OPTIMIZATION BY VECTOR SPACE METHODS

David G. Luenberger

Stanford University,
Stanford, California

John Wiley & Sons, Inc.
New York — London — Sydney — Toronto

CONTENTS

REFERENCES

§10.2. The contraction mapping theorem is the simplest of a variety of fixed-point theorems. For an introduction to other results, see Courant and Robbins [32], Graves [64], and Bonsall [24].

§10.3. Much of the theory of Newton's method in Banach spaces was developed by Kantorovich [78]. See also Kantorovich and Akilov [79], [80], Bartle [16], Antosiewicz and Rheinboldt [9], and Collatz [30]. Some interesting extensions and modifications of the method have been developed by Altman [6], [7], [8]. An interesting account of various applications of Newton's method is found in Bellman and Kalaba [18].

§10.4–5. For material on steepest descent, see Ostrowski [115], Rosenbloom [130], Nashed [108], Curry [33], and Kelley [84].

§10.6–8. For the original development of conjugate direction methods, see Hestenes and Stiefel [71] and Hestenes [70]. The underlying principle of orthogonalizing moments, useful for a number of computational problems, is discussed in detail by Vorobyev [150] and Faddeev and Faddeeva [50]. There are various extensions of the method of conjugate gradients to problems with nonquadratic objective functionals. The most popular of these are the Fletcher and Reeves [54] method and the PARTAN method in Shah, Buehler, and Kempthorne [138]. For another generalization and the derivation of the convergence rate see Daniel [34], [35]. For application to optimal control theory, see Lasdon, Mitter, and Waren [94] and Sinnott and Luenberger [140].

§10.9–11. The gradient-projection method is due to Rosen [128], [129]. A closely related method is that of Zoutendijk [158]. For an application of the primal-dual method, see Wilson [154]. The penalty function method goes back to a suggestion of Courant [31]. See also Butler and Martin [27], Kelley [84], Fabian [49], and Fiacco and McCormick [53]. For additional general material, bibliography, and review of optimization techniques, see Saaty and Bram [136], Wolfe [156], Zoutendijk [159], Leon [96], Wilde and Beightler [153], Dorn [43], and Spang [143].

§10.12. For extensions of Problem 9 to Banach space, see Altman [6] and the generalized inverse function theorem of Chapter 9. For Kantorovich's inequal-ity (Problem 11), see Greub and Rheinboldt [66]. The method introduced in Problem 20 has been studied extensively by Fiacco and McCormick [53].

7

[redacted]IZATION OF FUNCTIONALS

7.1 Introduction

In previous chapters, while developing the elements of functional analysis, we often considered minimum norm optimization problems. Although the availability of a large variety of different norms provides enough flexibility for minimum norm problems to be of importance, there are many important optimization problems that cannot be formulated in these terms. In this chapter we consider optimization of more general objective functionals. However, much of the theory and geometric insight gained while studying minimum norm problems are of direct benefit in considering these more general problems.

Our study is guided by two basic geometric representations of nonlinear functionals. Each of these representations has its own particular advantages, and often it is enlightening to view a given situation in both ways. The first, and perhaps most obvious, geometric representation of a nonlinear functional is in terms of its graph. Suppose f is a functional defined on a subset D of the vector space X. The space X is then imbedded in the product space $R \times X$ where R is the real line. Elements of this space consist of ordered pairs (r, x). The graph of f is the surface in $R \times X$ consisting of the points $(f(x), x)$ with $x \in D$. Usually the R axis (i.e., points of the form (r, θ)) is regarded as the vertical axis, and the value of the functional at x is then regarded as the vertical distance of the graph above the point $x \in X$. In this representation a typical functional is visualized as in Figure 7.1. This representation is certainly familiar for one- or two-dimensional space. The point here is that it is convenient conceptually even in infinite-dimensional space.

The second representation is an extension of the technique of representing a linear functional by a hyperplane. A functional is described by its contours in the space X. A typical representation is shown in Figure 7.2. If f is sufficiently smooth, it is natural to construct hyperplanes tangent to the contours and to define the gradient of f. Again, this technique is

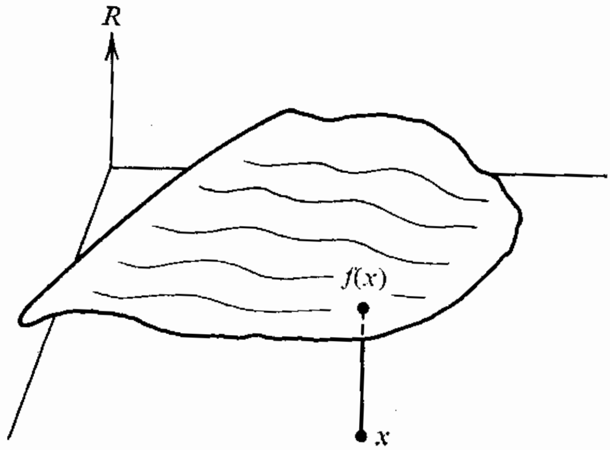

Figure 7.1 The graph of a functional

familiar in finite-dimensional space and is easily extended to infinite-dimensional space. The principal advantage of the second representation over the first is that for a given dimensionality of X, one less dimension is required for the representation.

The first half of this chapter deals with generalizations of the concepts of differentials, gradients, etc., to normed spaces and is essentially based on the second representation stated above. The detailed development is largely algebraic and manipulative in nature, although certain geometric interpretations are apparent. From this apparatus we obtain the local or variational theory of optimization paralleling the familiar theory in finite dimensions. The most elementary portion of the classical calculus of variations is used as a principal example of these methods.

The second half of the chapter deals with convex and concave functionals from which we obtain a global theory of optimization. This

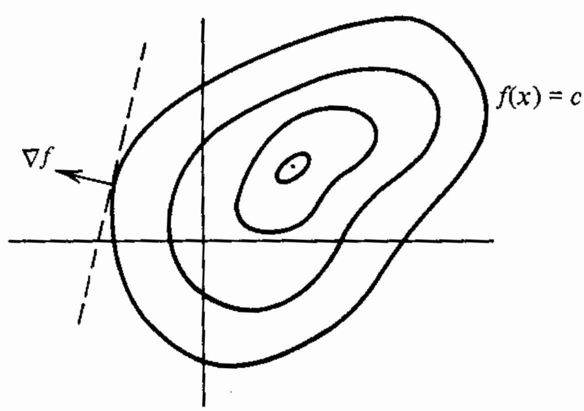

Figure 7.2 The contours of a functional

development, based essentially on the first representation for functionals, is largely geometric and builds on the theory of convex sets considered in Chapter 5. The interesting theory of conjugate functionals produces another duality theorem for a class of optimization problems.

LOCAL THEORY

7.2 Gateaux and Fréchet Differentials

In the following discussion let X be a vector space, Y a normed space, and T a (possibly nonlinear) transformation defined on a domain $D \subset X$ and having range $R \subset Y$.

Definition. Let $x \in D \subset X$ and let h be arbitrary in X. If the limit

$$\delta T(x; h) = \lim_{\alpha \to 0} \frac{1}{\alpha} [T(x + \alpha h) - T(x)] \tag{1}$$

exists, it is called the *Gateaux differential of T at x with increment h*. If the limit (1) exists for each $h \in X$, the transformation T is said to be *Gateaux differentiable at x*.

We note that it makes sense to consider the limit (1) only if $x + \alpha h \in D$ for all α sufficiently small. The limit (1) is, of course, taken in the usual sense of norm convergence in Y. We observe that for fixed $x \in D$ and h regarded as variable, the Gateaux differential defines a transformation from X to Y. In the particular case where T is linear, we have $\delta T(x; h) = T(h)$.

By far the most frequent application of this definition is in the case where Y is the real line and hence the transformation reduces to a (real-valued) functional on X. Thus if f is a functional on X, the Gateaux differential of f, if it exists, is

$$\delta f(x; h) = \frac{d}{d\alpha} f(x + \alpha h)\Big|_{\alpha = 0},$$

and, for each fixed $x \in X$, $\delta f(x; h)$ is a functional with respect to the variable $h \in X$.

Example 1. Let $X = E^n$ and let $f(x) = f(x_1, x_2, \ldots, x_n)$ be a functional on E^n having continuous partial derivatives with respect to each variable x_i. Then the Gateaux differential of f is

$$\delta f(x; h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i.$$

Example 2. Let $X = C[0, 1]$ and let $f(x) = \int_0^1 g(x(t), t)\, dt$, where it is assumed that the function g_x exists and is continuous with respect to x and t. Then

$$\delta f(x; h) = \frac{d}{d\alpha} \int_0^1 g(x(t) + \alpha h(t), t)\, dt \bigg|_{\alpha = 0}$$

Interchange of the order of differentiation and integration is permissible under our assumptions on g and hence

$$\delta f(x; h) = \int_0^1 g_x(x, t) h(t)\, dt.$$

Example 3. If $X = E^n$, $Y = E^m$, and T is a continuously differentiable mapping of X into Y, then $\delta T(x; h)$ exists. It is the vector in Y equal to the vector $h \in x$ multiplied by the matrix made up of partial derivatives of T at x.

The Gateaux differential generalizes the concept of directional derivative familiar in finite-dimensional space. The existence of the Gateaux differential is a rather weak requirement, however, since its definition requires no norm on X; hence, properties of the Gateaux differential are not easily related to continuity. When X is normed, a more satisfactory definition is given by the Fréchet differential.

Definition. Let T be a transformation defined on an open domain D in a normed space X and having range in a normed space Y. If for fixed $x \in D$ and each $h \in X$ there exists $\delta T(x; h) \in Y$ which is linear and continuous with respect to h such that

$$\lim_{\|h\| \to 0} \frac{\|T(x + h) - T(x) - \delta T(x; h)\|}{\|h\|} = 0,$$

then T is said to be *Fréchet differentiable at* x and $\delta T(x; h)$ is said to be the *Fréchet differential of* T *at* x *with increment* h.

We use the same symbol for the Fréchet and Gateaux differentials since generally it is apparent from the context which is meant.

Proposition 1. *If the transformation* T *has a Fréchet differential, it is unique.*

Proof. Suppose both $\delta T(x; h)$ and $\delta' T(x; h)$ satisfy the requirements of the last definition. Then

$$\begin{aligned} \|\delta T(x; h) - \delta' T(x; h)\| \le {} & \|T(x + h) - T(x) - \delta T(x; h)\| \\ & + \|T(x + h) - T(x) - \delta' T(x; h)\| \end{aligned}$$

or $\|\delta T(x; h) - \delta' T(x; h)\| = o(\|h\|)$. Since $\delta T(x; h) - \delta' T(x; h)$ is bounded and linear in h, it must be zero. ∎

Proposition 2. *If the Fréchet differential of T exists at x, then the Gateaux differential exists at x and they are equal.*

Proof. Denote the Fréchet differential by $\delta T(x; h)$. By definition we have for any h,

$$\frac{1}{\alpha}\|T(x+\alpha h) - T(x) - \delta T(x; \alpha h)\| \to 0, \qquad \text{as } \alpha \to 0.$$

Thus, by the linearity of $\delta T(x; \alpha h)$ with respect to α,

$$\lim_{\alpha \to 0} \frac{T(x+\alpha h) - T(x)}{\alpha} = \delta T(x; h). \quad \blacksquare$$

A final property is given by the following proposition.

Proposition 3. *If the transformation T defined on an open set D in X has a Fréchet differential at x, then T is continuous at x.*

Proof. Given $\varepsilon > 0$, there is a sphere about x such that for $x + h$ in this sphere

$$\|T(x+h) - T(x) - \delta T(x; h)\| < \varepsilon\|h\|.$$

Thus $\|T(x+h) - T(x)\| < \varepsilon\|h\| + \|\delta T(x; h)\| < M\|h\|$. $\blacksquare$

We are primarily concerned with Fréchet differentials rather than Gateaux differentials and often assume their existence or, equivalently, assume the satisfaction of conditions which are sufficient to imply their existence.

Example 4. We show that the differential

$$\delta f(x; h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i$$

in Example 1 (of the functional f on E^n) is a Fréchet differential. Obviously, $\delta f(x; h)$ is linear and continuous in h. It is therefore only necessary to verify the basic limit property required of Fréchet differentials.

Given $\varepsilon > 0$, the continuity of the partial derivatives implies that there is a neighborhood $S(x; \delta)$ such that

$$\left| \frac{\partial f(x)}{\partial x_i} - \frac{\partial f(y)}{\partial x_i} \right| < \frac{\varepsilon}{n}$$

for all $y \in S(x; \delta)$ and $i = 1, 2, \ldots, n$.

Define the unit vectors e_i in the usual way, $e_i = (0, 0, \ldots, 1, 0, \ldots, 0)$, and for $h = \sum_{i=1}^{n} h_i e_i$ define $g_0 = \theta$ and

$$g_k = \sum_{i=1}^{k} h_i e_i, \qquad \text{for } k = 1, 2, \ldots, n.$$

We note that $\|g_k\| \leq \|h\|$ for all k. Then

$$\left|f(x+h) - f(x) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i\right| = \left|\sum_{k=1}^{n} \left\{f(x+g_k) - f(x+g_{k-1}) - \frac{\partial f}{\partial x_k} h_k\right\}\right|$$
$$\leq \sum_{k=1}^{n} \left|f(x+g_k) - f(x+g_{k-1}) - \frac{\partial f}{\partial x_k} h_k\right|.$$

Now we examine the k-th term in the above summation. The vector $x + g_k$ differs from $x + g_{k-1}$ only in the k-th component. In fact, $x + g_k = x + g_{k-1} + h_k e_k$. Thus, by the mean value theorem for functions of a single variable,

$$f(x+g_k) - f(x+g_{k-1}) = \frac{\partial f}{\partial x_k}(x + g_k + \alpha e_k)h_k$$

for some α, $0 \leq \alpha \leq h_k$.

