

Intermediate Mathematical Economics (Econ 6170)

Fall 2013

Notes on Separation Theorems

1 A Basic Separation Theorem for a Closed Convex Set

The basic separation theorem covered in this section is concerned with the separation of a non-empty, *closed*, convex set from a point not belonging to the set with a hyperplane.

Proposition 1 *Let A be a non-empty, closed and convex subset of \mathbb{R}^n . Let $b \in \mathbb{R}^n$ be a point which does not belong to A . Then, there is $p \in \mathbb{R}^n$, $p \neq 0$, and $\alpha \in \mathbb{R}$ such that:*

$$\left. \begin{array}{l} (i) \quad px \leq \alpha \text{ for all } x \in A \\ (ii) \quad pb > \alpha \end{array} \right\} \quad (1)$$

Proof. Step 1:

The first step is to find a point in A that is “nearest” to b . This would be straightforward if A was bounded. But, that information is not given. So, what is called for is a modified application of Weierstrass theorem.

To this end, note that since A is non-empty, there is some $a \in A$. And, since $b \notin A$, we have $b \neq a$, and so $r \equiv d(a, b) > 0$, where d is the Euclidean metric on $\mathbb{R}^n \times \mathbb{R}^n$. Let $B = \{x \in \mathbb{R}^n : d(x, b) \leq r\}$, and define $C = A \cap B$. Notice that $a \in A$ and $a \in B$, so $a \in C$. Further, B is closed in \mathbb{R}^n (it is a closed ball with center b and radius r , and it is straightforward to check that it is a closed set in \mathbb{R}^n), and so is A . Thus, C is a closed set, since it is an intersection of two closed sets. [This, too, is straightforward to verify]. Finally, B is bounded, since it is contained in an open ball with center b and radius $(r + 1)$. Since $C \subset B$, C must be bounded too.

The function $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$, defined by $f(x) = d(x, b)$ is a continuous function on \mathbb{R}^n , and hence on C . Thus, by Weierstrass theorem, there is $c \in C$, such that:

$$d(x, b) \geq d(c, b) \text{ for all } x \in C \quad (2)$$

Since $a \in C$, we have from (2) that:

$$d(a, b) \geq d(c, b) \quad (3)$$

Now, if $x \in A$, but $x \notin C$, then by the definition of C , it must be the case that $x \notin B$. This means that $d(x, b) > r$, by definition of B . Thus, if $x \in A$ and $x \notin C$, we can infer from (3) that:

$$d(x, b) > r = d(a, b) \geq d(c, b) \quad (4)$$

Using (2) and (4), we have:

$$d(x, b) \geq d(c, b) \text{ for all } x \in A \quad (5)$$

That is, among the points in A , c is a “nearest point” to b .

Step 2:

The next step uses the point $c \in A$ to define a hyperplane, and to check that given the definition, property (1) is satisfied. Define:

$$p = (b - c) \text{ and } \alpha = pc \quad (6)$$

Note that $p \in \mathbb{R}^n$ and $p \neq 0$, while $\alpha \in \mathbb{R}$. The hyperplane is $H = \{x \in \mathbb{R}^n : px = \alpha\}$, and (1) then says that the set A lies in one half-space $H_1 = \{x \in \mathbb{R}^n : px \leq \alpha\}$ associated with this hyperplane, while the point b belongs to the other half-space $H_2 = \{x \in \mathbb{R}^n : px > \alpha\}$.

Using (6), we have:

$$pb = p(b - c) + pc = pc + \|p\|^2 > pc = \alpha \quad (7)$$

Thus (1)(ii) is satisfied, and it remains to verify (1)(i).

Let x be an arbitrary point in A . Since $c \in A$ and A is a convex set, $x(t) \equiv tx + (1 - t)c$ is in A for all $0 < t < 1$. Thus, by (5),

$$d(x(t), b) \geq d(c, b) \text{ for all } 0 < t < 1 \quad (8)$$

We can write:

$$\begin{aligned} [d(x(t), b)]^2 &= [t(x - b) + (1 - t)(c - b)][t(x - b) + (1 - t)(c - b)] \\ &= t^2\|x - b\|^2 + 2t(1 - t)(c - b)(x - b) + (1 - t)^2\|c - b\|^2 \end{aligned} \quad (9)$$

Using (8) and (9) we obtain for all $0 < t < 1$,

$$t^2\|x - b\|^2 + 2t(1 - t)(c - b)(x - b) + (t^2 - 2t)\|c - b\|^2 \geq 0 \quad (10)$$

Dividing through by $t > 0$ in (10), we obtain for all $0 < t < 1$:

$$t\|x - b\|^2 + 2(1 - t)(c - b)(x - b) + (t - 2)\|c - b\|^2 \geq 0 \quad (11)$$

Letting $t \rightarrow 0$ in (11),

$$(b - c)(x - b) \leq -\|c - b\|^2 \quad (12)$$

This can be rewritten as:

$$px \leq pb - pp = p(b - (b - c)) = pc = \alpha \quad (13)$$

Since x was an arbitrary point in A , (13) establishes (1)(i). ■

2 Separation Theorem for Convex Sets

The standard version of the separation theorem does not use *closedness* of the sets. We present a standard version of the separation theorem which will suffice for most applications. This is obtained by using the Proposition stated in the previous section. Naturally, the interest focuses on how the closedness assumption is dispensed with.

