Convex Programming and Duality in Normed Space

DAVID G. LUENBERGER, MEMBER, IEEE

Abstract—**This paper describes some extensions of classical convex programming to problems formulated on an abstract normed space. The utility of the normed-space formulation is illustrated by some simple applications.**

I. Introduction

THE THEORY of convex programming, including the duality results of Fenchel and the Lagrange multiplier theorem of Kuhn and Tucker, have corresponding generalizations in infinite-dimensional spaces. These more general results enable one to employ many of the standard methods for optimization of finite-dimensional problems to problems which are infinite-dimensional in character, such as problems of optimal control of systems governed by differential equations and other optimization problems in which one seeks a function defined on an interval, rather than simply a vector in a finite-dimensional space.

A number of extensions of the finite-dimensional theory have been made. They differ primarily in the nature (or level of abstraction) of the underlying vector spaces, the degree of continuity or differentiability required of the functions involved and the nature of the constraints. One early generalization was given by Hurwicz,[1] and a simplified version of this together with applications was given by Lack.[2] A recent generalization to the differentiable case is Varaiya.[3] See also Russell.[4] Duality results for finite-dimensional problems were exploited by Everett,[5] and the infinite-dimensional theory was developed by Rockafellar,[6] Dieter,[7] and Ritter.[8]

In this paper attention is given primarily to the extension of the Lagrange multiplier method and its application. The development is formulated in normed space which represents a rather comfortable intermediate level of abstraction, and which covers almost all important applications.

II. Mathematical Background

In this section we review briefly the basic elements of functional analysis that are required to present the theory.

A. Normed Space

Definition: A *vector space* X is a set of elements called vectors together with two operations, addition $(+)$ and scalar multiplication $(\cdot)$, satisfying the following axioms:

Manuscript received August 14, 1967; revised March 6, 1968. This research was supported in part by Joint Services Contract Nonr 225(S3) and in part by NSF Grant NsF GK 1683. This paper was first presented at the 1967 Systems Science and Cybernetics Conference, Boston, Mass., October, 1967.

The author is with the Stanford Electronics Laboratories, Stanford University, Stanford, Calif. 94305.

1) $x + y = y + x$ for all $x, y \in X$
2) $(x + y) + z = x + (y + z)$
3) there is a null vector $\theta \in X$ such that $x + \theta = x$ for all $x \in X$
4) $\alpha(x + y) = \alpha x + \alpha y$ for all $x, y \in X$ and all scalars α
5) $(\alpha + \beta) x = \alpha x + \beta x$
6) $(\alpha\beta) x = \alpha(\beta x)$
7) $0 \cdot x = \theta$, $1 \cdot x = x$.

Definition: A *normed space* X is a vector space together with a real-valued function $\|x\|$ defined on the elements of X and satisfying:

1) $\|x\| \geq 0 \ \forall x \in X$, $\|x\| = 0$ if and only if $x = \theta$;
2) $\|\alpha x\| = |\alpha| \cdot \|x\| \ \forall x \in X$ and $\forall \alpha$ scalars;
3) $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\| \ \forall x_1, x_2$.

Example 1: The space R consisting of the real numbers becomes a normed space if the norm of a number x is defined as $|x|$.

Example 2: Euclidean n space E_n consisting of n tuples $x = (\xi_1, \xi_2, \ldots, \xi_n)$ of real numbers with $\|x\| = \left(\sum_{i=1}^{n} \xi_i^2\right)^{1/2}$ is a normed space.

Example 3: The space $C[t_0, t_1]$ consisting of all continuous functions on the interval $[t_0, t_1]$ with

$$\|x\| = \max_{t_0 \leq t \leq t_1} |x(t)|$$

is a normed space.

Example 4: The space $L_1[t_1, t_2]$ consists of all real-valued functions $x(t)$ on the interval $[t_1, t_2]$ that are Lebesque measurable and for which

$$\int_{t_1}^{t_2} |x(t)| dt < \infty.$$

The norm of an element $x \in L_1[t_1, t_2]$ is defined as

$$\|x\| = \int_{t_1}^{t_2} |x(t)| dt.$$

In this space two functions are considered to represent the same vector if they differ only on a set of measure zero.

