in Section 2.5 on polyhedral sets. For further study of linear programming, see Bazaraa and Jarvis [1977], Charnes and Cooper [1961], Dantzig [1963], Hadley [1962], and Simonnard [1966].

Chapter 3
Convex Functions

Convex and concave functions have many special and important properties. For example, any local minimum of a convex function over a convex set is also a global minimum. In this chapter, we introduce the important topics of convex and concave functions and develop some of their properties. As we will learn in this and later chapters, these properties can be utilized in developing suitable optimality conditions and computational schemes for optimization problems that involve convex and concave functions.

The following is an outline of the chapter.

SECTION 3.1: Definitions and Basic Properties We introduce convex and concave functions and develop some of their basic properties. Continuity of convex functions is proved, and the concept of a directional derivative is introduced.

SECTION 3.2 Subgradients of Convex Functions A convex function has a convex epigraph and hence has a supporting hyperplane. This leads to the important notion of a subgradient of a convex function.

SECTION 3.3 Differentiable Convex Functions This section gives some characterizations of differentiable convex functions. These are helpful tools for checking convexity of simple differentiable functions.

SECTION 3.4: Minima and Maxima of Convex Functions This section is important, since it deals with the questions of minimizing and maximizing a convex function over a convex set. A necessary and sufficient condition for a minimum is developed. We also show that the maximum occurs at an extreme point. This fact is particularly important if the convex set is polyhedral.
SECTION 3.5: Generalizations of Convex Functions

Various relaxations of convexity and concavity are possible. We present quasiconvex and pseudoconvex functions and develop some of their properties. We then discuss various types of convexity at a point. These types of convexity are sometimes sufficient for optimality, as will be shown in Chapter 4.

3.1 Definitions and Basic Properties

This section deals with some basic properties of convex and concave functions. In particular, we investigate their continuity and differentiability properties.

3.1.1 Definition

Let \( f : S \to E_1 \), where \( S \) is a nonempty convex set in \( E_n \). The function \( f \) is said to be convex on \( S \) if

\[
    f(\lambda x_1 + (1 - \lambda) x_2) \leq \lambda f(x_1) + (1 - \lambda) f(x_2)
\]

for each \( x_1, x_2 \in S \) and for each \( \lambda \in (0, 1) \). The function \( f \) is called strictly convex on \( S \) if the above inequality is true as a strict inequality for each distinct \( x_1 \) and \( x_2 \) in \( S \) and for each \( \lambda \in (0, 1) \). The function \( f : S \to E_1 \) is called concave (strictly concave) on \( S \) if \(-f\) is convex (strictly convex) on \( S \).

Now let us consider the geometrical interpretation of convex and concave functions. Let \( x_1, x_2 \) be two distinct points in the domain of \( f \), and consider the point \( \lambda x_1 + (1 - \lambda) x_2 \) with \( \lambda \in (0, 1) \). Note that \( \lambda f(x_1) + (1 - \lambda) f(x_2) \) gives the weighted average of \( f(x_1) \) and \( f(x_2) \), while \( f(\lambda x_1 + (1 - \lambda) x_2) \) gives the value of \( f \) at the point \( \lambda x_1 + (1 - \lambda) x_2 \). So for a convex function \( f \), the value of \( f \) at points on the line segment \( \lambda x_1 + (1 - \lambda) x_2 \), is less or equal to the height of the cord joining the points \( [x_1, f(x_1)] \) and \( [x_2, f(x_2)] \). For a concave function, the cord is below the function itself. Figure 3.1 shows some examples of convex and concave functions.

![Figure 3.1 Examples of convex and concave functions.](image)

The following are some examples of convex functions. By taking the negatives of these functions, we get some examples of concave functions.

1. \( f(x) = 3x + 4 \)
2. \( f(x) = |x| \)
3. \( f(x) = x^2 - 2x \)
4. \( f(x) = -x^{1/2} \) if \( x \geq 0 \)
5. \( f(x_1, x_2) = 2x_1^2 + x_2^2 - 2x_1x_2 \)
6. \( f(x_1, x_2, x_3) = x_1^4 + 2x_2^2 + 3x_3^2 - 4x_1 - 4x_2x_3 \)

Note that in each of the above examples, except for part 4, the function \( f \) is convex over \( E_n \). In the example of part 4, the function is not defined for \( x < 0 \). One can readily construct examples of functions that are convex over a region but not over \( E_n \). For instance, \( f(x) = x^2 \) is not convex over \( E_1 \) but is convex over \( S = \{x : x \geq 0\} \).

From now on, we will concentrate only on convex functions. Results for concave functions can be easily obtained by noting that \( f \) is concave if and only if \(-f\) is convex.

Associated with a convex function \( f \) is the set \( S_\alpha = \{x \in S : f(x) \leq \alpha\} \), usually referred to as a level set. Lemma 3.1.2 below shows that \( S_\alpha \) is convex for each real number \( \alpha \).

3.1.2 Lemma

Let \( S \) be a nonempty convex set in \( E_n \) and let \( f : S \to E_1 \) be a convex function. Then the level set \( S_\alpha = \{x \in S : f(x) \leq \alpha\} \), where \( \alpha \) is a real number, is a convex set.

Proof

Let \( x_1, x_2 \in S_\alpha \). Thus, \( x_1, x_2 \in S \), and \( f(x_1) \leq \alpha \) and \( f(x_2) \leq \alpha \). Now let \( \lambda \in (0, 1) \) and \( x = \lambda x_1 + (1 - \lambda) x_2 \). By convexity of \( S, x \in S \). Furthermore, by convexity of \( f \),

\[
    f(x) \leq \lambda f(x_1) + (1 - \lambda) f(x_2) \leq \lambda \alpha + (1 - \lambda) \alpha = \alpha
\]

Hence \( x \in S_\alpha \), and therefore \( S_\alpha \) is convex.

Continuity of Convex Functions

An important property of convex and concave functions is that they are continuous on the interior of their domain. This fact is proved below.
3.1.3 Theorem

Let $S$ be a nonempty convex set in $E_n$, and let $f : S \to E_1$ be convex. Then $f$ is continuous on the interior of $S$.

Proof

Let $\bar{x} \in \text{int } S$. To prove continuity of $f$ at $\bar{x}$, we need to show that given $\varepsilon > 0$, there exists a $\delta > 0$ such that $\|x - \bar{x}\| \leq \delta$ implies that $|f(x) - f(\bar{x})| \leq \varepsilon$. Since $\bar{x} \in \text{int } S$, there exists a $\delta' > 0$ such that $\|x - \bar{x}\| \leq \delta'$ implies that $x \in S$. Construct $\theta$ as follows.

$$\theta = \max \{ \max \{ f(\bar{x} + \delta' e_i) - f(\bar{x}), f(\bar{x} - \delta' e_i) - f(\bar{x}) \} \}$$

(3.1)

where $e_i$ is a vector of zeros except for a 1 at the $i$th position. Note that $0 \leq \theta < \infty$. Let

$$\delta = \min \left( \frac{\delta'}{n}, \frac{\varepsilon}{n \theta} \right)$$

(3.2)

Choose an $x$ with $\|x - \bar{x}\| \leq \delta$. If $x_i - \bar{x}_i \geq 0$, let $z_i = \delta' e_i$; otherwise let $z_i = -\delta' e_i$. Then $x - \bar{x} = \sum_{i=1}^{n} \alpha_i z_i$, where $\alpha_i \geq 0$ for $i = 1, 2, \ldots, n$. Furthermore

$$\|x - \bar{x}\| = \delta' \left( \sum_{i=1}^{n} \alpha_i^2 \right)^{1/2}$$

(3.3)

From (3.2) and since $\|x - \bar{x}\| \leq \delta$, it follows that $\alpha_i \leq 1/n$ for $i = 1, 2, \ldots, n$. Hence, by convexity of $f$ and since $0 \leq n \alpha_i \leq 1$, we get

$$f(x) = f(\bar{x} + \sum_{i=1}^{n} \alpha_i z_i) = f\left( \frac{1}{n} \sum_{i=1}^{n} (\bar{x} + n \alpha_i z_i) \right) \leq \sum_{i=1}^{n} f(\bar{x} + n \alpha_i z_i) = \frac{1}{n} \sum_{i=1}^{n} f((1 - n \alpha_i)(\bar{x} + n \alpha_i z_i)) \leq \frac{1}{n} \sum_{i=1}^{n} f((1 - n \alpha_i)\bar{x} + n \alpha_i z_i)$$

Therefore $f(x) - f(\bar{x}) \leq \sum_{i=1}^{n} \alpha_i [f(\bar{x} + z_i) - f(\bar{x})]$. From (3.1) it is obvious that $f(\bar{x} + z_i) - f(\bar{x}) \leq \theta$ for each $i$, and since $\alpha_i \geq 0$, it follows that

$$f(x) - f(\bar{x}) \leq \theta \sum_{i=1}^{n} \alpha_i$$

(3.4)

3.1.4 Definition

Let $S$ be a nonempty set in $E_n$, and let $f : S \to E_1$. Let $\bar{x} \in S$ and $d$ be a nonzero vector such that $\bar{x} + \lambda d \in S$ for $\lambda > 0$ and sufficiently small. The directional derivative of $f$ at $\bar{x}$ along the vector $d$, denoted by $f'(\bar{x}; d)$, is given by the following limit if it exists:

$$f'(\bar{x}; d) = \lim_{\lambda \to 0^+} \frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda}$$

In particular, the limit in Definition 3.1.4, with the values $\infty$ and $-\infty$ permitted, exists for convex and concave functions as shown below.

3.1.5 Lemma

Let $S$ be a nonempty convex set in $E_n$, and let $f : S \to E_1$ be convex. Let $\bar{x} \in S$ and $d$ be a nonzero vector such that $\bar{x} + \lambda d \in S$ for $\lambda > 0$ and sufficiently small.
Then the following limit exists:
\[
\lim_{\lambda \to 0^+} \frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda}
\]

**Proof**
Let \(\lambda_2 > \lambda_1 > 0\) be sufficiently small. Noting convexity of \(f\), we have
\[
f(\bar{x} + \lambda_1 d) = \frac{\lambda_1}{\lambda_2} f(\bar{x} + \lambda_1 d) + \left(1 - \frac{\lambda_1}{\lambda_2}\right) f(\bar{x})
\]
This inequality implies that
\[
\frac{f(\bar{x} + \lambda_1 d) - f(\bar{x})}{\lambda_1} \leq \frac{f(\bar{x} + \lambda_2 d) - f(\bar{x})}{\lambda_2}
\]
Thus the difference quotient \(\frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda}\) is a nondecreasing function of \(\lambda > 0\), and hence the limit in the theorem exists and is given by
\[
\lim_{\lambda \to 0^+} \frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda} = \inf_{\lambda > 0} \frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda}
\]

### 3.2 Subgradients of a Convex Function

In this section, we introduce the important concept of subgradients of convex and concave functions. This can be done by establishing supporting hyperplanes to the epigraphs of convex functions and to the hypographs of concave functions.

**Epigraph and Hypograph of a Function**

A function \(f\) on \(S\) can be fully described by the set \([x, f(x)]: x \in S\) \(\subseteq E_{n+1}\), which is referred to as the *graph* of the function. One can construct two sets that are related to the graph of \(f\): the epigraph, which consists of points above the graph of \(f\), and the hypograph, which consists of points below the graph of \(f\). These notions are clarified by Definition 3.2.1 below.

#### 3.2.1 Definition

Let \(S\) be a nonempty set in \(E_n\), and let \(f: S \to E_1\). The **epigraph** of \(f\), denoted by \(\text{epi } f\), is a subset of \(E_{n+1}\) defined by
\[
\{(x, y) : x \in S, y \in E_1, y \geq f(x)\}
\]

The **hypograph** of \(f\), denoted by \(\text{hyp } f\), is a subset of \(E_{n+1}\) defined by
\[
\{(x, y) : x \in S, y \in E_1, y \leq f(x)\}
\]

**Figure 3.2** Examples of epigraphs and hypographs.

Figure 3.2 illustrates the epigraphs and hypographs of several functions. In Figure 3.2a, neither the epigraph nor the hypograph of \(f\) is a convex set. But in Figure 3.2b and c, respectively, the epigraph and hypograph of \(f\) are convex sets. It turns out that a function is convex if and only if its epigraph is a convex set and also that a function is concave if and only if its hypograph is a convex set.

#### 3.2.2 Theorem

Let \(S\) be a nonempty convex set in \(E_n\), and let \(f: S \to E_1\). Then \(f\) is convex if and only if \(\text{epi } f\) is a convex set.

