\boldsymbol{g}^{(1)T}\boldsymbol{d}^{(1)}}{\boldsymbol{d}^{(1)T}\boldsymbol{Q}\boldsymbol{d}^{(1)}} = 0.2187,$$

and

$$\boldsymbol{x}^{(2)} = \boldsymbol{x}^{(1)} + \alpha_1 \boldsymbol{d}^{(1)} = [0.9346, -0.1215, 0.1495]^T.$$

To perform the third iteration, we compute

$$\begin{aligned} \boldsymbol{g}^{(2)} &= \nabla f(\boldsymbol{x}^{(2)}) = [-0.04673, -0.1869, 0.1402]^T, \\ \beta_1 &= \frac{\boldsymbol{g}^{(2)T}\boldsymbol{Q}\boldsymbol{d}^{(1)}}{\boldsymbol{d}^{(1)T}\boldsymbol{Q}\boldsymbol{d}^{(1)}} = 0.07075, \\ \boldsymbol{d}^{(2)} &= -\boldsymbol{g}^{(2)} + \beta_1 \boldsymbol{d}^{(1)} = [0.07948, 0.1476, -0.1817]^T. \end{aligned}$$

Hence,

$$\alpha_2 = -\frac{\boldsymbol{g}^{(2)T}\boldsymbol{d}^{(2)}}{\boldsymbol{d}^{(2)T}\boldsymbol{Q}\boldsymbol{d}^{(2)}} = 0.8231,$$

and

$$\boldsymbol{x}^{(3)} = \boldsymbol{x}^{(2)} + \alpha_2 \boldsymbol{d}^{(2)} = [1.000, 0.000, 0.000]^T.$$

Note that

$$\boldsymbol{g}^{(3)} = \nabla f(\boldsymbol{x}^{(3)}) = \boldsymbol{0},$$

as expected, because f is a quadratic function of three variables. Hence, $\boldsymbol{x}^* = \boldsymbol{x}^{(3)}$. ∎

10.4 THE CONJUGATE GRADIENT ALGORITHM FOR NON-QUADRATIC PROBLEMS

In the previous section, we showed that the conjugate gradient algorithm is a conjugate direction method, and therefore minimizes a positive definite quadratic function of n variables in n steps. The algorithm can be extended to general nonlinear functions by interpreting $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{b}$ as a second-order Taylor series approximation of the objective function. Near the solution such functions behave approximately as