# Minimax Theorems in Optimization and Approximation

Shinji TANIMOTO

*Kochi Women's University, Kochi 780, Japan*

(Accepted November 18, 1985)

**Abstract.** A classical minimax theorem was first proved by von Neumann in 1937. Since then a lot of variants concerning this theorem have been presented by many authors. In this paper a generalized minimax theorem is proved by means of the concept of formal linear combinations which was introduced in the author's previous paper, and is further transformed into another version. Some applications of minimax theorems are also given to the areas of optimization theory and approximation theory.

## 1. Introduction

The purpose of this paper is to extend some results obtained in the author's recent papers[6),7)] in several aspects. In the first paper[6)], a generalized minimax theorem is proved and applied to uniform (or Chebyshev) approximation by elements of a finite-dimensional subspace. In the second one[7)], based on this, a duality theory is developed in the areas of optimization and best approximation. An alternative proof is also given for the generalized minimax theorem by means of formal linear combinations.

First, in the present paper, we replace the compactness condition imposed on a set by a weaker one and establish a minimax theorem which is also a generalization of the previous one.

We shall next proceed to a study of another form of the generalized minimax theorem. A special type of it has been discussed in detail[7)]. We have called it a dual problem and obtained a duality theorem. However, we have treated there a minimization problem by elements of a finite-dimensional closed convex set. Our corresponding result is valid without such a condition.

The remainder of the paper is devoted to the study of duality theory in optimization, optimal control and best approximation. We relate the duality with a minimax theorem that is proved by Pomerol[3)]. More explicitly, we shall consider again the duality to optimization problems of the minimax type which was discussed by the author[4),5)]. We reveal the reason why dual problems to optimization problems of the minimax type become problems of the max type and *vice versa*. Finally, we shall deal with a problem of approximating a given point by elements in a subspace of a normed linear space, and a duality theorem related to it.

## 2. A generalized minimax theorem

Let $U$ be a convex subset of a real linear topological space and $V$ an arbitrary non-empty set. As is shown by the auther[7)], the set $V$ can be embedded into a real linear space by means of the concept of *formal linear combinations*. The set of all formal linear combinations forms a linear space generated by elements of the set $V$.

In this section, we shall extend a minimax theorem[6], replacing compactness condition on $U$ by inf-compactness. A function $f$ is said to be *inf-compact*, if for all $\mu \in R$, the lower level sets $\{x \mid f(x) \leq \mu\}$ are relatively compact.

Let us define the sets $\overline{V}_n$ for all positive integers $n$ :

$$\overline{V}_n = \{ (\overline{\lambda}_n, \overline{v}_n) \mid \overline{\lambda}_n = (\lambda_1, \dots, \lambda_n),\ \overline{v}_n = (v_1, \dots, v_n), \sum_{i=1}^{n} \lambda_i = 1,$$
$$\lambda_i > 0,\ v_i \in V\ (i=1, \dots, n) \}.$$

THEOREM 1. *Let $U$ be a convex svbset of a real linear topological space and $V$ an arbitrary non-empty set. Assume that a real-valued function $J$ on $U \times V$ satisfies the following two conditions :*

(i) *$J(.,v) : U \to R$ is convex and lower semi-continuous or all $v \in V$ ;*

(ii) *for some $(\overline{\lambda}_n, \overline{v}_n) \in \overline{V}_n$, the function defined by*

$$\sum_{i=1}^{n} \lambda_i J(.,v_i) : U \to R$$

*is inf-compact. Then the equality*

$$\min_{u \in U} \sup_{v \in V} J(u,v) = \lim_{n \to \infty} \sup_{(\overline{\lambda}_n, \overline{v}_n) \in \overline{V}_n} \inf_{u \in U} \sum_{i=1}^{n} \lambda_i J(u, v_i)$$

*holds*.

