1 Notes for 640:621; Separation of Convex Sets

Our first major theorem, stating conditions for no arbitrage in a one period market, will rely on a separation theorem for convex sets. Separation theorems are at the heart of the theory of convex sets and functions, which in turn is very important to applied mathematics. Thus, it is worthwhile to learn a bit of separation theory. This is the purpose of these supplementary notes, which are based on material from R. Tyrrell Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

1.1 Mathematical setting and notation

The setting of our discussion will be the space $\mathbb{R}^n$ of n-vectors of real numbers. We will think of the elements of $\mathbb{R}^n$ as column vectors,

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

The zero vector is the vector consisting of all zero entries. It will be denoted by $\underline{0}$ to distinguish it from the number 0. To get the associated row vector we use the transpose, $x^* = (x_1, \ldots, x_n)$. The inner product (called the dot product in dimensions $n = 2$ and $n = 3$) between vectors in $\mathbb{R}^n$ is defined as

$$x \cdot y \stackrel{\triangle}{=} x^* y = \sum_{i=1}^{n} x_i y_i.$$

The inner product is used to define orthogonality of vectors in higher dimension: vector x is *orthogonal* to vector y if and only if $x \cdot y = 0$. The length $\|x\|$ of a vector x is defined by extending Pythagoras's formula to general dimensions:

$$\|x\| \stackrel{\triangle}{=} \sqrt{x \cdot x} = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}.$$

The distance between two vectors x and y is simply the length $\|y - x\|$ of their vector difference.

Our discussion will make use of the concept of a closed set in $\mathbb{R}^n$. A closed subset of $\mathbb{R}^n$ is one which contains its limit points. That is, K is closed if $x_n \to x$ as $n \to \infty$ and $x_n \in K$ for every n imply that $x \in K$ also. Alternately, a set is closed if its complement is open. This means that K is closed if, whenever x is a point **not** in K, there is an $\epsilon > 0$ so that all points a distance less than ϵ from x are also **not** in K. It can be shown that the two definitions of a closed set are equivalent.

1.2 Hyperplanes and separation.

With the inner product in hand, we can define hyperplanes in arbitrary dimension. Given a vector $\alpha \neq \underline{0}$ in $\mathbb{R}^n$ and a scalar β, the hyperplance determined by α and β is the set of solutions x in $\mathbb{R}^n$ to the linear equation

$$\alpha \cdot x = \beta. \tag{1}$$

(Written out, this equation is $\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta$.)

Let H denote the hyperplane determined by vector $\alpha \neq \underline{0}$ and scalar β. What does H look like? First, one may observe by a direct and simple calculation that $x_0 = (\beta/\|\alpha\|)\alpha$ lies in H:

$$\alpha \cdot x_0 = \alpha \cdot \left[\frac{\beta}{\alpha \cdot \alpha}\alpha\right] = (\frac{\beta}{\alpha \cdot \alpha})\alpha \cdot \alpha = \beta.$$

Now let y be another point in H, so that $\alpha \cdot y = \beta$. Consider the vector $y - x_0$ pointing from the head of x_0 to the head of y. Then

$$\alpha \cdot (y - x_0) = \alpha \cdot y - \alpha \cdot x_0 = \beta - \beta = 0.$$

The vector $y - x_0$ is thus orthogonal to α; writing $w = y - x_0$, we conclude that any y in H may be written $y = x_0 + w$ where w is orthogonal to α.

Conversely, let z be any vector orthogonal to α. then $\alpha \cdot (x_0 + z) = \alpha \cdot x_0 + \alpha \cdot z = \beta + 0 = \beta$. Thus $x_0 + z$ is in the hyperplane H.

We have thus shown that H is precisely all vectors of the form $x_0 + z$ where z is orthogonal to α. So, geometrically, H may be pictured as the $n - 1$ dimensional plane in $\mathbb{R}^n$ that passes through x_0 and is orthogonal to α. (Loosely speaking, H has dimension $n - 1$ because $n - 1$ of the variables in the equation $\alpha \cdot x = \beta$ can be freely chosen and their values uniquely determine the value of the remaining variable.) Hyperplanes are thus easy to visualize in dimensions two and three. In $\mathbb{R}^2$, the hyperplane determined by α and β is one of the straight lines orthogonal to α; in dimension $\mathbb{R}^3$, it is one of the 2 dimensional planes orthogonal to α.