**Proof**

Assume that \(f\) is convex, and let \((x_1, y_1)\) and \((x_2, y_2)\) \(\in \text{epi } f\); that is, \(x_1, x_2 \in S, y_1 \geq f(x_1)\), and \(y_2 \geq f(x_2)\). Let \(\lambda \in (0, 1)\). Then
\[
\lambda y_1 + (1 - \lambda) y_2 \geq \lambda f(x_1) + (1 - \lambda) f(x_2) \geq f(\lambda x_1 + (1 - \lambda) x_2)
\]
where the last inequality follows by convexity of \(f\). Note that \(\lambda x_1 + (1 - \lambda) x_2 \in S\). Thus \([\lambda x_1 + (1 - \lambda) x_2, \lambda y_1 + (1 - \lambda) y_2] \in \text{epi } f\), and hence \(\text{epi } f\) is convex. Conversely, assume that \(\text{epi } f\) is convex, and let \(x_1, x_2 \in S\). Then \([x_1, f(x_1)]\) and \([x_2, f(x_2)]\) belong to \(\text{epi } f\), and by convexity of \(\text{epi } f\), we must have
\[
[\lambda x_1 + (1 - \lambda) x_2, \lambda f(x_1) + (1 - \lambda) f(x_2)] \in \text{epi } f\quad \text{for } \lambda \in (0, 1)
\]
In other words, \(\lambda f(x_1) + (1 - \lambda) f(x_2) \geq f(\lambda x_1 + (1 - \lambda) x_2)\) for each \(\lambda \in (0, 1)\); that is, \(f\) is convex. This completes the proof.

The above theorem can be used to verify convexity or concavity of a given function \(f\). Making use of this result, it is clear that the functions illustrated in Figure 3.2 are (a) neither convex nor concave, (b) convex, and (c) concave.
Since the epigraph of a convex function and the hypograph of a concave function are convex sets, they have supporting hyperplanes at points of their boundary. These supporting hyperplanes lead to the notion of subgradients, which is defined below.

### 3.2.3 Definition

Let $S$ be a nonempty convex set in $E_m$ and let $f: S \to E_1$ be convex. Then $\xi$ is called a subgradient of $f$ at $\bar{x} \in S$ if

$$f(x) \geq f(\bar{x}) + \xi'(x - \bar{x}) \quad \text{for all } x \in S$$

Similarly, let $f: S \to E_1$ be concave. Then $\xi$ is called a subgradient of $f$ at $\bar{x} \in S$ if

$$f(x) \leq f(\bar{x}) + \xi'(x - \bar{x}) \quad \text{for all } x \in S$$

From the above definition, it immediately follows that the collection of subgradients of $f$ at $\bar{x}$ is a convex set. Figure 3.3 shows examples of subgradients of convex and concave functions. From the figure, we see that the function $f(\bar{x}) + \xi'(x - \bar{x})$ corresponds to a supporting hyperplane of the epigraph or the hypograph of the function $f$. The subgradient vector $\xi$ corresponds to the slope of the supporting hyperplane.

### 3.2.4 Example

Let $f(x) = \min\{f_1(x), f_2(x)\}$, where $f_1$ and $f_2$ are defined below.

- $f_1(x) = 4 - |x| \quad x \in E_1$
- $f_2(x) = 4 - (x - 2)^2 \quad x \in E_1$

Since $f_2(x) \geq f_1(x)$ for $1 \leq x \leq 4$, $f$ can be represented as follows:

$$f(x) = \begin{cases} 4 - x & 1 \leq x \leq 4 \\ 4 - (x - 2)^2 & \text{otherwise} \end{cases}$$

In Figure 3.4, the concave function $f$ is shown in dark lines. Note that $\xi = -1$ is a subgradient of $f$ at any point $x$ in the open interval $(1, 4)$. If $x < 1$ or $x > 4$, then $\xi = -2(x - 2)$ is a subgradient of $f$. At the points $x = 1$ and $x = 4$, the subgradients are not unique because many supporting hyperplanes exist.

The following theorem shows that every convex or concave function has at least one subgradient at points in the interior of its domain. The proof relies on the fact that a convex set has a supporting hyperplane at points of the boundary.

### 3.2.5 Theorem

Let $S$ be a nonempty convex set in $E_m$ and let $f: S \to E_1$ be convex. Then, for $\bar{x} \in \text{int } S$ there exists a vector $\xi$ such that the hyperplane

$$H = \{(x, y) : y = f(\bar{x}) + \xi'(x - \bar{x})\}$$
supports epi $f$ at $[\bar{x}, f(\bar{x})]$. In particular,
\[ f(x) \geq f(\bar{x}) + \xi'(x - \bar{x}) \text{ for each } x \in S \]
that is, $\xi$ is a subgradient of $f$ at $\bar{x}$.

Proof

By Theorem 3.2.2, epi $f$ is convex. Noting that $[\bar{x}, f(\bar{x})]$ belongs to the boundary of epi $f$, by Theorem 2.3.7 there exists a nonzero vector $(\xi_0, \mu) \in E_n \times E_1$ such that
\[ \xi_0' (x - \bar{x}) + \mu [y - f(\bar{x})] \leq 0 \quad \text{for all } (x, y) \in \text{epi } f \tag{3.7} \]
Note that $\mu$ is not positive because otherwise the above inequality will be contradicted by choosing $y$ sufficiently large. We now show that

By contradiction, suppose that $\mu > 0$. Then $\xi_0' (x - \bar{x}) \leq 0$ for all $x \in S$. Since $\bar{x} \in \text{int } S$, there exists a $\lambda > 0$ such that $\bar{x} + \lambda \xi_0 \in S$, and hence $\lambda \xi_0' \xi_0 \leq 0$. This implies that $\xi_0 = 0$ and $(\xi_0, \mu) = (0, 0)$, contradicting the fact that $(\xi_0, \mu)$ is a nonzero vector. Therefore, $\mu < 0$. Denoting $\xi_0/|\mu|$ by $\xi$ and dividing the inequality in (3.7) by $|\mu|$, we get
\[ \xi'(x - \bar{x}) - y + f(\bar{x}) \leq 0 \quad \text{for all } (x, y) \in \text{epi } f \tag{3.8} \]
In particular, the hyperplane $H = \{(x, y): y = f(\bar{x}) + \xi'(x - \bar{x})\}$ supports epi $f$ at $[\bar{x}, f(\bar{x})]$. By letting $y = f(x)$ in (3.8), we get $f(x) \geq f(\bar{x}) + \xi'(x - \bar{x})$ for all $x \in S$, and the proof is complete.

Corollary

Let $S$ be a nonempty convex set in $E_n$, and let $f: S \rightarrow E_1$ be strictly convex. Then, for $\bar{x} \in \text{int } S$ there exists a vector $\xi$ such that
\[ f(x) > f(\bar{x}) + \xi'(x - \bar{x}) \quad \text{for all } x \in S, x \neq \bar{x} \]

Proof

By the theorem, there exists a vector $\xi$ such that
\[ f(x) \geq f(\bar{x}) + \xi'(x - \bar{x}) \quad \text{for all } x \in S \tag{3.9} \]
By contradiction, suppose that there is an $x \neq \bar{x}$ such that $f(\bar{x}) = f(x) + \xi'(x - \bar{x})$. Then, noting strict convexity of $f$ for $\lambda \in (0, 1)$, we get
\[ f(\lambda \bar{x} + (1 - \lambda)x) < \lambda f(\bar{x}) + (1 - \lambda)f(x) = f(\bar{x}) + (1 - \lambda)f'(x - \bar{x}) \tag{3.10} \]
But letting $x = \lambda \bar{x} + (1 - \lambda)x$ in (3.9), we must have
\[ f(\lambda \bar{x} + (1 - \lambda)x) \geq f(\bar{x}) + (1 - \lambda)f'(x - \bar{x}) \]
contradicting (3.10). This proves the corollary.

The converse of the above theorem is not true in general. In other words, if corresponding to each point $x \in \text{int } S$ there is a subgradient of $f$, then $f$ is not necessarily a convex function. To illustrate, consider the following example, where $f$ is defined on $S = \{(x_1, x_2): 0 \leq x_1, x_2 \leq 1\}$:
\[ f(x_1, x_2) = \begin{cases} 0 & 0 \leq x_1 \leq 1, \ 0 < x_2 \leq 1 \\
\quad (1 - (x_1 - \frac{1}{2})^2 & 0 \leq x_1 \leq 1, \ x_2 = 0 \end{cases} \]
For each point in the interior of the domain, the zero vector is a subgradient of $f$. However, $f$ is not convex on $S$ since epi $f$ is clearly not a convex set. However, as the following theorem shows, $f$ is indeed convex on int $S$.

3.2.6 Theorem

Let $S$ be a nonempty convex set in $E_n$, and let $f: S \rightarrow E_1$. Suppose that for each point $\bar{x} \in \text{int } S$ there exists a subgradient $\xi$ of $f$ such that
\[ f(x) \geq f(\bar{x}) + \xi'(x - \bar{x}) \quad \text{for each } x \in S. \]
Then, $f$ is convex on int $S$.

Proof

Let $x_1, x_2 \in \text{int } S$, and let $\lambda \in (0, 1)$. By Corollary 1 to Theorem 2.2.2, int $S$ is convex, and we must have $\lambda x_1 + (1 - \lambda)x_2 \in \text{int } S$. By assumption, there exists a subgradient $\xi$ of $f$ at $\lambda x_1 + (1 - \lambda)x_2$. In particular the following two inequalities hold true:
\[ f(x_1) \geq f(\lambda x_1 + (1 - \lambda)x_2) + (1 - \lambda)\xi'(x_1 - x_2) \\
 f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2) + \lambda \xi'(x_2 - x_1) \]
Multiplying the above two inequalities by $\lambda$ and $(1 - \lambda)$, respectively, and adding, we obtain
\[ \lambda f(x_1) + (1 - \lambda)f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2) \]
and the result follows.

3.3 Differentiable Convex Functions

We now focus on differentiable convex and concave functions. First, consider the following definition of differentiability.

3.3.1 Definition

Let $S$ be a nonempty set in $E_n$, and let $f: S \rightarrow E_1$. Then $f$ is said to be differentiable at $\bar{x} \in \text{int } S$ if there exists a vector $\nabla f(\bar{x})$, called the gradient vector,
and a function \( \alpha : E_n \to E_1 \) such that 
\[
  f(x) = f(\bar{x}) + \nabla f(\bar{x})'(x - \bar{x}) + \|x - \bar{x}\| \alpha(\bar{x}; x - \bar{x})
\]
for each \( x \in S \)

where \( \lim_{x \to \bar{x}} \alpha(\bar{x}; x - \bar{x}) = 0 \). The function \( f \) is said to be differentiable on the open set \( S' \subset S \) if it is differentiable at each point in \( S' \).

Note that if \( f \) is differentiable at \( \bar{x} \), then there could only be one gradient vector, and this vector is given by 
\[
  \nabla f(\bar{x}) = \left( \frac{\partial f(\bar{x})}{\partial x_1}, \ldots, \frac{\partial f(\bar{x})}{\partial x_n} \right)
\]
where \( \frac{\partial f(\bar{x})}{\partial x_i} \) is the partial derivative of \( f \) with respect to \( x_i \) at \( \bar{x} \) (see Exercise 3.23).

The following lemma shows that a differentiable convex function has only one subgradient, namely the gradient vector. Hence, the results of the previous section can be easily specialized to the differentiable case, in which the gradient vector replaces subgradients.

### 3.3.2 Lemma

Let \( S \) be a nonempty convex set in \( E_n \) and let \( f : S \to E_1 \) be convex. Suppose that \( f \) is differentiable at \( \bar{x} \in \text{int} S \). Then the collection of subgradients of \( f \) at \( \bar{x} \) is the singleton set \( \{\nabla f(\bar{x})\} \).

