9-1-1971

Duality Theory of Convex Programming for Infinite Horizon Economic Models

Martin L. Weitzman

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut

COWLES FOUNDATION DISCUSSION PAPER NO. 317

DUALITY THEORY OF CONVEX PROGRAMMING FOR INFINITE HORIZON ECONOMIC MODELS

Martin L. Weitzman

September 27, 1971

DUALITY THEORY OF CONVEX PROGRAMMING FOR INFINITE HORIZON ECONOMIC MODELS*

by

Martin L. Weitzman

Summary

The present state of convex programming theory for infinite horizon free endpoint economic models is not entirely satisfactory. Roughly speaking, classical duality principles can be shown to apply to finite subsections of an optimal trajectory and this avoids classical inefficiencies of the finite horizon variety. But it has never been completely clear how to avoid the kind of non-optimality which results from piling up too much "left over" capital in the limit. While certain rule of thumb "transversality conditions" have been proposed by analogy with finite horizon models, they have not in general been put on a rigorous footing and it is not clear which of them are valid under what circumstances. In this paper a rigorous treatment of the subject is undertaken. Under a set of general axioms, a certain limiting transversality condition in conjunction with other duality conditions is shown to be necessary and sufficient for infinite horizon optimality.

---

*The research described in this paper was carried out while I was a research visitor at the Computing Center of the Soviet Academy of Sciences. It was supported by a grant from the Guggenheim Foundation.

The research described in this paper was also partially supported under grants from the National Science Foundation and from the Ford Foundation.

Introduction

An important subclass of convex programming models of special interest to mathematical economists and operations researchers can be characterized by the Markovian property: the choice of options available at any particular time depends only on the values of the state variables at that time. In other words, all of the influence of past history on the present is summarized by current state variable levels.

With such programming models, it is often not clear how to appropriately fashion an "end" to the underlying economic process. For concreteness, this dilemma is illustrated by means of the standard model of optimal economic growth (although all remarks could be given a more general character). Any attempt to maximize "utility" ("gain") on an arbitrary finite interval collides with the problem of evaluating the capital stock (state variables) at the end of that interval. Insofar as the worth of capital is defined by the utility of consumption to which it gives rise, precise evaluation of this sort must await the solution of an analogous problem on a second interval. Repeated application of this reasoning leads to an infinite regress. The only way out of this regress would seem to be in recognizing that the future does not have a definite and forseeable end, and consequently optimization must be undertaken over an infinite horizon. Paradoxically, it is often easier to solve an optimization problem modeled on an infinite interval of time than it is to solve the corresponding problem on a finite interval with arbitrary end conditions.

Unfortunately, infinite horizon convex programming models with a free endpoint introduce some new headaches which are not present in their finite dimensional counterparts. For example, the very notion of an "optimal solution" for the infinite horizon case is somewhat vague and must be carefully and precisely defined. For this purpose we use a "classical" generalization of the usual finite dimensional criterion based on comparing convergent infinite sums. The chief headache in the infinite horizon case concerns the existence and form of strong (necessary and sufficient) duality relations. Duality is of course extremely useful, even essential, for characterizing an optimal solution. For finite dimensional convex programming models, as is well known, strong duality relations can be derived. Expressed with the aid of efficiency prices, the duality theorem in the finite Markovian case takes the form of an intemporal profit maximization condition between periods plus a specific type of transversality condition on stocks left over after the last period. As we shall see, analogous necessary and sufficient conditions can be derived for the infinite horizon case, with a transversality condition in the limit at infinity playing a key role.[1]

---

[1]From a technical point of view, showing that a certain transversality condition at infinity is a <u>necessary</u> optimality condition is perhaps the basic contribution of this paper. Roughly speaking, it has been more or less well known that intertemporal profit maximization plus a certain type of transversality condition at infinity are sufficient for an optimum. It is also more or less well known that intertemporal profit maximization is a necessary condition, but not sufficient by itself. What is <u>not</u> yet known, so far as I am aware, is how to state the appropriate transversality conditions at infinity (whether or not one or another form is a necessary condition under which circumstances). In many cases of interest in economics, all of the proposed transversality conditions turn out in fact to be valid, but so far it has been necessary to verify this separately in each case. In the present paper it is established under fairly general assumptions that the specific transversality condition which along with intertemporal profit maximization is sufficient, is also necessary. Hopefully this will help to clear the mystery about the role of transversality conditions at infinity.

Definitions and Assumptions

In what follows the index t, a positive integer, will denote the period of time from instant $t-1$ to t. The phase vector $x_{t-1} \in E^n$ is an n component column vector denoting the state of the system during period t. At the beginning, x_0 is considered given and denoted by $\bar{x}_0$. In many economic applications, x_{t-1} is understood as a vector whose i^{th} component represents the amount of capital of type i available for use at time $t-1$ and throughout the t^{th} period. The "gain" in period t is denoted by u_t. By "gain" might be understood "utility," "profit," "income," etc. depending on the specific features of the problem under consideration. Gains in each period are expressed in comparable units. In other words, all gains are measured as payout values discounted back to the first period. This is important because economic performance will be evaluated by the sum of single period gains.