Also, $x + g_{k-1} + \alpha e_k \in S(x; \delta)$ if $\|h\| < \delta$. Thus

$$\left|f(x+g_k) - f(x+g_{k-1}) - \frac{\partial f(x)}{\partial x_k} h_k\right| < \frac{\varepsilon}{n}\|h\|.$$

Finally, it follows that

$$\left|f(x+h) - f(x) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_k} h_i\right| < \varepsilon\|h\|$$

for all h, $\|h\| < \delta$.

Example 5. We show that the differential

$$\delta f(x; h) = \int_0^1 g_x(x, t)h(t)\, dt$$

in Example 2 is a Fréchet differential. We have

$$|f(x+h) - f(x) - \delta f(x; h)|$$
$$= \left|\int_0^1 \{g(x+h, t) - g(x, t) - g_x(x, t)h(t)\}\, dt\right|.$$

For a fixed t we have, by the one-dimensional mean value theorem,

$$g(x(t) + h(t), t) - g[x(t), t] = g_x(\bar{x}(t), t)h(t)$$

where $|x(t) - \bar{x}(t)| \leq |h(t)|$. Given $\varepsilon > 0$, the uniform continuity of g_x in x and t implies that there is a $\delta > 0$ such that for $|h| < \delta$, $|g_x(x+h, t) - g_x(x, t)| < \varepsilon$. Therefore, we have

$$|f(x+h) - f(x) - \delta f(x, h)| = \left|\int_0^1 (g_x(\bar{x}, t) - g_x(x, t))h(t)\, dt\right| \leq \varepsilon\|h\|$$

for $\|h\| < \delta$. The result follows.

Example 6. Let $X = C^n[0, 1]$, the space of continuous n-vector functions on $[0, 1]$. Let $Y = C^m[0, 1]$ and define $T : X \to Y$ by

$$T(x) = \int_0^t F[x(\tau), \tau]\, d\tau$$

where $F = (f_1, f_2, \ldots, f_m)$ has continuous partial derivatives with respect to its arguments. The Gateaux differential of T is easily seen to be

$$\delta T(x; h) = \int_0^t F_x[x(\tau), \tau]h(\tau)\, d\tau.$$

By combining the analyses of Examples 4 and 5, we can conclude that this is actually a Fréchet differential. Note also that since the partial derivatives of F are continuous, $\delta T(x; h)$ is continuous in the variable x.

7.3 Fréchet Derivatives

Suppose that the transformation T defined on an open domain $D \subset X$ is Fréchet differentiable throughout D. At a fixed point $x \in D$ the Fréchet differential $\delta T(x; h)$ is then, by definition, of the form $\delta T(x; h) = A_x h$, where A_x is a bounded linear operator from X to Y. Thus, as x varies over D, the correspondence $x \to A_x$ defines a transformation from D into the normed linear space $B(X, Y)$; this transformation is called the *Fréchet derivative* T' *of* T. Thus we have, by definition, $\delta T(x; h) = T'(x)h$.

If the correspondence $x \to T'(x)$ is continuous at the point x_0 (i.e., if given $\varepsilon > 0$ there is $\delta > 0$ such that $\|x - x_0\| < \delta$ implies $\|T'(x) - T'(x_0)\| < \varepsilon$), we say that the Fréchet derivative of T is *continuous* at x_0. This should not be confused with the statement that $T'(x_0)$ is a continuous mapping from X to Y, a property that is basic to the definition of the Fréchet derivative. If the derivative of T is continuous on some open sphere S, we say that T is *continuously Fréchet differentiable* on S. All of the examples of Fréchet differentiable transformations in the last section are in fact continuously Fréchet differentiable.

In the special case where the original transformation is simply a functional f on the space X, we have $\delta f(x; h) = f'(x)h$ where $f'(x) \in X^*$ for each x. The element $f'(x)$ is called the gradient of f at x and is sometimes denoted $\nabla f(x)$ rather than $f'(x)$. We sometimes write $\langle h, f'(x)\rangle$ for $\delta f(x; h)$ since $f'(x) \in X^*$, but usually we prefer $f'(x)h$ which is consistent with the notation for differentials of arbitrary transformations.

Much of the theory of ordinary derivatives can be generalized to Fréchet derivatives. For instance, the implicit function theorem and Taylor series have very satisfactory extensions. The interested reader should consult the references cited at the end of the chapter. In this section we discuss the elementary properties of Fréchet derivatives used in later sections.

It follows immediately from the definition that if T_1 and T_2 are Frechét differentiable at $x \in D$, then $\alpha_1 T_1 + \alpha_2 T_2$ is Fréchet differentiable at x and $(\alpha_1 T_1 + \alpha_2 T_2)'(x) = \alpha_1 T_1'(x) + \alpha_2 T_2'(x)$. We next show that the chain rule applies to Fréchet derivatives.

Proposition 1. *Let S be a transformation mapping an open set $D \subset X$ into an open set $E \subset Y$ and let P be a transformation mapping E into a normed space Z. Put $T = PS$ and suppose S is Fréchet differentiable at $x \in D$ and P is Fréchet differentiable at $y = S(x) \in E$. Then T is Fréchet differentiable at x and $T'(x) = P'(y)S'(x)$.*

Proof. For $h \in X$, $x + h \in D$, we have

$$T(x + h) - T(x) = P[S(x + h)] - P[S(x)] = P(y + g) - P(y)$$

where $g = S(x + h) - S(x)$. Thus $\|T(x + h) - T(x) - P'(y)g\| = o(\|g\|)$. Since, however,

$$\|g - S'(x)h\| = o(\|h\|),$$

we obtain

$$\|T(x + h) - T(x) - P'(y)S'(x)h\| = o(\|h\|) + o(\|g\|).$$

Since, according to Proposition 3 of Section 7.2, S is continuous at x, we conclude that $\|g\| = O(\|h\|)$ and hence

$$T'(x)h = P'(y)S'(x)h. \blacksquare$$

We now give a very useful inequality that replaces the mean value theorem for ordinary functions.

Proposition 2. *Let T be Fréchet differentiable on an open domain D. Let $x \in D$ and suppose that $x + \alpha h \in D$ for all α, $0 \leq \alpha \leq 1$. Then*

$$\|T(x + h) - T(x)\| \leq \|h\| \sup_{0 < \alpha < 1} \|T'(x + \alpha h)\|.$$

Proof. Let y^* be a nonzero element of Y^* aligned with the element $T(x + h) - T(x)$. The function $\varphi(\alpha) = y^*[T(x + \alpha h)]$ is defined on the interval $[0, 1]$ and, by the chain rule, has derivative

$$\varphi'(\alpha) = y^*[T'(x + \alpha h)h].$$

By the mean value theorem for functions of a real variable, we have

$$\varphi(1) - \varphi(0) = \varphi'(\alpha_0), \qquad 0 < \alpha_0 < 1,$$

and hence

$$|y^*[T(x+h) - T(x)]| \leq \|y^*\| \sup_{0<\alpha<1} \|T'(x+\alpha h)\| \|h\|,$$

and since y^* is aligned with $T(x+h) - T(x)$,

$$\|T(x+h) - T(x)\| \leq \|h\| \sup_{0<\alpha<1} \|T'(x+\alpha h)\|. \blacksquare$$

If $T: X \to Y$ is Fréchet differentiable on an open domain $D \subset X$, the derivative T' maps D into $B(X, Y)$ and may itself be Fréchet differentiable on a subset $D_1 \subset D$. In this case the Fréchet derivative of T' is called the second Fréchet derivative of T and is denoted by T''.

Example 1. Let f be a functional on $X = E^n$ having continuous partial derivatives up to second order. Then $f''(x_0)$ is an operator from E^n to E^n having matrix form

$$f''(x_0) = \left[\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right]_{x=x_0},$$

where x_i is the i-th component of x.

The following inequality can be proved in a manner paralleling that of Proposition 2.

Proposition 3. *Let T be twice Fréchet differentiable on an open domain D. Let $x \in D$ and suppose that $x + \alpha h \in D$ for all α, $0 \leq \alpha \leq 1$. Then*

$$\|T(x+h) - T(x) - T'(x)h\| \leq \tfrac{1}{2}\|h\|^2 \sup_{0<\alpha<1} \|T''(x+\alpha h)\|.$$

7.4 Extrema

It is relatively simple to apply the concepts of Gateaux and Fréchet differentials to the problem of minimizing or maximizing a functional on a linear space. The technique leads quite naturally to the rudiments of the calculus of variations where, in fact, the abstract concept of differentials originated. In this section, we extend the familiar technique of minimizing a function of a single variable by ordinary calculus to a similar technique based on more general differentials. In this way we obtain analogs of the classical necessary conditions for local extrema and, in a later section, the Lagrange technique for constrained extrema.

Definition. Let f be a real-valued functional defined on a subset Ω of a normed space X. A point $x_0 \in \Omega$ is said to be a *relative minimum* of f on Ω if there is an open sphere N containing x_0 such that $f(x_0) \leq f(x)$ for all $x \in \Omega \cap N$. The point x_0 is said to be a *strict relative minimum* of f on Ω

if $f(x_0) < f(x)$ for all $x \neq x_0$, $x \in \Omega \cap N$. *Relative maxima* are defined similarly.

We use the term *extremum* to refer to either a maximum or a minimum over any set. A *relative extremum* (over a subset of a normed space) is also referred to as a *local extremum*. The set Ω on which an extremum problem is defined is sometimes called the *admissible set*.

Theorem 1. *Let the real-valued functional f have a Gateaux differential on a vector space X. A necessary condition for f to have an extremum at $x_0 \in X$ is that $\delta f(x_0; h) = 0$ for all $h \in X$.*

Proof. For every $h \in X$, the function $f(x + \alpha h)$ of the real variable α must achieve an extremum at $\alpha = 0$. Thus, by the ordinary calculus,

$$\frac{d}{d\alpha} f(x_0 + \alpha h)\Big|_{\alpha=0} = 0. \quad \blacksquare$$

A point at which $\delta f(x; h) = 0$ for all h is called a *stationary point*; hence, the above theorem merely states that extrema occur at stationary points. It should be noted that a similar result holds for a local extremum of a functional defined on an open subset of a normed space, the proof being identical for both cases.

The simplicity of Theorem 1 can be misleading for it is a result of great utility. Much of the calculus of variations can be regarded as a simple consequence of this one result. Indeed, many interesting problems are solved by careful identification of an appropriate vector space X and some algebraic manipulations to obtain the differential. There are a number of useful generalizations of Theorem 1. We offer one of the simplest.

Theorem 2. *Let f be a real-valued functional defined on a vector space X. Suppose that x_0 minimizes f on the convex set $\Omega \subset X$ and that f is Gateaux differentiable at x_0. Then*

$$\delta f(x_0; x - x_0) \geq 0$$

for all $x \in \Omega$.

Proof. Since Ω is convex, $x_0 + \alpha(x - x_0) \in \Omega$ for $0 \leq \alpha \leq 1$ and hence, by ordinary calculus,

$$\frac{d}{d\alpha} f(x_0 + \alpha(x - x_0))\Big|_{\alpha=0} \geqslant 0$$

for a minimum at x_0. $\blacksquare$

*7.5 Euler-Lagrange Equations

The classical problem in the calculus of variations is that of finding a function x on the interval $[t_1, t_2]$ minimizing an integral functional of the form

$$J = \int_{t_1}^{t_2} f[x(t), \dot{x}(t), t]\, dt.$$

To specify the problem completely, we must agree on the class of functions within which we seek the extremum—the so-called admissible set. We assume that the function f is continuous in x, $\dot{x}$, and t and has continuous partial derivatives with respect to x and $\dot{x}$. We seek a solution in the space $D[t_1, t_2]$. In the simplest version of the problem, we assume that the end points $x(t_1)$ and $x(t_2)$ are fixed. This further restricts the admissible set.

Starting with a given admissible vector x, we consider vectors of the form $x + h$ that are admissible. The class of such vectors h is called the class of *admissible variations*. In the case of fixed end points, it is clear that the class of admissible variations is the subspace of $D[t_1, t_2]$, consisting of functions which vanish at t_1 and t_2. The necessary condition for the extremum problem is that for all such h, $\delta J(x; h) = 0$.