**Proof**

By Theorem 3.2.5, the set of subgradients of \( f \) at \( \bar{x} \) is not empty. Now let \( \xi \) be a subgradient of \( f \) at \( \bar{x} \). As a result of Theorem 3.2.5 and the differentiability of \( f \) at \( \bar{x} \), for any vector \( d \) and for \( \lambda \) sufficiently small, we get
\[
  f(\bar{x} + \lambda d) = f(\bar{x}) + \lambda \xi d
\]
\[
  f(\bar{x} + \lambda d) = f(\bar{x}) + \lambda \nabla f(\bar{x})' d + \lambda \|d\| \alpha(\bar{x}; \lambda d)
\]

Subtracting the equation from the inequality, we obtain
\[
  0 \geq \lambda [\xi - \nabla f(\bar{x})]'d - \lambda \|d\| \alpha(\bar{x}; \lambda d)
\]

If we divide by \( \lambda > 0 \) and let \( \lambda \to 0 \), it follows that \( [\xi - \nabla f(\bar{x})]'d \leq 0 \). Choosing \( d = \xi - \nabla f(\bar{x}) \), the last inequality implies that \( \xi = \nabla f(\bar{x}) \). This completes the proof.

In the light of the above lemma, we give the following important characterization of differentiable convex functions. The proof is immediate from Theorems 3.2.5 and 3.2.6 and Lemma 3.3.2.

### 3.3.3 Theorem

Let \( S \) be a nonempty open convex set in \( E_n \) and let \( f : S \to E_1 \) be differentiable on \( S \). Then \( f \) is convex if and only if for any \( \bar{x} \in S \), we have
\[
  f(x) \geq f(\bar{x}) + \nabla f(\bar{x})'(x - \bar{x})
\]
for each \( x \in S \).

Similarly, \( f \) is strictly convex if and only if for each \( \bar{x} \in S \), we have
\[
  f(x) > f(\bar{x}) + \nabla f(\bar{x})'(x - \bar{x})
\]
for each \( x \not= \bar{x} \) in \( S \).

The following theorem gives another necessary and sufficient characterization of differentiable convex functions. For a function of one variable, the characterization reduces to the gradient vector being nondecreasing.

### 3.3.4 Theorem

Let \( S \) be a nonempty open convex set in \( E_n \) and let \( f : S \to E_1 \) be differentiable on \( S \). Then \( f \) is convex if and only if for each \( x_1, x_2 \in S \), we have
\[
  [\nabla f(x_2) - \nabla f(x_1)]'(x_2 - x_1) \geq 0
\]
Similarly, \( f \) is strictly convex if and only for each distinct \( x_1, x_2 \in S \), we have
\[
  [\nabla f(x_2) - \nabla f(x_1)]'(x_2 - x_1) > 0
\]

**Proof**

Assume that \( f \) is convex, and let \( x_1, x_2 \in S \). By Theorem 3.3.3, we have
\[
  f(x_1) \geq f(x_2) + \nabla f(x_1)'(x_1 - x_2)
\]
\[
  f(x_2) \geq f(x_1) + \nabla f(x_1)'(x_2 - x_1)
\]

Adding the two inequalities, we get \([\nabla f(x_2) - \nabla f(x_1)]'(x_2 - x_1) \geq 0\). To show the converse, let \( x_1, x_2 \in S \). By the mean value theorem,
\[
  f(x_2) - f(x_1) = \nabla f(x_1)'(x_2 - x_1)
\]
(3.11)

where \( x = \lambda x_1 + (1 - \lambda) x_2 \) for some \( \lambda \in (0, 1) \). By assumption,\([\nabla f(x) - \nabla f(x_1)]'(x_2 - x_1) \geq 0\); that is, \((1 - \lambda)[\nabla f(x) - \nabla f(x_1)]'(x_2 - x_1) \geq 0\). This implies that \([\nabla f(x_2)' - \nabla f(x_1)'](x_2 - x_1) \geq 0\). By (3.11), we get \( f(x_2) \geq f(x_1) + \nabla f(x_1)'(x_2 - x_1) \). By Theorem 3.3.3, \( f \) is convex. The strict case is similar and the proof is complete.

Even though Theorems 3.3.3 and 3.3.4 provide necessary and sufficient characterizations of convex functions, checking these conditions is difficult from a computational standpoint. A simple and more manageable characterization, at least for quadratic functions, can be obtained, provided that the function is twice differentiable.
Twice Differentiable Convex and Concave Functions

A function $f$ that is differentiable at $\bar{x}$ is said to be twice differentiable at $\bar{x}$ if the representation of Definition 3.3.5 below is true.

### 3.3.5 Definition

Let $S$ be a nonempty set in $E_n$, and let $f:S \rightarrow E_1$. Then $f$ is said to be twice differentiable at $\bar{x} \in S$ if there exist a vector $\nabla f(\bar{x})$, and $n \times n$ symmetric matrix $H(\bar{x})$, called the Hessian matrix, and a function $\alpha:E_n \rightarrow E_1$ such that

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})' (x - \bar{x}) + \frac{1}{2} (x - \bar{x})' H(\bar{x}) (x - \bar{x}) + \alpha(x; \bar{x})$$

for each $x \in S$, where $\lim_{x \to \bar{x}} \alpha(x; x - \bar{x}) = 0$. The function $f$ is said to be twice differentiable on the open set $S' \subset S$ if it is twice differentiable at each point in $S'$.

It may be noted that the entry in row $i$ and column $j$ of the Hessian matrix $H(\bar{x})$ is the second partial derivative $\frac{\partial^2 f(\bar{x})}{\partial x_i \partial x_j}$.

The following theorem shows that if $f$ is convex on $S$, then for any $\bar{x} \in S$, we must have $\nabla^2 f(\bar{x}) \succeq 0$ for all $x \in E_n$.

### 3.3.6 Theorem

Let $S$ be a nonempty open convex set in $E_n$, and let $f:S \rightarrow E_1$ be twice differentiable on $S$. Then $f$ is convex if and only if its Hessian matrix is positive semidefinite at each point in $S$.

**Proof**

Suppose that $f$ is convex and let $\bar{x} \in S$. We need to show that $\nabla^2 f(\bar{x}) \succeq 0$ for each $x \in E_n$. Since $S$ is open, then for any given $x \in E_n$, $\bar{x} + \lambda x \in S$ for $\lambda$ sufficiently small. By Theorem 3.3.3 and by twice differentiability of $f$, we get the following two expressions:

$$f(\bar{x} + \lambda x) \succeq f(\bar{x}) + \nabla f(\bar{x})' (x - \bar{x}) + \frac{1}{2} \lambda^2 \nabla^2 f(\bar{x}) (x - \bar{x})$$

Subtracting (3.13) from (3.12), we get

$$\frac{1}{2} \lambda^2 \nabla^2 f(\bar{x}) (x - \bar{x}) \succeq 0$$

Dividing by $\lambda^2$ and letting $\lambda \to 0$, it follows that $\nabla^2 f(\bar{x}) \succeq 0$. Conversely, suppose that the Hessian matrix is positive semidefinite at each point in $S$. Consider $x$ and $\bar{x}$ in $S$. Then, by the mean value theorem, we have

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})' (x - \bar{x}) + \frac{1}{2} (x - \bar{x})' H(\bar{x}) (x - \bar{x})$$

where $\bar{x} = \lambda \bar{x} + (1 - \lambda) x$ for some $\lambda \in (0, 1)$. Note that $\bar{x} \in S$, and hence, by assumption, $H(\bar{x})$ is positive semidefinite. Therefore $(x - \bar{x})' H(\bar{x}) (x - \bar{x}) \succeq 0$, and from (3.14), we conclude that

$$f(x) \succeq f(\bar{x}) + \nabla f(\bar{x})' (x - \bar{x})$$

Since the above inequality is true for each $x$, $\bar{x} \in S$, $f$ is convex by Theorem 3.3.3. This completes the proof.

The above theorem is useful in checking convexity or concavity of a twice differentiable function. In particular, if the function is quadratic, then the Hessian matrix is independent of the point under consideration. Hence, checking convexity reduces to checking the nonnegativity of the eigenvalues of a single matrix.

### 3.3.7 Example

Let $f(x_1, x_2) = 3x_1^2 - 4x_2^2 + 2x_1x_2$. We want to check whether $f$ is convex or concave in $x_1$ and $x_2$. We rewrite $f$ in the following, more convenient form:

$$f(x_1, x_2) = (2, 6)^T x_1 + \frac{1}{2} (x_1, x_2)^T \begin{bmatrix} 4 & 4 \\ 4 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

In order to check whether the Hessian matrix $H$ is positive semidefinite or negative semidefinite or neither, we compute the eigenvalues by solving the following system:

$$0 = \det (H - \lambda I) = \det \begin{bmatrix} -4 - \lambda & 4 \\ 4 & -6 - \lambda \end{bmatrix} = (-4 - \lambda)(-6 - \lambda) - 16 = \lambda^2 + 10\lambda + 8$$

The solutions of this equation are $\lambda_1 = -5 + \sqrt{17}$ and $\lambda_2 = -5 - \sqrt{17}$. Since both $\lambda_1$ and $\lambda_2$ are negative, then $H$ is negative semidefinite and hence $f$ is concave.

Results analogous to Theorem 3.3.6 can be obtained for the strict convex and concave cases. It turns out that if the Hessian matrix is positive definite at each point in $S$, then the function is strictly convex. In other words, if for any given point $\bar{x}$ in $S$, we have $x' H(\bar{x}) x > 0$ for all $x \neq 0$ in $E_n$, then $f$ is strictly convex. However, if $f$ is strictly convex, then its Hessian matrix is positive semidefinite. It is not necessarily true that the Hessian matrix is positive definite everywhere in $S$. To illustrate, consider the strictly convex function defined by $f(x) = x^4$. The Hessian matrix $H(x) = 12x^2$ is positive definite for all nonzero $x$ but is positive semidefinite at $x = 0$. The following theorem is therefore obvious.

### 3.3.8 Theorem

Let $S$ be a nonempty open convex set in $E_n$, and let $f:S \rightarrow E_1$ be twice differentiable on $S$. If the Hessian matrix is positive definite at each point in $S$, then $f$ is strictly convex.
2. If \( f \) is strictly convex, then the Hessian matrix is positive semidefinite at each point in \( S \).

3.4 Minima and Maxima of Convex Functions

In this section, we consider the problems of minimizing and maximizing a convex function over a convex set and develop necessary conditions for optimality.

Minimizing a Convex Function

The case of maximizing a concave function is similar to that of minimizing a convex function. We develop the latter in detail and ask the reader to draw the analogous results for the concave case.

3.4.1 Definition

Let \( f: E \to E \), and consider the problem to minimize \( f(x) \) subject to \( x \in S \). A point \( x \in S \) is called a feasible solution to the problem. If \( \bar{x} \in S \) and \( f(\bar{x}) \leq f(x) \) for each \( x \in S \), then \( \bar{x} \) is called an optimal solution, a global optimal solution, or simply a solution to the problem. If \( \bar{x} \in S \) and there exists an \( \varepsilon \)-neighborhood \( N_\varepsilon(\bar{x}) \) around \( \bar{x} \) such that \( f(x) \leq f(\bar{x}) \) for each \( x \in S \cap N_\varepsilon(\bar{x}) \), then \( \bar{x} \) is called a local optimal solution.

Theorem 3.4.2 below shows that each local minimum of the above problem is also a global minimum. This fact is quite useful in the optimization process, since it enables us to stop with a global optimal solution if the search in the vicinity of a feasible point does not lead to an improved feasible solution.

3.4.2 Theorem

Let \( S \) be a nonempty convex set in \( E \), and \( f: S \to E \). Consider the problem to minimize \( f(x) \) subject to \( x \in S \). Suppose that \( \bar{x} \in S \) is a local optimal solution to the problem.
1. If \( f \) is convex, then \( \bar{x} \) is a global optimal solution.
2. If \( f \) is strictly convex, then \( \bar{x} \) is the unique global optimal solution.

Proof

To prove part (1), suppose that \( f \) is convex. Since \( \bar{x} \) is a local optimal solution, there exists an \( \varepsilon \)-neighborhood \( N_\varepsilon(\bar{x}) \) around \( \bar{x} \) such that

\[
 f(x) \geq f(\bar{x}) \quad \text{for each } x \in S \cap N_\varepsilon(\bar{x})
\]

(3.15)

By contradiction, suppose that \( \bar{x} \) is not a global optimal solution so that \( f(\bar{x}) < f(\bar{x}) \) for some \( \bar{x} \in S \). By convexity of \( f \), the following is true for each \( \lambda \in (0, 1) \)

\[
 f(\lambda \bar{x} + (1-\lambda)\bar{x}) \leq \lambda f(\bar{x}) + (1-\lambda)f(\bar{x}) < \lambda f(\bar{x}) + (1-\lambda)f(\bar{x}) = f(\bar{x})
\]

But for \( \lambda > 0 \) and sufficiently small, \( \lambda \bar{x} + (1-\lambda)\bar{x} \in S \cap N_\varepsilon(\bar{x}) \). Hence, the above inequality contradicts (3.15), and part (1) is proved.

