8

Convex Sets, Separation Theorems, and Non-Convex Sets in R^N

Definition A set of points S in $\mathbf{R}^N$ is said to be convex if the line segment between any two points of the set is completely included in the set, that is, S is ***convex*** if $x, y \in S$ implies $\{z \mid z = \alpha x + (1-\alpha)y, 0 \leq \alpha \leq 1\} \subseteq S$.

S is said to be strictly convex if $x, y{\in}S, x{\neq}y, 0{<}\alpha{<}1$ implies $\alpha x{+}(1{-}\alpha)y \in$ interior S.

The notion of convexity is that a set is convex if it is connected, has no holes on the inside, and has no indentations on the boundary. Figure 2.3 displays convex and nonconvex sets. A set is strictly convex if it is convex and has a continuous strict curvature (no flat segments) on the boundary.

Properties of convex sets Let C_1 and C_2 be convex subsets of $\mathbf{R}^N$. Then

$C_1 \cap C_2$ is convex,
$C_1 + C_2$ is convex,
$\overline{C}_1$ is convex.

Proof See Exercise 8.1.

The concept of convexity of a set in $\mathbf{R}^N$ is essential in mathematical economic analysis. This reflects the importance of continuous point-valued optimizing behavior. To understand the importance of convexity, consider for a moment what will happen when it is absent. Suppose widgets are consumed only in discrete lots of 100. The insistence on discrete lots is a nonconvexity. Suppose a typical widget eater at some prices to be indiffer-

Former Figure 2.3 goes here

Fig. 8.1. Convex and nonconvex sets.

ent between buying a lot of 100 and buying 0. He will definitely not buy a fractional lot. At a low price, he will want to buy a lot of 100. As prices increase he will become indifferent at some price, say at p^*, between 0 and 100. At still higher prices, he will demand 0. The demand curve has a gap at p^*. Demand is set-valued (consisting of the two points 0 and 100) and appears discontinuous[1] at p^* . With a gap that big in the demand curve, it is clear that there may be no intersection of supply and demand and hence no equilibrium. It is to prevent this family of difficulties that we will focus on convexity (until Chapter 25 and the concluding sections of this chapter and Chapter 22).

Strict convexity typically will assure uniqueness (point-valuedness) of maxima. Conversely, when opportunity sets or preferences are nonconvex (not convex), optimizing behavior of firms or households may jump between discrete noncontiguous points as prices vary.

8.1 Separation theorems

The Separating Hyperplane Theorem says that if we have two disjoint convex sets in $\mathbf{R}^N$ we can find a (hyper)plane between them so that one of the two sets is above the plane and the other below. The plane separates the convex sets. Because the plane is linear, it is defined by an equation that looks like a price system for N commodities. The Bounding Hyperplane Theorem leads to a similar interpretation. When the economy is described by the convex sets representing tastes (convex upper contour sets) or technology, we can use the separation theorems to characterize an efficient allocation as sustained by a price system. We'll see this in Chapters 18 and 22.

All of the sets and vectors we treat here will be in $\mathbf{R}^N$. Let $p \in \mathbf{R}^N, p \neq 0$. Then we define a hyperplane with normal p and constant k to be a set of the form $H \equiv \{x \mid x \in \mathbf{R}^N, p \cdot x = k\}$, where k is a real number. Note that for any two vectors, x and y, in $H, p \cdot (x - y) = 0$. H divides $\mathbf{R}^N$ into two subsets, the portion "above" H and the portion "below" as measured by the dot product of p with points of $\mathbf{R}^N$. The closed half space above H is defined as the set $\{x \mid x \in \mathbf{R}^N, p \cdot x \geq k\}$. The closed half space below H is defined as $\{x|x \in \mathbf{R}^N, p \cdot x \leq k\}$. H is said to be bounding for $S \subset \mathbf{R}^N$ if S is a subset of one of the two half spaces defined by H.

Lemma 8.1 Let K be a nonempty closed convex subset of $\mathbf{R}^N$, and let $z \in \mathbf{R}^N$, $z \notin K$. Then there is $y \in K$ and $p \in \mathbf{R}^N, p \neq 0$, so that $p{\cdot}z < k =$

[1] The set-valued demand function in this case is upper hemi-continuous but not convex-valued. This is a concept developed in chapters 23, 24, and 25.

Former Figure 2.7 goes here

Fig. 8.2. Bounding and separating hyperplanes for convex sets.

$p{\cdot}y \leq p{\cdot}x$ for all $x \in K$.

The lemma says that for a nonempty, closed, convex set K (not including the whole space) there is a hyperplane separating K from a point outside the set.

