AN "ECONOMICS PROOF" OF A SEPARATING HYPERPLANE THEOREM

by

Martin L. Weitzman

Discussion Paper No. 1881
October 1999

Harvard Institute of Economic Research
Harvard University
Cambridge, Massachusetts

An “Economics Proof’ of a Separating Hyperplane Theorem

Martin L. Weitzman

Department of Economics
Harvard University
(October 4, 1999)

Abstract

Based centrally on the economic concept of a cost function, an “economics proof’ (by induction) is given of the supporting hyperplane theorem.

Keywords: Duality; Cost functions: Separating hyperplane theorem

JEL classification: A2; C6

1. Introduction

Separation theorems for convex sets are a basic mathematical tool that find enormously widespread use throughout economics. As has been recognized in the literature. all of the standard separating hyperplane theorems are readily derivable from the supporting hyperplane version of the theorem. (The most difficult case to prove is where the hvo convex sets are “kissing” along their common boundary, which is essentially the case treated here of the supporting hy-perplane theorem for a closed convex set.) In a sense, we are really talking about one unified “family” of separating (or supporting) hyperplane theorems. So far as I have ever seen. the standard proofs of the family of separating (or supporting) hyperplane theorems found in the literature consist of either formal abstract mathematics’ or, if motivated at all. are geometrically motivated.[2] Although the separating hyperplane theorems are ultimately loaded

[1] See, e.g., Gale [1960], page 44.

[2] See, e.g., Debreu [1959].

with economic content, I have never encountered a proof that showed clearly and directly *within the proof itself* the basic connection between the "economics of convexity" and the "mathematics of convexity."

This paper contains a proof of one of the main separating hyperplane theorems that starts with a rigorously stated, but economically intuitive, economic proposition about cost functions, and derives from it the proof of the supporting hyperplane theorem. The proof is by induction on the number of "inputs" as arguments of a concave production function. The proof is *constructive* in the sense that it indicates precisely how to construct the *n-plus-first* price (the "output" price), given the previous n prices (the "input" prices). The proof is also economically intuitive in the sense of being based on a fundamental economic idea about the possibility of being able to decentralize (production) decisions in a convex technology. Some variant of this basic idea finds applications almost everywhere throughout economics, being, e.g., the subject of the celebrated first essay ("Allocation of Resources and the Price System") in Koopmans [1957].

Of course, the separating hyperplane theorem itself has been stated and proved, in its many versions, a long time ago. The only novelty of this paper is that the proof, which is by induction, is based upon economic concepts. The proof is taken from lecture notes to a graduate course I teach on "optimization for economists." Students studying formal economics, in this course or elsewhere, may wish (or even need) to see a rigorous proof of a separating hyperplane theorem. In this context, I then asked myself the following question. Why not expose such students to a proof that in itself reveals the close connection between the economics and the mathematics of convexity – by being based on the language and ideas of economics directly? In the hopes that such a proof may perhaps be found to have some modest interest, pedagogic or otherwise, I offer it here to a wider audience than that of my own classroom.

2. Strategy of Proof

Because the goal of this proof is to emphasize economic content and a connection with economics, before diving into the details I want first to give an overview of the strategy.

Let y stand for a single "output," while x represents a n-dimensional vector of "inputs." The convex (n+l)-dimensional closed "production set" is denoted S. Thus. the "output-input" combination $(y, \boldsymbol{x})$ is feasible if and only if

$$(y,\ \boldsymbol{x})\ \varepsilon\ S\ . \tag{1}$$

Let $(y^*, \boldsymbol{x}^*)$ represent a boundary point of S. In economics the interesting boundary points are identified with the "efficient" point of S. but formally we are treating here any boundary point. The proof is going to be by induction on the number of "inputs." Suppose, then. the *induction assumption:* there exists a n-dimensional price vector of "inputs," denoted $\boldsymbol{p_x}$, which "supports" $\boldsymbol{x}^*$ at y^*. In economics. a price vector "supporting" an input combination means a particular set of prices that would cause *this particular* input combination to be demanded by the cost-minimizing producer. Mathematically, it means that $\boldsymbol{x}^*$ is a solution at given prices $\boldsymbol{p_x}$ of the following problem of cost minimization:

minimize

$$\boldsymbol{p_x} \cdot \boldsymbol{x}\ . \tag{2}$$

subject to

$$(y^*,\ x)\ \varepsilon\ s\ . \tag{3}$$

Now we wish to ask, and answer, the following question. Is there a (n+l)-dimensional price vector $\boldsymbol{P}=(P_y, \boldsymbol{P_x})$ that "supports," as the soiution of a decentralized profit-maximizing problem the "input" *and* "output" combination $(y^*, \boldsymbol{x}^*)$? Intuitively, the answer is yes. With a convex production structure, if you can decentralize (via cost minimization) the choice of inputs for a given output by naming input prices, then there should additionally exist a price of output that will allow you to further decentralize (via profit maximization) the choice of output *and* inputs. Furthermore, we can readily intuit how *we* may *construct* this output price, as follows.