Now assume \( f \) is strictly convex. Since strict convexity implies convexity, then by part (1) \( \bar{x} \) is a global optimal solution. By contradiction suppose that \( \bar{x} \) is not the unique global optimal solution so that there exists a \( x \in S, x \neq \bar{x} \), such that \( f(x) = f(\bar{x}) \). By strict convexity,

\[
 f(\frac{1}{2}x + \frac{1}{2}\bar{x}) < \frac{1}{2}f(x) + \frac{1}{2}f(\bar{x}) = f(\bar{x})
\]

By convexity of \( S, \frac{1}{2}x + \frac{1}{2}\bar{x} \in S \), and the above inequality violates global optimality of \( \bar{x} \). This completes the proof.

We now develop a necessary and sufficient condition for the existence of a global solution. If such an optimal solution does not exist, then either inf \( f(x) : x \in S \) is finite but not achieved at any point in \( S \) or it is equal to \(-\infty\).

3.4.3 Theorem

Let \( f: E \to E \) be a convex function, and \( S \) be a nonempty convex set in \( E \). Consider the problem to minimize \( f(x) \) subject to \( x \in S \). The point \( \bar{x} \in S \) is an optimal solution to this problem if and only if \( f \) has a subgradient \( \xi \) at \( \bar{x} \) such that \( \xi(x - \bar{x}) \geq 0 \) for all \( x \in S \).

Proof

Suppose that \( \xi'(x - \bar{x}) \geq 0 \) for all \( x \in S \), where \( \xi \) is a subgradient of \( f \) at \( \bar{x} \). By convexity of \( f \), we have

\[
 f(x) \geq f(\bar{x}) + \xi'(x - \bar{x}) \geq f(\bar{x}) \quad \text{for } x \in S
\]

and hence \( \bar{x} \) is an optimal solution of the problem.

To show the converse, suppose that \( \bar{x} \) is an optimal solution to the problem, and construct the following two sets in \( E_{n+1} \):

\[
 \Lambda_1 = \{ (x - \bar{x}, y) : x \in E, y > f(x) - f(\bar{x}) \}
\]

\[
 \Lambda_2 = \{ (x - \bar{x}, y) : x \in S, y \leq 0 \}
\]

The reader may easily verify that both \( \Lambda_1 \) and \( \Lambda_2 \) are convex sets. Also \( \Lambda_1 \cap \Lambda_2 = \emptyset \) because otherwise there would exist a point \((x, y)\) such that \( x \in S \quad 0 \geq y > f(x) - f(\bar{x}) \) contradicting the assumption that \( \bar{x} \) is an optimal solution of the problem. By
Theorem 3.2.8, there is a hyperplane that separates \( A_1 \) and \( A_2 \); that is, there exist a nonzero vector \((\xi_0, \mu)\) and a scalar \( \alpha \) such that

\[
\xi_0(x - \bar{x}) + \mu y \leq \alpha \quad x \in E_n, \quad y > f(x) - f(\bar{x})
\]

(3.16)

\[
\xi_0(x - \bar{x}) + \mu y \geq \alpha \quad x \in S, \quad y \leq 0
\]

(3.17)

If we let \( x = \bar{x} \) and \( y = 0 \) in (3.17), it follows that \( \alpha \leq 0 \). Next, letting \( x = \bar{x} \) and \( y = \varepsilon > 0 \) in (3.16), it follows that \( \mu \varepsilon \leq \alpha \). Since this is true for every \( \varepsilon > 0 \), then \( \mu \leq 0 \) and also \( \alpha \geq 0 \). To summarize, we have shown that \( \mu = 0 \) and \( \alpha = 0 \). If \( \mu = 0 \), then from (3.16) \( \xi_0(x - \bar{x}) = 0 \) for each \( x \in E_n \). If we let \( x = \bar{x} + \xi_0 \), it follows that

\[
0 = \xi_0(x - \bar{x})^2 = ||\xi_0||^2
\]

and hence \( \xi_0 = 0 \). Since \( (\xi_0, \mu) \neq (0, 0), \mu < 0 \). Dividing (3.16) and (3.17) by \( -\mu \), and denoting \( \xi_0/\mu \) by \( \xi \), we get the following inequalities:

\[
y \geq \xi'(x - \bar{x}) \quad x \in E_n, \quad y > f(x) - f(\bar{x})
\]

(3.18)

\[
\xi'(x - \bar{x}) - y \geq 0 \quad x \in S, \quad y \leq 0.
\]

(3.19)

By letting \( y = 0 \) in (3.19), we get \( \xi'(x - \bar{x}) \geq 0 \) for all \( x \in S \). From (3.18), it is obvious that

\[
f(x) \geq f(\bar{x}) + \xi'(x - \bar{x}) \quad \text{for all } x \in E_n.
\]

Therefore \( \xi \) is a subgradient of \( f \) at \( \bar{x} \) with the property that \( \xi'(x - \bar{x}) \geq 0 \) for all \( x \in S \), and the proof is complete.

**Corollary 1**

Under the assumptions of the theorem, if \( S \) is open, then \( \bar{x} \) is an optimal solution to the problem if and only if there exists a zero subgradient of \( f \) at \( \bar{x} \). In particular, if \( S = E_n \), then \( \bar{x} \) is a global minimum if and only if there exists a zero subgradient of \( f \) at \( \bar{x} \).

**Proof**

By the theorem, \( \bar{x} \) is an optimal solution if and only if \( \xi'(x - \bar{x}) \geq 0 \) for each \( x \in S \), where \( \xi \) is a subgradient of \( f \) at \( \bar{x} \). Since \( S \) is open, then \( x = \bar{x} - \lambda \xi \in S \) for some positive \( \lambda \). Therefore \( -\lambda ||\xi||^2 \geq 0 \), that is, \( \xi = 0 \).

**Corollary 2**

In addition to the assumptions of the theorem, suppose that \( f \) is differentiable. Then \( \bar{x} \) is an optimal solution if and only if \( \nabla f(\bar{x})' (x - \bar{x}) \geq 0 \) for all \( x \in S \). Furthermore, if \( S \) is open then \( \bar{x} \) is an optimal solution if and only if \( \nabla f(\bar{x}) = 0 \).

Note the important implications of the above theorem. First, the theorem gives a necessary and sufficient characterization of optimal points. This characterization reduces to the well-known condition of vanishing derivatives if \( f \) is differentiable and \( S \) is open. Another important implication is that if we reach a suboptimal point \( \bar{x} \), where \( \nabla f(\bar{x})' (x - \bar{x}) < 0 \) for some \( x \in S \), then there is an obvious way to proceed to an improved solution. This can be achieved by moving from \( \bar{x} \) in the direction \( x - \bar{x} \). The size of the step is to be determined by a one-dimensional minimization subproblem of the form: minimize \( f[\bar{x} + \lambda (x - \bar{x})] \) subject to \( \lambda \geq 0 \) and \( \bar{x} + \lambda (x - \bar{x}) \in S \). This procedure is called the method of feasible directions and is discussed in more detail in Chapter 10.

### 3.4.4 Example

Minimize \( (x_1 - \frac{1}{2})^2 + (x_2 - 5)^2 \)

subject to

\[
-2x_1 + 3x_2 \leq 11
\]

\[
-x_1 \leq 0
\]

\[
-x_2 \leq 0
\]

Clearly \( f(x_1, x_2) = (x_1 - \frac{1}{2})^2 + (x_2 - 5)^2 \) is a convex function, which gives the square of the distance from the point \((\frac{1}{2}, 5)\). The convex polyhedral set \( S \) is represented by the above four inequalities. The problem is depicted in Figure 3.5. From the figure, clearly the optimal point is \((1, 3)\). The gradient vector of \( f \) at the point \((1, 3)\) is \( \nabla f(1, 3) = (-1, -4)' \). We see geometrically that the vector \((-1, -4)\) makes an angle of \( \leq 90^\circ \) with each vector of the form \((x_1 - 1, x_2 - 3)\), where \((x_1, x_2) \in S \). Thus, the optimality condition of Theorem 3.4.3 is verified.

![Figure 3.5 Illustration of Example 3.4.4.](image)
Now, suppose that it is claimed that \((0,0)\) is an optimal point. By Theorem 3.4.3 we can easily verify that this is not the case. Note that \(\nabla f(0,0) = (-3,-10)\) and actually for nonzero \(x \in S\), we have \(-3x_1 - 10x_2 < 0\). Hence the origin could not be an optimal point. Moreover, we can improve \(f\) by moving from \(0\) in the direction \(x - 0\) for any \(x \in S\). In this case, the best local direction is \(-Vf(0,0)\), that is, the direction \((3,10)\). In Chapter 10, we discuss methods for finding a particular direction among many alternatives.

Maximizing a Convex Function

We now develop a necessary condition for a maximum of a convex function over a convex set. Unfortunately, however, this condition is not sufficient. Therefore, it is possible, and actually not unlikely, that several local maxima satisfying the condition of Theorem 3.4.5 exist. Unlike the minimization case, there exists no local information that could lead us to better points. Hence, maximizing a convex function is usually a much harder task than minimizing a convex function. Again, minimizing a concave function is similar to maximizing a convex function, and hence the development for the concave case is left to the reader.

3.4.5 Theorem

Let \(f: E_n \rightarrow E_1\) be a convex function, and let \(S\) be a nonempty convex set in \(E_n\). Consider the problem to maximize \(f(x)\) subject to \(x \in S\). If \(\bar{x} \in S\) is a local optimal solution then \(\bar{x}^f(x - \bar{x}) \leq 0\) for each \(x \in S\), where \(\bar{x}\) is any subgradient of \(f\) at \(\bar{x}\).

Proof

Suppose that \(\bar{x} \in S\) is a local optimal solution. Then there exists an \(\epsilon\)-neighborhood \(N_\epsilon(\bar{x})\) such that \(f(x) \leq f(\bar{x})\) for each \(x \in S \cap N_\epsilon(\bar{x})\). Let \(x \in S\), and note that \(\bar{x} + \lambda (x - \bar{x}) \in S \cap N_\epsilon(\bar{x})\) for \(\lambda > 0\) and sufficiently small. Hence,

\[
[f(\bar{x}) + \lambda (x - \bar{x})] - f(\bar{x}) \geq \lambda \bar{x}^f(x - \bar{x})
\]

Let \(\bar{x}\) be a subgradient of \(f\) at \(\bar{x}\), and by convexity of \(f\), we have

\[
f(\bar{x} + \lambda (x - \bar{x})) - f(\bar{x}) = \lambda \bar{x}^f(x - \bar{x})
\]

The above inequality, together with (3.20), implies that \(\lambda \bar{x}^f(x - \bar{x}) \leq 0\), and dividing by \(\lambda > 0\), the result follows.

Corollary

In addition to the assumptions of the theorem, suppose that \(f\) is differentiable. If \(\bar{x} \in S\) is a local optimal solution then \(\nabla f(\bar{x})^f(x - \bar{x}) \leq 0\) for all \(x \in S\).

Note that the above result is, in general, necessary but not sufficient for optimality. To illustrate, let \(f(x) = x^2\) and \(S = \{x: 1 - x \leq 2\}\). The maximum of \(f\) over \(S\) is equal to 4 and is achieved at \(x = 2\). However, at \(\bar{x} = 0\), we have \(\nabla f(\bar{x}) = 0\), and hence \(\nabla f(x) = 0\) for each \(x \in S\). Clearly the point \(\bar{x} = 0\) is not even a local maximum. Referring to Example 3.4.4 discussed earlier, we have two local maxima, namely \((0,0)\) and \((1/2,0)\). Both points satisfy the necessary condition of Theorem 3.4.5. If we are currently at the local optimal point \((0,0)\), unfortunately there exists no local information that will lead us toward the global maximum point \((1/2,0)\). Also, if we are at the global maximum point \((1/2,0)\), there is no convenient local criterion that tells us that we are at the optimal point.

Theorem 3.4.6 shows that a convex function achieves a maximum over a compact polyhedral set at an extreme point. This result has been utilized by several computational schemes for solving such problems. Theorem 3.4.6 could be extended to the case where the convex feasible region is not polyhedral.

3.4.6 Theorem

Let \(f: E_m \rightarrow E_1\) be a convex function, and let \(S\) be a nonempty compact polyhedral set in \(E_n\). Consider the problem to maximize \(f(x)\) subject to \(x \in S\). There then exists an optimal solution \(\bar{x}\) to the problem, where \(\bar{x}\) is an extreme point of \(S\).