Proof of Lemma 8.1 Choose $y \in K$ as the closest point in K to z. That is, y minimizes $|x - z|$ for all $x \in K$ (continuity of the Euclidean norm and closedness of K ensure that a minimizer exists). Now we define $p = y - z$ and $k = p{\cdot}y$.

We must demonstrate that $p{\cdot}z < k$ and that $p{\cdot}x \geq k$ for all $x \in K$. The first of these follows directly: $p{\cdot}z = p{\cdot}z - p{\cdot}y + p{\cdot}y = -p{\cdot}p + p{\cdot}y < k$. Consider $x \in K$. We must show that $p{\cdot}x \geq k$. Since K is convex, we know that every point w on the line segment between x and y, $w = \alpha x + (1-\alpha)y, 1 \geq \alpha \geq 0$, is an element of K. We will show that the proposition $p{\cdot}x < k$ leads to a contradiction. $w = y + \alpha(x - y)$. Consider

$$
\begin{aligned}
|z-y|^2 - |z-w|^2 &= |z-y|^2 - |(z-y) - \alpha(x-y)|^2 \\
&= (z-y){\cdot}(z-y) - [(z-y){\cdot}(z-y) - 2\alpha(z-y){\cdot}(x-y) \\
&\quad -\alpha^2(x-y){\cdot}(x-y)] \\
&= -2\alpha p{\cdot}(x-y) - \alpha^2(x-y){\cdot}(x-y) \\
&= -\alpha[2p{\cdot}(x-y) + \alpha(x-y){\cdot}(x-y)].
\end{aligned}
$$

Recall that $p{\cdot}y = k$. Suppose, contrary to hypothesis, that $p{\cdot}x < k$. Then $p{\cdot}(x-y) = p{\cdot}x - p{\cdot}y < 0$. Then for α sufficiently small, $|z-y|^2 - |z-w|^2 > 0$ and hence $|z - y| > |z - w|$. But this is a contradiction. The point y was chosen as the element of K closest to z. There can be no w in K closer to z than y.

The contradiction proves the lemma. QED

Theorem 8.1 (Bounding Hyperplane Theorem (Minkowski)) Let K be convex, $K \subset \mathbf{R}^N$. There is a hyperplane H through z and bounding for K if z is not interior to K.

Proof If $z \notin \overline{K}$, then the existence of H follows directly from the lemma. If $z \in$ boundary K, then consider a sequence $z^\nu \notin \overline{K}, z^\nu \to z$. Let p^ν be the corresponding sequence of normals to the supporting hyperplane, chosen to have length unity. The sequence is in a closed bounded set (the unit sphere). It thus has a convergent subsequence, whose limit is the required normal. QED

Theorem 8.2 (Separating Hyperplane Theorem) Let $A, B \subset \mathbf{R}^N$; let A and B be nonempty, convex, and disjoint, that is $A \cap B = \phi$. Then there is $p \in \mathbf{R}^N, p \neq 0$, so that $p \cdot x \geq p \cdot y$, for all $x \in A, y \in B$.

Proof Consider $K = A - B$. K is convex. Since A and B are disjoint, $0 \notin K$. Then, by the lemma, there is p so that $p \cdot z \geq p \cdot 0 = 0$ for all $z \in K$. If we let $z = x - y$ then $p \cdot x \geq p \cdot y$. QED

The hyperplane with normal p is said to separate A and B. Bounding and separating hyperplanes are presented in Figure 8.2.

8.2 The Shapley-Folkman Theorem

Properties of convex sets are developed above in this chapter and in Chapter 9. We'll find throughout the rest of this book how useful the convexity property is. However, not all economic relations can conveniently be described using convex sets. Some relations (typically involving economies of scale or specialization in consumption or production) are best described using nonconvex sets. There is a remarkable family of results, the Shapley-Folkman Theorem, that tells us that the sum of a large number of nonconvex sets — though still nonconvex — is approximately convex. The nonconvexities do not compound each other indefinitely.

The overwhelming majority of results in mathematical general equilibrium theory follow from the study of convex sets (above) and from the fixed point theorems that apply in convex settings (chapter 9). The results on nonconvex sets below are a bit technical — the first-time reader may skip them. They are useful in dealing with small scale economies and preferences for concentrated consumption (chapter 25) and for the most general proofs of convergence of the core of an economy (chapter 22, section 22.4).

8.2.1 Nonconvex sets and their convex hulls

A typical nonconvex set contains a hole or indentation.

Example Consider $V^1 = \{x \in R^2 | 3 \leq |x| \leq 10\}$. V^1 is a disk in R^2 with a hole in the center. The hole makes it nonconvex. Let $V^2 = \{x \in R^2| \; |x| \leq 10; x_1 \geq 0 \text{ or } x_2 \geq 0\}$. V^2 is the disk of radius 10 centered at the origin with the lower left quadrant omitted. The indentation at the lower left makes V^2 nonconvex.