A hyperplane H determined by α and β divides $\mathbb{R}^n$ into two closed half spaces, each half space consisting of the points that lie to one or the other side of H: mathematically these half spaces are the sets of points

$$\{x : \alpha \cdot x \leq \beta\} \qquad \text{and} \qquad \{x : \alpha \cdot x \geq \beta\}.$$

The intersection of these two half spaces is clearly H itself. H is said to *separate* two subsets A and B of $\mathbb{R}^n$ if A lies in one of these half spaces and B in the other. Unfortunately this definition does not preclude both A and B being contained in H itself, which is an uninteresting case. So we say that H *properly separates* A and B if A lies in one of the half spaces determined by H and B in the other and both are not contained in H. For example let A be a line segment on the x_1-axis in $\mathbb{R}^2$ and let B consist just of one point not on the x_1-axis. Then the x_1-axis properly separates A and B.

Let us next turn to a completely algebraic formulation of proper separation. Suppose H is determined by vector α and scalar β and that $A \subset \{x : \alpha \cdot x \leq \beta\}$, that $B \subset \{x : \alpha \cdot x \geq \beta\}$ and that at least one of A and B is not wholly contained in H. Then

$$\begin{aligned} \alpha \cdot x \leq \beta \leq \alpha \cdot y \qquad & \text{for every } x \in A \text{ and every } y \in B, \\ \alpha \cdot \tilde{x} < \alpha \cdot \tilde{y} \qquad & \text{for at least one } \tilde{x} \text{ in } A \text{ and one } \tilde{y} \text{ in B.} \end{aligned}$$

Eliminating the middle man β we can write this as

$$\alpha \cdot x \leq \alpha \cdot y \qquad \text{for every } x \in A \text{ and every } y \in B, \tag{2}$$

$$\alpha \cdot \tilde{x} < \alpha \cdot \tilde{y} \qquad \text{for at least one } \tilde{x} \text{ in } A \text{ and one } \tilde{y} \text{ in B.} \tag{3}$$

Conversely, if (2) and (3) hold for some vector α, then there is a hyperplane properly separating A and B; to see this take β to be the supremum of $\alpha \cdot x$ over all values x in A. Thus, (2) and (3) give an equivalent characterization of separation without explicit mention of β.

1.3 Convex sets

Let x and y be vectors in $\mathbb{R}^n$. A convex combination of x and y is a vector of the form $tx + (1-t)y$, where $t \in [0,1]$. The set of convex combinations of x and y is simply the line segment connecting x to y, when x and y are interpreted as points in $\mathbb{R}^n$.

A subset C of $\mathbb{R}^n$ is convex if, whenever x and y are in C so also is any convex combination of x and y; in other words, C contains the line segment connecting x and y if it contains both x and y. The next figure illustrates a simple convex region in the plane.

A *cone* in $\mathbb{R}^n$ is a subset K satisfying the property that $x \in K$ implies λx is also in K for all scalars $\lambda > 0$. Note that a non-empty closed cone, that is a cone that is also closed, must necessarily contain the zero vector $\underline{0}$. To see this, take x in a closed cone K and a sequence $\{\lambda_k\}$ of positive numbers such that $\lambda_k \downarrow 0$ as $k \to \infty$. Then $\lambda_k x$ is in K for each k and $\lambda_k x \to \underline{0}$ as $k \to \infty$. Hence $\underline{0}$ is a limit of elements in K, and since K is closed, must also be in K. Notice that any vector subspace of $\mathbb{R}^n$, for example a line through the origin, is a cone. The next few pictures give examples of cones in the plane. The cone in the last figure is a sector in the plane including its bounding rays. Notice that it is also closed and convex. The separation theorem we will directly use for stating no arbitrage conditions concerns closed convex cones.

1.4 Separation of convex sets and convex cones

If you draw two, non-intersecting, bounded convex sets in the plane, it is clear by visual inspection that they can be separated by a hyperplane. We can even allow the two convex sets to intersect at boundary points and it is clear we are still able to separate them.