Proof

By Theorem 3.1.3 note that \(f\) is continuous. Since \(S\) is compact, \(f\) assumes a maximum at \(x^* \in S\). If \(x^*\) is an extreme point of \(S\), then the result is at hand. Otherwise, by Theorem 2.5.7, \(x^* = \sum_{j=1}^k \lambda_j x_j\), where \(\sum_{j=1}^k \lambda_j = 1\), \(\lambda_j > 0\), and \(x_j\) is an extreme point of \(S\) for \(j = 1, \ldots, k\). By convexity of \(f\), we have

\[
f(x^*) = f\left(\sum_{j=1}^k \lambda_j x_j\right) \leq \sum_{j=1}^k \lambda_j f(x_j)
\]

But since \(f(x^*) \geq f(x_j)\) for \(j = 1, \ldots, k\), the above inequality implies that \(f(x^*) = f(x_j)\) for \(j = 1, \ldots, k\). Thus the extreme points \(x_1, \ldots, x_k\) are optimal solutions to the problem and the proof is complete.

3.5 Generalizations of a Convex Function

In this section, we present various types of functions that are similar to convex and concave functions but that share only some of their desirable properties. As we will learn, many of the results presented later in the book do not require the restrictive assumption of convexity but rather the less restrictive assumptions of quasiconvexity, pseudoconvexity, and convexity at a point.
Quasiconvex Functions

Definition 3.5.1 introduces quasiconvex functions. From the definition, it is apparent that every convex function is also quasiconvex.

3.5.1 Definition

Let \( f: S \rightarrow \mathbb{R} \), where \( S \) is a nonempty convex set in \( \mathbb{R}^n \). The function \( f \) is said to be quasiconvex if, for each \( x_1 \) and \( x_2 \) in \( S \), the following inequality is true:

\[
[f(\lambda x_1 + (1-\lambda)x_2) \leq \max \{f(x_1), f(x_2)\}] \quad \text{for each } \lambda \in (0, 1)
\]

The function \( f \) is said to be quasiconcave if \(-f\) is quasiconvex.

From the above definition, a function \( f \) is quasiconvex if whenever \( f(x_2) \geq f(x_1) \), then \( f(x_2) \) is greater than or equal to \( f \) at all convex combinations of \( x_1 \) and \( x_2 \). A function \( f \) is quasiconcave if whenever \( f(x_2) \geq f(x_1) \), then \( f \) at all convex combinations of \( x_1 \) and \( x_2 \) is greater than or equal to \( f(x_1) \). Figure 3.6 shows some examples of quasiconvex and quasiconcave functions. We will concentrate only on quasiconvex functions. The reader is advised to draw all the parallel results for quasiconcave functions.

We have learned in Section 3.2 that a convex function can be characterized by convexity of its epigraph. We now learn that a quasiconvex function can be characterized by convexity of its level sets. This result is given in Theorem 3.5.2.

3.5.2 Theorem

Let \( f: S \rightarrow \mathbb{R} \), where \( S \) is a nonempty convex set in \( \mathbb{R}^n \). The function \( f \) is quasiconvex if and only if \( S_\alpha = \{x \in S: f(x) \leq \alpha\} \) is convex for each real number \( \alpha \).

Proof

Suppose that \( f \) is quasiconvex, and let \( x_1, x_2 \in S_\alpha \). Therefore \( x_1, x_2 \in S \) and maximum \( \{f(x_1), f(x_2)\} \leq \alpha \). Let \( \lambda \in (0, 1) \), and let \( x = \lambda x_1 + (1-\lambda)x_2 \). By convexity of \( S \), \( x \in S \). Furthermore, by quasiconvexity of \( f \), \( f(x) \leq \max \{f(x_1), f(x_2)\} \leq \alpha \). Hence \( x \in S_\alpha \), and thus \( S_\alpha \) is convex. Conversely, suppose that \( S_\alpha \) is convex for each real number \( \alpha \). Let \( x_1, x_2 \in S \). Furthermore, let \( \lambda \in (0, 1) \) and \( x = \lambda x_1 + (1-\lambda)x_2 \). Note that \( x_1, x_2 \in S_\alpha \) for \( \alpha = \max \{f(x_1), f(x_2)\} \). By assumption, \( S_\alpha \) is convex, so that \( x \in S_\alpha \). Therefore, \( f(x) \leq \alpha = \max \{f(x_1), f(x_2)\} \). Hence, \( f \) is quasiconvex, and the proof is complete.

We now give a result analogous to Theorem 3.4.6. Theorem 3.5.3 shows that the maximum of a quasiconvex function over a compact polyhedral set occurs at an extreme point.

3.5.3 Theorem

Let \( S \) be a nonempty compact polyhedral set in \( \mathbb{R}^n \), and let \( f: E_n \rightarrow \mathbb{R} \) be quasiconvex and continuous on \( S \). Consider the problem to maximize \( f(x) \) subject to \( x \in S \). Then there exists an optimal solution \( \hat{x} \) to the problem, where \( \hat{x} \) is an extreme point of \( S \).

Proof

Note that \( f \) is continuous on \( S \) and hence attains a maximum, say, at \( x' \in S \). If there is an extreme point whose objective is equal to \( f(x') \), then the result is at hand. Otherwise, let \( x_1, \ldots, x_k \) be the extreme points of \( S \), and assume that \( f(x') > f(x_j) \) for \( j = 1, \ldots, k \). By Theorem 2.5.7, \( x' \) can be represented as

\[
x' = \sum_{j=1}^{k} \lambda_j x_j
\]

\[
\sum_{j=1}^{k} \lambda_j = 1
\]

Since \( f(x') > f(x_j) \) for each \( j \), then

\[
f(x') > \max_{1 \leq j \leq k} f(x_j) = \alpha
\]

(3.21)

Now consider the set \( S_\alpha = \{x: f(x) \leq \alpha\} \). Note that \( x_j \in S_\alpha \) for \( j = 1, \ldots, k \) and by quasiconvexity of \( f, S_\alpha \) is convex. Hence, \( x' = \sum_{j=1}^{k} \lambda_j x_j \) belongs to \( S_\alpha \). This implies that \( f(x') \leq \alpha \), which contradicts (3.21). This contradiction shows that \( f(x') = f(x_j) \) for some extreme point \( x_j \), and the proof is complete.

Differentiable Quasiconvex Functions

The following theorem gives a necessary and sufficient characterization of a differentiable quasiconvex function.
3.5.4 Theorem

Let $S$ be a nonempty open convex set in $E_n$, and let $f: S \to E_1$ be differentiable on $S$. Then $f$ is quasiconvex if and only if either one of the following equivalent statements hold:

1. If $x_1, x_2 \in S$ and $f(x_1) \leq f(x_2)$, then $\nabla f(x_2)'(x_1 - x_2) \leq 0$.
2. If $x_1, x_2 \in S$ and $\nabla f(x_2)'(x_1 - x_2) > 0$, then $f(x_1) < f(x_2)$.

Proof

Obviously statements (1) and (2) are equivalent. We shall prove part (1). Let $f$ be quasiconvex, and let $x_1, x_2 \in S$ be such that $f(x_1) \leq f(x_2)$. By differentiability of $f$ at $x_2$, for $\lambda \in (0, 1)$, we have

$$f(\lambda x_1 + (1 - \lambda)x_2) - f(x_2) = \lambda \nabla f(x_2)'(x_1 - x_2) + \lambda \|x_1 - x_2\| \alpha(x_2; \lambda(x_1 - x_2)) \leq 0$$

where $\alpha(x_2; \lambda(x_1 - x_2)) \to 0$ as $\lambda \to 0$. By quasiconvexity of $f$, we have $f(\lambda x_1 + (1 - \lambda)x_2) \leq f(x_2)$, and hence the above equation implies that $\lambda \nabla f(x_2)'(x_1 - x_2) + \lambda \|x_1 - x_2\| \alpha(x_2; \lambda(x_1 - x_2)) \leq 0$.

Dividing by $\lambda$ and letting $\lambda \to 0$, we get $\nabla f(x_2)'(x_1 - x_2) \leq 0$.

Conversely, suppose that $x_1, x_2 \in S$ be such that $f(x_1) \leq f(x_2)$. We need to show that $f(\lambda x_1 + (1 - \lambda)x_2) \leq f(x_2)$ for each $\lambda \in (0, 1)$. We do this by showing that

$$L = \{x : x = \lambda x_1 + (1 - \lambda)x_2, \lambda \in (0, 1), f(x) > f(x_2)\}$$

is empty. By contradiction, suppose that there exists an $x' \in L$. Therefore $x' = \lambda x_1 + (1 - \lambda)x_2$ for some $\lambda \in (0, 1)$ and $f(x') > f(x_2)$. Since $f$ is differentiable, it is continuous, and there must exist a $\delta \in (0, 1)$ such that

$$f(\mu x' + (1 - \mu)x_2) > f(x_2) \quad \text{for each } \mu \in [\delta, 1]$$

and $f(x') > f(\delta x' + (1 - \delta)x_2)$. By this inequality and the mean value theorem, we must have

$$0 < f(x') - f(\delta x' + (1 - \delta)x_2) = (1 - \delta)\nabla f(\hat{x})'(x' - x_2)$$

where $\hat{x} = \mu x' + (1 - \mu)x_2$ for some $\mu \in (\delta, 1)$. From (3.22), it is clear that $f(\hat{x}) > f(x_2)$. Dividing (3.23) by $1 - \delta > 0$, it follows that $\nabla f(\hat{x})'(x' - x_2) > 0$, which in turn implies that

$$\nabla f(\hat{x})'(x_1 - x_2) > 0$$

But on the other hand, $f(\hat{x}) > f(x_2) \geq f(x_1)$, and $\hat{x}$ is a convex combination of $x_1$ and $x_2$, say, $\hat{x} = \lambda x_1 + (1 - \lambda)x_2$ where $\lambda \in (0, 1)$. By the assumption of the theorem, $\nabla f(\hat{x})'(x_1 - \hat{x}) \leq 0$, and thus we must have

$$0 \geq \nabla f(\hat{x})'(x_1 - \hat{x}) = (1 - \lambda)\nabla f(\hat{x})'(x_1 - x_2)$$

The above inequality is not compatible with (3.24). Therefore $L$ is empty, and the proof is complete.

To illustrate the above theorem, let $f(x) = x^3$. To check quasiconvexity, suppose that $f(x_1) \leq f(x_2)$, that is, $x_1^3 \leq x_2^3$. This is true only if $x_1 \leq x_2$. Now consider $\nabla f(x_2)(x_1 - x_2) = 3(x_1 - x_2)x_2^2$. Since $x_1 \leq x_2$, $3(x_1 - x_2)x_2^2 \leq 0$. Therefore, $f(x_1) \leq f(x_2)$ implies that $\nabla f(x_2)(x_1 - x_2) \leq 0$, and by the theorem, $f$ is quasiconvex. As another illustration, let $f(x_1, x_2) = x_1^3 + x_2^3$. Let $x_1 = (2, -2)'$ and $x_2 = (1, 0)'$. Note that $f(x_1) = 0$ and $f(x_2) = 1$, so that $f(x_1) < f(x_2)$. But, on the other hand, $\nabla f(x_2)'(x_1 - x_2) = (3, 0)'(1, -2) = 3$. By the necessary part of the theorem, $f$ is not quasiconvex. This also shows that the sum of two quasiconvex functions is not necessarily quasiconvex.

### Strictly Quasiconvex Functions

Strictly quasiconvex and strictly quasiconcave functions are especially important in nonlinear programming because they ensure that a local minimum and a local maximum over a convex set are respectively a global minimum and a global maximum.

3.5.5 Definition

Let $f: S \to E_1$, where $S$ is a nonempty convex set in $E_n$. The function $f$ is said to be strictly quasiconvex if for each $x_1, x_2 \in S$ with $f(x_1) \neq f(x_2)$, we have

$$f(\lambda x_1 + (1 - \lambda)x_2) < \max \{f(x_1), f(x_2)\} \quad \text{for each } \lambda \in (0, 1).$$

The function $f$ is called strictly quasiconcave if $-f$ is strictly quasiconvex.

Note from the above definition that every convex function is strictly quasiconvex. Figure 3.7 below gives an example of a strictly quasiconvex function and an example of a strictly quasiconcave function.
The following theorem shows that a local minimum of a strictly quasiconvex function over a convex set is also a global minimum. This property is not enjoyed by quasiconvex functions as seen from Figure 3.6a.