The differential of J is

$$\delta J(x; h) = \frac{d}{d\alpha} \int_{t_1}^{t_2} f(x + \alpha h, \dot{x} + \alpha \dot{h}, t)\, dt \Big|_{\alpha = 0}$$

or, equivalently,

$$\text{(1)} \qquad \delta J(x; h) = \int_{t_1}^{t_2} f_x(x, \dot{x}, t) h(t)\, dt + \int_{t_1}^{t_2} f_{\dot{x}}(x, \dot{x}, t) \dot{h}(t)\, dt,$$

and it is easily verified that this differential is actually Fréchet. Equating this differential to zero and assuming that the function $f_{\dot{x}}$ has a continuous derivative with respect to t when the optimal solution is substituted for x, we may integrate by parts to obtain

$$\delta J(x; h) = \int_{t_1}^{t_2} [f_x(x, \dot{x}, t) - \frac{d}{dt} f_{\dot{x}}(x, \dot{x}, t)] h(t)\, dt + f_{\dot{x}}(x, \dot{x}, t) h(t) \Big|_{t_1}^{t_2} = 0.$$

The boundary terms vanish for admissible h and thus the necessary condition is

$$\int_{t_1}^{t_2} \left[f_x(x, \dot{x}, t) - \frac{d}{dt} f_{\dot{x}}(x, \dot{x}, t) \right] h(t)\, dt = 0$$

for all $h \in D[t_1, t_2]$ vanishing at t_1 and t_2.

Since the term multiplying $h(t)$ in the integrand is continuous, it readily follows (see Lemma 1 below) that it must vanish identically on $[t_1, t_2]$. Thus we conclude that the extremal x must satisfy the Euler-Lagrange equation

$$f_x(x, \dot{x}, t) - \frac{d}{dt} f_{\dot{x}}(x, \dot{x}, t) = 0. \tag{2}$$

The above derivation of the Euler-Lagrange equations suffers from the weakness of the assumption that

$$\frac{d}{dt} f_{\dot{x}}$$

is continuous at the optimal solution. Actually we have no basis on which to assume that the solution is in fact smooth enough for this assumption to hold. The alternate derivation, given after the following three lemmas, avoids this drawback.

Lemma 1. *If* $\alpha(t)$ *is continuous on* $[t_1, t_2]$ *and* $\int_{t_1}^{t_2} \alpha(t)h(t)\,dt = 0$ *for every* $h \in D[t_1, t_2]$ *with* $h(t_1) = h(t_2) = 0$, *then* $\alpha(t) \equiv 0$ *on* $[t_1, t_2]$.

Proof. Assume that $\alpha(t)$ is nonzero, say positive, for some $t \in [t_1, t_2]$. Then $\alpha(t)$ is positive on some interval $[t_1', t_2'] \subset [t_1, t_2]$. Let

$$h(t) = \begin{cases} (t - t_1')^2(t_2' - t)^2 & t \in [t_1', t_2'] \\ 0 & \text{otherwise.} \end{cases}$$

The function h satisfies the hypotheses of the lemma and

$$\int_{t_1}^{t_2} \alpha(t)h(t)\,dt > 0. \quad \blacksquare$$

Lemma 2. *If* $\alpha(t)$ *is continuous in* $[t_1, t_2]$ *and* $\int_{t_1}^{t_2} \alpha(t)\dot{h}(t)\,dt = 0$ *for every* $h \in D[t_1, t_2]$ *with* $h(t_1) = h(t_2) = 0$, *then* $\alpha(t) \equiv c$ *in* $[t_1, t_2]$ *where* c *is a constant.*

Proof. Let c be the unique constant satisfying $\int_{t_1}^{t_2} [\alpha(t) - c]\,dt = 0$ and let

$$h(t) = \int_{t_1}^{t} [a(\tau) - c]\,d\tau.$$

Then

$$\begin{aligned} \int_{t_1}^{t_2} [\alpha(t) - c]^2\,dt &= \int_{t_1}^{t_2} [\alpha(t) - c)]\dot{h}(t)\,dt \\ &= \int_{t_1}^{t_2} \alpha(t)\dot{h}(t)\,dt - c[h(t_2) - h(t_1)] = 0, \end{aligned}$$

and hence $\alpha(t) \equiv c$. $\blacksquare$

Lemma 3. *If $\alpha(t)$ and $\beta(t)$ are continuous in $[t_1, t_2]$ and*

$$\int_{t_1}^{t_2} [\alpha(t)h(t) + \beta(t)\dot{h}(t)]\, dt = 0 \tag{3}$$

for every $h \in D[t_1, t_2]$ with $h(t_1) = h(t_2) = 0$, then β is differentiable and $\dot{\beta}(t) \equiv \alpha(t)$ in $[t_1, t_2]$.

Proof. Define

$$A(t) = \int_{t_1}^{t} \alpha(\tau)\, d\tau.$$

Then by integration by parts we have

$$\int_{t_1}^{t_2} \alpha(t)h(t)\, dt = -\int_{t_1}^{t_2} A(t)\dot{h}(t)\, dt.$$

Therefore (3) becomes

$$\int_{t_1}^{t_2} [-A(t) + \beta(t)]\dot{h}(t)\, dt = 0$$

which by Lemma 2 implies

$$\beta(t) = A(t) + c$$

for some constant c. Hence, by the definition of A, $\dot{\beta}(t) = \alpha(t)$. ∎

Now, in view of Lemma 3, it is clear that the Euler-Lagrange equation (2) follows directly from equation (1) without an *a priori* assumption of the differentiability of $f_{\dot{x}}$.

Example 1. (Minimum Arc Length) Given t_1, t_2 and $x(t_1)$, $x(t_2)$, let us employ the Euler-Lagrange equations to determine the curve in $D[t_1, t_2]$ connecting these points with minimum arc length. We thus seek to minimize

$$J = \int_{t_1}^{t_2} \sqrt{1 + (\dot{x})^2}\, dt.$$

From (2) we obtain immediately

$$\frac{d}{dt}\frac{\partial}{\partial \dot{x}} \sqrt{1 + (\dot{x})^2} = 0$$

or, equivalently,

$$\dot{x} = \text{const.}$$

Thus the extremizing arc is the straight line connecting the two points.

Example 2. (Estate Planning) What is the lifetime plan of investment and expenditure that maximizes total enjoyment for a man having a fixed quantity of savings S? We assume that the man has no income other than that obtained through his investment. His rate of enjoyment (or utility) at a given time is a certain function U of r, his rate of expenditure. Thus, we assume that it is desired to maximize

$$\int_0^T e^{-\beta t} U[r(t)]\, dt$$

where the $e^{-\beta t}$ term reflects the notion that future enjoyment is counted less today.

If $x(t)$ is the total capital at time t, then

$$\dot{x}(t) = \alpha x(t) - r(t)$$

where α is the interest rate of investment. Thus the problem is to maximize

$$\int_0^T e^{-\beta t} U[\alpha x(t) - \dot{x}(t)]\, dt$$

subject to $x(0) = S$ and $x(T) = 0$ (or some other fixed value). Of course, there is the additional constraint $x(t) \geq 0$, but, in the cases we consider, this turns out to be satisfied automatically.

From the Euler-Lagrange equation (2), we obtain

$$\alpha e^{-\beta t} U'[\alpha x(t) - \dot{x}(t)] + \frac{d}{dt} e^{-\beta t} U'[\alpha x(t) - \dot{x}(t)] = 0$$

where U' is the derivative of the utility function U. This becomes

$$\frac{d}{dt} U'[\alpha x(t) - \dot{x}(t)] = (\beta - \alpha) U'[\alpha x(t) - \dot{x}(t)]. \tag{4}$$

Hence, integrating (4), we find that

$$U'[r(t)] = U'[r(0)] e^{(\beta - \alpha)t}. \tag{5}$$

Hence, the form of the time dependence of r can be obtained explicitly once U is specified.

A simple utility function, which reflects both the intuitive notions of diminishing marginal enjoyment (i.e., U' decreasing) and infinite marginal enjoyment at zero expenditure (i.e., $U'[0] = \infty$), is $U[r] = 2r^{1/2}$. Substituting this in (5), we obtain

$$\alpha x(t) - \dot{x}(t) = r(t) = r(0) e^{2(\alpha - \beta)t}.$$

Integrating this last equation, we have

$$
\begin{aligned}
x(t) &= e^{\alpha t}x(0) + \frac{r(0)}{\alpha - 2\beta}\,[e^{\alpha t} - e^{2(\alpha-\beta)t}] \\
&= \left\{x(0) - \frac{r(0)}{2\beta - \alpha}\right\}e^{\alpha t} + \frac{r(0)}{2\beta - \alpha}\,e^{2(\alpha-\beta)t}.
\end{aligned}
\tag{6}
$$

Assuming $\alpha > \beta > \alpha/2$, we may find $r(0)$ from (6) by requiring $x(T) = 0$. Thus

$$
r(0) = \frac{(2\beta - \alpha)x(0)}{1 - e^{-(2\beta-\alpha)T}}.
$$

The total capital grows initially and then decreases to zero.

*7.6 Problems with Variable End Points

The class of admissible functions for a given problem is not always a linear variety of functions nor even a convex set. Such problems may, however, generally be approached in essentially the same manner as before. Basically the technique is to define a continuously differentiable one-parameter family of admissible curves that includes the optimal curve as a member. A necessary condition is obtained by equating the derivative of the objective functional, with respect to the parameter, equal to zero. In the case where the admissible class is a linear variety, we consider the family $x + \varepsilon h$ and differentiate with respect to ε. In other cases, a more general family $x(\varepsilon)$ of admissible curves is considered.

An interesting class of problems that can be handled in this way is calculus of variations problems having variable end points. Specifically, suppose we seek an extremum of the functional

$$
J = \int_{t_1}^{t_2} f(x, \dot{x}, t)\,dt
$$

where the interval $[t_1, t_2]$, as well as the function $x(t)$, must be chosen. In a typical problem the two end points are constrained to lie on fixed curves in the $x - t$ plane. We assume here that the left end point is fixed and allow the right end point to vary along a curve S described by the function $x = g(t)$, as illustrated in Figure 7.3.

Suppose $x(\varepsilon, t)$ is a one-parameter family of functions emanating from the point 1 and terminating on the curve S. The termination point is $x(\varepsilon, t_2(\varepsilon))$ and satisfies $x(\varepsilon, t_2(\varepsilon)) = g(t_2(\varepsilon))$. We assume that the family is defined for $\varepsilon \in [-a, a]$ for some $a > 0$ and that the desired extremal is the

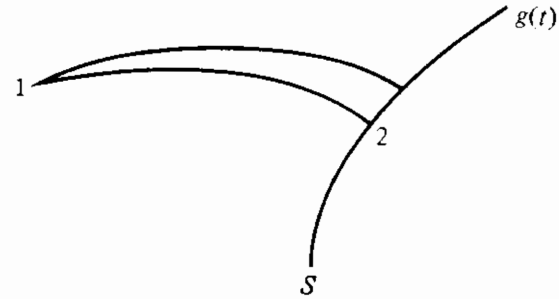

Figure 7.3 A variable end-point problem

curve corresponding to $\varepsilon = 0$. The variation δJ of the functional J is defined as the first derivative of J with respect to ε and must vanish at $\varepsilon = 0$.

Defining, in the obvious way,

$$\delta x(t) = \frac{d}{d\varepsilon} x(\varepsilon, t)\bigg|_{\varepsilon=0}$$

$$\delta t_2 = \frac{d}{d\varepsilon} t_2(\varepsilon)\bigg|_{\varepsilon=0},$$

we have

$$J(\varepsilon) = \int_{t_1}^{t_2(\varepsilon)} f[x(t, \varepsilon), \dot{x}(t, \varepsilon), t]\, dt$$

$$\delta J = f[x(t_2), \dot{x}(t_2), t_2]\delta t_2 + \int_{t_1}^{t_2} \{f_x(x, \dot{x}, t)\delta x + f_{\dot{x}}(x, \dot{x}, t)\delta \dot{x}\}\, dt = 0.$$

Arguments similar to those of the last section lead directly to the necessary conditions

(1) $$f_x(x, \dot{x}, t) = \frac{d}{dt} f_{\dot{x}}(x, \dot{x}, t)$$

and

(2) $$f[x(t_2), \dot{x}(t_2), t_2]\delta t_2 + f_{\dot{x}}[x(t_2), \dot{x}(t_2), t_2]\delta x(t_2) = 0.$$

Condition (1) is again the Euler-Lagrange equation, but in addition we have the transversality condition (2) which must be employed to determine the termination point. The transversality condition (2) must hold for all admissible $\delta x(t_2)$ and δt_2. These two quantities are not independent, however, since

(3) $$x[\varepsilon, t_2(\varepsilon)] = g[t_2(\varepsilon)].$$

Upon differentiating (3) with respect to ε, we obtain the relation

$$\delta x(t_2) + \dot{x}(t_2)\delta t_2 = \dot{g}(t_2)\delta t_2$$

and hence the complete set of necessary conditions is (1) and the transversality condition

$$\text{(4)} \qquad \{f(x, \dot{x}, t) + [\dot{g} - \dot{x}]f_{\dot{x}}(x, \dot{x}, t)\}\big|_{t=t_2} = 0.$$

Example 1. As an example of the transversality condition, we consider the simple problem of finding the differentiable curve of minimum are length from the origin (0, 0) to the curve g. Thus we seek to extremize the functional

$$J = \int_0^{t_2} \sqrt{1 + (\dot{x})^2}\, dt,$$

where t_2 is the point of intersection of the line.