In a normed space one introduces several topological properties including open and closed sets, convergence, completeness, etc. For the purposes of what follows, we need only the following two topological concepts.

Definition: An infinite sequence $\{x_i\}$ in a normed space X is said to *converge* to a limit $x \in X$ if $\|x_i - x\| \to 0$ as $i \to \infty$.

Definition: A subset S of a normed space is said to be *closed* if every convergent sequence from S has its limit in S; i.e., if $x_i \to x$ with $x_i \in S$ implies $x \in S$.

B. Dual Spaces

Given a normed space X, we denote by X^* the space of all continuous[1] linear functionals on X. The space X^* is normed by the equation

$$\|x^*\| = \sup_{\|x\|=1.} x^*(x).$$

Example 1: If $X = E_n$, then $X^* = E_n$ in the sense that corresponding to an element $x^* \in X^*$ there is a vector $y = (\eta_1, \eta_2, \ldots, \eta_n)$ such that for $x = (\xi_1, \xi_2, \ldots, \xi_n)$ we have

$$x^*(x) = \sum_{i=1}^{n} \eta_i \xi_i \text{ and } \|x^*\| = \|y\|.$$

Conversely, every $y \in E_n$ defines a unique element of X^* in this way.

Example 2: If $X = C[t_1, t_2]$, then every element of X^* can be represented by a function v of bounded variation in the sense that

$$x^*(x) = \int_{t_1}^{t_2} x(t)dv(t)$$

and $\|x^*\| =$ total variation of v.

C. Separating Hyperplane Theorem

A fundamental result, used extensively in optimization theory, is that under mild conditions two disjoint convex sets can be separated by a hyperplane.

Definition: A set K in a vector space is called *convex* if given $x, y \in K$, all elements of the form $\alpha x + (1 - \alpha) y \in K$ for $0 < \alpha < 1$.

Definition: A point x_0 is said to be an *interior point* of a set K in a normed space X if there is a sphere with center at x_0 contained in K, i.e., if there is an $\epsilon > 0$ such that $\|x - x_0\| < \epsilon \Longrightarrow x \in K$.

Definition: A *closed hyperplane* in a normed space is any set of the form $H = \{x \in X : x^*(x) = c\}$ where x^* is a fixed nonzero element of X^* and c is a fixed constant.

This definition of a hyperplane generalizes the familiar notion in finite-dimensional spaces. It can be shown, for example, that a closed hyperplane divides the space X into two disjoint regions.

With these definitions we now state the separating hyperplane theorem.

Theorem 1: Let K and L be convex sets in a real normed space X. Assume that K contains an interior point and that no interior point of K is contained in L. Then there is a hyperplane separating K and L; i.e., there is an element $x^* \in X^*$ and real constant c such that $x^*(k) \leq c \leq x^*(l)$ for all $k \in K, l \in L$.

D. Cones and Inequalities

By introducing a cone of vectors distinguished as the positive vectors of a space, it is possible to consider inequality problems in abstract vector spaces.

[1] A real-valued function f on a normed space is continuous if, given a sequence $\{x_i\}$ with $x_i \rightarrow x$, we have $f(x_i) \rightarrow f(x)$.

Definition: A set C in a vector space is a *cone* if $x \in C$ implies $\alpha X \in C$ for all $\alpha \geq 0$. A *convex cone* is a set that is both a cone and is convex.

A convex cone can be used to define the positive elements of a vector space. For example, in E_n the convex cone

$$P = \{x \in E_n : x = (\xi_1, \xi_2, \ldots, \xi_n); \xi_1 \geq 0 \text{ all } i\} \quad (1)$$

defines the ordinary positive orthant. In a space of functions defined on an interval of the real line, say $[t_1, t_2]$, it is natural to define the positive cone as consisting of all functions in the space that are non-negative everywhere on the interval $[t_1, t_2]$.

In a normed space it is sometimes important to define the positive cone by a closed convex cone. For example, in E_n the cone defined by (1) is closed, but if one or more of the inequalities in the description of the set were changed to strict inequality, the resulting cone would not be closed.