Let the marginal cost at input prices $\boldsymbol{p_x}$ of producing an incremental unit of output at the level y * be denoted $MC(y^*|\boldsymbol{p_x})$. Then, intuitively, the output "price"

$$p_y \equiv MC(y^*|\boldsymbol{p_x}) \tag{4}$$

is giving exactly the marginal tradeoff between output and the cost of inputs. This price should "support" the output level y *. The augmented *(n* + 1) *-dimensional* price vector that should competitively "support" $(y^*, \boldsymbol{x}^*)$ is then

$$\boldsymbol{P} \equiv (p_y, -\boldsymbol{p_x}) \ . \tag{5}$$

Slightly more formally, our economic intuition tells us that $(y^*, \boldsymbol{x}^*)$ should emerge as the decentralized solution of the following profit maximization problem:

maximize

$$p_y \, y \ - \ \boldsymbol{p_x} \cdot \boldsymbol{x} \ , \tag{6}$$

subject to

$$(y, \boldsymbol{x}) \ \varepsilon \ S \ . \tag{7}$$

We are now about to show that *the above statement is just exactly the separating hyperplane theorem.* More exactly, we will show that a version of the separating hy-perplane theorem (from which all of the other main versions may be derived) can be proved by induction on the number of dimensions of what we are interpreting as "inputs" by just paralleling rigorously the logical sequence outlined above.

Note the extremely intuitive economic content of the above statement. Suppose an ideal (or idealized) central planner wishes to produce a given output at minimum cost. Suppose, further, that the planner can decentralize the choice of inputs by naming the input prices and instructing the managers to choose inputs to minimize costs for the given output level. The theorem says that the pianner can *further* decentralize (by one more dimension) the input *and* output decisions. by additionally naming the marginal cost of production at the desired output level as the relevant output price, and then instructing the managers to maximize profits. The separating hyperplane theorem is essentially the statement that this can always be done with any concave production function.

In most economic problems of interest the relevant prices, outputs, and inputs will be non-negative. The mode of proof we are about to employ, however, while having the above cost-function interpretation. does not anywhere in the proof actually force us to commit ourselves on the issue of whether the entities under discussion are. in fact, non-negative or not.

3. Three Lemmas for the initial Case $n=1$

We will assume without proof the following three intuitive lemmas, which constitute the initial induction step.

Lemma I. *If* $A \subset E^1$ *is a closed convex set on the one-dimensional line. then A has one of the following three possible forms:* $(-\infty, b_1], [b_2, \infty), [b_3, b_4]$, *where* b,, b,, b,, b_4 *represent finite) boundary points of A and* $b_3 \leq b_4$.

Lemma 2 (one-dimensional version of separating or supporting hyperplane theorem). If A is a closed convex set on the one-dimensional line, with boundary b, then there exists a scalar $\gamma \neq 0$ *such that*

$$x \varepsilon A \Rightarrow \gamma x \leq \gamma b \quad . \tag{8}$$

Lemma 3. *Let f(y) be a convex function defined over the closed one-dimensional convex set A.*

Choose any $a \varepsilon A$*. Then, there exists a number called* $f'(a)$ *such that for all* $y \varepsilon A$

$$f(y) - f(a) \geq f'(a)(y-a) \ . \tag{9}$$

Armed with these three intuitive lemmas for the initial case $n=1$, we are now ready to state and prove a basic member of the family of separating hyperplane results, which is sometimes called the "supporting hyperplane theorem."

4. The Supporting Hyperplane Theorem

Theorem. Let $S \subset E^m$ *be a closed convex set. Let* z^* *be any boundary point of S. Then there exists a m-dimensional vector* $P \neq 0$ *such that*

$$z \ \varepsilon \ S \ \Rightarrow \ P \cdot z \leq P \cdot z^* \ . \tag{10}$$

Proof: By Lemma 2, the theorem is true for m=l. Now suppose it is true for $m=n$, with n an arbitrary positive integer. Then we must prove that the theorem is also true for $m=n+l$.

Let z be a (n+l)-dimensional vector, x be a n-dimensional vector, and y be a 1-dimensional scalar. Think of z being decomposed into y and x as follows:

$$z = (y, x) \ . \tag{11}$$

Define parametrically the n-dimensional set

$$X(y) \equiv \{x \mid (y, x) \varepsilon S\} \ . \tag{12}$$

As economists, we naturally gravitate towards the *interpretation* that y is the single "output" and x is the vector of n "inputs," while $X(y)$ is the "input set" of all possible "inputs" that can produce y. Note, however, that there is nothing in the formal mathematics of the proof that *requires* such an interpretation. We simply choose to *make* such an interpretation because it turns out to be natural. helpful. and reassuring. As economists, such connections seem reassuring because they make the mathematics come alive and "speak" to us in familiar economic terms – so that we both trust the logical conclusions of the formal mathematical apparatus more, and at the same time fear less our struggles with this austere opponent.