PROOF. We denote by $\boldsymbol{V}$ the set of all formal linear combinations, that is, all possible expressions of the form :

$$\boldsymbol{v} = \sum_{i=1}^{n} \lambda_i v_i, \qquad \lambda_i \in R,\ v_i \in V\ (i=1, \dots, n),$$

where $n$ ranges over all positive integers. Moreover, we denote by $\boldsymbol{V}_0$ the subset of $\boldsymbol{V}$, which contains all possible expressions of the form :

$$\boldsymbol{v} = \sum_{i=1}^{n} \lambda_i v_i, \quad \sum_{i=1}^{n} \lambda_i = 1, \quad \lambda_i \geqq 0,\ v_i \in V\ (i = 1, \dots, n).$$

Thus, $\boldsymbol{V}_0$ becomes a convex subset of the linear space $\boldsymbol{V}$. Next we define the real-valued function on $U \times \boldsymbol{V}_0$ by

$$\boldsymbol{J}(u, \boldsymbol{v}) = \sum_{i=1}^{n} \lambda_i J(u, v_i),$$

when $u \in U$, $\boldsymbol{v} = \sum_{i=1}^{n} \lambda_i v_i \in \boldsymbol{V}_0$. Then $\boldsymbol{J}(u, \boldsymbol{v})$ satisfies all the conditions of the "lopsided minimax theorem"[2]. Therefore, we obtain the equality

$$\min_{u \in U} \sup_{\boldsymbol{v} \in \boldsymbol{V}_0} \sum_{i=1}^{n} \lambda_i J(u, v_i) = \min_{u \in U} \sup_{v \in V} J(u,v)$$
$$= \sup_{\boldsymbol{v} \in \boldsymbol{V}_0} \inf_{u \in U} \sum_{i=1}^{n} \lambda_i J(u, v_i)$$

$$= \lim_{n\to\infty} \sup_{(\bar{\lambda}_n, \bar{v}_n) \epsilon \bar{V}_n} \inf_{u \epsilon U} \sum_{i=1}^{n} \lambda_i J(u, v_i)$$

and, moreover, the minimum of the left-hand side is really attained at some point $u \epsilon U$. This completes the proof.

## 3. Alternative form of the minimax theorem

The notation is the same as in the preceding section. We assume now that the set $U$ is compact. Then, the result mentioned in Theorem 1 remains valid. With $U$ a cmpact set, it was first proved by the author[6]. Based on this, we shall establish another form of it. The underlying idea comes from a theorem in approximation theory, which indicates an algorithm for approximating a continuous function by polynomials of a prescribed degree. Its generalization has been proved by the author[7]. In this section, we shall prove one more result related to it.

THEOREM 2. *Let $U$ be a compact convex subset of a real linear topological space and $V$ a non-empty arbitrary set. Let $J$ be a function defined on $U\times V$ satisfying condition* (i) *of Theorem* 1. *Then the equality*

$$\min_{u \epsilon U} \sup_{v \epsilon V} J(u,v) = \sup_{\bar{v} \epsilon \boldsymbol{F}} (\min_{u \epsilon U} \max_{v_i \epsilon \bar{v}} J(u, v_i))$$

*holds true, where $\boldsymbol{F}$ denotes the set of all finite subsets $\bar{v} = \{v_i\}_{i=1}^{n}$ of $V$.*

PROOF. With $\bar{v} = \{v_i\}_{i=1}^{n}$ in place of $V$ in Theorem 1, we have

$$\min_{u \epsilon U} \max_{v_i \epsilon \bar{v}} J(u, v_i) = \max_{\bar{\lambda}_n \epsilon \Lambda_n} \min_{u \epsilon U} \sum_{i=1}^{n} \lambda_i J(u, v_i),$$

where, for each positive integer $n$, $\Lambda_n$ denotes the subset of $\boldsymbol{R}^n$ defined by

$$\Lambda_n = \{\bar{\lambda}_n = (\lambda_1, \ldots, \lambda_n) \mid \sum_{i=1}^{n} \lambda_i = 1, \ \lambda_i \geqq 0 \text{ for all } i\}.$$

Therefore, The right-hand side of the equality in this Theorem takes on the form

$$\lim_{n\to\infty} \sup_{(\bar{\lambda}_n, \bar{v}_n) \epsilon \bar{V}_n} \min_{u \epsilon U} \sum_{i=1}^{n} \lambda_i J(u, v_i).$$

Hence, by "Theorem 1"[6] we have the the desired equality. This completes the proof.