### 3.5.6 Theorem

Let \( f: E_n \to E_1 \) be strictly quasiconvex. Consider the problem to minimize \( f(x) \) subject to \( x \in S \), where \( S \) is a nonempty convex set in \( E_n \). If \( \bar{x} \) is a local optimal solution, then \( \bar{x} \) is also a global optimal solution.

**Proof**

Assume, on the contrary, that there exists an \( \bar{x} \in S \) with \( f(\bar{x}) < f(\bar{x}) \). By convexity of \( S \), \( \lambda \bar{x} + (1 - \lambda)\bar{x} \in S \) for each \( \lambda \in (0, 1) \). Since \( \bar{x} \) is a local minimum by assumption, then \( f(\bar{x}) = f(\lambda \bar{x} + (1 - \lambda)\bar{x}) \) for all \( \lambda \in (0, \delta) \) and for some \( \delta \in (0, 1) \). But since \( f \) is strictly quasiconvex, and \( f(\bar{x}) < f(\bar{x}) \), then \( f(\lambda \bar{x} + (1 - \lambda)\bar{x}) < f(\bar{x}) \) for each \( \lambda \in (0, 1) \). This contradiction shows that such an \( \bar{x} \) does not exist, and the proof is complete.

As seen from Definition 3.1.1, every strictly convex function is indeed a quasiconvex function. But every strictly quasiconvex function is not quasiconvex. To illustrate, consider the following function given by Karamardian [1967].

\[
f(x) = \begin{cases} 
1 & \text{if } x = 0 \\
0 & \text{if } x \neq 0 
\end{cases}
\]

By Definition 3.5.5, \( f \) is strictly quasiconvex. However, \( f \) is not quasiconvex, since for \( x_1 = 1 \) and \( x_2 = -1 \), \( f(x_1) = f(x_2) = 0 \), but \( f(\frac{1}{2}x_1 + \frac{1}{2}x_2) = f(0) = 1 > f(x_2) \). If \( f \) is lower semicontinuous, then \( f \) is quasiconvex.

### 3.5.7 Lemma

Let \( S \) be a nonempty convex set in \( E_n \), and let \( f: S \to E_1 \) be strictly quasiconvex and lower semicontinuous. Then \( f \) is quasiconvex.

**Proof**

Let \( x_1 \) and \( x_2 \in S \). If \( f(x_1) \neq f(x_2) \), then by strict quasiconvexity of \( f \), we must have \( f(\lambda x_1 + (1 - \lambda) x_2) < \max \{ f(x_1), f(x_2) \} \) for each \( \lambda \in (0, 1) \). Now suppose that \( f(x_1) = f(x_2) \). To show that \( f \) is quasiconvex, we need to show that \( f(\lambda x_1 + (1 - \lambda) x_2) \leq f(x) \) for each \( \lambda \in (0, 1) \). By contradiction, suppose that \( f(\mu x_1 + (1 - \mu) x_2) > f(x) \) for some \( \mu \in (0, 1) \). Denote \( x_1 + (1 - \mu)x_2 \) by \( x \). Since \( f \) is lower semicontinuous, there exists a \( \lambda \in (0, 1) \) such that

\[
f(x) > f(\lambda x_1 + (1 - \lambda)x) > f(x_1) = f(x_2)
\]

(3.25)

Note that \( x \) can be represented as a convex combination of \( \lambda x_1 + (1 - \lambda)x \) and \( x_2 \). Hence by strict quasiconvexity of \( f \) and since \( f(\lambda x_1 + (1 - \lambda)x) > f(x_2) \), then \( f(x) < f(\lambda x_1 + (1 - \lambda)x) \), contradicting (3.25). This completes the proof.

### Strongly Quasiconvex Functions

From Theorem 3.5.6 it followed that a local minimum of a strictly quasiconvex function over a convex set is also a global optimal solution. However, strict quasiconvexity does not assert uniqueness of the global optimal solution. We shall define here another version of quasiconvexity, called strong quasiconvexity, which assures uniqueness of the global minimum.

#### 3.5.8 Definition

Let \( S \) be a nonempty convex set in \( E_n \), and let \( f: S \to E_1 \). The function \( f \) is said to be **strongly quasiconvex** if for each \( x_1, x_2 \in S \) with \( x_1 \neq x_2 \), we have

\[
f(\lambda x_1 + (1 - \lambda)x_2) < \max \{ f(x_1), f(x_2) \}
\]

for each \( \lambda \in (0, 1) \). The function \( f \) is said to be **strongly quasiconcave** if \(-f\) is strongly quasiconvex.

From Definition 3.5.8 above, and from Definitions 3.1.1, 3.5.1, and 3.5.5 the following statements hold:

1. Every strictly convex function is strongly quasiconvex.
2. Every strongly quasiconvex function is strictly quasiconvex.
3. Every strongly quasiconvex function is quasiconvex even in the absence of semicontinuity assumption.

Figure 3.7a illustrates a case where the function is both strongly quasiconvex and strictly quasiconvex, whereas the function represented in Figure 3.7b is strictly quasiconvex but not strongly quasiconvex.

### 3.5.9 Theorem

Let \( f: E_n \to E_1 \) be strongly quasiconvex. Consider the problem to minimize \( f(x) \) subject to \( x \in S \), where \( S \) is a nonempty convex set in \( E_n \). If \( \bar{x} \) is a local optimal solution, then \( \bar{x} \) is the unique global optimal solution.
Proof
Since \( \bar{x} \) is a local optimal solution, then there exists an \( \varepsilon \)-neighborhood \( N_{\varepsilon}(x) \) around \( \bar{x} \) such that \( f(\bar{x}) \leq f(x) \) for all \( x \in \mathbb{S} \cap N_{\varepsilon}(x) \). Suppose by contradiction to the conclusion of the theorem that there exists a point \( x \in \mathbb{S} \) such that \( f(\bar{x}) \neq f(x) \) and \( f(\bar{x}) \leq f(x) \). By strong quasiconvexity, it follows that

\[
[(1-\lambda)x] \leq \max \{ f(\bar{x}), f(x) \} = f(\bar{x})
\]

for all \( \lambda \in (0, 1) \). But for \( \lambda \) small enough, \( Ax + (1-A)x \in \mathbb{S} \) so that the above inequality violates local optimality of \( \bar{x} \). This completes the proof.

### Pseudoconvex Functions

From the definition of a pseudoconvex function given below, the reader can easily verify that if \( \nabla f(\bar{x}) = 0 \), then \( \bar{x} \) is a global minimum of \( f \). This property is not shared by differentiable strongly or strictly quasiconvex functions as shown in Figure 3.8b.

**3.5.10 Definition**

Let \( \mathbb{S} \) be a nonempty open set in \( E_{n} \), and let \( f: \mathbb{S} \rightarrow E_{1} \) be differentiable on \( \mathbb{S} \). The function \( f \) is said to be pseudoconvex if for each \( x_{1}, x \in \mathbb{S} \) with \( \nabla f(x_{1})'(x_{2} - x_{1}) \geq 0 \) we have \( f(x_{2}) \geq f(x_{1}) \), or equivalently, if \( f(x_{2}) < f(x_{1}) \) then \( \nabla f(x_{1})'(x_{2} - x_{1}) < 0 \). The function \( f \) is said to be pseudoconcave if \(-f\) is pseudoconvex.

The function \( f \) is said to be strictly pseudoconvex if for each distinct \( x_{1}, x_{2} \in \mathbb{S} \) satisfying \( \nabla f(x_{1})'(x_{2} - x_{1}) \geq 0 \) we have \( f(x_{2}) > f(x_{1}) \), or equivalently, if for each distinct \( x_{1}, x_{2} \in \mathbb{S} \), \( f(x_{2}) < f(x_{1}) \) implies that \( \nabla f(x_{1})'(x_{2} - x_{1}) < 0 \). The function \( f \) is said to be strictly pseudoconcave if \(-f\) is strictly pseudoconvex.

Figure 3.8a illustrates a pseudoconvex function. This is also strictly quasi-convex, which is true in general as shown by Theorem 3.5.11 below. The reader may note that the figure in 3.8b is not pseudoconvex but is strictly quasi-convex.

**3.5.11 Theorem**

Let \( \mathbb{S} \) be a nonempty open convex set in \( E_{n} \), and let \( f: \mathbb{S} \rightarrow E_{1} \) be a differentiable pseudoconvex function on \( \mathbb{S} \). Then \( f \) is both strictly quasi-convex and quasi-convex.

**Proof**
We first show that \( f \) is strictly quasi-convex. By contradiction, suppose that there exists \( x_{1}, x_{2} \in \mathbb{S} \) such that \( f(x_{1}) \neq f(x_{2}) \) and \( f(x_{2}) \geq \max \{ f(x_{1}), f(x_{2}) \} \), where \( x = \lambda x_{1} + (1-\lambda)x_{2} \) for some \( \lambda \in (0, 1) \). Without loss of generality, assume that \( f(x_{1}) < f(x_{2}) \), so that

\[
f(x_{2}) = f(x_{2}) > f(x_{1})
\]

(3.26)

Note by pseudoconvexity of \( f \) that \( \nabla f(x'_{1})(x_{2} - x'_{1}) < 0 \). Now, since \( \nabla f(x_{1})'(x_{2} - x_{1}) < 0 \) and \( x_{2} - x_{1} = - (1-\lambda)(x_{2} - x_{1})/\lambda \), then \( \nabla f(x_{1})'(x_{2} - x_{1}) > 0 \), and hence by pseudoconvexity of \( f \), we must have \( f(x_{2}) \geq f(x_{1}) \). Therefore, by (3.26), we get \( f(x_{2}) = f(x_{1}) \). Also, since \( \nabla f(x_{1})'(x_{2} - x_{2}) > 0 \), there exists a point \( \xi = \mu x + (1-\mu)x_{2} \) with \( \mu \in (0, 1) \) such that

\[
f(x_{1}) > f(x_{2})
\]

Again, by pseudoconvexity of \( f \) we have \( \nabla f(\xi'(x_{2} - \xi)) < 0 \). Similarly, \( \nabla f(\xi'(x_{2} - \xi)) < 0 \). Summarizing, we must have

\[
\nabla f(\xi'(x_{2} - \xi)) < 0
\]

Note that \( x_{2} - \xi = \mu(x_{2} - x_{2})/(1-\mu) \), and hence above two inequalities are not compatible. This contradiction shows that \( f \) is strictly quasi-convex. By Lemma 3.5.7, then, \( f \) is also quasi-convex, and the proof is complete.

In Theorem 3.5.12 below we see that every strictly pseudoconvex function is strongly quasi-convex.

**3.5.12 Theorem**

Let \( \mathbb{S} \) be a nonempty open convex set in \( E_{n} \), and let \( f: \mathbb{S} \rightarrow E_{1} \) be a differentiable strictly pseudoconvex function on \( \mathbb{S} \). Then \( f \) is strongly quasi-convex.

**Proof**
By contradiction suppose that there exist distinct \( x_{1}, x_{2} \in \mathbb{S} \) and \( \lambda \in (0, 1) \) such that \( f(x) \geq \max \{ f(x_{1}), f(x_{2}) \} \), where \( x = \lambda x_{1} + (1-\lambda)x_{2} \). Since \( f(x_{2}) = f(x_{1}) \), we get \( f(x_{2}) = f(x_{1}) = f(x) \), contradicting the strictness of the pseudoconvexity of \( f \). This completes the proof.
by strict pseudoconvexity of \( f \), then \( \nabla f(x)'(x_1 - x) < 0 \) and, hence,

\[
\nabla f(x)'(x_1 - x_2) < 0 \tag{3.27}
\]

Likewise, since \( f(x_2) \leq f(x) \), then

\[
\nabla f(x)'(x_2 - x_1) < 0 \tag{3.28}
\]

The two inequalities (3.27) and (3.28) are not compatible, and hence \( f \) is strongly quasiconvex.

Convexity at a Point

Another useful concept in optimization is the notion of convexity or concavity at a point. In some cases, the requirement of a convex or concave function may be too strong and really not essential. Instead, convexity or concavity at a point may be all that is needed.

3.5.13 Definition

Let \( S \) be a nonempty convex set in \( E^m \) and \( f: S \rightarrow E_1 \). The following are relaxations of various forms of convexity presented in this chapter.

**Convexity at \( \bar{x} \).** The function \( f \) is said to be convex at \( \bar{x} \in S \) if

\[
f(\lambda \bar{x} + (1 - \lambda)x) \leq \lambda f(\bar{x}) + (1 - \lambda)f(x)
\]

for each \( \lambda \in (0, 1) \) and each \( x \in S \).