Now, given a convex cone P, taken to be the positive cone in a vector space, we write $x \geq y$ if $x - y \in P$, and the positive cone defines a partial ordering in the space. In the obvious way we define the negative cone N as $-P$ and write $y \leq x$ for $y - x \in N$.

In the case of a normed space we write $x > \theta$ if x is an interior point of the positive cone P. For many applications, it is essential that P possess an interior point so that the separating hyperplane theorem can be applied. Nevertheless, this is not possible in many common normed spaces. For instance, if $X = L_1[t_1, t_2]$ and P is taken as the subset of non-negative functions on the interval $[t_1, t_2]$, one can easily show that P contains no interior point. On the other hand, in $C[t_1, t_2]$ the cone of non-negative functions does possess interior points, and for this reason the space $C[t_1, t_2]$ is of particular interest for problems involving inequalities.

Given a normed space X together with a positive convex cone $P \subset X$, it is natural to define a corresponding positive convex cone $P^\oplus$ in the dual space X^* by

$$P^\oplus = \{x^* \in X^* : x^*(x) \geq 0 \quad \text{for all } x \in P\}. \quad (2)$$

Example 1: If P is taken as given by (1) in E_n, then $P^\oplus$ consists of all linear functionals represented by elements of E_n with non-negative components.

Example 2: If, in the space $C[t_1, t_2]$, P is taken as the set of all non-negative continuous functions on $[t_1, t_2]$, then $P^\oplus$ consists of all linear functionals on $C[t_1, t_2]$ represented by functions v of bounded variation and nondecreasing on $[t_1, t_2]$.

One can easily show that $P^\oplus$ is closed even if P is not. If P is closed, P and $P^\oplus$ are related through the following result, which can be proved using the separating hyperplane theorem.

Lemma 1: Let the positive cone P in the normed space X be closed. If $x \in X$ satisfies

$$x^*(x) \geq 0 \quad \text{for all } x^* \geq \theta$$

then $x \geq \theta$.

III. Convex Programming Problems

A. The Basic Theorem

Definition: Let X, Z be linear vector spaces and suppose that there is a cone $P \subset Z$ defined as the positive cone in Z. A function F defined on a convex subset $C \subset X$ with range in Z is said to be *convex* if $F(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha F(x_1) + (1 - \alpha)F(x_2)$ for all $x_1, x_2 \in C$ and all α, $0 \leq \alpha \leq 1$.

In the case of a real-valued function (i.e., the case where Z is the real line), a function f is convex if $f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha) f(x_2)$, where inequality has the usual interpretation for real numbers. For example, any norm defined on a vector space is a real convex function.

The basic problem considered in the remainder of this paper is to

$$\text{minimize } f(x) \tag{3}$$

subject to

$$x \in \Omega, \; G(x) \leq \theta$$

where Ω is a convex subset of a vector space X, f is a real-valued convex function on Ω, and G is a convex function from Ω into a normed space Z having positive cone P.

To solve this problem, let us embed it in a family of problems in the following way.

Let $\Gamma \subset Z$ be defined as $\Gamma = \{z: \exists x \in \Omega, G(x) \leq z\}$. It is easily verified that Γ is a convex set. On the set Γ, we define the primal function $\omega(z)$ by $\omega(z) = \inf \{f(x): x \in \Omega, G(x) \leq z\}$. The original problem (3) can be regarded as that of determining the single value $\omega(\theta)$.

Lemma 2: The functional $\omega(z)$ is convex.

Proof:

$$\begin{aligned}
\omega(\alpha z_1 + (1 - \alpha) z_2) &= \inf \{f(x): x \in \Omega, G(x) \leq \alpha z_1 + (1 - \alpha) z_2\} \\
&\leq \inf \{f(\alpha x_1 + (1 - \alpha)x_2): x_1, x_2 \in \Omega, G(x_1) \leq z_1, \\
&\qquad G(x_2) \leq z_2\} \\
&\leq \alpha \inf \{f(x_1): x_1 \in \Omega, G(x_1) \leq z_1\} \\
&\qquad + (1 - \alpha) \inf \{f(x_2): x_2 \in \Omega, \\
&\qquad G(x_2) \leq z_2\} \leq \alpha\omega(z_1) \\
&\qquad + (1 - \alpha)\, \omega(z_2).
\end{aligned}$$

Lemma 3: The functional $\omega(z)$ is decreasing, i.e., if $z_1 \geq z_2$. then $\omega(z_1) \leq \omega(z_2)$.