The amount of gain u_t attainable in period t naturally depends on the initial and terminal states for that period, x_{t-1} and x_t. The $2n+1$ dimensional set of "transition possibilities" for period t, denoted Q_t, consists of all realizable triples of the form (x_{t-1}, u_t, x_t). In other words a transition which yields gain u_t can be made from state x_{t-1} at the beginning of period t to state x_t at the end of that period if and only if

$$(x_{t-1}, u_t, x_t) \in Q_t . \qquad (1)$$

A program $\{u_t, x_t\}$ is called feasible if for each t it satisfies (1) and

$$x_0 = \bar{x}_0 . \tag{2}$$

It is supposed that for all t the set Q_t obeys the following stipulations.

°1 If $(x,u,y) \in Q_t$, then $x \geqq 0$, $y \geqq 0$.

°2 If $(x,u,y) \in Q_t$ and if $x' \geqq x$, then $(x',u,y) \in Q_t$.

°3 If $(x,u,y) \in Q_t$ and $(x', u', y') \in Q_t$, then $(\lambda x + (1-\lambda)x', \lambda u + (1-\lambda)u', \lambda y + (1-\lambda)y') \in Q_t$ for all λ, $0 \leq \lambda \leq 1$.

°4 $(0,0,0) \in Q_t$.

These four conditions are reasonably standard. The first requires that the state variables (capital) be non-negative. The second is a free disposal type proposition. Condition °3 is the usual convexity assumption. The fourth condition is a "nothing ventured nothing gained" statement that it is possible to start with no capital, do nothing, gain nothing, and end up once again with no capital.[2]

---

[2] In cases where the level of gain is arbitrary, condition °4 amounts to a normalization convention under which zero gain is always attainable in any state. For the utility function in the theory of optimal growth, this means shifting its level so that zero utility is an absolute floor. In such cases °4 is not necessarily an empty condition because in order for the theory developed in this paper to be applicable, $\sum_{t=1}^{\infty} u_t$ must converge; tampering with the level of utility can in theory alter the convergence properties of $\sum_{t=1}^{\infty} u_t$. In the standard case of discounting the same utility function by a constant rate (i.e. $u_t = \alpha^t U(C)$ where $\alpha < 1$, U is a concave utility function and C is the consumption vector) it is easy to see that maximizing $\sum_{t=1}^{\infty} \alpha^t [U(C) - U(0)]$ gives the same solution as maximizing

A fifth condition will guarantee that a strictly positive state vector is always "reachable" at any time. This kind of a productivity stipluation is needed as a means of insuring enough regularity so that meaningful dual prices can be formed.

$^{o}5$ For each t there exists an $\hat{x}_t > 0$ with corresponding $\{x^t_\tau, u^t_\tau\}_{1 \leq \tau \leq t}$ satisfying

$$(x^t_{\tau-1}, u^t_\tau, x^t_\tau) \in Q_\tau \qquad 1 \leq \tau \leq t$$

$$x^t_0 = \bar{x}_0$$

$$x^t_t = \hat{x}_t$$

Let S be the class of all summable infinite sequences $(\{s_t\} \in S$ iff $\lim_{T\to\infty} \sum_{t=1}^{T} s_t$ exists). A program $\{u_t, x_t\}$ is said to be allowable if

---

$\sum_{t=1}^{\infty} \alpha^t U(C)$ and that one sum converges if and only if the other does. This trick of taking first differences with the worst possible utility level in order to satisfy $^{o}4$ will not work for the Ramsey case of no discounting $(\alpha = 1)$ because shifting utility to the form $u_t = U(C) - U(0)$ will in general result in an unbounded objective function ($\sum_{t=1}^{\infty} u_t$ does not converge). In such cases the theory developed in this paper is irrelevant and although a pricing theory can be worked out by other means, the transversality condition may not hold.

it is feasible (satisfies (1), (2)) and if $\{u_t\} \in S$. The effect of limiting attention to programs with summable gains is to introduce a complete preference ordering on programs. A program $\{u_t^*, x_t^*\}$ is called <u>optimal</u> if it is allowable and if for any other allowable program $\{\tilde{u}_t, \tilde{x}_t\}$,

$$\sum_{t=1}^{\infty} u_t^* \geqq \sum_{t=1}^{\infty} \tilde{u}_t .$$

It is typically much more difficult to prove that in theory an optimal program must exist for an infinite horizon model than it is for its finite horizon counterpart.[3] Nevertheless, it seems to be empirically true that the definition of optimality used here is broad enough so that optimal programs turn up for a great many infinite horizon models of interest.