**Strict convexity at \( \bar{x} \).** The function \( f \) is said to be strictly convex at \( \bar{x} \in S \) if

\[
f(\lambda \bar{x} + (1 - \lambda)x) < \lambda f(\bar{x}) + (1 - \lambda)f(x)
\]

for each \( \lambda \in (0, 1) \) and for each \( x \in S \), \( x \neq \bar{x} \).

**Quasiconvexity at \( \bar{x} \).** The function \( f \) is said to be quasiconvex at \( \bar{x} \in S \) if

\[
f(\lambda \bar{x} + (1 - \lambda)x) \leq \max \{f(x), f(\bar{x})\}
\]

for each \( \lambda \in (0, 1) \) and each \( x \in S \).

**Strict quasiconvexity at \( \bar{x} \).** The function \( f \) is said to be strictly quasiconvex at \( \bar{x} \in S \) if

\[
f(\lambda \bar{x} + (1 - \lambda)x) < \max \{f(x), f(\bar{x})\}
\]

for each \( \lambda \in (0, 1) \) and each \( x \in S \), \( x \neq \bar{x} \).

**Strong quasiconvexity at \( \bar{x} \).** The function \( f \) is said to be strongly quasiconvex at \( \bar{x} \in S \) if

\[
f(\lambda \bar{x} + (1 - \lambda)x) < \max \{f(x), f(\bar{x})\}
\]

for each \( \lambda \in (0, 1) \) and each \( x \in S \), \( x \neq \bar{x} \).

**Pseudoconvexity at \( \bar{x} \).** The function \( f \) is said to be pseudoconvex at \( \bar{x} \in S \) if \( \nabla f(\bar{x})'(x - \bar{x}) \geq 0 \) for \( x \in S \) implies that \( f(x) \geq f(\bar{x}) \).

**Strict pseudoconvexity at \( \bar{x} \).** The function \( f \) is said to be strictly pseudoconvex at \( \bar{x} \in S \) if \( \nabla f(\bar{x})'(x - \bar{x}) > 0 \) for \( x \in S \), \( x \neq \bar{x} \), implies that \( f(x) > f(\bar{x}) \).

Various types of convexity at a point can be stated in a similar fashion. Figure 3.10 shows some types of convexity at a point. As the figure suggests, these types of convexity at a point represent a significant relaxation of the concept of convexity.

Figure 3.9 Relationship among various types of convexity.

Figure 3.10 Types of convexity at a point.
We shall specify below some important results related to convexity of a function $f$ at a point, where $f: S \to E_1$ and $S$ is a nonempty convex set in $E_n$. Of course, not all the results we developed throughout this chapter will hold true. However, several of these results hold true, and are summarized below. The proofs are similar to the corresponding theorems in this chapter.
Exercises

3.1 Let $S$ be a nonempty convex set in $E_n$, and let $f: S \rightarrow E_1$. Show that $f$ is concave if and only if $-f$ is convex.

3.2 Let $S$ be a nonempty convex set in $E_n$, and let $f: S \rightarrow E_1$. Show that $f$ is convex if and only if for any integer $k \geq 2$, the following holds true: $x_1, \ldots, x_k \in S$ implies that $f(x_1) + \ldots + f(x_k) \leq kf\left(\frac{x_1 + \ldots + x_k}{k}\right)$, where $f(0) = 0$ for $f = 1, \ldots, k$.

3.3 Which of the following functions are convex, concave, or neither? Why?
   a. $f(x_1, x_2) = x_1^2 + 2x_1x_2 - 10x_1 + 5x_2$
   b. $f(x_1, x_2) = x_1e^{-x_2}$
   c. $f(x_1, x_2) = -x_1^2 - 5x_2^2 + 2x_1x_2 + 10x_1 - 10x_2$
   d. $f(x_1, x_2, x_3) = x_1x_2 + x_1^2 + 2x_1^2 - 6x_1x_3$

3.4 Show that a function $f: E_1 \rightarrow E_1$ is affine if and only if $f$ is both convex and concave. (A function $f$ is affine if it is of the form $f(x) = ax + c$, where $a$ is a scalar and $c$ is a given vector.)

3.5 Prove or disprove concavity of the following function defined over $S = \{(x_1, x_2): -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1\}$
   
   $$f(x_1, x_2) = 10 - 2(x_1 - x_1^2)^2$$

3.6 Let $F$ be a cumulative distribution function for a random variable $b$, that is, $F(y) = \text{Prob}(b \leq y)$. Show that $f(z) = F'-F(y)$ dy is a convex function. Is $f$ convex for any nondecreasing function $F$?

3.7 Let $g: E_1 \rightarrow E_1$ be a concave function, and let $f$ be defined by $f(x) = 1/g(x)$. Show that $f$ is convex over $S = \{x: g(x) > 0\}$.

3.8 Let $f_1, f_2, \ldots, f_k : E_1 \rightarrow E_1$ be convex functions. Consider the function $f$ defined by $f(x) = \max\{f_1(x), f_2(x), \ldots, f_k(x)\}$. Show that $f$ is convex. State and prove a similar result for concave functions.

3.9 Let $f_1, f_2, \ldots, f_k : E_1 \rightarrow E_1$ be convex functions. Consider the function $f$ defined by $f(x) = \sum_{i=1}^{k} a_i f_i(x)$, where $a_i > 0$ for $i = 1, 2, \ldots, k$. Show that $f$ is convex. State and prove a similar result for concave functions.

3.10 Let $S$ be a nonempty convex set in $E_n$, and let $f: E_n \rightarrow E_1$ be defined as follows:
   
   $$f(y) = \inf\{\|y-x\|: x \in S\}$$

Note that $f(y)$ gives the distance from $y$ to the set $S$ and is called the distance function. Prove that $f$ is convex.

3.11 Let $S = \{(x_1, x_2): x_1^2 + x_2^2 \leq 1\}$. Let $f$ be the distance function defined in Exercise 3.10. Find the function $f$ explicitly.

3.12 Let $f: E_n \rightarrow E_1$ be a convex function, and let $g: E_1 \rightarrow E_1$ be a nondecreasing convex function. Consider the composite function $h: E_n \rightarrow E_1$ defined by $h(x) = g(f(x))$. Show that $h$ is convex.

3.13 Let $S$ be a nonempty, bounded convex set in $E_n$, and let $f: E_n \rightarrow E_1$ be defined as follows:
   
   $$f(y) = \sup\{\langle y, x \rangle : x \in S\}$$

The function $f$ is called the support function of $S$. Prove that $f$ is convex. Also show that if $f(y) = y^T\bar{x}$, where $\bar{x} \in S$, then $\bar{x}$ is a subgradient of $f$ at $y$.

3.14 Let $S = A \cup B$, where
   
   $$A = \{(x_1, x_2): x_1 < 0, x_1^2 + x_2^2 \leq 1\}$$
   $$B = \{(x_1, x_2): x_1 \geq 0, -1 \leq x_2 \leq 1\}$$

Find the support function defined in Exercise 3.13 explicitly.

3.15 A function $f: E_n \rightarrow E_1$ is called a gauge function if it satisfies the following equality:
   
   $$f(\lambda x) = \lambda f(x)$$

for all $x \in E_n$ and all $\lambda \geq 0$

Further, a gauge function is said to be subadditive if it satisfies the following inequality:
   
   $$f(x) + f(y) \leq f(x+y)$$

for all $x, y \in E_n$.

Prove that subadditivity is equivalent to convexity of gauge functions.

3.16 Let $f: E_1 \rightarrow E_1$ be convex. Show that $\xi$ is a subgradient of $f$ at $x$ if and only if the hyperplane $\{(x, y): y = f(x) + \xi(x-x')\}$ supports $f$ at $[x, f(x)]$. State and prove a similar result for concave functions.

3.17 Let $f$ be a convex function on $E_n$. Prove that the set of subgradients of $f$ at a given point forms a closed convex set.

3.18 Consider the function $\theta$ defined by the following optimization problem:
   
   $$\theta(u_1, u_2) = \text{Minimum} \quad x_1(1-u_1) + x_2(1-u_2)$$

subject to $x_1^2 + x_2^2 \leq 1$

a. Show that $\theta$ is concave.
   b. Evaluate $\theta$ at the point $(1, 1)$.
   c. Find the collection of subgradients of $\theta$ at $(1, 1)$.

3.19 Let $f: E_1 \rightarrow E_1$ be defined by $f(x) = \|x\|$. Prove that subgradients of $f$ are characterized as follows:
   
   If $x = 0$, then $\xi$ is a subgradient of $f$ at $x$ if and only if $\|\xi\| \leq 1$; on the other hand, if $x \neq 0$, then $\xi$ is a subgradient of $f$ at $x$ if and only if $\|\xi\| = 1$ and $\xi^T x = \|x\|$.

Use this result to show that $f$ is differentiable at each $x \neq 0$, and characterize the gradient vector.

3.20 Let $f_1, f_2: E_1 \rightarrow E_1$ be differentiable convex functions. Consider the function $f$ defined by $f(x) = \max\{f_1(x), f_2(x)\}$. Let $\xi$ be such that $f(x) = f_1(\xi) = f_2(\xi)$. Show that $\xi$ is a subgradient of $f$ at $\xi$ if and only if $\xi = \lambda_1\nabla f_1(\xi) + (1-\lambda)\nabla f_2(\xi)$ where $\lambda \in [0, 1]$.

Generalize the result to several convex functions and state a similar result for concave functions.

3.21 Consider the function $\theta$ defined by the following optimization problem, where $X$ is a compact polyhedral set.
   
   $$\theta(u) = \text{Minimum} \quad c^T x + u^T(Ax - b)$$

subject to $x \in X$

a. Show that $\theta$ is concave.
   b. Characterize the subgradients of $\theta$ at any given $u$.

3.22 In reference to Exercise 3.21, find the function $\theta$ explicitly and describe the set of subgradients at each point $u \geq 0$ if:
   
   $$A = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad c = \begin{bmatrix} -1 \\ -3 \end{bmatrix}$$

$$X = \{(x_1, x_2): 0 \leq x_1, 2, 0 \leq x_2 \leq 3\}$$
Convex Functions

3.23 Let \( f: E_1 \rightarrow E_1 \) be a differentiable function. Show that the gradient vector is given by

\[
\nabla f(x) = \left( \frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n} \right)
\]

3.24 Let \( f: E_1 \rightarrow E_1 \) be a differentiable function. The linear approximation of \( f \) at a given point \( \bar{x} \) is given by

\[
f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x})
\]

If \( f \) is twice differentiable at \( \bar{x} \), then the quadratic approximation of \( f \) at \( \bar{x} \) is given by

\[
f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x}) + \frac{1}{2}(x - \bar{x})^T H(\bar{x})(x - \bar{x})
\]

Let \((x_1, x_2) = e^{x_1^2+x_2^2-5x_1+10x_2} \). Give the linear and quadratic approximations of \( f \) at \((0,1)\). Are these approximations convex, concave, or neither? Why?

3.25 Consider the following problem:

Minimize \((x_1 - 4)^2 + (x_2 - 6)^2\)

subject to \(x_2 \geq x_1^2\)
\(x_2 \leq 4\)

Write a necessary condition for optimality and verify that it is satisfied by the point \((2,4)\). Is this the optimal point? Why?

3.26 Use Theorem 3.4.3 to prove that every local minimum of a convex function over a convex set is also a global minimum.

3.27 Consider the following problem:

Maximize \(c^T x + \frac{1}{2}x^T H x\)

subject to \(Ax \leq b\)
\(x \geq 0\)

where \( H \) is a symmetric negative definite matrix, \( A \) is an \( m \times n \) matrix, \( c \) is an \( n \) vector, and \( b \) is an \( m \) vector. Write the necessary and sufficient condition for optimality of Theorem 3.4.3, and simplify it using the special structure of this problem.

3.28 Consider the problem to minimize \( f(x) \) subject to \( x \in S \), where \( f: E_n \rightarrow E_1 \) is a differentiable convex function, and \( S \) is a nonempty convex set in \( E_n \). Prove that \( \bar{x} \) is an optimal solution if and only if \( \nabla f(\bar{x})^T(x - \bar{x}) \geq 0 \) for each \( x \in S \). Also state and prove a similar result for the maximization of a concave function.

(This result was proved in the text as Corollary 2 to Theorem 3.4.3. In this exercise, the reader is asked to give a direct proof without resort to subgradients.)