In a previous paper[7] the author showed that if $U$ lies in an $n$-dimensional space, one can take $\boldsymbol{F}$ as the set of all $(n+1)$-tuples of elements of $V$. Moreover, the author discussed relations between optimal solutions of both sides of the equality appearing in Theorem 2. More explicitly, it was shown that an optimal solution of either side is also an optimal solution of the other. Unfortunately, such an assertion cannot be deduced in the present case.

If $V$ is a non-empty convex subset of a real linear space and $J(u, .)$ is a concave function on $V$ for each $u \epsilon U$, then one can let $\boldsymbol{F}$ consist of singletons of the set $V$. This follows from a well-known classical minimax theorem.[2]

A similar equality presented in the above theorem is also proved by Aubin andE keland[2] under the additional condition that $J(u, .)$ is a concave function for each $u \epsilon U$. Their equality is, however, different from ours in essence in one point, that is, the right-hand side in ours is replaced by the form "inf sup inf" in their equality.

## 4. Duality theory

In a previous paper[5)] a duality theorem was obtained for an optimal control problem of the minimax type. The problem took on the form:

$$\max_{v(.)} \min_{u(.)} \int_{t_0}^{t_1} \{f(t,x(t))+g(t,u(t),v(t))\}\,dt$$

subject to

$$\dot{x}(t)=A(t)x(t)+h(t,u(t),v(t)),\ (\dot{x}=dx/dt)$$
$$x(t_0)=x_0,\ u(t)\in U,\ u(t)\in V,$$

where $x\in R^n$, $u\in R^p$, $v\in R^q$ and $A(t)$ is an $n\times n$ continuous matrix. It was assumed that $f(t,x)$ is a real-valued cotinuous function on $R\times R^n$, convex and continuously differentiable with respect to $x$. The controls $u(.)$ and $v(.)$ are measurable functions of $t$ and assume their values in compact subsets $U\subset R^p$, $V\subset R^q$, respectively.

The Hamiltonian for the above problem is given by

$$k(t,p)\ =\ \max_{v\in V}\ \min_{u\in U}\ [g(t,u,v)+ph(t,u,v)]$$

for each $p\in R^n$ ($p$ is assumed to be a row vector).

We referred to the following problem as the dual problem to the above one:

$$\max_{x(.)} \int_{t_0}^{t_1} \{k(t,p(t))-p(t)(\dot{x}(t)-A(t)x(t))+f(t,x(t))\}\,dt$$

subject to

$$-\dot{p}(t)=p(t)A(t)+f_x(t,x(t)),\ \ p(t_1)=0,\ \ x(t_0)=x_0.$$

Here $f_x$ denotes the partial derivatives of $f$ with respect to $x$ and the supremum is taken over all pairs $(x(.),p(.))$, where $x(.)$ is an absolutely continuous function safisfying $x(t_0)=x_0$ and $p(.)$ is the corresponding solution with condition $p(t_1)=0$.

We proved the following duality theorem ; if a pair of controls $(u^*(.),v^*(.))$ gives a saddle point to the first (minimax) problem and $x^*$ is the corresponding optimal trajectory, then $x^*$ becomes an optimal control of the dual problem, and that both extremal values are equal. It was also shown that even if there are no such controls $(u^*(.),v^*(.))$, the extremal value of the dual problem provides a lower bound for the original (or primal) problem.