Theorems stating conditions under which two convex sets may be separated are fundamental to the theory of convex sets and functions. The first separation theorem we give is rather specialized, as it applies only to closed convex cones. However it is stated directly in the form we need for obtaining a mathematical characterization of no arbitrage. The full proof of this theorem is worked out in a series of exercises in the following section.

Theorem 1 *Let K and M be closed convex cones in $\mathbb{R}^n$. Then*

(i) K *and* M *intersect only in the origin.*

(ii) K *does not contain a line through the origin.*

if and only if there is a non-zero $\alpha \in \mathbb{R}^n$ such that

$$\alpha \cdot x > 0 \qquad \text{for every } x \in K,\ x \neq \underline{0}; \tag{4}$$

$$\alpha \cdot y \leq 0 \qquad \text{for every } y \in M. \tag{5}$$

Exercise. One part of the proof of Theorem 1 is easy. Show that if K and M are closed convex cones and conditions (4) and (5) hold, then K and M satisfy (i) and (ii).

The pictures below illustrate separating hyperplanes (lines actually) for convex cones in the plane satisfying the hypotheses of the theorem. Notice that in one example, M and the separating hyperplane coincide.

Theorem 1 is specialized. To show the reader that convex sets can indeed be separated under fairly minimal assumptions, we will state a general separation theorem in Chapter 11 of the Rockafellar, Convex Analyis.. Roughly, the theorem says that we can separate two convex sets as long as they do not intersect at "interior" points. But we must first be careful to define what an interior point of a convex set is. Ordinarily, a point x_0 is said to be in the interior of a subset A of $\mathbb{R}^n$ if there is some ball of non-zero radius centered at x_0 and totally contained in A. By this definition, the line segment A in the figure below has no interior points as a subset of the plane. However, it is clear that if B is any other convex set in the plane that intersects A at most in one of the endpoints of A, A and B can be separated. Thus we would like to say the all the points of A except the endpoints are interior points *relative* to A.

Taking our cue from this discussion, we will define the *relative interior* of a convex set A. To do this, first consider for two distinct vector x and y in $\mathbb{R}^n$, the map

$$I(t) = y + t(x - y).$$

When $t = 0$, $I(0) = y$; when $t = 1$, $I(1) = x$; and as t moves from 0 to 1, the values of $I(t)$ trace out the line segment connecting y to x. As t increases past a, $I(t)$ moves beyond x in the direction of the vector $x-y$. Now let A be a convex set in $\mathbb{R}^n$ and let x be in A. We will say that x is in the relative boundary of A if there exists at least one y, also in A and not equal to x, so that $y+t(x-y)$ is not in A for $t>1$. This means that, if we start at this particular y in A and move toward x along the line connecting y to x, we will exit A as soon as we pass x. Hence it is appropriate to call x a boundary point. The relative interior of A is the set of points in A which are not relative boundary points. (Note: Rockafellar defines relative interior using the notion of the affine hull of a convex set; to avoid introducing affine hulls we have used this more direct, but equivalent, definition.) Recall that proper separation of two sets is equivalent to the existence of a vector α for which conditions (2) and (3) hold. The next theorem states the general criterion for separation of convex sets.

Theorem 2 *Let A and B be non-empty convex sets. Then A and B can be properly separated if and only if the relative interiors of A and B do not intersect.*

Notice that in this statement no assumption of boundedness is made.

1.5 Exercises towards a proof of Theorem 1.

Some of these exercises are very elementary. Others have varying levels of difficulty, but do require knowing that closed, bounded sets in $\mathbb{R}^n$ are compact and that every sequence of points in a compact set contains a converging subsequence. The proof of Theorem 1 is developed in problem 5; lemmas for this proof are treated in problems 3 and 4. Problem 4 is already of independent interest as it is a rather general theorem on the separation of closed, convex sets. The reader may wish to read the statements of problems 3 and 4 and use them to do problem 5, before actually tackling problems 3 and 4. The proof developed in problem 5 comes from Appendix B of D. Duffie, *Dynamic Asset Pricing*, Princeton University Press, 2001.