Duality Theory

Under the five axioms listed in the last section, the following duality theorem holds.

<u>Theorem</u>: For the allowable program $\{u_t^*, x_t^*\}$ to be optimal, it is necessary and sufficient that there exist a sequence of non-negative n-dimensional price row vectors $\{p_t\}$ satisfying

1^o for $t = 1$,

[3]Roughly speaking, the inherent discount rate used to discount (undiscounted) payouts into (discounted) gains must be at least as high as the inherent potential growth rate of (undiscounted) payouts. Otherwise the sum of one period gains (each one of which is normalized so that zero gain is always attainable in any state) may not converge.

$$u_1^* + p_1 x_1^* \geqq u_1 + p_1 x_1$$

$$(\bar{x}_0, u_1, x_1) \in Q_1$$

2° for $t \geq 2$,

$$u_t^* + p_t x_t^* - p_{t-1} x_{t-1}^* \geqq u_t + p_t x_t - p_{t-1} x_{t-1}$$

$$(x_{t-1}, u_t, x_t) \in Q_t$$

3° for $t = \infty$

$$\lim_{t \to \infty} p_t x_t^* = 0 .$$

Conditions 1° and 2° have the obvious interpretation that an optimal transition maximizes imputed profits.[4] The transversality condition 3° is a special feature of infinite horizon programming models with a free endpoint. It can be interpreted as saying that in an optimal program the present value of "left over" capital must eventually go to zero. This limitation on the rate at which capital ought to be accumulated is very important in problem solving applications. In order to specially emphasize the non-trivial character of 3° and to examine some other issues, the following artificial problem is proposed.

[4] There are various other equivalent ways of writing the optimality conditions 1°, 2°. For example, they could be expressed as a discrete version of the so-called Maximum Principle (cf. Weitzman and Schmidt). The present formulation is chosen because it seems the simplest.

An Example

Suppose it is desired to maximize

$$\sum_{t=1}^{\infty} \frac{c_t}{(1+\beta)^t}$$

under the conditions

$$c_t + z_t \leq (1+\beta)z_{t-1}$$

$$c_t, \quad z_t \geq 0$$

$$z_0 = 1$$

for $\beta > -1$. In this problem z is a kind of fictitious circulating capital and c is consumption. The coefficient β is simultaneously a discount factor on consumption and the capital rate of reproduction (decay if $\beta < 0$) . Gain in period t is $u_t \equiv c_t/(1+\beta)^t$. The set of transition possibilities for period t is

$$Q_t \equiv \{z_{t-1}, u_t, z_t | u_t = c_t/(1+\beta)^t, \; c_t + z_t \leq (1+\beta)z_{t-1}, \; c_t \geq 0, \; z_{t-1} \geq 0, \; z_t \geq 0\} .$$

It is easy to verify that conditions $^o1$-$^o5$ are fulfilled. Let $q_t \geq 0$ be the dual "price" of z_t . Applying 1^o and 2^o, an optimal program if it exists must for all t

$$\text{maximize} \quad \frac{c_t}{(1+\beta)^t} + q_t z_t - q_{t-1} z_{t-1} \tag{3}$$

under the constraints

$$c_t + z_t \leqq (1+\beta)z_{t-1} \tag{4}$$

$$c_t,\ z_t,\ z_{t-1} \geqq 0 \tag{5}$$

and the additional constraint

$$z_0 = 1 \tag{6}$$

for $t = 1$.

It follows immediately that for an optimal solution inequality (4) can be replaced for all t by the strict equality

$$c_t + z_t = (1+\beta)z_{t-1}\ . \tag{7}$$

If an optimal solution of (3)-(6) is to exist, it must be true for all t that

$$q_{t+1} \leqq q_t/(1+\beta) \tag{8}$$

(if not, "blowing up" z_{t+1} and z_t leads to an unbounded objective function).

It is easy to verify that

$$q_t \geqq \frac{1}{(1+\beta)^t}\ . \tag{9}$$

If (9) does not hold, an optimal solution of (3)-(6) will not exist ("blowing up" c_{t+1} and z_t would lead to an unbounded objective).

Now suppose that (8) does not hold for all t with full equality.