As in Example 1 of the last section, the Euler-Lagrange equation leads to $\dot{x} = \text{const}$ and, hence, the extremal must be a straight line. The transversality condition in this case is

$$\left\{\sqrt{1 + (\dot{x})^2} + (\dot{g} - \dot{x})\frac{\dot{x}}{\sqrt{1 + (\dot{x})^2}}\right\}\Bigg|_{t=t_2} = 0,$$

or

$$\dot{x}(t_2) = -\frac{1}{\dot{g}(t_2)}.$$

Thus the extremal arc must be orthogonal to the tangent of g at t_2. (See Figure 7.4.)

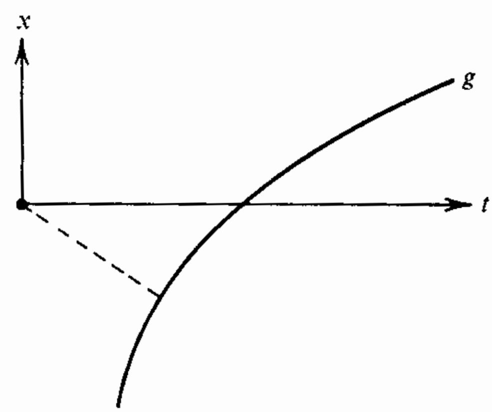

Figure 7.4 Minimum distance to a curve

7.7 Problems with Constraints

In most optimization problems the optimal vector is required to satisfy some type of constraint. In the simplest calculus of variations problem, for instance, the end points of the curve are constrained to fixed points.

Such problems, and those of greater complexity, often can be resolved (in the sense of establishing necessary conditions) by considering a one-parameter family of vectors satisfying the constraints which contains the optimal vector as a member. A complication arises, however, when the constraint set is defined implicitly in terms of a set of functional equations rather than explicitly as a constraint surface in the space. In such cases the one-parameter family must be constructed implicitly. In this section we carry out the necessary construction for a finite number of functional constraints and arrive at our first Lagrange multiplier theorem.

A more general discussion of constrained optimization problems, including the material of this section as a special case, is given in Chapter 9. The later discussion parallels, but is much deeper, than the one given here.

The problem investigated in this section is that of optimizing a functional f subject to n nonlinear constraints given in the implicit form:

$$
\begin{aligned}
g_1(x) &= 0 \\
g_2(x) &= 0 \\
&\vdots \\
g_n(x) &= 0.
\end{aligned} \tag{1}
$$

These n equations define a region Ω in the space X within which the optimal vector x_0 is constrained to lie. Throughout this section it is assumed that all functionals f, g_i are continuous and Fréchet differentiable on the normed space X.

Before embarking on the general resolution of the problem, let us briefly consider the geometry of the problem. The case where X is two dimensional is depicted in Figure 7.5. If x_0 is optimal, the functional f has an extremum at x_0 with respect to small displacements along Ω. Under sufficient smoothness conditions, it seems that f has an extremum at x_0 with respect to small displacements along T, the tangent to Ω at x_0.

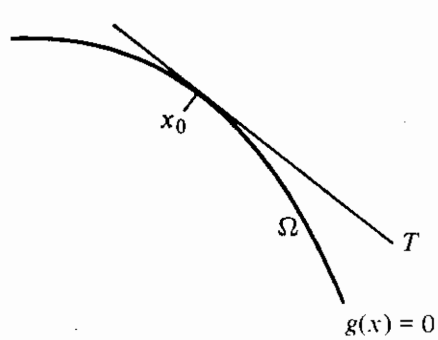

Figure 7.5 Constrained extremum

The utility of this observation is that the exact form of the surface Ω near x_0 is replaced by the simple description of its tangent in the expression for the necessary conditions. In order for the procedure to work, however, it must be possible to express the tangent in terms of the derivatives of the constraints. For this reason we introduce the definition of a regular point.

Definition. A point x_0 satisfying the constraints (1) is said to be a *regular point* of these constraints if the n linear functionals $g_1'(x_0), g_2'(x_0), \ldots, g_n'(x_0)$ are linearly independent.

The following theorem gives the necessary conditions satisfied by the solution of the constrained extremum problem.

Theorem 1. *If x_0 is an extremum of the functional f subject to the constraints $g_i(x) = 0$, $i = 1, 2, \ldots, n$; and if x_0 is a regular point of these constraints, then*

$$\delta f(x_0; h) = 0$$

for all h satisfying $\delta g_i(x_0; h) = 0$, $i = 1, 2, \ldots, n$.

Proof. Choose $h \in X$ such that $\delta g_i(x_0; h) = 0$ for $i = 1, 2, \ldots, n$. Let $y_1, y_2, \ldots, y_n \in X$ be n linearly independent vectors chosen so that the $n \times n$ matrix

$$M = \begin{bmatrix} \delta g_1(x_0; y_1) & \delta g_1(x_0; y_2) & \cdots & \delta g_1(x_0; y_n) \\ \delta g_2(x_0; y_1) & \delta g_2(x_0; y_2) & \cdots & \delta g_2(x_0; y_n) \\ \vdots & & & \\ \delta g_n(x_0; y_1) & & \cdots & \delta g_n(x_0; y_n) \end{bmatrix}$$

is nonsingular. The existence of such a set of vectors y_i follows directly from the regularity of the point x_0 (see Problem 11).

We now introduce the $n + 1$ real variables $\varepsilon, \varphi_1, \varphi_2, \ldots, \varphi_n$ and consider the n equations

$$\begin{aligned} &g_1(x_0 + \varepsilon h + \varphi_1 y_1 + \varphi_2 y_2 + \cdots + \varphi_n y_n) = 0 \\ &g_2(x_0 + \varepsilon h + \varphi_1 y_1 + \varphi_2 y_2 + \cdots + \varphi_n y_n) = 0 \\ &\vdots \\ &g_n(x_0 + \varepsilon h + \varphi_1 y_1 + \varphi_2 y_2 + \cdots + \varphi_n y_n) = 0. \end{aligned} \tag{2}$$

The Jacobian of this set with respect to the variables φ_i, at $\varepsilon = 0$, is just the determinant of M and is therefore nonzero by assumption. Hence, the implicit function theorem (see, for example, Apostol [10] and also Section 9.2) applies and guarantees the existence of n functions $\varphi_i(\varepsilon)$ satisfying (2) and defined in some neighborhood of $\varepsilon = 0$.

Denote by $y(\varepsilon)$ the vector $\sum_{i=1}^{n} \varphi_i(\varepsilon) y_i$ and by $\varphi(\varepsilon)$ the n-dimensional vector having components $\varphi_i(\varepsilon)$. For each i we have

$$(3) \qquad \begin{aligned} 0 &= g_i\left(x_0 + \varepsilon h + \sum_j \varphi_j y_j\right) \\ &= g_i(x_0) + \varepsilon \delta g_i(x_0; h) + \delta g_i[x_0, y(\varepsilon)] + o(\varepsilon) + o[\|y(\varepsilon)\|]. \end{aligned}$$

Or, writing all n equations simultaneously and taking into account the fact that the first two terms on the right side of (3) are zero, we have, after taking the norm of the result,

$$(4) \qquad 0 = \|M\varphi(\varepsilon)\| + o(\varepsilon) + o[\|y(\varepsilon)\|].$$

Since M is nonsingular, there are constants $c_1 > 0$, $c_2 > 0$ such that $c_1\|\varphi(\varepsilon)\| \leq \|M\varphi(\varepsilon)\| \leq c_2\|\varphi(\varepsilon)\|$; and since the y_i's are linearly independent, there are constants $d_1 > 0$, $d_2 > 0$ such that $d_1\|y(\varepsilon)\| \leq \|\varphi(\varepsilon)\| \leq d_2\|y(\varepsilon)\|$. Hence, $c_1 d_1\|y(\varepsilon)\| \leq \|M\varphi(\varepsilon)\| \leq c_2 d_2\|y(\varepsilon)\|$ and, therefore (4) implies $\|y(\varepsilon)\| = o(\varepsilon)$. Geometrically, this result states that if one moves along the tangent plane of the constraint surface Ω from x_0 to $x_0 + \varepsilon h$, it is possible to get back to Ω by moving to $x_0 + \varepsilon h + y(\varepsilon)$, where $y(\varepsilon)$ is small compared with εh.

The points $x_0 + \varepsilon h + y(\varepsilon)$ define a one-parameter family of admissible vectors. Considering the functional f at these points, we must have

$$\frac{d}{d\varepsilon} f[x_0 + \varepsilon h + y(\varepsilon)]\bigg]_{\varepsilon=0} = 0.$$

Thus, since $\|y(\varepsilon)\| = o(\varepsilon)$, $\delta f(x_0; h) = 0$. ∎

From Theorem 1 it is easy to derive a finite-dimensional version of the Lagrange multiplier rule by using the following lemma.

Lemma 1. *Let $f_0, f_1, \ldots, f_n$ be linear functionals on a vector space X and suppose that $f_0(x) = 0$ for every $x \in X$ satisfying $f_i(x) = 0$ for $i = 1, 2, \ldots, n$. Then there are constants $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that*

$$f_0 + \lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_n f_n = 0.$$

Proof. See Problem 16, Chapter 5 (which is solved by use of the Hahn-Banach theorem). ∎

Theorem 2. *If x_0 is an extremum of the functional f subject to the constraints*

$$g_i(x) = 0, \qquad i = 1, 2, \ldots, n,$$

and x_0 is a regular point of these constraints, then there are n scalars, λ_i, $i = 1, 2, \ldots, n$, that render the functional

$$f(x) + \sum_{i=1}^{n} \lambda_i g_i(x)$$

stationary at x_0.

Proof. By Theorem 1 the differential $\delta f(x_0; h)$ is zero whenever each of the differentials $\delta g_i(x_0; h)$ is zero. The result then follows immediately from Lemma 1. ▌

Example 1. Constrained problems of the form above are called isoperimetric problems in the calculus of variations since they were originally studied in connection with finding curves of given perimeter which maximized some objective such as enclosed area. Here we seek the curve in the $t-x$ plane having end points $(-1, 0)$, $(1, 0)$; length l; and enclosing maximum area between itself and the t-axis. Thus we wish to maximize[1]

$$\int_{-1}^{1} x(t)\, dt$$

subject to

$$\int_{-1}^{1} \sqrt{\dot{x}^2 + 1}\, dt = l.$$

Therefore, we seek a stationary point of

$$J(x) = \int_{-1}^{1} (x + \lambda\sqrt{\dot{x}^2 + 1})\, dt = \int_{-1}^{1} f(x, \dot{x}, t)\, dt.$$

Applying the Euler equations, we require

$$f_x - \frac{df_{\dot{x}}}{dt} = 0$$

or

$$1 - \lambda \frac{d}{dt} \frac{\dot{x}}{\sqrt{\dot{x}^2 + 1}} = 0$$

or

$$\frac{\dot{x}}{\sqrt{\dot{x}^2 + 1}} = \frac{1}{\lambda} t + c.$$

[1] See Problem 12 for a different formulation of this problem.

It is easily verified that a solution to this first-order differential equation is given by the arc of a circle

$$(x - x_1)^2 + (t - t_1)^2 = r^2.$$

The parameters x_1, t_1, and r are chosen to satisfy the boundary conditions and the condition on total length.

GLOBAL THEORY

7.8 Convex and Concave Functionals

Definition. A real-valued functional f defined on a convex subset C of a linear vector space is said to be *convex* if

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2)$$

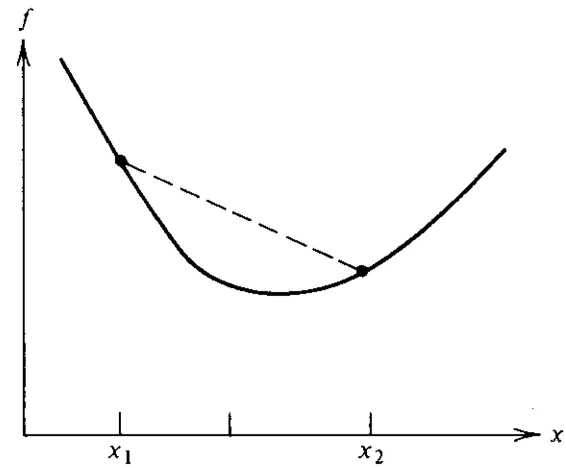

Figure 7.6 A convex function

for all $x_1, x_2 \in C$ and all α, $0 < \alpha < 1$. If strict inequality holds whenever $x_1 \neq x_2$, f is said to be *strictly convex*. A functional g defined on a convex set is said to be (*strictly*) *concave* if $-g$ is (strictly) convex.

Examples of convex functions in one dimension are $f(x) = x^2$; $f(x) = e^x$ for $x > 0$; and the discontinuous function

$$f(x) = \begin{cases} 1 & x = 0 \\ x^2 & x > 0 \end{cases}$$

defined on $[0, \infty)$. The functional

$$f(x) = \int_0^1 \{x^2(t) + |x(t)|\}\, dt$$

defined on $L_2[0, 1]$ is convex and continuous (the reader may wish to verify this). Any norm is a convex functional.

A convex functional defined on an infinite-dimensional normed space may be discontinuous everywhere since, for example, any linear functional is convex, and we constructed discontinuous linear functionals earlier.

Convex functionals play a special role in the theory of optimization because most of the theory of local extrema for general nonlinear functionals can be strengthened to become global theory when applied to convex functionals. Conversely, results derived for minimization of convex functionals often have analogs as local properties for more general problems. The study of convex functionals leads then not only to an aspect of optimization important in its own right but also to increased insight for a large portion of optimization theory.

