Proposition (Second-order sufficient conditions—unconstrained case). *Let $f \in C^2$ be a function defined on a region in which the point $\mathbf{x}^*$ is an interior point. Suppose in addition that*

i) $\nabla f(\mathbf{x}^*) = \mathbf{0}$ (7)

ii) $\mathbf{F}(\mathbf{x}^*)$ *is positive definite.* (8)

Then $\mathbf{x}^*$ is a strict relative minimum point of f.

oof. Since $\mathbf{F}(\mathbf{x}^*)$ is positive definite, there is an $a > 0$ such that for all $\mathbf{d}$, $(\mathbf{x}^*)\mathbf{d} \geqslant a|\mathbf{d}|^2$. Thus by the Taylor's Theorem (with remainder)

$$f(\mathbf{x}^* + \mathbf{d}) - f(\mathbf{x}^*) = \tfrac{1}{2}\mathbf{d}'\mathbf{F}(\mathbf{x}^*)\mathbf{d} + o(|\mathbf{d}|^2)$$
$$\geqslant (a/2)|\mathbf{d}|^2 + o(|\mathbf{d}|^2).$$

small $|\mathbf{d}|$ the first term on the right dominates the second, implying that h sides are positive for small $\mathbf{d}$. ∎

CONVEX AND CONCAVE FUNCTIONS

order to develop a theory directed toward characterizing global, rather n local, minimum points, it is necessary to introduce some sort of vexity assumptions. This results not only in a more potent, although e restrictive, theory but also provides an interesting geometric inter-ation of the second-order sufficiency result derived above.

Definition. A function f defined on a convex set Ω is said to be *convex* if, for every $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ and every α, $0 \leqslant \alpha \leqslant 1$, there holds

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leqslant \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2).$$

If, for every $0 < \alpha < 1$ and $\mathbf{x}_1 \neq \mathbf{x}_2$, there holds

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) < \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2),$$

then f is said to be *strictly convex.*

Several examples of convex or nonconvex functions are shown in Fig. Geometrically, a function is convex if the line joining two points on its h lies nowhere below the graph, as shown in Fig. 6.2a, or, thinking of a tion in two dimensions, it is convex if its graph is bowl shaped. Next we turn to the definition of a concave function.

Definition. A function g defined on a convex set Ω is said to be *concave* f the function $f = -g$ is convex. The function g is *strictly concave* if $-g$ is strictly convex.