Note that since S is a convex set. then X(y) is a convex set for all given y because

$$(\lambda y+(1-\lambda)y, \lambda \boldsymbol{x}+(1-\lambda)\boldsymbol{x}') \in S \ . \tag{13}$$

Furthermore, since S is closed. then $X(y)$ is also closed. and. since $(y, \boldsymbol{x}^*)$ is a boundary point of S. then $\boldsymbol{x}^*$ is aiso a boundary point of $X(y^*)$. It follows by the supporting-hyperplane induction assumption for dimension n, that there exists a "input" price vector $\boldsymbol{p_x} \neq \boldsymbol{0}$ such that

$$\boldsymbol{x} \in X(y^*) \ \Rightarrow \ \boldsymbol{p_x}\cdot\boldsymbol{x} \ \geq \ \boldsymbol{p_x}\cdot\boldsymbol{x}^* \ . \tag{14}$$

Consider now the one-dimensional set

$$A \ \equiv \{y | \ X(y) \ \textit{is non-empty}) \ . \tag{15}$$

Then it is readily shown that A is convex and closed, and thus is of the form described by Lemma 1. Next, the *cost function induced bv input prices* $\boldsymbol{p_x}$ is defined for $y \in A$ by the equation

$$C_{p_x}(y) \ \equiv \ \min_{x \in X(y)} \boldsymbol{p_x}\cdot\boldsymbol{x} \ . \tag{16}$$

It follows directly from applying the definition (16) to condition (14) that

$$C_{p_x}(y^*) \ = \ \boldsymbol{p_x}\cdot\boldsymbol{x}^* \ . \tag{17}$$

The function $C_{p_x}(y)$ is convex because if $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{x}}'$ are solutions of (16) corresponding, respectively, to y and y $'$, then

$$(\lambda y+(1-\lambda)y', \lambda\hat{\boldsymbol{x}}+(1-\lambda)\hat{\boldsymbol{x}}') \;\varepsilon\; s \quad , \qquad \textbf{(18)}$$

for $0\leq\lambda\leq 1$, implying, because $C_{\boldsymbol{p}_x}(y)$ is a minimum function, that

$$C_{\boldsymbol{p}_x}(\lambda y+(1-\lambda)y') \leq \boldsymbol{p}_x\cdot(\lambda\hat{\boldsymbol{x}}+(1-\lambda)\hat{\boldsymbol{x}}') = \lambda\boldsymbol{p}_x\cdot\hat{\boldsymbol{x}}+(1-\lambda)\boldsymbol{p}_x\cdot\hat{\boldsymbol{x}}' = \lambda C_{\boldsymbol{p}_x}(y)+(1-\lambda)C_{\boldsymbol{p}_x}(y') \quad . \qquad (19)$$

With (16) defining a cost function, by Lemma 3 there is some number, which might be called the "marginal cost of producing y^*" and is denoted here as p_y, that satisfies for all $y\varepsilon A$ the condition

$$C_{\boldsymbol{p}_x}(y) - C_{\boldsymbol{p}_x}(y^*) \;\geq\; p_y\,(y - y^*) \quad . \qquad \textbf{(20)}$$

We are now practically at the point where we have the conclusion that we want. We just need now to repackage (16), (17), (20) as the desired result.

We know from the definition (16) that

$$(y, \boldsymbol{x}) \;\varepsilon\; S \;\Rightarrow\; C_{\boldsymbol{p}_x}(y) \leq \boldsymbol{p}_x\cdot\boldsymbol{x} \;. \qquad \textbf{(21)}$$

Combining (21) with (20) with (17), we then have

$$(y, \boldsymbol{x}) \;\varepsilon\; S \;\Rightarrow\; \boldsymbol{p}_x\cdot\boldsymbol{x} - \boldsymbol{p}_x\cdot\boldsymbol{x}^* \geq p_y(y - y^*) \;. \qquad \textbf{(22)}$$

Define the $(n+1)$-dimensional price vector

$$\boldsymbol{P} \;\equiv\; (p_y, -\boldsymbol{p}_x) \;. \qquad \textbf{(23)}$$

Because $\boldsymbol{p}_x\neq\boldsymbol{0}$, it must also be true that $\boldsymbol{P}\neq\boldsymbol{0}$.

Then, with the decompositions (11) and (23), condition (22) can be translated as saying that

$$z \varepsilon S \Rightarrow \boldsymbol{P}\cdot z \leq \boldsymbol{P}\cdot z^{*} , \qquad \textbf{(24)}$$

which is equivalent to (8) holding for $m = \boldsymbol{n} + 1$. ■

5. Conclusion

All of the other main separating hyperpiane results are derivable from (8). Thus, a proof by induction has been provided here of the family of separating hyperplane theorems, which uses directly the language and intuition of one of the most powerful concepts in economic theory -- the idea that there exist shadow prices that support efficient production outcomes by decentralized profit maximization. This is one more piece of evidence that the mathematical and economic consequences of convexity run in deep parallel courses

References

Debreu, G., 1959, Separation theorems for convex sets, in T. C. Koopmans and A. F. Bausch, Selected topics in economics involving mathematical reasoning, SIAM Review 1. 79-148.

Gale, D., 1960, The theory of linear economic models (McGraw-Hill).

Koopmans, T. C., 1957, Three essays on the state of economic science (McGraw-Hill).