The following proposition illustrates the global nature of results for minimization problems involving convex functionals.

Proposition 1. *Let f be a convex functional defined on a convex subset C of a normed space. Let $\mu = \inf_{x \in C} f(x)$. Then*

1. *The subset Ω of C where $f(x) = \mu$ is convex.*
2. *If x_0 is a local minimum of f, then $f(x_0) = \mu$ and, hence x_0* is a *global minimum.*

Proof.

1. Suppose $x_1, x_2 \in \Omega$. Then for $x = \alpha x_1 + (1 - \alpha)x_2$, $0 < \alpha < 1$, we have $f(x) \le \alpha f(x_1) + (1 - \alpha) f(x_2) = \mu$. But for any $x \in C$ necessarily $f(x) \ge \mu$. Thus, $f(x) = \mu$.
2. Suppose N is a neighborhood about x_0 in which x_0 minimizes f. For any $x_1 \in C$, there is an $x \in N$ such that $x = \alpha x_0 + (1 - \alpha)x_1$ for some α, $0 < \alpha < 1$. We have $f(x_0) \le f(x) \le \alpha f(x_0) + (1 - \alpha) f(x_1)$. Or $f(x_0) \le f(x_1)$. ∎

The study of convex functionals is quickly and effectively reduced to the study of convex sets by considering the region above the graph of the function.

Definition. In correspondence to a convex functional f defined on a convex set C in a vector space X, we define the convex set $[f, C]$ in $R \times X$ as

$$[f, C] = \{(r, x) \in R \times X : x \in C, f(x) \le r\}.$$

Usually we think of the space $R \times X$ as being oriented so that the R axis, i.e., all vectors of the form (r, θ), is the vertical axis. Then the set $[f, C]$ can be thought of as the region above the graph of f, as illustrated in Figure 7.7. This set $[f, C]$ is sometimes called the *epigraph* of f over C.

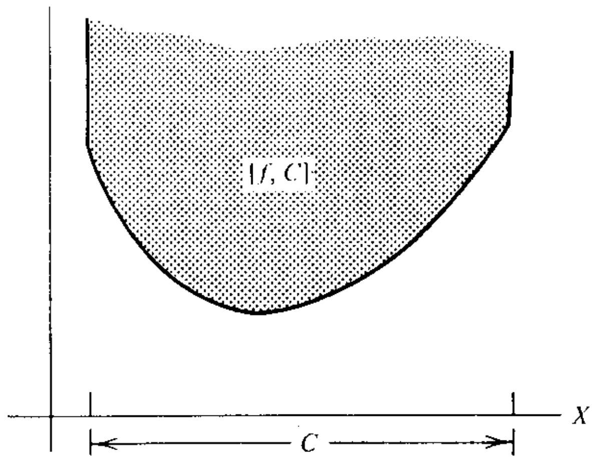

Figure 7.7 The convex region above the graph

Although we have no occasion to do so, we could imagine forming the set $[f, C]$ as in the above definition even if f were not convex. In any case, however, we have the following proposition.

Proposition 2. *The functional f defined on the convex domain C is convex if and only if $[f, C]$ is a convex set.*

The major portion of our analysis of convex functionals is based on consideration of the corresponding convex set of the last definition. To apply our arsenal of supporting hyperplane theorems to this set, however, it must be determined under what conditions the set $[f, C]$ contains interior points. The next section is devoted primarily to an analysis of this question. Since the results are so favorable—namely continuity of the functional at a single point guarantees an interior point—the reader need only glance through the proposition statements at first reading and proceed to the next section.

*7.9 Properties of the Set $[f, C]$

Proposition 1. *If f is a convex functional on the convex domain C in a normed space and C has nonempty relative interior $\mathring{C}$, then the convex set $[f, C]$ has a relative interior point (r_0, x_0) if and only if f is continuous at the point $x_0 \in \mathring{C}$.*

Proof. First assume that f is continuous at a point $x_0 \in \mathring{C}$. Denote by $N(x_0, \delta)$ the open spherical neighborhood of x_0 with radius δ. We note that $v([f, C])$, the linear variety generated by $[f, C]$, is equal to $R \times v(C)$. Given ε, $0 < \varepsilon < 1$, there is a $\delta > 0$ such that for $x \in N(x_0, \delta) \cap v(C)$ we have $x \in \mathring{C}$ and $|f(x) - f(x_0)| < \varepsilon$. Let $r_0 = f(x_0) + 2$. Then the point $(r_0, x_0) \in [f, C]$ is a relative interior point of $[f, C]$ since $(r, x) \in [f, C]$ for $|r - r_0| < 1$ and $x \in N(x_0, \delta) \cap v(C)$.

Now suppose that (r_0, x_0) is a relative interior point of $[f, C]$. Then there is $\varepsilon_0 > 0$, $\delta_0 > 0$ such that for $x \in N(x_0, \delta_0) \cap v(C)$ and $|r - r_0| < \varepsilon_0$ we have $r \geq f(x)$. Thus f is bounded above by $f(x_0) + \varepsilon_0$ on the neighborhood $N(x_0, \delta_0) \cap v(C)$.

We show now that the above implies that f is continuous at x_0. Without loss of generality we may assume $x_0 = \theta$ and $f(x_0) = 0$. For any ε, $0 < \varepsilon < 1$, and for any $x \in N(x_0, \varepsilon\delta_0) \cap v(C)$, we have

$$f(x) = f\left[(1 - \varepsilon)\theta + \varepsilon\left(\frac{1}{\varepsilon}x\right)\right] \leq (1 - \varepsilon)f(\theta) + \varepsilon f\left(\frac{1}{\varepsilon}x\right) \leq \varepsilon\varepsilon_0$$

where ε_0 is the bound on f in $N(x_0, \delta_0) \cap v(C)$. Furthermore,

$$0 = f(\theta) = f\left[\frac{1}{1+\varepsilon}x + \left(1 - \frac{1}{1+\varepsilon}\right)\left(-\frac{1}{\varepsilon}x\right)\right] \leq \frac{1}{1+\varepsilon}f(x) + \left(1 - \frac{1}{1+\varepsilon}\right)f\left(-\frac{1}{\varepsilon}x\right)$$

or

$$f(x) \geq -\varepsilon f\left(-\frac{1}{\varepsilon}x\right) \geq -\varepsilon\varepsilon_0.$$

Therefore, for $x \in N(x_0, \varepsilon\delta_0) \cap v(C)$, we have $|f(x)| \leq \varepsilon\varepsilon_0$. Thus f is continuous at x_0. ∎

Convex functionals enjoy many of the properties of linear functionals. As an example, the following proposition is a generalization of Proposition 1, Section 5.2.

Proposition 2. *A convex functional f defined on a convex domain C and continuous at a single point in the relative interior $\mathring{C}$ of C is continuous throughout $\mathring{C}$.*

Proof. Without loss of generality we may assume that f is continuous at $\theta \in \mathring{C}$ and that $f(\theta) = 0$. Furthermore, by restricting attention to $v(C)$, we may assume that C has interior points rather than relative interior points.

Let y be an arbitrary point in $\mathring{C}$. Since $\mathring{C}$ is (relatively) open, there is a $\beta > 1$ such that $\beta y \in \mathring{C}$. Given $\varepsilon > 0$, let $\delta > 0$ be such that $\|x\| < \delta$ implies $|f(x)| < \varepsilon$. Then for $\|z - y\| < (1 - \beta^{-1})\delta$, we have

$$z = y + (1 - \beta^{-1})x = \beta^{-1}(\beta y) + (1 - \beta^{-1})x$$

for some $x \in \mathring{C}$ with $\|x\| < \delta$. Thus $z \in C$ and

$$f(x) \le \beta^{-1}f(\beta y) + (1 - \beta^{-1})f(x) < \beta^{-1}f(\beta y) + (1 - \beta^{-1})\varepsilon.$$

Thus f is bounded above in the sphere $\|z - y\| < (1 - \beta^{-1})\delta$. It follows that for sufficiently large r the point (r, y) is an interior point of $[f, C]$; hence, by Proposition 1, f is continuous at y. ∎

The proof of the following important corollary is left to the reader.

Corollary 1. *A convex functional defined on a finite-dimensional convex set C is continuous throughout* $\mathring{C}$.

Having established the simple relation between continuity and interior points, we conclude this section by noting a property of f which holds *if* $[f, C]$ happens to be closed. As illustrated in Figure 7.8, closure of $[f, C]$ is related to the continuity properties of f on the boundary of C.

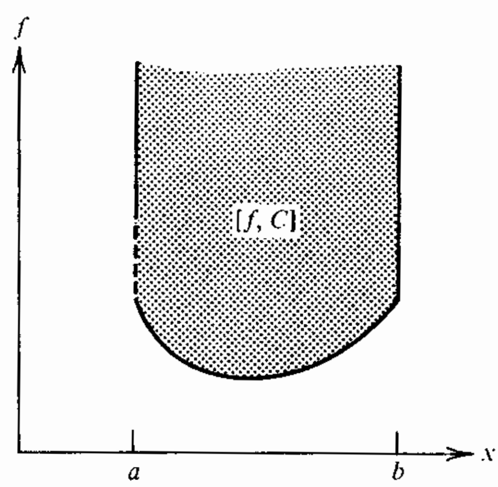

Figure 7.8 A nonclosed epigraph

Proposition 3. *If* $[f, C]$ *is closed, then* f *is lower semicontinuous on* C.

Proof. The set $\{(a, x) \in R \times X : x \in X\}$ is obviously closed for each $a \in R$. Hence, if $[f, C]$ is closed, so is

$$[f, C] \cap \{(a, x) : x \in X\} = \{(a, x) : x \in C, f(x) \le \mathrm{a}\}$$

for each $a \in R$. It follows that the set

$$T_a = \{x : x \in C, f(x) \leq a\}$$

is closed.

Now suppose $\{x_i\}$ is a sequence from C converging to $x \in C$. Let $b = \liminf_{x_i \to x} f(x_i)$. If $b = -\infty$, then $x \in \overline{T}_a = T_a$ for each $a \in R$ which is impossible. Thus $b > -\infty$ and $x \in \overline{T}_{b+\varepsilon} = T_{b+\varepsilon}$ for all $\varepsilon > 0$. In other words, $f(x) \leq \liminf_{x_i \to x} f(x_i)$ which proves that f is lower semicontinuous. ∎

Figure 7.9 shows the graph of a convex functional f defined on a disk C in E^2 that has closed $[f, C]$ but is discontinuous (although lower semicontinuous) at a point x.

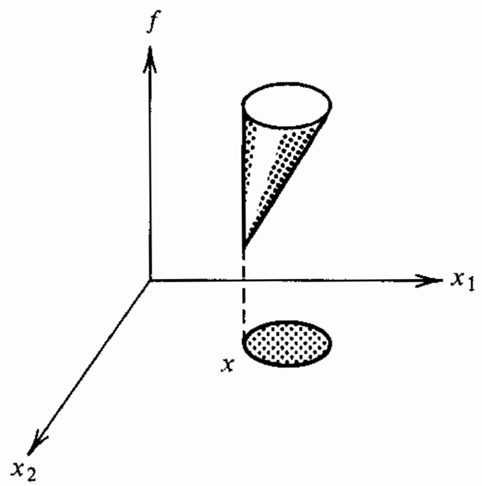

Figure 7.9

7.10 Conjugate Convex Functionals

A purely abstract approach to the theory of convex functionals, including a study of the convex set $[f, C]$ as in the previous section, leads quite naturally to an investigation of the dual representation of this set in terms of closed hyperplanes. The concept of conjugate functionals plays a natural and fundamental role in such a study. As an important consequence of this investigation, we obtain a very general duality principle for optimization problems which extends the earlier duality results for minimum norm problems.

Definition. Let f be a convex functional defined on a convex set C in a normed space X. The *conjugate set* C^* is defined as

$$C^* = \{x^* \in X^* : \sup_{x \in C} [\langle x, x^* \rangle - f(x)] < \infty\}$$

and the *functional f^* conjugate to f* is defined on C^* as

$$f^*(x^*) = \sup_{x \in C} [\langle x, x^* \rangle - f(x)].$$

Proposition 1. *The conjugate set C^* and the conjugate functional f^* are convex and $[f^*, C^*]$ is a closed convex subset of $R \times X^*$.*

Proof. For any $x_1^*, x_2^* \in X^*$ and any α, $0 < \alpha < 1$, we have

$$\begin{aligned}\sup_{x \in C} \{\langle x, \alpha x_1^* + (1-\alpha)x_2^* \rangle - f(x)\} &= \sup_{x \in C} \{\alpha[\langle x, x_1^* \rangle - f(x)] \\ &\qquad + (1-\alpha)[\langle x, x_2^* \rangle - f(x)]\} \\ &\leq \alpha \sup_{x \in C} [\langle x, x_1^* \rangle - f(x)] \\ &\qquad + (1-\alpha) \sup_{x \in C} [\langle x, x_2^* \rangle - f(x)]\end{aligned}$$

from which it follows immediately that C^* and f^* are convex.