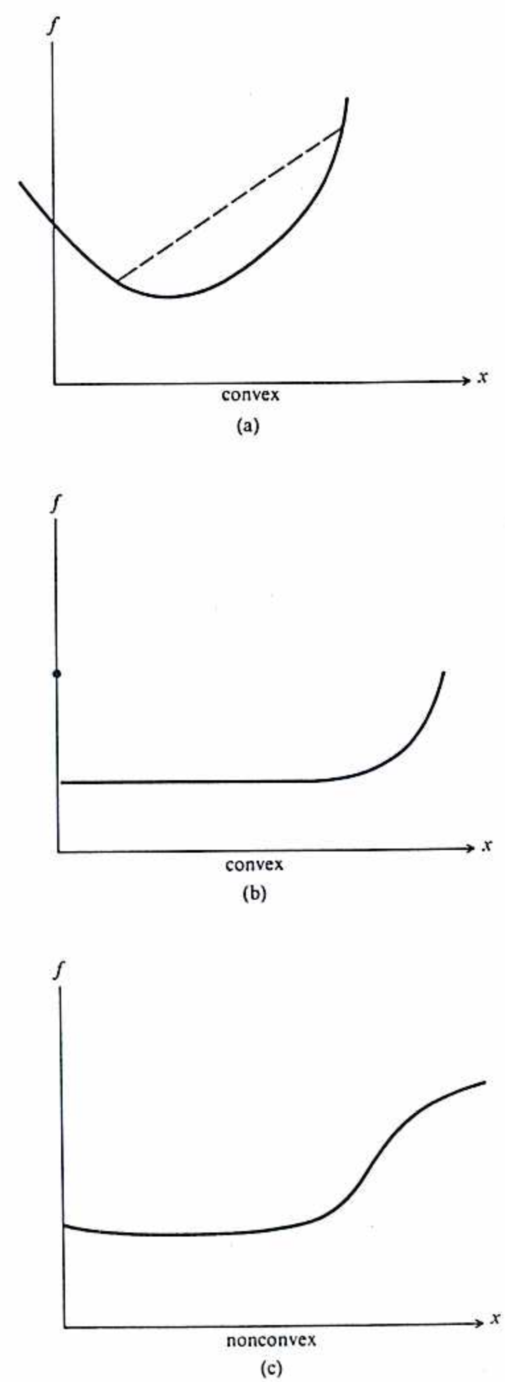

Fig. 6.2 Convex and nonconvex functions

Combinations of Convex Functions

We show that convex functions can be combined to yield new convex functions and that convex functions when used as constraints yield convex constraint sets.

Proposition 1. *Let f_1 and f_2 be convex functions on the convex set Ω. Then the function $f_1 + f_2$ is convex on Ω.*

Proof. Let $x_1, x_2 \in \Omega$, and $0 < \alpha < 1$. Then

$$f_1(\alpha x_1 + (1-\alpha)x_2) + f_2(\alpha x_1 + (1-\alpha)x_2)$$
$$\leqslant \alpha[f_1(x_1) + f_2(x_1)] + (1-\alpha)[f_1(x_2) + f_2(x_2)]. \blacksquare$$

Proposition 2. *Let f be a convex function over the convex set Ω. Then the function af is convex for any $a \geqslant 0$.*

Proof. Immediate.

Note that through repeated application of the above two propositions it follows that a positive combination $a_1 f_1 + a_2 f_2 + \cdots + a_m f_m$ of convex functions is again convex.

Finally, we consider sets defined by convex inequality constraints.

Proposition 3. *Let f be a convex function on a convex set Ω. The set $\Gamma_c = \{x : x \in \Omega, f(x) \leqslant c\}$ is convex for every real number c.*

Proof. Let $x_1, x_2 \in \Gamma_c$. Then $f(x_1) \leqslant c, f(x_2) \leqslant c$ and for $0 < \alpha < 1$,

$$f(\alpha x_1 + (1-\alpha)x_2) \leqslant \alpha f(x_1) + (1-\alpha)f(x_2) \leqslant c.$$

Thus $\alpha x_1 + (1-\alpha)x_2 \in \Gamma_c$. $\blacksquare$

We note that, since the intersection of convex sets is also convex, the set of points simultaneously satisfying

$$f_1(x) \leqslant c_1, \quad f_2(x) \leqslant c_2, \ldots, f_m(x) \leqslant c_m,$$

where each f_i is a convex function, defines a convex set.

Properties of Differentiable Convex Functions

If a function f is differentiable, then there are alternative characterizations of convexity.

Proposition 4. *Let $f \in C^1$. Then f is convex over a convex set Ω if and only if*

$$f(y) \geqslant f(x) + \nabla f(x)(y - x) \tag{9}$$

for all $x, y \in \Omega$.

Proof. First suppose f is convex. Then for all α, $0 \leqslant \alpha \leqslant 1$,

$$f(\alpha y + (1-\alpha)x) \leqslant \alpha f(y) + (1-\alpha)f(x).$$

Thus for $0 < \alpha \leqslant 1$

$$\frac{f(x + \alpha(y - x)) - f(x)}{\alpha} \leqslant f(y) - f(x).$$

Letting $\alpha \to 0$ we obtain

$$\nabla f(x)(y - x) \leqslant f(y) - f(x).$$

This proves the "only if" part.

Now assume

$$f(y) \geqslant f(x) + \nabla f(x)(y - x)$$

for all $x, y \in \Omega$. Fix $x_1, x_2 \in \Omega$ and α, $0 \leqslant \alpha \leqslant 1$. Setting $x = \alpha x_1 + (1-\alpha)x_2$ and alternatively $y = x_1$ or $y = x_2$, we have

$$f(x_1) \geqslant f(x) + \nabla f(x)(x_1 - x) \tag{10}$$
$$f(x_2) \geqslant f(x) + \nabla f(x)(x_2 - x). \tag{11}$$

Multiplying (10) by α and (11) by $(1-\alpha)$ and adding, we obtain

$$\alpha f(x_1) + (1-\alpha)f(x_2) \geqslant f(x) + \nabla f(x)[\alpha x_1 + (1-\alpha)x_2 - x].$$

But substituting $x = \alpha x_1 + (1-\alpha)x_2$, we obtain

$$\alpha f(x_1) + (1-\alpha)f(x_2) \geqslant f(\alpha x_1 + (1-\alpha)x_2). \blacksquare$$

The statement of the above proposition is illustrated in Fig. 6.3. It can be regarded as a sort of dual characterization of the original definition illustrated in Fig. 6.2. The original definition essentially states that linear interpolation between two points overestimates the function, while the above proposition states that linear approximation based on the local derivative underestimates the function.