When either $u$ or $v$ is not involved, we have the corresponding duality theorems, one of which was stated in that paper[5)]. We consider here another case when $u$ is not involved and obtain a duality theorem for a minimization control problem with *concave* $f$. For this purpose, rewrite the primal problem as follows :

$$\max_{v(.)} \int_{t_0}^{t_1} \{f(t,x(t))+g(t,v(t))\}\,dt$$

$$= -\min_{v(.)} \int_{t_0}^{t_1} \{-f(t,x(t))-g(t,v(t))\}\,dt$$

subject to

$$\dot{x}(t) = A(t)x(t)+h(t,v(t)),\ \ x(t_0)=x_0,\ \ v(t) \in V.$$

The dual problem for this reduces to the following :

$$\max_{x(.)} \int_{t_0}^{t_1} \{k(t,p(t))-p(t)(\dot{x}(t)-A(t)x(t))+f(t,x(t))\}\,dt$$

$$= -\min_{x(.)} \int_{t_0}^{t_1} \{-k(t,p(t))+p(t)(\dot{x}(t)-f(t,x(t))\}\,dt$$

subject to

$$-\dot{p}(t)=p(t)A(t)+f_x(t,x(t)),\ \ p(t_1)=0, x(t_0)=a_0.$$

The Hamiltonian for this problem is defined by

$$k(t,p)=\max_{v\in V}\ \{g(t,v)+ph(t,v)\}$$

for all $p \in R^n$.

Rewriting $-f$ by $F$ and $-g$ by $G$, these problems can be reformulated as follows :

(Primal Problem) $$\min_{v(.)} \int_{t_0}^{t_1} \{F(t,x(t))+G(t,v(t))\}\,dt$$

subject to $\dot{x}(t)=A(t)x(t)+h(t,v(t)),\ \ x(t_0)=x_0,\ \ v(t) \in V$ ;

(Dual Problem) $$\min_{x(.)} \int_{t_0}^{t_1} \{K(t,p(t))+p(t)(\dot{x}(t)-A(t)x(t))+F(t,x(t))\}\,dt$$

subject to $-\dot{p}(t)=p(t)A(t)-F_x(t,x(t)),\ \ p(t_1)=0,\ \ x(t_0)=x_0,$

where

$$K(t,p)= \min_{v\in V}\ \{G(t,v)-ph(t,v)\}$$

for each $p \in R^n$.

Applying the above-mentioned duality theorem, we can obtain the next proposition.

PROPOSITION 3. *Let $F(t,x)$ be a real-valued continuous functionon on $R\times R^n$, concave and continuously differentiable with respect to $x$. If the primal problem has an optimal control, then the corresponding optimal trajectory $x^*(.)$ provides an optimal (absolutely continuous) control. Moreover, both extremal values are equal.*

In case that the primal problem has no optimal control, the extremal value of the dual problem provides an upper bound for the primal problem.

## 5. Remarks on duality

We assume again that $f(t,x)$ is a continuous real-valued function on $R\times R^n$ and that it is convex and continuously differentiable with respect to $x$. If the control variable $v$ is not involved in the optimal control problem considered in the preceding section, the primal and dual problems take on the forms :

(Primal Problem) $\min_{u(.)} \int_{t_0}^{t_1} \{f(t, x(t))+g(t, u(t))\}\, dt$

subject to $\dot{x}(t)=A(t)x(t)+h(t,u(t)),\ x(t_0)=x_0,\ u(t) \in U$ ;

(Dual Problem) $\max_{x(.)} \int_{t_0}^{t_1} \{k(t,p(t))-p(t)(\dot{x}(t)-A(t)x(t))+f(t,x(t))\}\, dt$

subject to $-\dot{p}(t)=p(t)A(t)+f_x(t,x(t)),\ x(t_0)=x_0,\ p(t_1)=0,$

where $k(t,p)$ is defined by

$$k(t,p)=\min_{u \in U} \{g(t,u)+ph(t,u)\}$$

for each $p \in R^n$.

The following duality relation in this circumstance was established in a previous paper [5] as a corollary of its main duality theorem.

PROPOSITION 4. *If the primal problem has an optimal control $u^*(.)$, then the corresponding optimal trajectory $x^*(.)$ provides an optimal (absolutely continuous) control for the dual problem. Moreover, both extremal values are equal.*

It should be noted that in case that the primal problem has no optimal controls, the extremal value of the dual one gives a lower bound for the primal one.

The above proposition suggests thal a dual problem of max-min type is associated with a minimization problem involving a control variable. On the other hand, it was also shown that dual maximization problems involving a control variable are associated with problems of min-max type[4]. We shall illustrate these relations by a simple example.