Let τ be the first time that (8) holds with strict inequality. That is,

$$q_{t+1} = q_t/(1+\beta) \qquad 1 \leq t < \tau$$

$$q_{\tau+1} < q_\tau/(1+\beta) \; . \tag{10}$$

From (9), $q_{\tau+1} \geq 1/(1+\beta)^{\tau+1}$. In conjunction with (10), this implies that

$$q_\tau > \frac{1}{(1+\beta)^\tau} \; . \tag{11}$$

Inducting (11) backward on (8),

$$q_t > \frac{1}{(1+\beta)^t} \qquad 1 \leq t \leq \tau \; .$$

The corresponding optimal solution of (3)-(6) is $c_t = 0$, $z_t = (1+\beta)z_{t-1}$ (i.e. $z_t = (1+\beta)^t$) for $1 \leq t \leq \tau$. However, from (9) and (10),

$$\frac{1}{(1+\beta)^{\tau+1}} \leq q_{\tau+1} < q_\tau/(1+\beta)$$

to which the corresponding optimal solution of (3)-(6) is $0 = c_{\tau+1} = z_{\tau+1} = z_\tau$. This contradicts $z_\tau = (1+\beta)^\tau$. It follows that

$$q_{t+1} = q_t/(1+\beta) \tag{12}$$

which has the solution

$$q_t = \frac{q_0}{(1+\beta)^t} \tag{13}$$

with $q_0 \geq 1$ (from (9)).

Suppose now that $q_0 > 1$. With (13) holding, the corresponding solution of (3)-(6) is $c_t = 0$, $z_t = (1+\beta)z_{t-1}$ (i.e. $z_t = (1+\beta)^t$). Using (13),

$$q_t z_t = q_0$$

which, in the limit as $t \to \infty$, is a direct violation of 3^o.

It has been shown that if an optimal program exists, the dual prices must have the form

$$q_t = \frac{1}{(1+\beta)^t} . \tag{14}$$

Actually this result is intuitively obvious from some simple considerations of capital theory.

With the prices (14), a solution of (3)-(6) would be any solution of (4)-(6) which satisfies (7). But this does not mean that <u>any</u> solution of (5)-(7) is an optimal program because it would only be satisfying 1^o and 2^o. For example, the program $c_t = 0$, $z_t = (1+\beta)^t$ is hardly optimal --it satisfies 1^o and 2^o, but not 3^o. An optimal program can be completely characterized as follows. A program satisfying (5)-(7) will be optimal if and only if it satisfies 3^o, i.e. iff

$$\lim_{t\to\infty} \frac{z_t}{(1+\beta)^t} = 0 .$$

Further attention will now be restricted to paths where consumption and new circulating capital are split in the proportions $(1-\alpha)/\alpha$. That is,

$$z_t = \alpha(1+\beta)z_{t-1}$$
$$c_t = (1-\alpha)(1+\beta)z_{t-1}$$

for $0 \leq \alpha \leq 1$. It follows from $z_0 = 1$ that

$$z_t = \alpha^t(1+\beta)^t \tag{15}$$
$$c_t = (1-\alpha)\alpha^t(1+\beta)^t .$$

As was indicated, such a program satisfies 1^o and 2^o for any α, $0 \leq \alpha \leq 1$. But from $q_t z_t = \alpha^t$ it will only be optimal for $\alpha < 1$ since otherwise 3^o will be violated. It can be seen directly that $\sum_{t=1}^{\infty} u_t = 1$ for $0 \leq \alpha < 1$, but that $\sum_{t=1}^{\infty} u_t = 0$ for $\alpha = 1$.

The above example with $0 \leq \alpha < 1$ is useful in showing the wide variety of optimal limiting behavior which can be handled by the theory developed here. For $\beta > 0$ and $\alpha(1+\beta) > 1$, from (14), (15) in the limit as $t \to \infty$

$$z_\infty = \infty , \quad q_\infty = 0 . \tag{i}$$

For $\beta > 0$ and $\alpha(1+\beta) = 1$

$$z_\infty = 1 , \quad q_\infty = 0 . \tag{ii}$$

For $\beta > 0$ and $\alpha(1+\beta) < 1$

$$z_\infty = 0 , \quad q_\infty = 0 . \tag{iii}$$

For $\beta = 0$, $\alpha(1+\beta)$ is necessarily < 1 implying

$$z_\infty = 0\,, \qquad q_\infty = 1\,. \tag{iv}$$

For $-1 < \beta < 0$, $\alpha(1+\beta)$ is necessarily < 1 implying

$$z_\infty = 0\,, \qquad q_\infty = \infty\,. \tag{v}$$

As is obvious from this set of examples, either quantities or prices can go to infinity in the limit without in any way impairing the theory developed in this paper. The common aspect of these five different optimal programs is the limit of zero for $p_\infty x_\infty$. Note that (iv) and (v) definitely contradict the notion that

$$\lim_{t\to\infty} p_t = 0$$

is any kind of a generally valid transversality condition.