The convex hull of a set S will be the smallest convex set containing S. The convex hull of S will be denoted con(S). We can define con(S) , for $S \subset R^N$ as follows

$$\text{con(S)} \equiv \{x \mid x = \sum_{i=0}^{N} \alpha^i x^i, \text{where } x^i \in \text{S}, \alpha^i \geq 0 \text{ all i, and } \sum_{i=0}^{N} \alpha^i = 1\}.$$

or equivalently as

$$\text{con(S)} \equiv \bigcap_{S \subset \; T; T \; convex} T.$$

That is con(S) is the smallest convex set in R^N containing S.

Example con$(V^1) = \{x \in R^2 | |x| \leq 10\}$, and con$(V^2) = \{x \in R^2 | |x| \leq 10$ for $x_1 \geq 0$ or $x_2 \geq 0$; for x_1, $x_2 \leq 0$, $x_1 + x_2 \geq -10\}$. Taking the convex hull of a set means filling in the holes just enough to make the amended set convex.

8.2.2 The Shapley-Folkman Lemma

Most economic analysis uses convex sets. We'd like a means to formalize the distinction between economic behavior characterized by convex sets versus nonconvex sets. One way to represent this distinction is to look at the discrepancy between a nonconvex set and its convex hull, $con(S) \setminus S$. This focus leads to the Shapley-Folkman Theorem. We'll now confine attention to compact sets. The Theorem tells us that the result of summing up a large number of compact nonconvex sets is an approximately convex set. The theorem makes the approximation more precise.

Lemma (Shapley-Folkman): Let S^1, S^2, S^3, ...,S^m, be nonempty compact subsets of R^N. Let $x \in con(S^1 + S^2 + S^3 + \ldots + S^m)$. Then for each i=1,2,...,m, there is $y^i \in con(S^i)$ so that $\sum_{i=1}^m y^i = x$ and with at most N exceptions, $y^i \in S^i$. Equivalently: Let F be a finite family of nonempty compact sets in R^N and let $y \in con(\sum_{S \in F} S)$. Then there is a partition of F into two disjoint subfamilies F' and F'' with the number of elements in $F' \leq N$ so that $y \in \sum_{S \in F'} con(S) + \sum_{S \in F''} S$.

To see how the lemma works, let's take a simple example. Let's start with ten identical subsets of R^2. Let $S^i = \{(0,0),(0,1),(1,0),(1,1)\}$ for i = 1, 2, ..., 10. Each of the sets S^i consists of four points, the four corners of a square in R^2 with one corner at the origin and sides lying on the co-ordinate axes. Now consider $con(S^1 + S^2 + S^3 + \ldots + S^{10})$. $con(S^1 + S^2 + S^3 + \ldots + S^{10}) = \{x | x \in R^2, 0 \leq x_1, x_2 \leq 10\}$. Choose a typical point in $con(S^1 + S^2 + S^3 + \ldots + S^{10})$, say $x = (5.5, 5.7)$. The lemma says that x can be represented as a sum of points in the convex hulls of the original sets, $con(S^1)$, $con(S^2)$, $\ldots$, $con(S^{10})$. More important, the theorem says that x can be represented in this way as a sum of points most (all but two in R^2) coming from the original sets S^1, S^2, S^3, $\ldots S^{10}$, not from points of their convex hulls that were not part of the original sets S^i. In this example, there are many choices of x^i that will fulfill the theorem. For example, Let $x^1 = (0.5, 0) \in con(S^1)$, $x^2 = (0, 0.7) \in con(S^2)$, $x^3 = (1,1) \in S^3$, $x^4 = (1,1) \in S^4$, $x^5 = (1,1) \in S^5$, $x^6 = (1,1) \in S^6$, $x^7 = (1,1) \in S^7$, $x^8 = (0,0) \in S^8$, $x^9 = (0,0) \in S^9$, $x^{10} = (0,0) \in S^{10}$. Then $x = \sum_{i=1}^{10} x^i$, all $x^i \in con(S^i)$ and with only two exceptions $x^i \in S^i$. This is just what the Shapley-Folkman Lemma asserts.

8.2.3 Measuring Non-Convexity, The Shapley-Folkman Theorem

We now introduce a scalar measure of the size of a non-convexity.

Definition: The radius of a compact set S is defined as
$rad(S) \equiv \inf_{x \in R^N} \sup_{y \in S} |x - y|$.

That is, $rad(S)$ is the radius of the smallest closed ball containing S.