Proof: This is immediate.

A typical $\omega(z)$ for Z one dimensional is shown in Fig. 1.

Conceptually, the Lagrange multiplier theorem follows from the simple observation that since ω is convex, there is a hyperplane tangent to $\omega(z)$ at $z = \theta$ and lying below $\omega(z)$ throughout its region of definition. Our task is merely to translate this observation into a precise mathematical result. The proof for a general normed space Z exactly parallels the proof for finite-dimensional space.

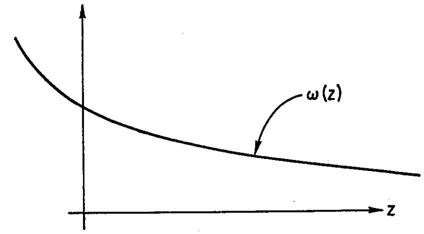


Fig. 1.

In the proof of the main theorem we consider the normed space $R \times Z$, the product of the real line with the normed space Z. This space consists of all ordered pairs (r, z) with $r \in R$, $z \in Z$. The norm on the space is defined as $\|(r, z)\| = |r| + \|z\|$.

Theorem 2: Let X be a linear vector space, Z a normed space, Ω a convex subset of Z, and P the positive cone in Z. Assume that P contains an interior point.

Let f be a real-valued convex function on Ω, and let G be a convex function from Ω into Z. Assume the existence of a point $x_1 \in \Omega$ for which $G(x_1) < \theta$, i.e., $G(x_1)$ is an interior point of $N = -P$.

Let

$$\mu_0 = \inf f(x)$$

subject to $x \in \Omega$, $G(x) \leq \theta$ and assume μ_0 is finite. Then there is an element $z_0^* \geq \theta$ in Z^* such that

$$\mu_0 = \inf_{x \in \Omega} \{f(x) + z_0^* G(x)\}.$$

Furthermore, if the infimum is achieved in the first equation by an $x_0 \in \Omega$, $G(x_0) \leq \theta$, it is achieved by x_0 in the second and

$$z_0^* G(x_0) = 0.$$

Proof: In the space $W = R \times Z$ define the sets

$$\begin{aligned}
A &= \{(r, z): r \geq f(x), z \geq G(x) \quad \text{for some } x \in \Omega\} \\
B &= \{(r, z): r \leq \mu_0, z \leq \theta\}.
\end{aligned}$$

Since f and G are convex, both A and B are convex sets. (It should be noted that the set A is the convex region above the graph of the primal function ω.) The definition of μ_0 implies that A contains no interior points of B. Also, since N contains an interior point, the set B contains an interior point. Thus according to the separating hyperplane theorem, there is a nonzero element $w_0^* = (r_0, z_0^*) \in W^*$ such that

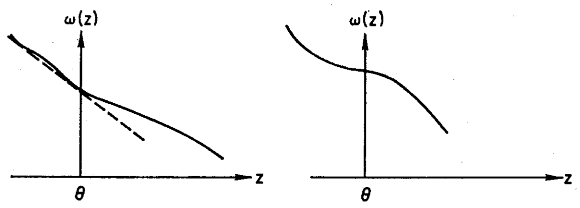
$$r_0 r_1 + z_0^*(z) \geq r_0 r_2 + z_0^*(z_2)$$

for $(r_1, z_1) \in A$, $(r_2, z_2) \in B$.

From the nature of B it follows immediately that $w_0^* \geq \theta$ or, equivalently, that

$$r_0 \geq 0$$

$$z_0^* \geq \theta.$$



(a) (b)

Fig. 2.