Next we prove that $[f^*, C^*]$ is closed. Let $\{(s_i, x_i^*)\}$ be a convergent sequence from $[f^*, C^*]$ with $(s_i, x_i^*) \to (s, x^*)$. We show now that $(s, x^*) \in [f^*, C^*]$. For every i and every $x \in C$, we have

$$s_i \geq f^*(x_i^*) \geq \langle x, x_i^* \rangle - f(x).$$

Taking the limit as $i \to \infty$, we obtain

$$s \geq \langle x, x^* \rangle - f(x)$$

for all $x \in C$. Therefore,

$$s \geq \sup_{x \in C} [\langle x, x^* \rangle - f(x)]$$

from which it follows that $x^* \in C^*$ and $s \geq f^*(x^*)$. ∎

We see that the conjugate functional defines a set $[f^*, C^*]$ which is of the same type as $[f, C]$; therefore we write $[f, C]^* = [f^*, C^*]$. Note that if $f = 0$, the conjugate functional f^* becomes the support functional of C.

Example 1. Let $X = C = E^n$ and define, for $x = (x_1, x_2, \ldots, x_n)$, $f(x) = 1/p \sum_{i=1}^n |x_i|^p$, $1 < p < \infty$. Then for $x^* = (\xi_1, \xi_2, \ldots, \xi_n)$,

$$f^*(x^*) = \sup \left[\langle x, x^* \rangle - \frac{1}{p} \sum_{i=1}^n |x_i|^p\right] = \sup \left[\sum_{i=1}^n \xi_i x_i - \frac{1}{p} \sum_{i=1}^n |x_i|^p\right].$$

The supremum on the right is achieved by some x since the problem is finite dimensional. We find, by differentiation, the solution

$$\xi_i = |x_i|^{p-1} \operatorname{sgn} x_i$$

$$f^*(x^*) = \sum_{i=1}^{n} |x_i|^p \left(1 - \frac{1}{p}\right) = \frac{1}{q} \sum_{i=1}^{n} |\xi_i|^q$$

where $1/p + 1/q = 1$.

Let us investigate the relation of the conjugate functional to separating hyperplanes. On the space $R \times X$, closed hyperplanes are represented by an equation of the form

$$sr + \langle x, x^* \rangle = k$$

where s, k, and x^* determine the hyperplane. Recalling that we agreed to refer to the R axis as vertical, we say that a hyperplane is nonvertical if it intersects the R axis at one and only one point. This is equivalent to the requirement that the defining linear functional (s, x^*) have $s \neq 0$. If attention is restricted to nonvertical hyperplanes, we may, without loss of generality, consider only those linear functionals of the form $(-1, x^*)$. Any nonvertical closed hyperplane can then be obtained by appropriate choice of x^* and k.

To develop a geometric interpretation of the conjugate functional, note that as k varies, the solutions (r, x) of the equation

$$\langle x, x^* \rangle - r = k$$

describe parallel closed hyperplanes in $R \times X$. The number $f^*(x^*)$ is the supremum of the values of k for which the hyperplane intersects $[f, C]$. Thus the hyperplane $\langle x, x^* \rangle - r = f^*(x^*)$ is a support hyperplane of $[f, C]$.

In the terminology of Section 5.13, $f^*(x^*)$ is the support functional $h[(-1, x^*)]$ of the functional $(-1, x^*)$ for the convex set $[f, C]$. The special feature here is that we only consider functionals of the form $(-1, x^*)$ on $R \times X$ and thereby eliminate the need of carrying an extra variable.

For the application to optimization problems, the most important geometric interpretation of the conjugate functional is that it measures vertical distance to the support hyperplane. The hyperplane

$$\langle x, x^* \rangle - r = f^*(x^*)$$

intersects the vertical axis (i.e., $x = \theta$) at $(-f^*(x^*), \theta)$. Thus, $-f^*(x^*)$ is the vertical height of the hyperplane above the origin. (See Figure 7.10.)

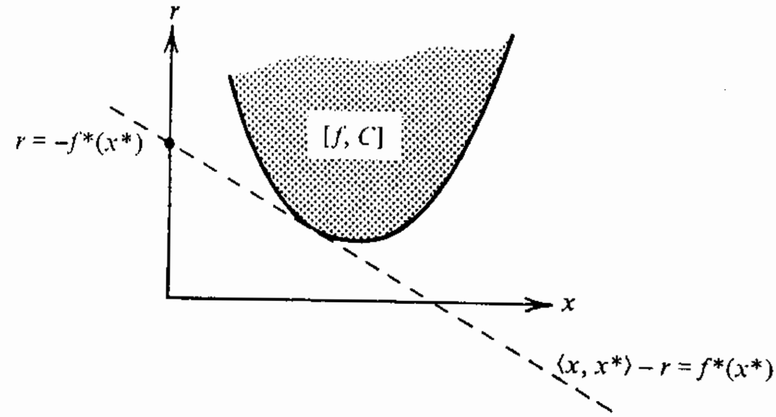

Figure 7.10 A conjugate convex functional

Another interpretation more clearly illuminates the duality between $[f, C]$ and $[f^*, C^*]$ in terms of the dual representation of a convex set as a collection of points or as the intersection of half-spaces. Given the point $(s, x^*) \in R \times X^*$, let us associate the half-space consisting of all $(r, x) \in R \times X$ satisfying

$$\langle x, x^* \rangle - r \leq s.$$

Then the set $[f^*, C^*]$ consists of those (nonvertical) half-spaces that contain the set $[f, C]$. Hence $[f^*, C^*]$ is the dual representation of $[f, C]$.

Beginning with an arbitrary convex functional φ defined on a convex subset Γ of a dual space X^*, we may, of course, define the conjugate of φ in X^{**} or, alternatively, following the standard pattern for duality relations (e.g., see Section 5.7), define the set ${}^*\Gamma$ in X as

$${}^*\Gamma = \{x : \sup_{x^* \in \Gamma} [\langle x, x^* \rangle - \varphi(x^*)] < \infty\}$$

and the convex functional

$${}^*\varphi(x) = \sup_{x^* \in \Gamma} [\langle x, x^* \rangle - \varphi(x^*)]$$

on ${}^*\Gamma$. We then write ${}^*[\varphi, \Gamma] = [{}^*\varphi, {}^*\Gamma]$. With these definitions we have the following characterization of the duality between a convex functional and its conjugate.

Proposition 2. *Let f be a convex functional on the convex set C in a normed space X. If $[f, C]$ is closed, then $[f, C] = {}^*[[f, C]^*]$.*

Proof. We show first that $[f, C] \subset {}^*[f^*, C^*] = {}^*[[f, C]^*]$. Let $(r, x) \in [f, C]$; then for all $x^* \in C^*$, $f^*(x^*) \geq \langle x, x^* \rangle - f(x)$. Hence, we have $r \geq f(x) \geq \langle x, x^* \rangle - f^*(x^*)$ for all $x^* \in C^*$. Thus

$$r \geq \sup_{x^* \in C^*} [\langle x, x^* \rangle - f^*(x^*)]$$

and $(r, x) \in {}^*[f^*, C^*]$.

We prove the converse by contraposition. Let $(r_0, x_0) \notin [f, C]$. Since $[f, C]$ is closed, there is a hyperplane separating (r_0, x_0) and $[f, C]$. Thus there exist $x^* \in X^*$, s, and c such that

$$sr + \langle x, x^* \rangle \leq c < sr_0 + \langle x_0, x^* \rangle$$

for all $(r, x) \in [f, C]$. It can be shown that, without loss of generality, this hyperplane can be assumed to be nonvertical and hence $s \neq 0$ (see Problem 16). Furthermore, since r can be made arbitrarily large, we must have $s < 0$. Thus we take $s = -1$. Now it follows that $\langle x, x^* \rangle - f(x) \leq c$ for all $x \in C$, which implies that $(c, x^*) \in [f^*, C^*]$. On the other hand, $c < \langle x_0, x^* \rangle - r_0$ implies $\langle x_0, x^* \rangle - c > r_0$, which implies that $(r_0, x_0) \notin {}^*[f^*, C^*]$. ∎

7.11 Conjugate Concave Functionals

A development similar to that of the last section applies to concave functionals. It must be stressed, however, that we *do not* treat concave functionals by merely multiplying by -1 and then applying the theory for convex functionals. There is an additional sign change in part of the definition. See Problem 15.

Given a concave functional g defined on a convex subset D of a vector space, we define the set

$$[g, D] = \{(r, x) : x \in D, r \leq g(x)\}.$$

The set $[g, D]$ is convex and all of the results on continuity, interior points, etc., of Section 7.9 have direct extensions here.

Definition. Let g be a concave functional on the convex set D. The *conjugate set* D^* is defined as

$$D^* = \{x^* \in X^* : \inf_{x \in D} [\langle x, x^* \rangle - g(x)] > -\infty\},$$

and the *functional* g^* *conjugate to* g is defined as

$$g^*(x^*) = \inf_{x \in D} [\langle x, x^* \rangle - g(x)].$$

We can readily verify that D^* is convex and that g^* is concave, We write $[g, D]^* = [g^*, D^*]$.

Since our notation does not completely distinguish between the development for convex and concave functionals, it is important to make clear which is being employed in any given context. This is particularly true when the original function is linear, since either definition of the conjugate functional might be employed and, in general, they are not equal.

The geometric interpretation for concave conjugate functionals is similar to that for convex conjugate functionals. The hyperplane $\langle x, x^* \rangle - r = g^*(x^*)$ supports the set $[g, D]$. Furthermore, $-g^*(x^*)$ is the intercept of that hyperplane with the vertical axis. The situation is summarized in Figure 7.11.

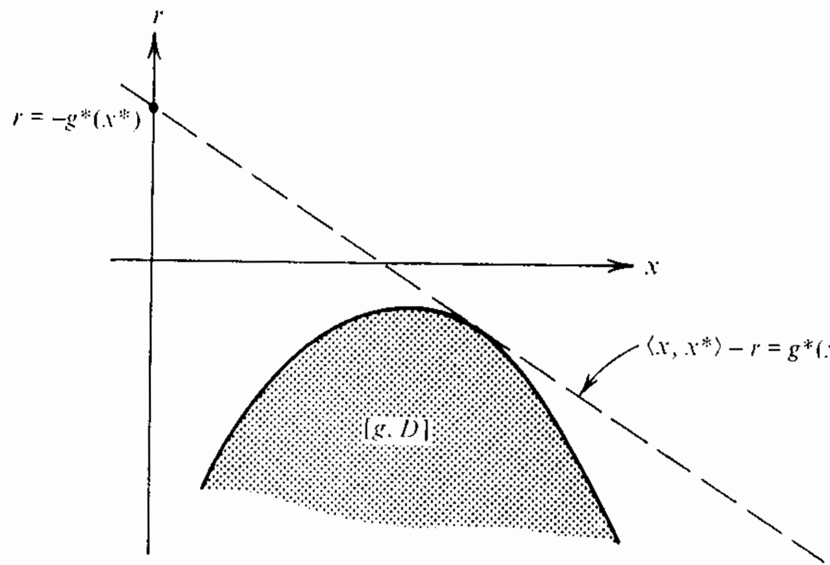

Figure 7.11 A conjugate concave functional

7.12 Dual Optimization Problems

We come now to the application of conjugate functionals to optimization. Suppose we seek to minimize a convex functional over a convex domain. Or more generally, if f is convex over C and g is concave over D, suppose we seek

$$\inf_{C \cap D} [f(x) - g(x)].$$

In standard minimization problems, g is usually zero. But as we shall see, the present generalization is conceptually helpful. The general problem is illustrated in Figure 7.12. The problem can be interpreted as that of finding the minimum vertical separation of the sets $[f, C]$ and $[g, D]$. It is reasonably clear, from geometric intuition, that this distance is equal to the maximum vertical separation of two parallel hyperplanes separating $[f, C]$ and $[g, D]$. This relation, between a given minimization problem and an equivalent maximization problem, is a generalization of the duality principle for minimum norm problems.

Conjugate functionals are precisely what is needed for expressing this duality principle algebraically. Since $-f^*(x^*)$ is the vertical distance to a

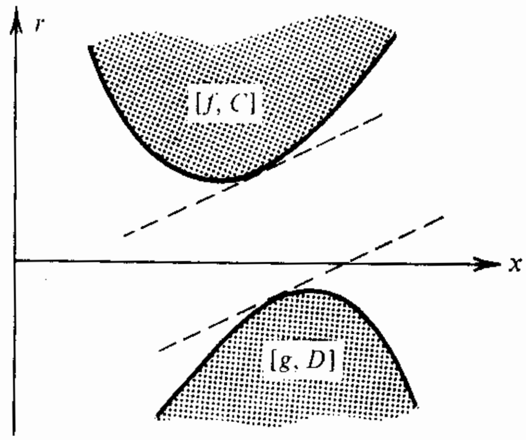

Figure 7.12

support hyperplane below $[f, C]$, and $-g^*(x^*)$ is the vertical distance to the parallel support hyperplane above $[g, D]$, $g^*(x^*) - f^*(x^*)$ is the vertical separation of the two hyperplanes. The duality principle is stated in detail in the following theorem.

Theorem 1. (*Fenchel Duality Theorem*) *Assume that f and g are, respectively, convex and concave functionals on the convex sets C and D in a normed space X. Assume that $C \cap D$ contains points in the relative interior of C and D and that either $[f, C]$ or $[g, D]$ has nonempty interior. Suppose further that $\mu = \inf_{x \in C \cap D} \{f(x) - g(x)\}$ it finite. Then*

$$\mu = \inf_{x \in C \cap D} \{f(x) - g(x)\} = \max_{x^* \in C^* \cap D^*} \{g^*(x^*) - f^*(x^*)\}$$

where the maximum on the right is achieved by some $x_0^* \in C^* \cap D^*$.