Proof of the Duality Theorem

Sufficiency:

Let $\{\tilde{u}_t, \tilde{x}_t\}$ be any allowable program.

$$\sum_{t=1}^{T} u_t^* - \sum_{t=1}^{T} \tilde{u}_t = (u_1^* + p_1 x_1^*) - (\tilde{u}_1 + p_1\tilde{x}_1) + \sum_{t=2}^{T} [(u_t^* + p_t x_t^* - p_{t-1}x_{t-1}^*)$$

$$- (\tilde{u}_t + p_t\tilde{x}_t - p_{t-1}\tilde{x}_{t-1})] + p_T(\tilde{x}_T - x_T^*)$$

$$\geqq p_T(\tilde{x}_T - \tilde{x}_T^*) \qquad \text{(from } 1^\circ \text{ and } 2^\circ\text{)}$$

Passing to the limit as $T \to \infty$, from 3° and the fact that $p_T\tilde{x}_T \geqq 0$ it follows that

$$\sum_{t=1}^{\infty} u_t^* \geqq \sum_{t=1}^{\infty} \tilde{u}_t .$$

Necessity:

Let E_+^n be the n-dimensional Euclidean non-negative orthant, i.e. $E_+^n \equiv \{x | x \in E^n, x \geqq 0\}$.

Consider for each t the following improper function φ_t which maps E_+^n into the extended real halfline $[0, +\infty]$

$$\varphi_t(x) \equiv \sup \sum_{\tau=t+1}^{\infty} u_\tau \tag{16}$$

$$\text{subject to } \{u_\tau\}_{\tau \geq t+1} \in S \tag{17}$$

$$(x_{\tau-1}, u_\tau, x_\tau) \in Q_\tau \qquad \tau \geq t+1 \tag{18}$$

$$x_t = x . \tag{19}$$

For each $x \in E_+^n$, $\varphi_t(x)$ is well defined and non-negative (although it might be infinite) because from $^o2$ and $^o4$ the following is a solution of (17)-(19): $x_t = x$, $x_\tau = 0$, $u_\tau = 0$ for $\tau \geq t+1$.

From $^o2$, $\varphi_t(x)$ is non-decreasing in x . From $^o3$, it is easy to verify that $\varphi_t(x)$ is an (improper) concave function of x . ($\varphi_t(x)$ is improper concave if for any $x', x'' \in E_+^n$, $0 < \lambda < 1$, (a) $\varphi_t(x')$, $\varphi_t(x'') < \infty$ implies $\varphi_t(\lambda x' + (1-\lambda)x'') \geq \lambda\varphi_t(x') + (1-\lambda)\varphi_t(x'')$, (b) $\varphi_t(x') = \infty$ or $\varphi_t(x'') = \infty$ implies $\varphi_t(\lambda x' + (1-\lambda)x'') = \infty$.) In addition, it is not difficult to verify that the function sequence $\{\varphi_t(x)\}$

must satisfy the following fundamental equation of dynamic programming

$$\varphi_t(x) = \sup_{\substack{u,\, y \\ (x,u,y)\in Q_{t+1}}} \{u + \varphi_{t+1}(y)\} . \qquad (20)$$

(If not, there is an immediate contradiction with the definition (16)-(19).) Note that for $x = x_t^*$ the operator <u>max</u> can replace the operator <u>sup</u> in (20) and a solution is (u_{t+1}^*, x_{t+1}^*) .

From the fact that $\hat{x}_t$ is attainable in period t (cf. °5) it follows that $\varphi_t(\hat{x}_t) < \infty$. (Otherwise there is a contradiction with optimality of the program $\{u_t^*, x_t^*\}$.) Let x' be any vector satisfying $x' \geqq 0$, $x' \neq \hat{x}_t$. From $\hat{x}_t > 0$, it follows that there exists a μ , $0 < \mu < 1$, such that $\hat{x}_t \geqq \mu x'$. Then $\hat{x}_t = \mu x' + (1-\mu)x''$, where $x'' \equiv (\hat{x}_t - \mu x')/(1-\mu) \geqq 0$. From the (improper) concavity of $\varphi_t(x)$ it follows that

$$\varphi_t(\hat{x}_t) \geqq \mu\varphi_t(x') + (1-\mu)\varphi_t(x'') .$$

(Here $\mu\infty = (1-\mu)\infty = \infty$, $r+\infty = \infty+r = \infty$ for any real r , $\infty+\infty = \infty$, and $\infty \geqq \infty$.) From $\varphi_t(x'') \geqq 0$ and $\varphi_t(\hat{x}_t) < \infty$ it follows that $\varphi_t(x') < \infty$. In other words $\varphi_t(x)$ is an ordinary (proper = finite) concave function on E_+^n .

A slight detour is now taken in order to prove a set of lemmas which will culminate in allowing the construction of duality prices for the functions $\{\varphi_t(x)\}$.

<u>Definition</u>: Let $f(y)$ be a function defined on E_+^n . Let y^* be any point of E_+^n . The function $f(y)$ is said to have <u>finite steepness</u> at the point y^* if

$$\sup_{\substack{y \in E^n_+ \\ y \neq y^*}} \frac{|f(y) - f(y^*)|}{\|y - y^*\|} < \infty .$$

Here and in what follows the norm $\|\cdot\|$ is understood to be Euclidean.