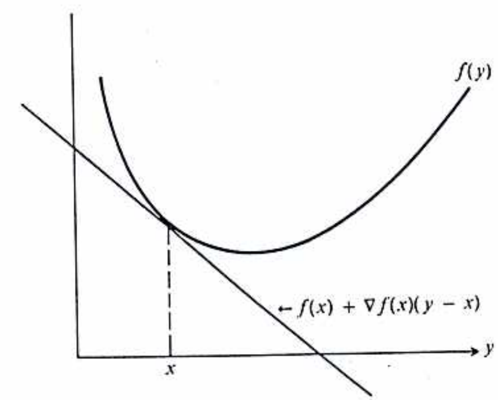

Fig. 6.3 Illustration of Proposition

For twice continuously differentiable functions, there is another characterization of convexity.

Proposition 5. *Let $f \in C^2$. Then f is convex over a convex set Ω containing an interior point if and only if the Hessian matrix* **F** *of f is positive semidefinite throughout Ω.*

Proof. By Taylor's theorem we have

$$f(\mathbf{y}) = f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{y}-\mathbf{x}) + \tfrac{1}{2}(\mathbf{y}-\mathbf{x})'\mathbf{F}(\mathbf{x}+\alpha(\mathbf{y}-\mathbf{x}))(\mathbf{y}-\mathbf{x}) \quad (12)$$

for some α, $0 \leqslant \alpha \leqslant 1$. Clearly, if the Hessian is everywhere positive semidefinite, we have

$$f(\mathbf{y}) \geqslant f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{y}-\mathbf{x}) \quad (13)$$

which in view of Proposition 4 implies that f is convex.

Now suppose the Hessian is not positive semidefinite at some point $\mathbf{x} \in \Omega$. By continuity of the Hessian it can be assumed, without loss of generality, that **x** is an interior point of Ω. There is a $\mathbf{y} \in \Omega$ such that $(\mathbf{y}-\mathbf{x})'\mathbf{F}(\mathbf{x})(\mathbf{y}-\mathbf{x}) < 0$. Again by the continuity of the Hessian, **y** may be selected so that for all α, $0 \leqslant \alpha \leqslant 1$,

$$(\mathbf{y}-\mathbf{x})'\mathbf{F}(\mathbf{x}+\alpha(\mathbf{y}-\mathbf{x}))(\mathbf{y}-\mathbf{x}) < 0.$$

This in view of (12) implies that (13) does not hold; which in view of Proposition 4 implies that f is not convex. ▮

The Hessian matrix is the generalization to E^n of the concept of the curvature of a function, and correspondingly, positive definiteness of the Hessian is the generalization of positive curvature. Convex functions have positive (or at least nonnegative) curvature in every direction. Motivated by these observations, we sometimes refer to a function as being *locally convex* if its Hessian matrix is positive semidefinite in a small region, and *locally strictly convex* if the Hessian is positive definite in the region. In these terms we see that the second-order sufficiency result of the last section requires that the function be locally strictly convex at the point $\mathbf{x}^*$. Thus, even the local theory, derived solely in terms of the elementary calculus, is actually intimately related to convexity—at least locally. For this reason we can view the two theories, local and global, not as disjoint parallel developments but as complementary and interactive. Results that are based on convexity apply even to nonconvex problems in a region near the solution, and conversely, local results apply to a global minimum point.

6.4 MINIMIZATION AND MAXIMIZATION OF CONVEX FUNCTIONS

We turn now to the three classic results concerning minimization or maximization of convex functions.

Theorem 1. *Let f be a convex function defined on the convex set Ω. Then the set Γ where f achieves its minimum is convex, and any relative minimum of f is a global minimum.*

Proof. If f has no relative minima the theorem is valid by default. Assume now that c_0 is the minimum of f. Then clearly $\Gamma = \{\mathbf{x} : f(\mathbf{x}) \leqslant c_0, \mathbf{x} \in \Omega\}$ and this is convex by Proposition 3 of the last section.