We shall show that $r_0 > 0$. The point (μ_0, θ) is in B, hence

$$r_0 r + z_0^*(z) \geq r_0 \mu_0$$

for all $(r, z) \in A$. If r_0 were zero, it would follow in particular that

$$z_0^* G(x_1) \geq 0$$

and that $z_0^* \neq \theta$. However, since $G(x_1)$ is an interior point of N and $z_0^* \geq \theta$, it follows that $z_0^* G(x_1) < 0$, which is a contradiction. Therefore, $r_0 > 0$, and without loss of generality we may assume $r_0 = 1$.

Since the point (μ_0, θ) is arbitrarily close to A and B, we have (with $r_0 = 1$)

$$\mu_0 = \inf_{(r,z) \in A} [r + z_0^*(z)] \leq \inf_{x \in \Omega} [f(x) + z_0^* G(x)]$$
$$\leq \inf_{\substack{x \in \Omega \\ G(x) \leq \theta}} f(x) = \mu_0.$$

Hence the first part of the theorem is proved.

If there exists an x_0 such that $G(x_0) \leq \theta$, $\mu_0 = f(x_0)$, then

$$\mu_0 \leq f(x_0) + z_0^* G(x_0) \leq f(x_0) = \mu_0$$

and hence $z_0^* G(x_0) = 0$.

Theorem 2 is a geometric version of the Kuhn–Tucker theorem for convex problems. An equivalent algebraic formulation of the results is given by the following saddle-point result.

Corollary: Let everything be as in the preceding and assume that x_0 achieves the constrained minimization. Then there is a $z_0^* \geq \theta$ such that the Lagrangian

$$L(x, z^*) = f(x) + z^* G(x)$$

has a saddle point at x_0, z_0^*, i.e.,

$$L(x_0, z^*) \leq L(x_0, z_0^*) \leq L(x, z_0^*)$$

for all $x \in \Omega$, $z^* \geq \theta$.

Proof: We leave this to the reader.

B. Sufficiency Conditions

The conditions of convexity and existence of interior points cannot be omitted if we are to guarantee the existence of a separating hyperplane in the space $R \times Z$. If, however, the appropriate hyperplane does exist, despite nonconvexity, the Lagrange technique for location of the optimal still applies. The situation is illustrated in Fig. 2, where

$$\omega(z) = \inf \{f(x) : x \in \Omega,\ G(x) \leq z\}$$

is again plotted but the convexity of f and G is not assumed. If, as in Fig. 2(a), an appropriate hyperplane exists, it is fairly clear that $f(x) + z_0^* G(x)$ attains a minimum at x_0. In Fig. 2(b) no supporting hyperplane exists at $z = \theta$, and the Lagrange statement cannot be made.

These observations lead to the following sufficiency theorems.

Theorem 3: Let f be a real-valued function defined on a subset Ω of a linear space X. Let G be a mapping from Ω into the normed space Z having closed nonempty positive cone P.

Suppose there exists a $z_0^* \in Z^*$, $z_0^* \geq \theta$ and an element $x_0 \in \Omega$ such that

$$f(x_0) + z_0^* G(x_0) \leq f(x) + z_0^* G(x)$$

all $x \in \Omega$. Then x_0 minimizes $f(x)$ subject to $G(x) \leq G(x_0)$, $x \in \Omega$.

Proof: Suppose there is an $x_1 \in \Omega$ with $f(x_1) < f(x_0)$, $G(x_1) \leq G(x_0)$. Then since $z_0^* \geq \theta$, it follows that

$$z_0^* G(x_1) \leq z_0^* G(x_0)$$

and hence that

$$f(x_1) + z_0^* G(x_1) < f(x_0) + z_0^* G(x_0),$$

which contradicts the hypothesis of the theorem.

Theorem 4: Let X, Z, Ω, P, f, G be as above. Suppose there exists a $z_0^* \in Z^*$, $z_0^* \geq \theta$ and $x_0 \in \Omega$ such that the Lagrangian $L(x, z^*) = f(x) + z^* G(x)$ possesses a saddle point at x_0, z_0^*, i.e.,

$$L(x_0, z^*) \leq L(x_0, z_0^*) \leq L(x, z_0^*)$$

all $x \in \Omega$, $z^* \geq \theta$. Then x_0 minimizes $f(x)$ subject to $G(x) \leq \theta$, $x \in \Omega$.