Theorem 8.3 (Shapley - Folkman): Let F be a finite family of compact subsets $S \subset R^N$ and $L > 0$ so that $rad(S) \leq L$ for all $S \in F$. Then for any $x \in con(\sum_{S \in F} S)$ there is $y \in \sum_{S \in F} S$ so that $|x - y| \leq L\sqrt{N}$.

The significance of the Shapley-Folkman theorem is that the sum of a large number of compact non-convex sets is approximately convex. We start with a family of sets F whose elements $S \in F$ are of $rad(S)$, the measure of size, less than or equal to L. The measure of the size of a nonconvexity suggested here is the distance between a point of the convex hull and the nearest point of the underlying set. Adding a few sets together may increase the size of the nonconvexity in the sum; but eventually the radius of the nonconvexity is limited by an upper bound of $L\sqrt{N}$. As additional sets are added, their nonconvexities do not compound one another; the nonconvexity of the sum does not become progressively larger. The size of the holes or indentations in the summation does not grow as additional summands are added. As additional sets are added, the sum of the sets will typically become larger, but nonconvexities in the sum are bounded above; they do not grow. Speaking imprecisely, we could say that the sum becomes approximately convex (as a proportion of the size of the sum) as the number of sets in the summation becomes large.

8.2.4 Corollary: A tighter bound

Definition: We define the inner radius of $S \subset R^N$ as
$r(S) \equiv \sup_{x \in con(S)} \inf_{T \subset S; x \in con(T)} rad(T)$

Corollary 8.1 Corollary to the Shapley-Folkman Theorem: Let F be a finite family of compact subsets $S \subset R^N$ and $L > 0$ so that $r(S) \leq L$ for all $S \in F$. Then for any $x \in con(\sum_{S \in F} S)$ there is $y \in \sum_{S \in F} S$ so that $|x - y| \leq L\sqrt{N}$.

The Corollary and its interpretation here are very similar to the Shapley-Folkman Theorem. The Theorem is stated in terms of the radius of spheres circumscribing the summands. The Corollary is stated in terms of the radius of spheres inscribed in the nonconvexities of the summands. Again, the interpretation is that after a finite number of sets are added, the addition of more sets to the summation will not increase the size of the nonconvexities while it increases the size of the summation. Thus, as a proportion of the size of the sum, or the number of summands, the sum of sets becomes approximately convex as the number of summands grows.

8.3 Bibliographic Note

Chapter 1 of Debreu (1959), provides an excellent concise survey of the mathematical results presented here and in Chapter 23. Green and Heller (1981) provide a very thorough treatment of convexity. Separation theorems are well expounded in Hildenbrand and Kirman (1988). A complete statement of the Shapley-Folkman Lemma, Theorem, and corollary together with their proofs is available in Arrow and Hahn (1971), Appendix B. The theorem and proof, due to L.S. Shapley and J.H. Folkman, was first published in Starr(1969).

Exercises

8.1 Demonstrate the following properties of convex sets in $\mathbf{R}^N$. Let A and B be convex subsets of $\mathbf{R}^N$. Then $A \cap B$ is convex, $A + B$ is convex, and $\overline{A}$ is convex.

8.2 Consider a closed square (two-dimensional cube) in $\mathbf{R}^2$ with side [0, 2]:

$$C = [0, 2] \times [0, 2] = \{(x, y)|0 \leq x \leq 2, 0 \leq y \leq 2\}.$$

Demonstrate that C is a convex set. That is, let (x^1, y^1) and $(x^2, y^2) \in C$. Let $0 \leq \alpha \leq 1$. Let $z = \alpha(x^1, y^1) + (1 - \alpha)(x^2, y^2)$. Show that $z \in C$.

8.3 Recall the Separating Hyperplane Theorem (Theorem 8.2):
Let $A, B \subset \mathbf{R}^N$, where A and B are nonempty convex sets, with disjoint interiors. Then there is $p \in \mathbf{R}^N, p \neq 0$, so that $p{\cdot}x \geq p{\cdot}y$ for all $x \in A, y \in B$.

(i) Show by (counter)example (a well-drawn figure is sufficient) that the convexity of both A and B are typically required to ensure this result. That is, show that if either of A or B is nonconvex then there may be no separating hyperplane.

(ii) Let $A, B \subset \mathbf{R}^2$. Let $A = \{(x, y) \mid x^2 + y^2 \leq 1\}$, the closed disk of radius one centered at the origin, and let $B = \{(x, y) \mid (x - 2)^2 + y^2 \leq 1\}$, the closed disk of radius one centered at (2,0). Show that A and B fulfill the conditions of the Separating Hyperplane Theorem and specify a separating hyperplane, including its normal.