Suppose now that $\mathbf{x}^* \in \Omega$ is a relative minimum point of f, but that there is another point $\mathbf{y} \in \Omega$ with $f(\mathbf{y}) < f(\mathbf{x}^*)$. On the line $\alpha\mathbf{y} + (1-\alpha)\mathbf{x}^*$, $0 < \alpha < 1$ we have

$$f(\alpha\mathbf{y} + (1-\alpha)\mathbf{x}^*) \leqslant \alpha f(\mathbf{y}) + (1-\alpha) f(\mathbf{x}^*) < f(\mathbf{x}^*)$$

contradicting the fact that $\mathbf{x}^*$ is a relative minimum point. ▮

We might paraphrase the above theorem as saying that for convex functions, all minimum points are located together (in a convex set) and all relative minima are global minima. The next theorem says that if f is continuously differentiable and convex, then satisfaction of the first order necessary conditions are both necessary and sufficient for a point to be a global minimizing point.

Theorem 2. *Let $f \in C^1$ be convex on the convex set Ω. If there is a point $\mathbf{x}^* \in \Omega$ such that, for all $\mathbf{y} \in \Omega$, $\nabla f(\mathbf{x}^*)(\mathbf{y}-\mathbf{x}^*) \geqslant 0$, then $\mathbf{x}^*$ is a global minimum point of f over Ω.*

Proof. We note parenthetically that since $\mathbf{y} - \mathbf{x}^*$ is a feasible direction at $\mathbf{x}^*$, the given condition is equivalent to the first-order necessary condition stated in Section 6.1. The proof of the proposition is immediate, since by Proposition 4 of the last section

$$f(\mathbf{y}) \geqslant f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)(\mathbf{y}-\mathbf{x}^*) \geqslant f(\mathbf{x}^*). ▮$$

Next we turn to the question of maximizing a convex function over a convex set. There is, however, no analog of Theorem 1 for maximization indeed, the tendency is for the occurrence of numerous non-global relative maximum points. Nevertheless, it is possible to prove one important result It is not used in subsequent chapters, but it is useful for some areas o optimization.

Theorem 3. *Let f be a convex function defined on the bounded, close convex set Ω. If f has a maximum over Ω it is achieved at an extreme poin of Ω.*

Proof. Suppose f achieves a global maximum at $\mathbf{x}^* \in \Omega$. We show first tha this maximum is achieved at some boundary point of Ω. If $\mathbf{x}^*$ is itself boundary point, then there is nothing to prove, so assume $\mathbf{x}^*$ is not boundary point. Let L be any line passing through the point $\mathbf{x}^*$. The inte

For twice continuously differentiable functions, there is another characterization of convexity.

Proposition 5. *Let $f \in C^2$. Then f is convex over a convex set Ω containing an interior point if and only if the Hessian matrix $\mathbf{F}$ of f is positive semidefinite throughout Ω.*

Proof. By Taylor's theorem we have

$$f(\mathbf{y}) = f(\mathbf{x}) = \nabla f(\mathbf{x})(\mathbf{y} - \mathbf{x}) + \tfrac{1}{2}(\mathbf{y} - \mathbf{x})^T\mathbf{F}(\mathbf{x} + \alpha(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x}) \qquad (12)$$

for some α, $0 \leqslant \alpha \leqslant 1$. Clearly, if the Hessian is everywhere positive semidefinite, we have

$$f(\mathbf{y}) \geqslant f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{y} - \mathbf{x}), \qquad (13)$$

which in view of Proposition 4 implies that f is convex.