Proof: The saddle-point condition with respect to z^* gives

$$z^* G(x_0) \leq z_0^* G(x_0)$$

for all $z^* \geq \theta$. Hence in particular for all $z_1^* \geq \theta$

$$(z_1^* + z_0^*) G(x_0) \leq z_0^* G(x_0)$$

or

$$z_1^* G(x_0) \leq 0.$$

We conclude by Lemma 2 that $G(x_0) \leq \theta$. The saddle-point condition, therefore, implies $z_0^* G(x_0) = 0$.

Assume now that $x_1 \in \Omega$ and that $G(x_1) \leq \theta$. Then by the saddle-point condition with respect to x

$$f(x_0) = f(x_0) + z_0^* G(x_0) \leq f(x_1) + z_0^* G(x_1) \leq f(x_1).$$

Thus x_0 minimizes $f(x)$ subject to $x \in \Omega$, $G(x) \leq \theta$.

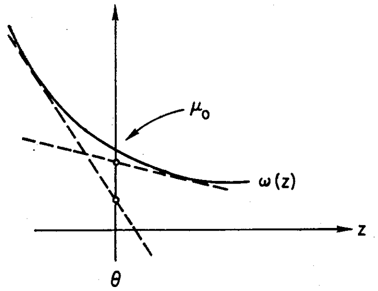


Fig. 3.

C. Duality

Associated with the primal problem (3) is a dual optimization problem formulated in the space Z^*. This dual problem is most easily visualized in terms of a family of hyperplanes in the space $R \times Z$.

As before, let

$$\mu_0 = \inf f(x)$$

subject to $G(x) \leq \theta$, $x \in \Omega$, where f and G are convex. The duality theorem is equivalent to the statement that μ_0 is the maximum intercept of all closed hyperplanes that lie below the function $\omega(z)$. This is illustrated in Fig. 3.

Theorem 5: Let f be a real-valued convex function defined on a convex subset Ω of a vector space X, and let G be a convex mapping of X into the normed space Z. Suppose there exists an x_1 such that $G(x_1) < \theta$. Then

$$\inf_{\substack{G(x) \leq \theta \\ x \in \Omega}} f(x) = \mu_0 = \max_{z^* \geq \theta} \inf_{x \in \Omega} \{f(x) + z^*G(x)\}$$

and the maximum on the right is achieved by some $z_0^* \geq \theta$.

If the infimum on the left is achieved by some $x_0 \in \Omega$, then

$$z_0^*G(x_0) = 0$$

and x_0 minimizes $f(x) + z_0^*G(x)$, $x \in \Omega$.

Proof: For any $z^* \geq \theta$ we have

$$\inf_{x \in \Omega} [f(x) + z^*G(x)] \leq \inf_{\substack{x \in \Omega \\ G(x) \leq \theta}} [f(x) + z^*G(x)]$$

$$\leq \inf_{\substack{x \in \Omega \\ G(x) \leq \theta}} f(x)$$

$$= \mu_0.$$

Therefore, the right-hand side of the equation in the theorem statement is less than or equal to μ_0. However, Theorem 2 establishes the existence of an element z_0^* which gives equality. The remainder of the theorem statement is given in Theorem 2.

IV. Applications

In this section we discuss three applications of the results developed in the previous sections. The applications are of some interest in their own right in that they represent well-known or important problems, but their interest here is primarily as vehicles for illustrating the scope of applicability of programming results in normed space.

A. Linear Programming

First, we consider an example in finite-dimensional space. Theorem 5 can be employed to produce a short derivation of the duality theorem of linear programming. We consider the primal problem. Find

$$\mu_0 = \min b'x$$

subject to

$$Ax \geq c, x \geq \theta$$

where

- x unknown n vector
- b given n vector
- c given m vector
- A given $m \times n$ matrix.