1. The notion of convex combination can be generalized to more than two vectors. Let $x^{(1)},\ldots,x^{(m)}$ be vectors in $\mathbb{R}^n$. A linear combination of the form

$$\sum_{j=1}^{m} s_j x^{(j)} \qquad \text{where } s_1\geq 0,\ldots,s_m\geq 0 \text{ and } \textstyle\sum_{j=1}^m s_j=1,$$

is called a convex combination of $x^{(1)},\ldots,x^{(m)}$. Show that if C is convex, then a convex combination of any number of vectors in C is again in C.

2. Let B be a subset of $\mathbb{R}^n$, not necessarily convex. The convex hull of B is the smallest convex set that contains B. Show that the convex hull of B is given by the collection of all convex combinations of two or more elements of B.

3. Let $x^{(1)}, \ldots, x^{(m)}$ be vectors in $\mathbb{R}^n$ and consider the convex combination

$$y = \sum_{j=1}^{m} s_j x^{(j)} \qquad \text{where } s_1 \geq 0, \ldots, s_m \geq 0 \text{ and } \textstyle\sum_{j=1}^{m} s_j = 1, \tag{6}$$

Suppose $m > n+1$. Show that we can write y as a linear combination of at most $n+1$ of the vectors $x^{(1)}, \ldots, x^{(m)}$. This result is known as Caratheodory's theorem.

Hint: Consider the vectors

$$\begin{pmatrix} 1 \\ x^{(1)} \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ x^{(m)} \end{pmatrix}$$

in $\mathbb{R}^{n+1}$. Since $m > n+1$, this collection of linear vectors is necessarily linearly dependent, and so there are are coefficients, $a_1, \ldots, a_m$, not all zero, such that

$$\underline{0} = a_1 x^{(1)} + \cdots + a_m x^{(m)} \tag{7}$$

Of course we can multiply both sides by any scalar λ:

$$\underline{0} = \lambda a_1 x^{(1)} + \cdots + \lambda a_m x^{(m)} \tag{8}$$

Adding the equations (6) and 8) together yields

$$y = \sum_{1}^{m} (s_j + \lambda a_j)\, x^{(j)}.$$

Show that one can choose λ so that $a_i = 0$ for some i and that the remaining coefficients $s_j + \lambda a_j$ are non-negative and sum to 1. This will represent y as a convex combination of no more than $m-1$ of the vectors $x^{(1)}, \ldots, x^{(m)}$. Keep repeating the procedure until at most $n+1$ vectors are left.

4. Let S be a closed, bounded convex subset of $\mathbb{R}^n$. Show that the convex hull of S is also closed and bounded. (The boundedness of the convex hull is easy. To show that it is closed, one can use the result of problem 3 and the definition of closed set. Or consider the map

$$F(x^{(1)}, \ldots, x^{(n+1)}, s_1, \ldots, s_{n+1}) = s_1 x^{(1)} + \cdots + s_{n+1} x^{(n+1)}$$

on $B \times \cdots (n+1)\text{-times} \cdots \times B \times \triangle$, where $\triangle$ is the simplex

$$\left\{ (s_1, \ldots, s_{n+1}) \;:\; s_i \geq 0 \text{ for every } i, \textstyle\sum s_i = 1 \right\}.)$$

5. Let A and B be disjoint, closed convex sets and assume one of them, say A, is also bounded. Show that A and B can be *strictly* separated: there is a vector such that

$$\alpha \cdot x > \alpha \cdot y \qquad \text{for every } x \text{ in } A \text{ and } y \text{ in } B.$$

(Hint: First prove that there is a vector y_0 in B and a vector x_0 in A such that the length of $y_0 - x_0$ is the smallest possible of all vectors from a point in A to a point in B:

$$\|y_0 - x_0\| \le \|y - x\| \qquad \text{for every } x \text{ in } A \text{ and ever } y \in B.$$

This step uses the compactness of B. Show that $\alpha = y_0 - x_0$ is non-zero and supplies the desired separating vector.)