If the infimum on the left is achieved by some $x_0 \in C \cap D$, then

$$\max_{x \in C} [\langle x, x_0^* \rangle - f(x)] = \langle x_0, x_0^* \rangle - f(x_0)$$

and

$$\min_{x \in D} [\langle x, x_0^* \rangle - g(x)] = \langle x_0, x_0^* \rangle - g(x_0).$$

Proof. By definition, for all $x^* \in C^* \cap D^*$, $x \in C \cap D$,

$$f^*(x^*) \geq \langle x, x^* \rangle - f(x)$$
$$g^*(x^*) \leq \langle x, x^* \rangle - g(x).$$

Thus,

$$f(x) - g(x) \geq g^*(x^*) - f^*(x^*)$$

and hence

$$\inf_{C \cap D} [f(x) - g(x)] \geq \sup_{C^* \cap D^*} [g^*(x^*) - f^*(x^*)].$$

Therefore, the equality in the theorem is proved if an $x_0^* \in C^* \cap D^*$ can be found for which $\inf_{C \cap D} [f(x) - g(x)] = g^*(x_0^*) - f^*(x_0^*)$.

The convex set $[f - \mu, C]$ is a vertical displacement of $[f, C]$; by definition of μ the sets $[f - \mu, C]$ and $[g, C]$ are arbitrarily close but have disjoint relative interiors. Therefore, since one of these sets has nonempty interior, there is a closed hyperplane in $R \times X$ separating them. This hyperplane cannot be vertical because otherwise its vertical projection onto X would separate C and D. Since the hyperplane is not vertical, it can be represented as $\{(r, x) \in R \times X : \langle x, x_0^* \rangle - r = c\}$ for some $x_0^* \in X^*$ and $c \in R$. Now since $[g, D]$ lies below this hyperplane but is arbitrarily close to it, we have

$$c = \inf_{x \in D} [\langle x, x_0^* \rangle - g(x)] = g^*(x_0^*).$$

Likewise,

$$c = \sup_{x \in C} [\langle x, x_0^* \rangle - f(x) + \mu] = f^*(x_0^*) + \mu.$$

Thus $\mu = g^*(x_0^*) - f^*(x_0^*)$.

If the infimum μ on the left is achieved by some $x_0 \in C \cap D$, the sets $[f - \mu, C]$ and $[g, D]$ have the point $(g(x_0), x_0)$ in common and this point lies in the separating hyperplane. ∎

In typical applications of this theorem, we consider minimizing a convex functional f on a convex domain D; the set D representing constraints. Accordingly, we take $C = X$ and $g = 0$ in the theorem. Calculation of f^* is itself an optimization problem, but with $C = X$ it is unconstrained. Calculation of g^* is an optimization problem with a linear objective functional when $g = 0$.

Example 1. (An Allocation Problem) Suppose that there is a fixed quantity x_0 of some commodity (such as money) which is to be allocated among n distinct activities in such a way as to maximize the total return from these activities. We assume that the return associated with the i-th activity when allocated x_i units is $g_i(x_i)$ where g_i, due to diminishing marginal returns, is assumed to be an increasing concave function. In these terms the problem is one of finding $x = (x_1, x_2, \ldots, x_n)$ so as to

$$(1) \qquad \begin{cases} \text{maximize } g(x) = \sum_{i=1}^{n} g_i(x_i) \\ \text{subject to } \sum_{i=1}^{n} x_i = x_0 \qquad x_i \geq 0 \qquad = 1, 2, \ldots, n. \end{cases}$$

To solve this problem by conjugate functionals, we set D equal to the positive orthant, f equal to the zero functional, and $C = \{x : \sum_{i=1}^{n} x_i = x_0\}$. The functional conjugate to f is (for $y = (y_1, y_2, \ldots, y_n)$)

$$f^*(y) = \sup_{\Sigma x_i = x_0} y'x.$$

This is finite only if $y = \lambda(1, 1, \ldots, 1)$ in which case it is equal to λx_0. Thus

$$C^* = \{y : y = \lambda(1, 1, \ldots, 1)\}$$
$$f^*(\lambda(1, 1, \ldots, 1, 1)) = \lambda x_0 .$$

Also, for each i we define the conjugate functions (of a single variable)

(2) $$g_i^*(y_i) = \sup_{x_i \geq 0} [x_i y_i - g_i(x_i)]$$

and it is clear that

$$g^*(y) = \sum_{i=1}^{n} g_i^*(y_i).$$

The problem conjugate to (1) is therefore

(3) $$\min_{\lambda} \left[\lambda x_0 - \sum_{i=1}^{n} g_i^*(\lambda)\right].$$

In this form we note that to solve the allocation problem requires evaluation of the conjugate functions g_i^* and then solution of (3) which is minimization with respect to the single variable λ. Once the optimal value of λ is determined the x_i's which solve (1) can be found to be those which minimize (2) with each $y_i = \lambda$.

This analysis can be modified in order to apply to a multistage allocation problem where there is the possibility of investment growth of uncommitted resources. See Problem 19.

Example 2. (Horse-Racing Problem) What is the best way to place bets totaling x_0 dollars in a race involving n horses? Assume we know p_i, the probability that the i-th horse wins, and s_i, the amount that the rest of the public is betting on the i-th horse. The track keeps a proportion $1 - C$ of the total amount bet ($0 < 1 - C < 1$) and distributes the rest among the public in proportion to the amounts bet on the winning horse.

Symbolically, if we bet x_i on the i-th horse, $i = 1, 2, \ldots, n$, we receive

$$C\left(x_0 + \sum_{i=1}^{n} s_i\right) \frac{x_i}{s_i + x_i}$$

if the i-th horse wins. Thus the expected net return, R, is

$$R = C\left(x_0 + \sum_{i=1}^{n} s_i\right)\left\{\sum_{i=1}^{n} \frac{p_i x_i}{s_i + x_i}\right\} - x_0. \tag{4}$$

Our problem is to find x_i, $i = 1, 2, \ldots, n$, which maximize R subject to

$$\sum_{i=1}^{n} x_i = x_0, \qquad x_i \geq 0, \qquad i = 1, 2, \ldots, n \tag{5}$$

or, equivalently, to maximize

$$\sum_{i=1}^{n} g_i(x_i) \tag{6}$$

subject to (5), where

$$g_i(x_i) = \frac{p_i x_i}{s_i + x_i}.$$

This problem is exactly the type treated in Example 1 since each g_i is concave for positive x_i (as can be verified by observing that the second derivative is negative). Solution to the problem is obtained by calculating the functions conjugate to the g_i's.

A typical g_i is shown in Figure 7.13. Its slope at $x_i = 0$ is p_i/s_i and it approaches the value p_i as $x_i \to \infty$. The value of the conjugate functional g_i^* at λ is obtained by finding the lowest line of slope λ which lies above g_i. It is clear from the diagram that for $\lambda \geq p_i/s_i$, we have $g_i^*(\lambda) = 0$, and that for $\lambda < 0$, $g_i^*(\lambda)$ is not defined. For $0 < \lambda < p_i/s_i$, we have

$$g_i^*(\lambda) = \min_{x_i > 0} [\lambda x_i - g_i(x_i)]. \tag{7}$$

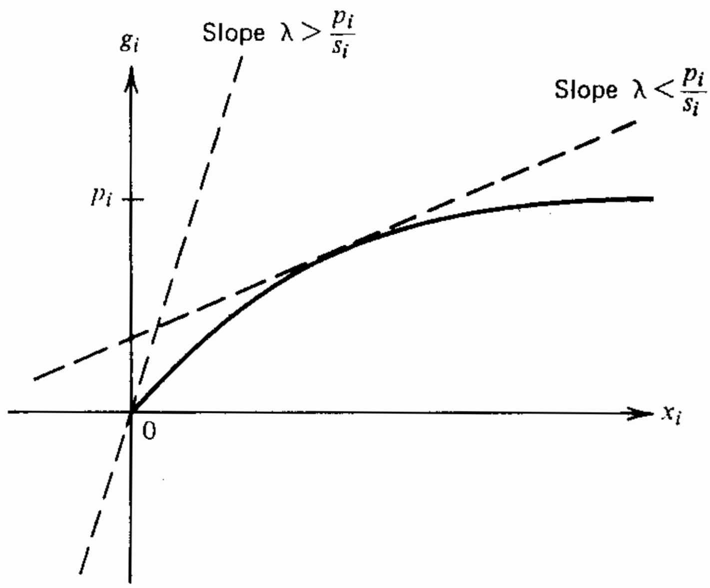

Figure 7.13

Performing the minimization by differentiation, we have the equation

$$\lambda = \frac{s_i p_i}{(s_i + x_i)^2}$$

or

(8) $$x_i = \sqrt{\frac{s_i p_i}{\lambda}} - s_i .$$

Substitution back into (7) yields

(9) $$g_i^*(\lambda) = \begin{cases} \dfrac{p_i x_i^2}{(s_i + x_i)^2} & \text{for } 0 < \lambda < \dfrac{p_i}{s_i} \\ 0 & \text{for } \lambda \geq \dfrac{p_i}{s_i} \end{cases}$$

where x_i is determined from equation (8).

We can now deduce the form of the answer. Suppose that λ has been found from (3). To simplify the notation, rearrange the indices so that $p_1/s_1 > p_2/s_2 > \cdots > p_n/s_n$ (assuming strict inequality). For the given λ, we define m as the largest index for which $p_i/s_i \geq \lambda$. Then, from equations (8) and (9), our solution may be written

(10) $$x_i = \begin{cases} \dfrac{\sqrt{s_i p_i}}{\lambda} - s_i & \text{for } i = 1, 2, \ldots, m \\ 0 & \text{for } i = m+1, \ldots, n. \end{cases}$$

The parameter λ in this section can be found from (3) or from the constraint (5). In other words, λ is chosen so that

(11) $$S(\lambda) = \sum_{p_i/s_i \geq \lambda} \left(\frac{\sqrt{s_i p_i}}{\lambda} \right) - s_i = x_0 .$$

Now $S(\lambda)$ is easily verified to be continuous and $S(0) = \infty$, $S(\infty) = 0$. Thus there is a λ satisfying (11).

Note that for small x_0 ($x_0 << \max s_i$), the total amount should be bet on a single horse, the horse corresponding to the maximum p_i/s_i, or equivalently, the maximum $p_i r_i$ where $r_i = C \sum_j s_j/s_i$ is the track odds.

Example 3. Consider the minimum energy control problem discussed in Sections 3.10 and 6.10. We seek the element $u \in L_2[0, 1]$ minimizing

$$f(u) = \tfrac{1}{2} \int_0^1 u^2(t)\, dt,$$

while satisfying the linear constraints

$$Ku = c$$

where $K : L_2[0, 1] \to E^n$.

In Theorem 1, let $C = L_2[0, 1]$, $g = 0$, and $D = \{u : Ku = c\}$. We assume that D is nonempty. Then we can readily verify that

$$C^* = L_2[0, 1]$$

and

$$f^*(u^*) = \tfrac{1}{2} \int_0^1 [u^*(t)]^2 \, dt.$$

Since $g = 0$, the conjugate functional of the set D is equal to the support functional. Thus

$$D^* = \{u^* : u^* = K^*a, \, a \in E^n\}$$

and $g^*(K^*a) = (a \mid c)$.

The dual problem is therefore the finite-dimensional problem in the vector a:

$$\max \{(a \mid c) - \tfrac{1}{2}(K^*a \mid K^*a)\}$$

which is solved by finding the n vector a satisfying

$$KK^*a = c$$

where KK^* is an $n \times n$ matrix.

Finally, the solution to the original problem can be found in terms of the solution of the dual by application of the second part of Theorem 1. Thus

$$u_0 = K^*a.$$

*7.13 Min-Max Theorem of Game Theory

In this section we briefly introduce the classical theory of games and prove the fundamental min-max theorem. Our purpose is to show that the min-max theorem can be regarded as an example of the Fenchel duality theorem.

Let X be a normed space and X^* its normed dual. Let A be a fixed subset of X and B a fixed subset of X^*. In the form of game that we consider, one player (player A) selects a vector from his strategy set A while his opponent (player B) selects a vector x^* from his strategy set B. When both players have selected their respective vectors, the quantity

$\langle x, x^* \rangle$ is computed and player A pays that amount (in some appropriate units) to player B. Thus A seeks to make his selection so as to minimize $\langle x, x^* \rangle$ while B seeks to maximize $\langle x, x^* \rangle$.

Assuming for the moment that the quantities

$$\mu^0 = \min_{x \in A} \max_{x^* \in B} \langle x, x^* \rangle$$

$$\mu_0 = \max_{x^* \in B} \min_{x \in A} \langle x, x^* \rangle$$

exist, we first take the viewpoint of A in this game. By selecting $x \in A$, he loses no more than $\max_{x^* \in B} \langle x, x^* \rangle$; hence, by proper choice of x, say x_0, he can be assured of losing no more than μ^0. On the other hand, player B by selecting $x^* \in B$, wins at least $\min_{x \in A} \langle x, x^* \rangle$; therefore, by proper choice of x^*, say x_0^*, he can be assured of winning at least μ_0. It follows that $\mu_0 \leq \langle x_0, x_0^* \rangle \leq \mu^0$, and the fundamental question that arises is whether $\mu_0 = \mu^0$ so that there is determined a unique pay-off value for optimal play by both players.