Theorem 1 *Let X be a non-empty convex set in \mathbb{R}^n , which is disjoint from $\Omega \equiv \mathbb{R}_{++}^n$. Then, there is $p \in \mathbb{R}_+^n$, $p \neq 0$, such that:*

$$\left. \begin{aligned} (i) \quad &px \leq 0 \text{ for all } x \in X \\ (ii) \quad &px > 0 \text{ for all } x \in \Omega \end{aligned} \right\} \quad (14)$$

Proof. Define $Y = X - \Omega \equiv \{y \in \mathbb{R}^n : y = x - \omega \text{ with } x \in X \text{ and } \omega \in \Omega\}$. Then Y is non-empty (since X is non-empty) and convex (since both X and Ω are convex). Further $0 \notin Y$. Otherwise there would be $x \in X$ and $\omega \in \Omega$ such that $0 = x - \omega$ and this would mean $x = \omega$, which contradicts the fact that X is disjoint from Ω . One could apply Proposition 1 to 0 and the set Y if Y was closed; but this information is not given. So we proceed as follows.

Define A to be the closure of Y ; that is, A consists of all the points of closure of Y . Then, A is a closed set, and it is non-empty, since $Y \subset A$. Convexity of A can be checked as follows. Let $a, a' \in A$ and $0 < \lambda < 1$; we have to show that $[\lambda a + (1 - \lambda)a'] \in A$. Given any $\varepsilon > 0$, there exist $y, y' \in Y$, such that $d(y, a) < \varepsilon$ and $d(y', a') < \varepsilon$. Since Y is convex, $[\lambda y + (1 - \lambda)y'] \in Y$, and:

$$\begin{aligned} d([\lambda y + (1 - \lambda)y'], [\lambda a + (1 - \lambda)a']) &= \|\lambda(y - a) + (1 - \lambda)(y' - a')\| \\ &\leq \lambda\|(y - a)\| + (1 - \lambda)\|(y' - a')\| \\ &< \lambda\varepsilon + (1 - \lambda)\varepsilon = \varepsilon \end{aligned}$$

using the triangle inequality for the Euclidean norm. Thus, $[\lambda a + (1 - \lambda)a']$ is also a point of closure of Y , and so $[\lambda a + (1 - \lambda)a'] \in A$.

Now, we could try to apply Proposition 1 to 0 and the set A . But while $0 \notin Y$, it is possible for 0 to belong to A (which consists of all the points of closure of Y), if 0 is a point of closure of Y (while not being an element of Y itself). So, we proceed as follows.

Let $t > 0$, and let u be the sum vector $(1, 1, \dots, 1)$ in \mathbb{R}^n . We claim that tu does not belong to A . For if $tu \in A$, then given $\varepsilon = (t/2) > 0$, there is some $y \in Y$ such that $d(y, tu) < \varepsilon = (t/2)$. This entails that $y_i > t - \varepsilon = (t/2) > 0$ for all $i \in \{1, \dots, n\}$. That is, $y \gg 0$ belongs to Y . But, then, there must exist $x \in X$ and $\omega \in \Omega$, such that $x - \omega = y$, and this implies that $x = \omega + y \gg 0$, which contradicts the fact that X is disjoint from Ω . This establishes our claim.

We can now apply Proposition 1 to tu and the set A (for every $t > 0$). We will then get $q(t) \in \mathbb{R}^n, q(t) \neq 0$, and $\beta(t) \in \mathbb{R}$ such that:

$$\left. \begin{array}{l} (i) \ q(t)a \leq \beta(t) \text{ for all } a \in A \\ (ii) \ q(t)(tu) > \beta(t) \end{array} \right\} \quad (15)$$

Note that if $q(t)$ has some component $q_j(t)$ negative, then by fixing $x \in X$, and fixing all but the j th component of ω , and taking ω_j sufficiently large, one would violate (15)(i). Thus $q(t) \in \mathbb{R}_+^n$.

Define $p(t) = q(t)/\|q(t)\|$ and $\alpha(t) = \beta(t)/\|q(t)\|$ for each $t > 0$. Then we get $p(t) \in \mathbb{R}_+^n, \|p(t)\| = 1$ and $\alpha(t) \in \mathbb{R}$ such that:

$$\left. \begin{array}{l} (i) \ p(t)a \leq \alpha(t) \text{ for all } a \in A \\ (ii) \ p(t)(tu) > \alpha(t) \end{array} \right\} \quad (16)$$

Combining (16)(i) and (16)(ii), one has $p(t) \in \mathbb{R}_+^n, \|p(t)\| = 1$ such that:

$$p(t)a \leq p(t)(tu) \text{ for all } a \in A \quad (17)$$

Now let t take the values $(1/m)$ for $m = 1, 2, 3, \dots$. Then, $\{p(1/m)\}_1^\infty$ is a bounded sequence in \mathbb{R}^n , and has a convergent subsequence converging to $p \in \mathbb{R}^n$. Then $p \in \mathbb{R}_+^n, \|p\| = 1$, and using (17), one has:

$$pa \leq 0 \text{ for all } a \in A \quad (18)$$

Since $Y \subset A$, one has $p \in \mathbb{R}_+^n, \|p\| = 1$, and:

$$py \leq 0 \text{ for all } y \in Y \quad (19)$$

Let $\omega = \lambda u$ for $\lambda > 0$, and let x be an arbitrary point in X . Then, $[x - \lambda u] \in Y$, and by using (19),

$$px \leq p\lambda u \text{ for all } x \in X \text{ and all } \lambda > 0 \tag{20}$$

Finally, letting $\lambda \rightarrow 0$ in (20), we obtain (14)(i). Since $p \in \mathbb{R}_+^n$ and $p \neq 0$, (14)(ii) clearly holds for all $x \in \Omega$. ■