We assume that there is an $x > 0$ with $Ax > c$. We associate $f(x) = b'x$, $G(x) = c - Ax$, $\Omega = P$ in the general duality theorem and obtain

$$\mu_0 = \max_{\lambda \geq \theta} \min_{x \geq \theta} \{b'x + \lambda'[c - Ax]\}.$$

The minimization over x is finite only if $A'\lambda \leq b$, in which case the minimum is $\lambda'c$. Hence the dual reduces to

$$\mu_0 = \max \lambda'c$$

subject to $A'\lambda \leq c$, $\lambda \geq \theta$, which is the standard dual linear programming problem.

B. Optimal Control

Consider a system governed by the set of differential equations

$$\dot{x}(t) = A(t)x(t) + b(t)u(t) \tag{4}$$

where

- $x(t)$ $n \times 1$ state vector
- $A(t)$ $n \times n$ matrix
- $b(t)$ $n \times 1$ distribution matrix
- $u(t)$ scalar control.

Given an initial state $x(t_0)$, we seek the control $u_0(t)$ on $[t_0, t_1]$ minimizing

$$J = \frac{1}{2}\int_{t_0}^{t_1} u^2(t)dt \tag{5}$$

while satisfying the terminal inequalities

$$x(t_1) \geq c$$

where c is a fixed $n \times 1$ vector.

We can write the solution to (4) in the form

$$x(t_1) = \Phi(t_1, t_0)x(t_0) + \int_{t_0}^{t_1} \Phi(t_1, t)b(t)u(t)dt \tag{6}$$

where Φ is the fundamental solution matrix of the corresponding homogeneous equation. With this observation the original problem can now be expressed as follows: minimize

$$J = \frac{1}{2}\int_{t_1}^{t_2} u^2(t)dt \tag{7}$$

subject to

$$Ku \geq d \tag{8}$$

where $d = c - \Phi(t_1, t_0)\, x(t_0)$ and K is the integral operator defined as

$$Ku = \int_{t_0}^{t_1} \Phi(t_1, t)b(t)u(t)dt.$$

This is a convex programming problem defined on, say, the vector space of square-integrable functions on the interval $[t_0, t_1]$ with the constraint space being finite dimensional. We show below that this infinite-dimensional problem can be reduced to a finite-dimensional one.

Denoting the minimum of (7) under the constraints (8) by μ_0, Theorem 5 gives

$$\mu_0 = \max_{\lambda \geq \theta} \min_u \{J(u) + \lambda'(d - Ku)\}$$

where λ is an $n \times 1$ vector. More explicitly,

$$\mu_0 = \max_{\lambda \geq 0} \min_u \int_{t_0}^{t_1} \left[\frac{1}{2} u^2(t) - \lambda'\Phi(t_1, t)b(t)u(t)\right]dt + \lambda' d \tag{9}$$

and hence

$$\mu_0 = \max_{\lambda \geq \theta} \lambda' Q \lambda + \lambda' d \tag{10}$$

where

$$Q = -\frac{1}{2}\int_{t_0}^{t_1} \Phi(t_1, t)b(t)b'(t)\Phi'(t_1, t)dt.$$

Problem (10) is a simple finite-dimensional maximization problem. Once the solution λ_0 is determined, the optimal control $u_0(t)$ is then given by the function that minimizes the corresponding term in (9). Thus

$$u_0(t) = \lambda_0'\Phi(t_1, t)b(t).$$

C. An Allocation Problem[2]

Consider a farmer who produces a single crop, say, wheat. After harvesting his crop, he may store it or reinvest it to produce more wheat. The farmer wishes to maximize the total amount stored up to time T.

Letting $x_1(t)$ be the rate of production, $x_2(t)$ the rate of reinvestment, and $x_3(t)$ the rate of storage, and assuming

$$\dot{x}_1(t) = x_2(t) \quad x_1(0) = x_0 > 0$$

$$x_2(t) + x_3(t) = x_1(t)$$

$$x_1(t) \geq 0, \quad x_2(t) \geq 0, \quad x_3(t) \geq 0,$$

the farmer wishes to operate so as to

$$\text{maximize} \int_0^T x_3(t)dt$$

subject to the constraints.