6. (Proof of Theorem 1.) We have already assigned the "if" part of the part of the proof as an exercise given immediately after the statement of Theorem 1. This problem does the "only if" part. Let K and M be given as in the statement of Theorem 1.

Let $S \triangleq K \cap \{x; |x| = 1\}$ be the intersection of K with the unit sphere in $\mathbb{R}^n$. Show that S is closed and bounded (easy).

Next let A be the convex hull of S. Show that $\underline{0} \notin A$ and that A and M are disjoint.

Now, using the results of problems 4 and 5, conclude there is a vector α which separates A and M. Show that this vector satisfies conditions (4) and (5) of Theorem 1.

2 Fundamental Theorem of Asset Pricing for the one-period, finite state model.

This discussion is modeled directly on Chapter 1, pp. 3-5, of D. Duffie, *Dynamic Asset Pricing*. In this discussion, if B is a matrix or a vector, B^* denotes its transpose.

We consider a one period model with p assets and m states $\{\omega_1, \ldots, \omega_m\}$. The prices of the assets at time 0 are denoted by $S_1(0), \ldots, S_p(0)$. It is convenient to gather them into a vector

$$A(0) = \begin{pmatrix} S_1(0) \\ \vdots \\ S_p(0) \end{pmatrix}.$$

The prices $S_i(1)(\omega_i)$, of the assets at time 1, if the state is ω_i, are collected in the vector

$$A(1)(\omega_i) = \begin{pmatrix} S_1(1)(\omega_i) \\ \vdots \\ S_p(1)(\omega_i) \end{pmatrix}.$$

This notation may seem a little overdone, but it is consistent with the notation we will need for the continuous-time, infinite state case, so setting up it up now will help us make the transition to the more complicated model.

A portfolio will be represented by a vector π in $I\!R^n$. Component π_i of π represents the number of units of asset i held in the portfolio. The value of the portfolio at time 0 is

$$\sum_{j=1}^{p} S_j(0)\pi_j = A^*(0)\pi.$$

Similarly, the value of a portfolio π at time 1, when the state is ω_i is $A^*(1)(\omega_i)\pi$. A portfolio π is an arbitrage if $A^*(0)\pi \leq 0$, $A_i^*(1)\pi \geq 0$ for every i, and at least one of these portfolio values is non-zero.

To undertake the analysis, we will collect the prices at time 1 into a matrix:

$$\mathcal{A} \stackrel{\triangle}{=} \begin{pmatrix} A(1)(\omega_1) & \cdots & A(1)(\omega_m) \end{pmatrix} = \begin{pmatrix} S_1(1)(\omega_1) & \cdots & S_1(1)(\omega_m) \\ S_2(1)(\omega_1) & \cdots & S_2(1)(\omega_m) \\ \vdots & \vdots & \vdots \\ S_p(1)(\omega_1) & \cdots & S_p(1)(\omega_m) \end{pmatrix}$$

This is a $p \times m$ matrix; row i lists the prices of asset i at time 1 for all possible states, and column j lists the prices of all the assets if the economy is in state ω_j. Observe that

$$\mathcal{A}^*\pi = \begin{pmatrix} A^*(1)(\omega_1)\pi \\ \vdots \\ A^*(1)(\omega_m)\pi \end{pmatrix}.$$

Let M be the set of all vectors in $\mathbb{R}^{m+1}$ of the form

$$\begin{pmatrix} -A^*(0)\pi \\ \mathcal{A}^*\pi \end{pmatrix} = \begin{pmatrix} -A^*(0)\pi \\ A^*(1)(\omega_1)\pi \\ \vdots \\ A^*(1)(1)(\omega_m)\pi \end{pmatrix}, \tag{9}$$

where π is a vector in $\mathbb{R}^p$. Notice that M is a subspace of $I\!R^{m+1}$ and hence it is a closed cone.

Let K be the set in $I\!R^{m+1}$ consisting of all vectors x, such that $x_i \geq 0$ for each component i, $1 \leq i \leq m+1$. K is also a closed, convex cone and K contains no lines.