The most interesting type of game that can be put into the form outlined above is the classical finite game. In a finite game each player has a finite set of strategies and the pay-off is determined by a matrix Q, the pay-off being q_{ij} if A uses strategy i and B uses strategy j. For example, in a simple coin-matching game, the players independently select either "heads" or "tails." If their choices match, A pays B 1 unit while, if they differ, B pays A 1 unit. The pay-off matrix in this case is

$$Q = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Finite games of this kind usually do not have a unique pay-off value. We might, however, consider a long sequence of such games and "randomized" strategies where each player determines his play in any one game according to a fixed probability distribution among his choices. Assuming that A has n basic strategies, he selects an n vector of probabilities $x = (x_1, x_2, \ldots, x_n)$ such that $x_i \geq 0$, $\sum_{i=1}^n x_i = 1$. Likewise, if B has m strategies, he selects $y = (y_1, y_2, \ldots, y_m)$ such that $y_i \geq 0$, $\sum_{i=1}^m y_i = 1$. The expected (or average) pay-off is then $(x \mid Qy)$.

Defining $A = \{x : x_i \geq 0, \sum_{i=1}^n x_i = 1\} \subset E_n$, and $B = \{x^* : x^* = Qy, y_i \geq 0, \sum_{i=1}^m y_i = 1\} \subset E^n$, the randomized game takes the standard form given at the beginning of the section. Note that A and B are both bounded closed convex sets. Other game-type optimization problems with bilinear objectives other than the classical randomized finite game also take this form.

We now give a simple proof of the min-max theorem based on duality. For simplicity, our proof is for reflexive spaces, although more general versions of the result hold. For the generalizations, consult the references at the end of the chapter.

Theorem 1. (Min-Max) *Let X be a reflexive normed space and let A and B be compact convex subsets of X and X^*, respectively. Then*

$$\min_{x \in A} \max_{x^* \in B} \langle x, x^* \rangle = \max_{x^* \in B} \min_{x \in A} \langle x, x^* \rangle.$$

Proof. Define the functional f on X by

$$f(x) = \max_{x^* \in B} \langle x, x^* \rangle.$$

The maximum exists for each $x \in X$ since B is compact. The functional is easily shown to be convex and continuous on X. We seek an expression for

$$\min_{x \in A} f(x)$$

which exists by the compactness of A and the continuity of f. We now apply the Fenchel duality theorem with the associations: $f \to f$, $C \to X$, $g \to 0$, $D \to A$. We have immediately:

$$D^* = X^* \tag{1}$$

$$g^*(x^*) = \min_{x \in A} \langle x, x^* \rangle. \tag{2}$$

We claim that furthermore

$$C^* = B \tag{3}$$

$$f^*(x^*) = 0. \tag{4}$$

To prove (3) and (4), let $x_1^* \notin B$, and by using the separating hyperplane theorem, let $x_1 \in X$ and α be such that $\langle x_1, x_1^* \rangle - \langle x_1, x^* \rangle > \alpha > 0$ for all $x^* \in B$. Then $\langle x, x_1^* \rangle - \max_{x^* \in B} \langle x, x^* \rangle$ can be made arbitrarily large by taking $x = kx_1$ with $k > 0$. Thus

$$\sup_x [\langle x, x_1^* \rangle - f(x)] = \infty$$

and $x_1^* \notin C^*$.

Conversely, if $x_1^* \in B$, then $\langle x, x_1^* \rangle - \max_{x^* \in B} \langle x, x^* \rangle$ achieves a maximum value of 0 at $x = \theta$. This establishes (3) and (4).

The final result follows easily from the equality

$$\min_{x \in A} f(x) = \max_{x^* \in B \cap X^*} g^*(x^*) = \max_{x^* \in B} \min_{x \in A} \langle x, x^* \rangle. \quad \blacksquare$$

An interesting special case of the min-max theorem is obtained by taking B to be the unit sphere in X^*. In that case we obtain

$$\min_{x \in A} \|x\| = \max_{\|x^*\| \le 1} - h(x^*)$$

where h is the support functional of the convex set A. This result is the duality theorem for minimum norm problems of Section 5.8.

7.14 Problems

1. On the vector space of continuous functions on [0, 1], define

$$f(x) = \max_{0 \le t \le 1} x(t).$$

Determine for which x the Gateaux differential $\delta f(x; h)$ exists and is linear in h.

2. Repeat Problem 1 for

$$f(x) = \int_0^1 |x(t)|\, dt.$$

3. Show that the functional

$$f(x_1, x_2) = \begin{cases} \dfrac{x_1 x_2^{\,2}}{(x_1^{\,2} + x_2^{\,4})} & \text{if } x_1 \ne 0 \\ 0 & \text{if } x_1 = 0 \end{cases}$$

is Gateaux differentiable but not continuous at $x_1 = x_2 = 0$.

4. On the space $X = C[0, 1]$, define the functional $f(x) = [x(\frac{1}{2})]^2$. Find the Fréchet differential and Fréchet derivative of f.
5. Let φ be a function of a single variable having a continuous derivative and satisfying $|\varphi(\xi)| < K|\xi|$. Find the Gateaux differential of the functional $f(x) = \sum_{i=1}^{\infty} \varphi(\xi_i)$ where $x = \{\xi_i\} \in l_2$. Is this also a Fréchet differential?
6. Suppose the real-valued functional f defined on an open subset D of a normed space has a relative minimum at $x_0 \in D$. Show that if f is twice Gateaux differentiable at x_0, then $\langle h, f''(x)h \rangle \ge 0$ for all $h \in X$.
7. Let f be a real-valued functional defined on an open region D in a normed space X. Suppose that at $x_0 \in D$ the first Fréchet differential vanishes identically on X and within a sphere $S(x_0, \varepsilon)$, $f''(x)$ exists and is uniformly continuous in x, and the lower bound of $\langle h, f''(x)h \rangle$ for $\|h\| = 1$ is positive. Show that f obtains a relative minimum at x_0.
8. Let A be a nonempty subset of a normed space X and let $x_0 \in A$. Denote by $C(A, x_0)$ the closed cone generated by $A - x_0$, i.e., the

intersection of all closed cones containing $A - x_0$. The *local closed cone* of A at x_0 is the set

$$LC(A, x_0) = \bigcap_{N \in \mathcal{N}} C(A \cap N, x_0)$$

where $\mathcal{N}$ is the class of all neighborhoods of x_0. Suppose that over A the Fréchet differentiable functional f achieves a minimum at x_0. Show that

$$f'(x_0) \in [LC(A, x_0)]^{\oplus}.$$

9. In some problems in the calculus of variations, it is necessary to consider a broader class of functions than usual. Suppose that we seek to extremize

$$J = \int_a^b F[x, \dot{x}, t]\, dt$$

among all functions x for which $x(a) = A$, $x(b) = B$ and which have continuous derivatives on $[a, b]$ except possibly at a single point in $[a, b]$. If the extremum has a continuous derivative except at $c \in [a, b]$, show that $F_x = d/dt F_{\dot{x}}$ on the intervals $[a, c)$ and $(c, b]$ and that the functions $F_{\dot{x}}$ and $F - \dot{x}F_{\dot{x}}$ are continuous at c. These are called the Weierstrass-Erdman corner conditions. Apply these considerations to the functional

$$J[x] = \int_{-1}^{1} x^2(1 - \dot{x})^2\, dt, \qquad x(-1) = 0, \qquad x(1) = 1.$$

10. Consider the effect of an additional source of constant income in the estate-planning problem.
11. Let $f_1, f_2, \ldots, f_n$ be linearly independent linear functionals on a vector space X. Show that there are n elements $x_1, x_2, \ldots, x_n$ in X such that the $n \times n$ matrix $[f_i(x_j)]$ is nonsingular.
12. Formulate and solve the isoperimetric problem of Example 1 of Section 7.7 by using polar coordinates.
13. Solve the candidates allocation problem described in Chapter 1.
14. Let $X = L_2[0, 1]$ and define $f(x) = \int_0^1 \{\frac{1}{2}x^2(t) + |x(t)|\}\, dt$ on X. Find the conjugate functional of f.
15. Exhibit a convex set C having the property that the convex conjugate functional of 0 over C is not equal to the negative of the concave conjugate functional of 0 over C.
16. Let M be a nonempty closed convex set in $R \times X$, and assume that there is at least one nonvertical hyperplane containing M in one of its half-spaces. Show that for any $x_0 \notin M$ there is a nonvertical hyper-

plane separating x_0 and M. Hint: Consider convex combinations of hyperplanes.

17. Let f be a convex functional on a convex set C in a normed space and let $[f^*, C^*] = [f, C]^*$. For $x \in C$, $x^* \in C^*$, deduce Young's inequality
$$\langle x, x^* \rangle \leq f(x) + f^*(x^*).$$
Apply this result to norms in L_p spaces.

18. Derive the minimum norm duality theorem (Theorem 1, Section 5.8) directly from the Fenchel duality theorem.

19. Suppose a fixed quantity x_0 of resource is to be allocated over a given time at n equally spaced time instants. Thus $x_1 \leq x_0$ is allocated first. The remaining resource $x_0 - x_1$ grows by a factor a so that at the second instant $x_2 \leq a(x_0 - x_1)$ may be allocated. In general, the uncommitted resource grows by the factor a between each step. Show that a sequence of allocations $\{x_i\}_{i=1}^n$, $x_i \geq 0$ is feasible if and only if
$$a^{n-1}x_1 + a^{n-2}x_2 + \cdots + ax_{n-1} + x_n \leq a^n x_0 .$$
Hence, show how to generalize the result of Example 1, Section 7.12, to multistage problems.

20. The owner of a small food stand at the beach is about to order his weekend supply of food. Mainly, he sells ice cream and hot dogs and wishes to optimize his allocation of money to these two items. He knows from past experience that the demands for these items depend on the weather in the following way:

	Hot Day	Cool Day
Ice cream	1000	200
Hot dogs	400	200

He believes that a hot or a cool day is equally probable. Anything he doesn't sell he may return for full credit. His profit on ice cream is 10 cents and on hot dogs it is 30 cents. He, of course, wishes to maximize his expected profit while remaining within his budget of \$100. Ice cream costs him 10 cents and hot dogs 20 cents.

(a) Formulate his problem and reduce it to the form
$$\begin{aligned}&\text{maximize } f_1(x_1) + f_2(x_2)\\ &\text{subject to } x_1 + x_2 \leq x_0,\ x_1 \geq 0,\ x_2 \geq 0.\end{aligned}$$
(b) Solve the problem using conjugate functionals.

21. Consider a control system governed by the n-dimensional set of differential equations
$$\dot{x}(t) = Ax(t) + bu(t)$$

which has solution

$$x(T) = \Phi(T)x(0) + \int_0^T \Phi(T-t)bu(t)\,dt,$$

where $\Phi(t)$ is a fundamental matrix of solutions. Using conjugate function theory and duality, find the control u minimizing

$$J = \tfrac{1}{2}\|x(T)\|^2 + \tfrac{1}{2}\int_0^T u^2(t)\,dt.$$

Hint: Assume first that $x(T)$ is known and reduce the problem to a finite-dimensional one. Next optimize $x(T)$. Alternatively, formulate the problem in $E^n \times L_2[0, T]$.

REFERENCES

§7.1–4. For general background on differentials, see Graves [62], [63], Hildebrandt and Graves [72], and Hille and Phillips [73, Chapters 3 and 26]. A fairly comprehensive account of several extensions of ordinary calculus is contained in Luisternik and Sobolev [101]. The inequality extension of the mean value theorem is discussed by Kantorovich and Akilov [79]; also see Antosiewicz and Reinboldt [9].

§7.5–6. The classic treatise on the calculus of variations is Bliss [22]. Also see Fox [55]. For a somewhat more modern approach to the subject, see Gelfand and Fomin [58] or Akhiezer [2].

§7.7. The implicit function theorem used here has generalizations to abstract spaces. See Hildebrandt and Graves [72] and also Section 9.2 of this book. There are a number of abstract versions of the Lagrange multiplier theorem, e.g., Goldstine [60], [61] and Blum [23].

§7.9–12. Conjugate functionals on finite-dimensional space were introduced by Fenchel [52]. For an excellent presentation of this topic, consult Karlin [52]. Some extensions and related discussions are to be found in Rockafellar [126], Rådström [121], and Whinston [151]. Extensions to more general linear topological spaces were made by Brondsted [25], Dieter [41], Rockafellar [126], and Moreau [107]. See Kretschmer [90] for an example of a case where the values of the primal and dual problems are unequal. An equivalent approach, the maximum transform, was developed by Bellman and Karush [19]. They treat the allocation problem. The horse-racing problem is due to Rufus Isaacs and was analyzed in detail by Karlin using the Neyman-Pearson lemma.

§7.13. See Edwards [46], Karlin [81], [82], [83], Dresher [44], and McKinsey [104]. For more general min-max theorems, see Fan [51] and Sion [141].

§7.14. For further developments along the lines of Problem 8, see Varaiya [149]. For a solution to Problem 16, see Brondsted [25].