The lemmas will end up proving that $\varphi_t(x)$ is of finite steepness at x^*_t. As Gale has shown (cf. Gale) this concept is of basic importance because it can be used to justify the use of dual Kuhn-Tucker prices with the usual desirable separation properties.

Lemma 1: Let $f(y)$ be a concave function defined on E^n_+ and bounded from below. Let y^* be any point of E^n_+. Then there exists a number γ such that $[f(y) - f(y^*)]/\|y - y^*\| \geqq \gamma$ for all $y \in E^n_+$, $y \neq y^*$.

Proof: Without loss of generality we take zero as the lower bound on $f(y)$, $y \in E^n_+$. In this proof, vector superscripts denote vector components.

Define $\delta > 0$ as follows:

$$\delta = \begin{cases} 1 & \text{if } y^* = 0 \\ \min\limits_{(y^*)^i > 0}\{(y^*)^i\} & \text{if } y^* \neq 0 \end{cases}$$

If $0 < \delta \leqq \|y - y^*\|$,

$$\frac{f(y) - f(y^*)}{\|y - y^*\|} \geqq - \frac{f(y^*)}{\delta} .$$

Define $\bar{y}$ by the following equation

$$\bar{y} \equiv y^* + \frac{(y - y^*)\delta}{\|y - y^*\|} .$$

Note that $\bar{y} \geqq 0$. (Examine the above definition component by component; if $(y^*)^i = 0$, trivial; if $(y^*)^i > 0$, then $[y^i - (y^*)^i]/\|y - y^*\| \geqq -1$ and $(y^*)^i - \delta \geqq 0$.)

Suppose $\delta > \|y - y^*\|$. From the identity

$$y = \bar{y}\,\frac{\|y - y^*\|}{\delta} + y^*\left(1 - \frac{\|y - y^*\|}{\delta}\right)$$

and concavity of $f(\cdot)$ it follows that

$$f(y) \geqq \frac{\|y - y^*\|}{\delta}\,f(\bar{y}) + \left(1 - \frac{\|y - y^*\|}{\delta}\right)f(y^*)$$

which implies

$$\frac{f(y) - f(y^*)}{\|y - y^*\|} \geqq -\frac{f(y^*)}{\delta}\ .$$

Setting $\gamma \equiv -f(y^*)/\delta$ concludes the lemma.

<u>Lemma 2</u>: If in addition to the hypotheses of lemma 1 there exists a strictly positive vector $\hat{y} > 0$ and a number M such that

$$\frac{f(\lambda\hat{y} + (1-\lambda)y^*) - f(y^*)}{\lambda} \leqq M$$

for all λ obeying $0 < \lambda \leq 1$, then $f(y)$ has finite steepness at the point y^* .

<u>Proof</u>: If $y^* = \hat{y}$, then y^* is interior to E^n_+ and the result is immediate. Suppose $y^* \neq \hat{y}$. From lemma 1, $[f(y) - f(y^*)]/\|y - y^*\|$ has a lower bound. We must now prove it has an upper bound as well. Suppose by contradiction that there is an infinite sequence $\{y_k\}$ where

$y_k \in E^n_+$, $y_k \neq y^*$ such that $[f(y_k) - f(y^*)]/\|y_k - y^*\| \geq k$. Without loss of generality we take $\|y_k - y^*\| \leq 1$ for all k. (If $\|y_k - y^*\| > 1$ we can in place of y_k use $\tilde{y}_k \equiv y^* + (y_k - y^*)/\|y_k - y^*\|$ since by concavity of $f(y)$, $(f(\tilde{y}_k) - f(y^*))/\|\tilde{y}_k - y^*\| \geq (f(y_k) - f(y^*))/\|y_k - y^*\|$.)

Define the set $R \equiv \{y | y \in E^n, \|y - y^*\| = 1\}$. Since R is compact and $y^* + (y_k - y^*)/\|y_k - y^*\| \in R$, there must exist a subsequence of $\{y_k\}$, denoted $\{y_j\}$ and a limit point $\bar{y} \in R$ such that

$$\left\| \frac{y_j - y^*}{\|y_j - y^*\|} + y^* - \bar{y} \right\| \leq \frac{1}{j}, \quad \frac{f(y_j) - f(y^*)}{\|y_j - y^*\|} \geq j .$$

Defining v_j by the equation

$$v_j \equiv \frac{y_j - y^*}{\|y_j - y^*\|} - \bar{y} + y^* ,$$

we have

$$y_j = y^* + \|y_j - y^*\| (\bar{y} - y^*) + v_j \|y_j - y^*\| \tag{21}$$

where $\|v_j\| \leq \frac{1}{j}$.