Let $(X,Y)$ and $(Z,W)$ be two dual pairs of linear spaces. The bilinear forms on $X,Y$ and $Z,W$ will be denoted by $\langle .,. \rangle$. Let $A$ be a continuous linear function from $X$ to $Z$ when $X$ is endowed with $\sigma(X,Y)$ topology and $Z$ with $\sigma(Z,W)$ topology. Let $P$ be the positive cone in $X$ and $Q$ a convex set in $Y$, whose interior is nonempty in the $\sigma(Y,X)$ topology. Define the equality constrained linear programming problem of min-max type as follows :

$$\operatorname*{minimize}_{x}\ \sup_{y}\ \langle x,y \rangle$$

subject to $Ax=b,\ x \in P,\ y \in Q$, where $b$ and $c$ are elements of $Z$ and $Y$, respectively. The feasible set is given by

$$F=\{x \in X \mid Ax=b,\ x \in P\}.$$

Assume that

(i) $\sup_{y \in Q}\ \inf_{x \in F}\ \langle x,y \rangle < +\infty,$

(ii) there exists a point $y_0 \in \operatorname{int} Q$ such that $\inf_{x \in F} \langle x,y_0 \rangle$ is finite ;

(iii) the sets of $Z \times R$

$$D_y= \{(Ax,\ \langle x,y \rangle) \mid x \in P\}$$

are closed for all $y \in Q$.

By a result due to Anderson[1], it is appropriate to call the following its dual problem :

$$\max \langle b, w \rangle$$
$$\text{subject to } -A^*w+y \in P^*,\ w \in W,\ y \in Q.$$

Here $A^*$ is the adjoint of $A$ and $P^*$ is the dual cone of $P$, that is,

$$P^* = \{y \in Y \mid \langle x, y \rangle \geqq 0 \text{ for all } x \text{ in } P\}.$$

Note that in the dual problem $y$ plays a role of control variable.

Combining Theorem 6 of Anderson[1] and Theorem 2.14 of Pomerol[3], one can obtain the following duality relation.

PROPOSITION 5. *Under the above conditions, both problems have the same extremal values.*

As anothar application of Theorem 2.14 of Pomerol, we consider a probem of approximating a given element $L_0$ in the dual $E_0$ of a normed linear space $E$ by elements in the annihilator $M^\perp$ of a closed subspace $M$ of $E$, that is, $M^\perp$ is the set of all linear functionals $L$ in $E^*$ that vanish on $M$. This problem can be written as follows :

$$\min_{L \in M^\perp} \| L - L_0 \| = \min_{L \in M^\perp} \sup_{\|x\| \leqq 1} \langle L - L_0, x \rangle.$$

Applying Theorem 2.14 of Pomerol, one obtain

$$\min_{L \in M^\perp} \| L - L_0 \| = \sup_{\|x\| \leqq 1} \inf_{L \in M^\perp} \langle L_0 - L, x \rangle$$

$$= \sup_{\substack{x \in M \\ \|x\| \leqq 1}} \inf_{L \in M^\perp} \langle L_0 - L, x \rangle = \sup_{\substack{x \in M \\ \|x\| \leqq 1}} |\langle L_0, x \rangle|.$$

Such a duality relation is treated by Ubhaya[8] in more detial.

## References


1) E. J Anderson : *J. Math. Anal. Appl.*, **97**, 380 (1983).
2) J. P. Aubin, I. Ekeland : *Nonlinear Applied Analysis* (Wiley-Interscience, 1984).
3) J. C. Pomerol : *J. Math. Anal. Appl.*, **103** 263 (1984).
4) S. Tanimoto : *J. Math. Anal. Appl.*, **79** 286 (1981).
5) S. Tanimoto : *IEEE Trans. on Automat. Contr.*, **AC-27** 1129 (1982).
6) S. Tanimoto : *J. Approx. Theory*, **45**, 1 (1985).
7) S. Tanimoto : *submitted for publication.*
8) V. A. Ubhaya : *J. Math. Anal. Appl.*, **58**, 419 (1977).


(高知女子大学　数学研究室)