3.29 Let \( f: E_1 \rightarrow E_1 \) be a convex function, and suppose that \( f(x + \lambda d) \leq f(x) \) for all \( \lambda \in (0, \delta) \), where \( \delta > 0 \). Show that \( f(x + \lambda d) \) is a nondecreasing function of \( \lambda \). In particular, show that \( f(x + \lambda d) \) is a strictly increasing function of \( \lambda \) if \( f \) is strictly convex.

3.30 A vector \( d \) is called a direction of descent of \( f \) at \( \bar{x} \) if there exists a \( \delta > 0 \) such that \( f(\bar{x} + \lambda d) < f(\bar{x}) \) for each \( \lambda \in (0, \delta) \). Suppose that \( f \) is convex. Show that \( d \) is a direction of descent if and only if \( f(\bar{x}; d) < 0 \). Does the result hold true without convexity of \( f \)?

3.31 Consider the problem to minimize \( f(x) \) subject to \( x \in S \), where \( f: E_n \rightarrow E_1 \) is convex, and \( S \) is a nonempty convex set in \( E_n \). The cone of feasible directions of \( S \) at \( \bar{x} \in S \) is defined by

\[
D = \{ d : \text{there exists a } \delta > 0 \text{ such that } \bar{x} + \lambda d \in S \text{ for } \lambda \in (0, \delta) \}
\]

Show that \( \bar{x} \) is an optimal solution of the problem if and only if \( f(\bar{x}; d) \geq 0 \) for each \( d \in D \). Compare this result with the necessary and sufficient condition of Theorem 3.4.3. Specialize the result to the case where \( S = E_n \).

3.32 Consider the following problem:

Maximize \( f(x) \)

subject to \( Ax = b \)
\(x \geq 0\)

where \( A \) is an \( m \times n \) matrix with rank \( m \), and \( f \) is a differentiable convex function. Consider the extreme point \((x_0, x_0) = (\bar{b}, \bar{d})\), where \( \bar{b} = B^{-1} \bar{b} \geq 0 \) and \( A = [B, N] \). Decompose \( \nabla f(x) \) accordingly into \( \nabla f_b(x) \) and \( \nabla f_N(x) \). Show that the necessary condition of Theorem 3.4.5 holds true if \( \nabla f_b(x) B^{-1} N \leq 0 \). If this condition holds, is it necessarily true that \( x \) is a local maximum? Prove or give a counterexample.

If \( \nabla f_b(x') - \nabla f_b(x) B^{-1} N \leq 0 \), choose a positive component \( j \) and increase its corresponding nonbasic variable \( x_j \) until a new extreme point is reached. Show that this process results in a new extreme point with a larger objective value. Does this method guarantee convergence to a global optimal solution? Prove or give a counterexample.

3.33 Apply the procedure of Exercise 3.32 to the following problem starting with the extreme point \((1, 3, 0, 0)\).

Maximize \((x_1 - \frac{3}{2})^2 + (x_2 - 5)^2\)

subject to \(-x_1 + x_2 + x_3 = 2\)
\(2x_1 + 3x_2 + x_4 = 11\)
\(x_1, x_2, x_3, x_4 \geq 0\)

3.34 Let \( e_1, e_2 \) be nonzero vectors in \( E_n \), and \( \alpha_1, \alpha_2 \) be scalars. Let \( S = \{ x : e_1^T x + \alpha_2 > 0 \} \). Consider the function \( f: S \rightarrow E_1 \) defined as follows:

\[
f(x) = \frac{c_1^T x + \alpha_1}{c_2^T x + \alpha_2}
\]

Show that \( f \) is both pseudoconvex and pseudoconcave. (Functions that are both pseudoconvex and pseudoconcave are called pseudolinear.)
Consider a quadratic function \( f: E \to E_1 \), and suppose that \( f \) is convex on \( S \), where \( S \) is a nonempty convex set in \( E_1 \). Show that:

a. The function \( f \) is convex on \( M(S) \), where \( M(S) \) is the affine manifold containing \( S \) defined by \( M(S) = \{ y : y = \sum_{k=1}^{n} \lambda_k x_k, \sum_{k=1}^{n} \lambda_k = 1, x_k \in S \text{ for all } k \geq 1 \} \).

b. The function \( f \) is convex on \( L(S) \), the linear subspace parallel to \( M(S) \), defined by \( L(S) = \{ y - x : y \in M(S) \text{ and } x \in S \} \).

This result is credited to Cottle [1967].

Consider the quadratic function \( f: E \to E_1 \) defined by \( f(x) = x' H x \). The function \( f \) is said to be positive subdefinite if \( x' H x \leq 0 \) for each \( x \in E_1 \). Prove that \( f \) is pseudoconvex on the nonnegative orthant, \( E_+ = \{ x \in E_1 : x \geq 0 \} \), if and only if it is positive subdefinite.

This result is credited to Martos [1969].

The function \( f \) defined in Exercise 3.36 is said to be strictly positive subdefinite if \( x' H x \leq 0 \) implies \( H x = 0 \) for each \( x \in E_1 \). Prove that \( f \) is pseudoconvex on the nonnegative orthant excluding \( x = 0 \) if and only if it is strictly positive subdefinite.

This result is credited to Martos [1969].

Let \( g: S \to E_1 \) and \( h: E_1 \to E_2 \), where \( S \) is a nonempty convex set in \( E_1 \). Consider the function \( f: S \to E_1 \) defined by \( f(x) = g(x)/h(x) \). Show that \( f \) is quasi-convex if the following two conditions hold true:

a. \( g \) is convex on \( S \) and \( g(x) \geq 0 \) for each \( x \in S \).

b. \( h \) is convex on \( S \) and \( h(x) > 0 \) for each \( x \in S \).

(Hint: Use Theorem 3.5.2.)

Consider the function \( f \) defined in Exercise 3.38 is quasi-convex if the following two conditions hold true:

a. \( g \) is convex on \( S \) and \( g(x) \leq 0 \) for each \( x \in S \).

b. \( h \) is convex on \( S \) and \( h(x) > 0 \) for each \( x \in S \).

Let \( g: S \to E_1 \) and \( h: S \to E_2 \), where \( S \) is a nonempty convex set in \( E_1 \). Consider the function \( f: S \to E_1 \) defined by \( f(x) = g(x)/h(x) \). Show that \( f \) is quasi-convex if the following two conditions hold true:

a. \( g \) is convex and \( g(x) \leq 0 \) for each \( x \in S \).

b. \( h \) is concave and \( h(x) > 0 \) for each \( x \in S \).

In each of the Exercises 3.38, 3.39, and 3.40, show that \( f \) is pseudoconvex provided that \( S \) is open and \( g \) and \( h \) are differentiable.

Let \( f: E_1 \to E_2 \) be convex, and let \( A \) be an \( m \times n \) matrix. Consider the function \( h: E_1 \to E_1 \) defined by \( h(x) = f(Ax) \). Show that \( h \) is convex.

Let \( f: E_1 \to E_2 \) and \( g: E_2 \to E_1 \) be differentiable and convex. Let \( \phi: E_2 \to E_1 \) satisfy the following inequality. If \( a_2 \geq a_1 \) and \( b_2 \geq b_1 \), then \( \phi(a_2,b_2) \geq \phi(a_1,b_1) \). Consider the function \( h: E_1 \to E_1 \) defined by \( h(x) = \phi(f(x),g(x)) \). Show the following:

a. If \( \phi \) is convex, then \( h \) is convex.

b. If \( \phi \) is pseudoconvex, then \( h \) is pseudoconvex.

c. If \( \phi \) is quasiconvex, then \( h \) is quasiconvex.

Let \( S \) be a nonempty convex set in \( E_1 \), and let \( f: E_1 \to E_1 \) and \( g: E_1 \to E_2 \) be convex functions. Consider the perturbation function \( \phi: E_2 \to E_1 \) defined below:

\[
\phi(y) = \inf \{ f(x) : g(x) \leq y, x \in S \}
\]

Show that \( \phi \) is convex.

Let \( f: E_1 \to E_1 \) be convex, and \( A \) be an \( m \times n \) matrix. Consider the function \( h: E_1 \to E_1 \) defined as follows:

\[
h(y) = \inf \{ f(x) : Ax = y \}
\]

Show that \( h \) is convex.

Let \( g_1, g_2: E_1 \to E_2 \), and let \( \alpha \in [0, 1] \). Consider the function \( g_\alpha: E_1 \to E_1 \) defined below:

\[
g_\alpha(x) = \frac{(1-\alpha)g_1(x) + \alpha g_2(x)}{1+\alpha} + \frac{\alpha g_2(x) - (1-\alpha)g_1(x)}{1+\alpha} - \alpha \frac{g_2(x) - g_1(x)}{1+\alpha}
\]

where 1 denotes the positive square root.

Show that \( g_\alpha \) is convex for \( \alpha \in [0, 1] \). Consider the function \( G_\alpha: E_1 \to E_1 \) defined by

\[
G_\alpha(x) = \frac{(1-\alpha)g_1(x) + \alpha g_2(x)}{1+\alpha} + \frac{\alpha g_2(x) - (1-\alpha)g_1(x)}{1+\alpha} - \alpha \frac{g_2(x) - g_1(x)}{1+\alpha}
\]

where 1 denotes the positive square root.

a. Show that \( G_\alpha(x) \geq 0 \) if and only if \( g_1(x) \geq 0 \) and \( g_2(x) \geq 0 \).

b. If \( g_1 \) and \( g_2 \) are differentiable, show that \( G_\alpha \) is differentiable at \( x \) for each \( \alpha \in (0, 1) \) provided that \( g_1(x), g_2(x) \neq 0 \).

c. Now suppose that \( g_1 \) and \( g_2 \) are concave. Show that \( G_\alpha \) is concave for \( \alpha \in [0, 1] \).

d. Suppose that \( g_1 \) and \( g_2 \) are quasiconcave. Show that \( G_\alpha \) is quasiconcave for \( \alpha = 1 \).

e. In some optimization problems, the restriction that the variable \( x = 0 \) or 1
In this chapter we deal with the important topic of convex and concave functions. The recognition of these functions is generally traced to Jensen [1905, 1906]. For earlier related works on the subject, see Hadamard [1893] and Hölder [1889].

In Section 3.1 several results related to continuity and directional derivatives of a convex function are presented. In particular, we showed that a convex function is continuous on the interior of the domain. See, for example, Rockafellar [1970]. In Section 3.2 we discuss subgradients of convex functions. Many of the properties of differentiable convex functions are retained by replacing the gradient vector by a subgradient. For this reason, subgradients have been used frequently in the optimization of nondifferentiable functions. See, for example, Held and Karp [1970], Held, Wolfe and Crowder [1974], and Wolfe [1976].

Section 3.3 gives some properties of differentiable convex functions. For further study of these topics as well as other properties of convex functions, refer to Eggleston [1958], Fenchel [1953], Roberts and Varberg [1973], and Rockafellar [1970].

Section 3.4 treats the subject of minima and maxima of convex functions over convex sets. For general functions, the study of minima and maxima is quite complicated. As shown in Section 3.4, however, every local minimum of a convex function over a convex set is also a global minimum, and the maximum of a convex function over a convex set occurs at an extreme point. For an excellent study of optimization of convex functions, see Rockafellar [1970].

In Section 3.5 we examine other classes of functions that are related to convex functions, namely quasiconvex and pseudoconvex functions. The class of quasiconvex functions was first studied by De Finetti [1949]. For further reading on this topic, refer to Fenchel [1953], Karamardian [1967], Mangasarian [1969], and Ponstein [1967]. For a good survey of quasiconvex and related functions, see the article of Greenberg and Pierskalla [1971]. Arrow and Enthoven [1961] derived necessary and sufficient conditions for quasiconvexity on the nonnegative orthant assuming twice differentiability. Their results were extended by Ferland [1972]. Note that a local minimum of a quasiconvex function over a convex set is not necessarily a global minimum. This result holds true, however, for a strictly quasiconvex function. Ponstein [1967] introduced the concept of strongly quasiconvex functions, which ensures that the global minimum is unique, a property that is not enjoyed by strictly quasiconvex functions. The notion of pseudoconvexity was first introduced by Mangasarian [1965]. The significance of the class of pseudoconvex functions stems from the fact that every point with a zero gradient is a global minimum. Matrix theoretic characterizations (see e.g., Exercises 3.36 and 3.37) of quadratic pseudoconvex and quasiconvex functions have been presented by Cottle and Ferland [1972] and by Martos [1965a, 1967b, 1969, 1975].