Because $\hat{y} > 0$, there must exist a μ, $0 < \mu < 1$, such that

$$y' \equiv \frac{\hat{y} - \mu\bar{y}}{1 - \mu} > 0 . \tag{22}$$

Define

$$\hat{y}_j \equiv y^* + \|y_j - y^*\| (\hat{y} - y^*) \tag{23}$$

$$y_j' \equiv \frac{\hat{y}_j - \mu y_j}{1-\mu} \tag{24}$$

$$= y^*(1 - \|y_j - y^*\|) + \|y_j - y^*\| \left(y' - \frac{\mu v_j}{1-\mu} \right) . \tag{25}$$

The expression (25) comes from substituting (21)-(23) in (24). Since $0 < \|y_j - y^*\| \leqq 1$, $\hat{y}_j$ and the first term of (25) are non-negative. The second term of (25) is non-negative for $j \geq$ some J because $y' > 0$ and $\|v_j\| \leqq \frac{1}{j}$. Thus $y_j' \geqq 0$ for $j \geqq J$.

From (24), $\hat{y}_j = \mu y_j + (1-\mu)y_j'$ where $\hat{y}_j, y_j, y_j' \in E_+^n$ for $j \geq J$. Applying concavity,

$$f(\hat{y}_j) \geqq \mu f(y_j) + (1-\mu)f(y_j') .$$

This implies (for $j \geq J$)

$$\frac{f(y_j) - f(y^*)}{\|y_j - y^*\|} \leqq \frac{1}{\mu} \frac{f(\hat{y}_j) - f(y^*)}{\|y_j - y^*\|} - \frac{1-\mu}{\mu} \frac{f(y_j') - f(y^*)}{\|y_j - y^*\|} . \tag{26}$$

By the hypothesis of this lemma, $[f(\hat{y}_j) - f(y^*)]/\|\hat{y}_j - y^*\| \leqq M$. From (23), $\|\hat{y}_j - y^*\| = \|y_j - y^*\| \cdot \|\hat{y} - y^*\|$. Combining,

$$\frac{1}{\mu} \frac{f(\hat{y}_j) - f(y^*)}{\|y_j - y^*\|} \leqq \frac{M\|\hat{y} - y^*\|}{\mu} . \tag{27}$$

Due to lemma 1, $[f(y_j') - f(y^*)]/\|y_j' - y^*\| \geq \gamma$. Since

$$y_j' - y^* = \frac{\|y_j - y^*\| (\hat{y} - y^*) - \mu(y_j - y^*)}{1-\mu} ,$$

we have

$$\|y_j' - y^*\| \leqq \|y_j - y^*\| \left(\frac{\|\hat{y} - y^*\| + u}{1-u}\right) .$$

If follows that

$$- \frac{(1-u)}{u} \frac{f(y_j') - f(y^*)}{\|y_j - y^*\|} \leq -\gamma\left(\frac{\|\hat{y} - y^*\| + u}{u}\right) .$$

Combining the above and (27) with (26),

$$\frac{f(y_j) - f(y^*)}{\|y_j - y^*\|} \leqq \frac{M\|\hat{y} - y^*\|}{u} - \gamma\left(\frac{\|\hat{y} - y^*\| + u}{u}\right)$$

which is a contradiction with $[f(y_j) - f(y^*)]/\|y_j - y^*\| \geq j$.

Lemma 3: There is a number M_t such that $[\varphi_t(\lambda\hat{x}_t + (1-\lambda)x^*) - \varphi_t(x^*)]/\lambda \leq M_t$ for $0 < \lambda \leq 1$.

Proof: Let

$$M_t \equiv \sum_{\tau=1}^{t} u_\tau^* - \sum_{\tau=1}^{t} u_\tau^t$$

where u_τ^t is defined in axiom $^o5$.

Suppose (by contradiction) that there exists a value of λ , $0 < \lambda \leq 1$, such that

$$\frac{\varphi_t(\lambda\hat{x}_t + (1-\lambda)x_t^*) - \varphi_t(x_t^*)}{\lambda} > M_t .$$

This would mean that

$$\lambda \sum_{\tau=1}^{t} u_\tau^t + (1-\lambda) \sum_{\tau=1}^{t} u_\tau^* + \varphi_t(\lambda\hat{x}_t + (1-\lambda)x_t^*) > \sum_{\tau=1}^{t} u_\tau^* + \varphi_t(x_t^*)$$

or that

$$\sum_{\tau=1}^{t}(\lambda u_{\tau}^{t} + (1-\lambda)u_{\tau}^{*}) + \varphi_t(\lambda\hat{x}_t + (1-\lambda)x_t^{*}) > \sum_{t=1}^{\infty} u_t^{*} \text{ .}$$

Using $^{o}3$ and equation (20), this is a contradiction with optimality of the program $\{u_t^{*}, x_t^{*}\}$.

Since the functions $\{\varphi_t(x)\}$ satisfy all the prerequisites of lemmas 1 and 2 under the correspondence $x_t \Longleftrightarrow y$, $\varphi_t(x) \Longleftrightarrow f(y)$, $x_t^{*} \Longleftrightarrow y^{*}$, $\hat{x}_t \Longleftrightarrow \hat{y}$, we can conclude that $\varphi_t(x)$ has finite steepness at x_t^{*} . The conclusion (cf. Gale) is that $\varphi_t(x)$ must have a supporting hyperplane at x_t^{*} with the usual separating properties familiar to the theory of convex programming.