The problem can be expressed entirely in terms of $x_2(t)$ as

$$\text{maximize} \int_0^T \left\{\int_0^t x_2(\tau)d\tau - x_2(t)\right\} dt$$

subject to

$$x_2(t) \geq 0$$

$$x_1(t) = x_0 + \int_0^t x_2(\tau)d\tau \geq x_2(t).$$

This is an infinite-dimensional linear programming problem, and, hence, we expect the solution to be on the boundary of the constraint set. The problem can be regarded as formulated with $x_2 \in C[0, T]$ and $Z = C[0, T]$. The interior point condition on the constraint $x_1(t) \geq x_2(t)$ is then satisfied. Employing the representation of the positive cone in the dual of $C[0, T]$, the Lagrange vector is a nondecreasing function v of bounded variation on $[0, T]$. The optimal v is the solution to the dual problem

$$\min_{v \nearrow} \sup_{x_2 \geq 0} L(x_2, v)$$

where

$$L(x_2, v) = \int_0^T \left\{\int_0^t x_2(\tau)d\tau - x_2(t)\right\} dt + \int_0^T \left\{x_0 + \int_0^t x_2(\tau)d\tau - x_2(t)\right\} dv(t).$$

Integration by parts reduces the above expression to

$$L(x_2, v) = \int_0^T \{(T - t - 1) + [v(T) - v(t)]\}\, x_2(t)dt + \int_0^T [x_0 - x_2(t)]dv(t).$$

The supremum of $L(x_2, v)$ over all $x_2 \geq 0$ is finite if and only if

$$\int_0^T [T - t - 1 + v(T) - v(t)]dt - \int_0^T dv(t)$$

is nonincreasing, in which case

$$\sup_{x_2 \geq 0} L(x_2, v) = \int_0^T x_0 dv(t).$$

[2] The problem considered in this section has been used by Berkovitz[9] as an example that can be solved by the maximum principle.

The minimizing v is thus found by inspection to satisfy

$$v(t) = 0, \quad t > T - 1$$

$$-\frac{dv(t)}{dt} = v(t) + t + 1 - T, v(T-1) = 0,$$

$$0 < t < T - 1.$$

There is no optimal $x_2 \in C[0, T]$, but the supremum is achieved by the discontinuous function

$$x_2(t) = \begin{cases} x_1(t) & 0 \leq t < T - 1 \\ 0 & T - 1 < t \leq T. \end{cases}$$

Hence the farmer reinvests everything until $t = T - 1$, after which he stores everything.

References

[1] L. Hurwicz, "Programming in linear spaces," in *Studies in Linear and Non-Linear Programming*, K. J. Arrow, L. Hurwicz, and H. Uzawa, Eds. Stanford, Calif.: Stanford University Press, 1958.

[2] G. N. T. Lack, "Optimization studies with applications to planning in the electrical power industry and optimal control theory," Institute in Engineering-Economic Systems, Stanford University, Stanford, Calif., Rept. CCS-5, August 1965.

[3] P. P. Varaiya, "Nonlinear programming in Banach space," *J. SIAM*, vol. 15, pp. 284–293, March 1967.

[4] D. L. Russell, "The Kuhn–Tucker conditions in Banach space with an application to control theory," *J. Math. Anal. and Appl.*, vol. 15, pp. 200–212, 1966.

[5] H. Everett, III, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Operations Research*, vol. 11, pp. 399–417, 1963.

[6] R. T. Rockafellar, "Extension of Fenchel's duality theorem for convex functions," *Duke Math. J.*, vol. 33, pp. 81–90, 1966.

[7] U. Dieter, "Dual extremal problems in locally convex linear spaces," *Proc. Colloq. on Convexity* (Copenhagen, Denmark, 1965), pp. 52–57.

[8] K. Ritter, "Duality for nonlinear programming in a Banach space," *J. SIAM*, vol. 15, pp. 294–302, March 1967.

[9] L. D. Berkovitz, unpublished lecture.

Reprinted from IEEE TRANSACTIONS ON SYSTEMS SCIENCE AND CYBERNETICS
Vol. SSC-4, No. 2, July 1968
pp. 182–188