Now here is the principal point of the construction so far. A portfolio π is an arbitrage if and only if

$$\begin{pmatrix} -A(0)\pi \\ \mathcal{A}^*\pi \end{pmatrix}$$

is a non-zero element of K. But this vector belongs to the set M. Thus arbitrage occurs if and only if the closed cones M and K intersect at a point other than the zero vector $\underline{0}$. Equivalently,

the market admits no arbitrage if and only if $K \cap M = \{\underline{0}\}$.

But Theorem 1 says that $K \cap L = \{\underline{0}\}$ if and only if there is a vector α in $\mathbb{R}^{m+1}$ such that

$$\alpha \cdot x > 0 \qquad \text{for every } x \in K,\ x \neq \underline{0}; \tag{10}$$

$$\alpha \cdot y \leq 0 \qquad \text{for every } y \in L. \tag{11}$$

Hence there is no arbitrage if and only if there is a vector α in $\mathbb{R}^{m+1}$ satisfying conditions (10) and (11). Thus we have reduced the characterization of no arbitrage to a mathematical condition, but to get a result in interesting and usable form, we need to explore condition (10) and (11) more. We will use them to prove the following result, which is the version for a one-period, finite state model of the so-called Fundamental Theorem of Asset Pricing.

Theorem 3 *The one-period market defined by $A(0)$ and $\mathcal{A}$ admits no arbitrage if and only if there is a vector ψ in $I\!R^m$ such that each entry of ψ is strictly positive and ψ solves*

$$A(0) = \mathcal{A}\psi. \tag{12}$$

The vector ψ is called a state-price vector. To understand what it does, consider what equation (12) says, entry by entry. For each asset i, it states that

$$A_i(0) = \sum_{j=1}^{m} \psi_j A_i(1)(\omega_j).$$

This is a representation of the price of asset i at time 0 as a weighted sum of its possible prices at time 1. The theorem says that no arbitrage holds if and only if there is a set of positive weights, namely the entries of ψ, that can be used to represent the price of each asset at time 0 in terms of its possible prices at time 1.

Theorem 3 provides a checkable condition: find all solutions to (12) (say, by Gaussian elimination) and check to see if there is one having all positive entries.

Proof: It is first of all easy to see that the existence of ψ satisfying (12) implies no arbitrage, because then

$$A^*(0)\pi = \psi^* \mathcal{A}^* \pi$$

If every entry of $\mathcal{A}^*\pi$ is non-negative, then, since the entries of ψ are all positive, it follows that $A^*(0)\pi \geq 0$ also.

Now suppose there is no arbitrage. Then we know there is a vector α in $\mathbb{R}^{m+1}$ satisfying the two conditions (10) and (11). For convenience in labeling, number the components of α starting from 0:

$$\alpha^* = (\alpha_0, \alpha_1, \ldots, \alpha_m).$$

We first observe the (10) implies that the entries of α are all strictly positive. Indeed, for any i, $0 \leq i \leq m$, let e_i be the vector with a 1 in component i and a zero in every other component. Then e_i is a non-zero vector in K and so, by (10) $0 < \alpha \cdot e_i = \alpha_i$.

Next, observe that, by the definition of M, (11) is equivalent to saying that

$$-\alpha_0 A^*(0)\pi + (\alpha_1, \ldots, \alpha_m)\mathcal{A}^*\pi \leq 0 \quad \text{for all } \pi \in \mathbb{R}^p. \tag{13}$$

But this implies in turn that actually

$$-\alpha_0 A^*(0)\pi + (\alpha_1, \ldots, \alpha_m)\mathcal{A}^*\pi = 0 \quad \text{for all } \pi \in \mathbb{R}^p. \tag{14}$$

Indeed suppose there is a π for which

$$-\alpha_0 A^*(0)\pi + (\alpha_1, \ldots, \alpha_m)\mathcal{A}^*\pi < 0.$$

Then for the portfolio $-\pi$, $-\alpha_0 A^*(0)(-\pi) + (\alpha_1, \ldots, \alpha_m)\mathcal{A}^*(-\pi) > 0$, and this contradicts (13).

Finally, (14) can hold only if $\alpha_0 A(0) = \mathcal{A} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}$. This implies that the vector $\psi = \alpha_0^{-1} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}$ satisfies (12). $\diamond$