By the definition of $\varphi_1(x)$ ((16)-(19)), it is clear that

$$u_1^{*} + \varphi_1(x_1^{*}) = \max_{\substack{u_1, x_1 \\ (\bar{x}_0, u_1, x_1) \in Q_1}} \{u_1 + \varphi_1(x_1)\}$$

From the theory of convex programming there must exist an n-dimensional row price vector p_1 satisfying

$$\varphi_1(x_1^{*}) - p_1 x_1^{*} \geqq \varphi_1(x_1) - p_1 x_1 \qquad (27)$$

$$x_1 \geqq 0$$

$$u_1^{*} + p_1 x_1^{*} \geqq u_1 + p_1 x_1 \qquad (28)$$

$$(\bar{x}_0, u_1, x_1) \in Q_1$$

Because $\varphi_1(x)$ is non-decreasing in x , from (27) p_1 must be non-negative.

Now suppose (by induction) that for a given t there exist non-negative price vectors $\{p_\tau\}_{1 \le \tau \le t}$ satisfying

(a) for $1 \le \tau \le t$: $\varphi_\tau(x_\tau^*) - p_\tau x_\tau^* \geqq \varphi_\tau(x_\tau) - p_\tau x_\tau$

$$x_\tau \geqq 0$$

(b) for $\tau = 1$: $u_1^* + p_1 x_1^* \geqq u_1 + p_1 x_1$

$$(\bar{x}_0, u_1, x_1) \in Q_1$$

for $2 \le \tau \le t$: $u_\tau^* + p_\tau x_\tau^* - p_{\tau-1} x_{\tau-1}^* \geqq u_\tau + p_\tau x_\tau - p_{\tau-1} x_{\tau-1}$

$$(x_{\tau-1}, u_\tau, x_\tau) \in Q_\tau$$

From (27) and (28), (a) and (b) are satisfied for $t = 1$. We must show that if (a), (b) are valid for arbitrary t, they must also hold for $t+1$.

Combining (a) for $\tau = t$ with equation (20),

$$\varphi(x_t^*) - p_t x_t^* = \max_{x \geqq 0} \{\varphi_t(x) - p_t x\}$$

$$\Big[\sup_{\substack{u, y \\ (x_t^*, u, y) \in Q_{t+1}}} \{u + \varphi_{t+1}(y)\}\Big] - p_t x_t^* = \max_{x \geqq 0} \Big\{\Big[\sup_{\substack{u, y \\ (x, u, y) \in Q_{t+1}}} \{u + \varphi_{t+1}(y)\}\Big] - p_t x\Big\}$$

$$u_t^* + \varphi_{t+1}(x_{t+1}^*) - p_t x_t^* = \max_{(x, u, y) \in Q_{t+1}} \{u + \varphi_{t+1}(y) - p_t x\} . \qquad (29)$$

Using the theory of convex programming applied to $\varphi_{t+1}(x)$ in (29), there must exist an n-dimensional row price vector p_{t+1} which satisfies

$$\varphi_{t+1}(x_{t+1}^*) - p_{t+1} x_{t+1}^* \geqq \varphi_{t+1}(x) - p_{t+1} x \qquad (30)$$

$$x \geqq 0$$

$$u_t^* + p_{t+1}x_{t+1}^* - p_t x_t^* \geqq u + p_{t+1}y - p_t x \qquad (31)$$

$$(x,u,y) \in Q_{t+1}$$

From (30), p_{t+1} must be non-negative because $\varphi_{t+1}(x)$ is non-decreasing in x.

Thus (30) and (31) show that (a) and (b) hold for $t+1$. This proves 1° and 2°.

To prove the transversality condition set $x_t = 0$ in (a), yielding

$$\varphi_t(x_t^*) - \varphi_t(0) \geq p_t x_t^* .$$

From $\varphi_t(0) \geqq 0$, $p_t x_t^* \geqq 0$, and

$$\lim_{t\to\infty} \varphi_t(x_t^*) = \lim_{t\to\infty} \sum_{\tau=t+1}^{\infty} u_\tau^* = 0 ,$$

condition 3° directly follows. This concludes the proof.

REFERENCES

[1] Gale, David. "A Geometric Duality Theorem with Economic Applications," Review of Economic Studies, January 1967.

[2] Weitzman, M.L. and A.G. Schmidt. "Printsip Maximuma dlya Discretnix Ekonomicheskikh Protsessov na Beskonyechnom Intervalye Bremeni" (The Maximum Principle for Discrete Economic Processes on an Infinite Time Interval) to be published in Kibernetika.