Now suppose the Hessian is not positive semidefinite at some point $\mathbf{x} \in \Omega$. By continuity of the Hessian it can be assumed, without loss of generality, that $\mathbf{x}$ is an interior point of Ω. There is a $\mathbf{y} \in \Omega$ such that $(\mathbf{y} - \mathbf{x})^T\mathbf{F}(\mathbf{x})(\mathbf{y} - \mathbf{x}) < 0$. Again by the continuity of the Hessian, $\mathbf{y}$ may be selected so that for all α, $0 \leqslant \alpha \leqslant 1$,

$$(\mathbf{y} - \mathbf{x})^T\mathbf{F}(\mathbf{x} + \alpha(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x}) < 0.$$

This in view of (12) implies that (13) does not hold; which in view of Proposition 4 implies that f is not convex. ∎

The Hessian matrix is the generalization to E^n of the concept of the curvature of a function, and correspondingly, positive definiteness of the Hessian is the generalization of positive curvature. Convex functions have positive (or at least nonnegative) curvature in every direction. Motivated by these observations, we sometimes refer to a function as being *locally convex* if its Hessian matrix is positive semidefinite in a small region, and *locally strictly convex* if the Hessian is positive definite in the region. In these terms we see that the second-order sufficiency result of the last section requires that the function be locally strictly convex at the point $\mathbf{x}^*$. Thus, even the local theory, derived solely in terms of the elementary calculus, is actually intimately related to convexity—at least locally. For this reason we can view the two theories, local and global, not as disjoint parallel developments but as complementary and interactive. Results that are based on convexity apply even to nonconvex problems in a region near the solution, and conversely, local results apply to a global minimum point.

6.5 MINIMIZATION AND MAXIMIZATION OF CONVEX FUNCTIONS

We turn now to the three classic results concerning minimization or max-

Theorem 1. *Let f be a convex function defined on the convex set Ω. Then the set Γ where f achieves its minimum is convex, and any relative minimum of f is a global minimum.*

Proof. If f has no relative minima the theorem is valid by default. Assume now that c_0 is the minimum of f. Then clearly $\Gamma = \{\mathbf{x} : f(\mathbf{x}) \leqslant c_0, \mathbf{x} \in \Omega\}$ and this is convex by Proposition 3 of the last section.

Suppose now that $\mathbf{x}^* \in \Omega$ is a relative minimum point of f, but that there is another point $\mathbf{y} \in \Omega$ with $f(\mathbf{y}) < f(\mathbf{x}^*)$. On the line $\alpha\mathbf{y} + (1 - \alpha)\mathbf{x}^*$, $0 < \alpha < 1$ we have

$$f(\alpha\mathbf{y} + (1 - \alpha)\mathbf{x}^*) \leqslant \alpha f(\mathbf{y}) + (1 - \alpha)f(\mathbf{x}^*) < f(\mathbf{x}^*),$$

contradicting the fact that $\mathbf{x}^*$ is a relative minimum point. ∎

We might paraphrase the above theorem as saying that for convex functions, all minimum points are located together (in a convex set) and all relative minima are global minima. The next theorem says that if f is continuously differentiable and convex, then satisfaction of the first-order necessary conditions are both necessary and sufficient for a point to be a global minimizing point.

Theorem 2. *Let $f \in C^1$ be convex on the convex set Ω. If there is a point $\mathbf{x}^* \in \Omega$ such that, for all $\mathbf{y} \in \Omega$, $\nabla f(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}^*) \geqslant 0$, then $\mathbf{x}^*$ is a global minimum point of f over Ω.*

Proof. We note parenthetically that since $\mathbf{y} - \mathbf{x}^*$ is a feasible direction at $\mathbf{x}^*$, the given condition is equivalent to the first-order necessary condition stated in Section 6.1. The proof of the proposition is immediate, since by Proposition 4 of the last section

$$f(\mathbf{y}) \geqslant f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}^*) \geqslant f(\mathbf{x}^*). \;∎$$

Next we turn to the question of maximizing a convex function over a convex set. There is, however, no analog of Theorem 1 for maximization; indeed, the tendency is for the occurrence of numerous nonglobal relative maximum points. Nevertheless, it is possible to prove one important result. It is not used in subsequent chapters, but it is useful for some areas of optimization.

Theorem 3. *Let f be a convex function defined on the bounded, closed convex set Ω. If f has a maximum over Ω it is achieved at an extreme point of Ω.*

Proof. Suppose f achieves a global maximum at $\mathbf{x}^* \in \Omega$. We show first that this maximum is achieved at some boundary point of Ω. If $\mathbf{x}^*$ is itself a boundary point, then there is nothing to prove, so assume $\mathbf{x}^*$ is not a boundary point. Let L be any line passing through the point $\mathbf{x}^*$. The intersection of this line with Ω is an interval of the line L having end points