MA3484—Methods of Mathematical Economics
School of Mathematics, Trinity College
Hilary Term 2017
Lecture 21 (March 10, 2017)

David R. Wilkins

5.5. A Separating Hyperplane Theorem

Definition

A subset K of $\mathbb{R}^m$ is said to be *convex* if $(1-\mu)\mathbf{x} + \mu\mathbf{x}' \in K$ for all elements $\mathbf{x}$ and $\mathbf{x}'$ of K and for all real numbers μ satisfying $0 \leq \mu \leq 1$.

It follows from the above definition that a subset K of $\mathbb{R}^{>}$ is a convex subset of $\mathbb{R}^m$ if and only if, given any two points of K, the line segment joining those two points is wholly contained in K.

Theorem 5.9

Let m be a positive integer, let K be a closed convex set in $\mathbb{R}^m$, and let $\mathbf{b}$ be a vector in $\mathbb{R}^m$, where $\mathbf{b} \notin K$. Then there exists a linear functional $\varphi\colon \mathbb{R}^m \to \mathbb{R}$ and a real number c such that $\varphi(\mathbf{x}) > c$ for all $\mathbf{x} \in K$ and $\varphi(\mathbf{b}) < c$.

Proof

It follows from Lemma 5.8 that there exists a point $\mathbf{g}$ of K such that $|\mathbf{x} - \mathbf{b}| \geq |\mathbf{g} - \mathbf{b}|$ for all $\mathbf{x} \in K$. Let $\mathbf{x} \in K$. Then $(1 - \lambda)\mathbf{g} + \lambda\mathbf{x} \in K$ for all real numbers λ satisfying $0 \leq \lambda \leq 1$, because the set K is convex, and therefore

$$|(1 - \lambda)\mathbf{g} + \lambda\mathbf{x} - \mathbf{b}| \geq |\mathbf{g} - \mathbf{b}|$$

for all real numbers λ satisfying $0 \leq \lambda \leq 1$. Now

$$(1 - \lambda)\mathbf{g} + \lambda\mathbf{x} - \mathbf{b} = \mathbf{g} - \mathbf{b} + \lambda(\mathbf{x} - \mathbf{g}).$$

It follows by a straightforward calculation from the definition of the Euclidean norm that

$$\begin{aligned} |\mathbf{g} - \mathbf{b}|^2 &\leq |(1-\lambda)\mathbf{g} + \lambda \mathbf{x} - \mathbf{b}|^2 \\ &= |\mathbf{g} - \mathbf{b}|^2 + 2\lambda(\mathbf{g} - \mathbf{b})^T(\mathbf{x} - \mathbf{g}) \\ &\qquad + \lambda^2 |\mathbf{x} - \mathbf{g}|^2 \end{aligned}$$

for all real numbers λ satisfying $0 \leq \lambda \leq 1$. In particular, this inequality holds for all sufficiently small positive values of λ, and therefore

$$(\mathbf{g} - \mathbf{b})^T(\mathbf{x} - \mathbf{g}) \geq 0$$

for all $\mathbf{x} \in K$.

Let

$$\varphi(\mathbf{x}) = (\mathbf{g} - \mathbf{b})^T \mathbf{x}$$

for all $\mathbf{x} \in \mathbb{R}^m$. Then $\varphi \colon \mathbb{R}^m \to \mathbb{R}$ is a linear functional on $\mathbb{R}^m$, and $\varphi(\mathbf{x}) \geq \varphi(\mathbf{g})$ for all $\mathbf{x} \in K$. Moreover

$$\varphi(\mathbf{g}) - \varphi(\mathbf{b}) = |\mathbf{g} - \mathbf{b}|^2 > 0,$$

and therefore $\varphi(\mathbf{g}) > \varphi(\mathbf{b})$. It follows that $\varphi(\mathbf{x}) > c$ for all $\mathbf{x} \in K$, where $c = \frac{1}{2}\varphi(\mathbf{b}) + \frac{1}{2}\varphi(\mathbf{g})$, and that $\varphi(\mathbf{b}) < c$. The result follows. ∎

5.6. Convex Cones

Definition

Let m be a positive integer. A subset C of $\mathbb{R}^m$ is said to be a *convex cone* in $\mathbb{R}^m$ if $\lambda\mathbf{v} + \mu\mathbf{w} \in C$ for all $\mathbf{v}, \mathbf{w} \in C$ and for all real numbers λ and μ satisfying $\lambda \geq 0$ and $\mu \geq 0$.

Lemma 5.10

Let m be a positive integer. Then every convex cone in $\mathbb{R}^m$ is a convex subset of $\mathbb{R}^m$.

Proof

Let C be a convex cone in $\mathbb{R}^m$ and let $\mathbf{v}, \mathbf{w} \in C$. Then $\lambda\mathbf{v} + \mu\mathbf{w} \in C$ for all non-negative real numbers λ and μ. In particular $(1-\lambda)\mathbf{w} + \lambda\mathbf{v} \in C$. whenever $0 \leq \lambda \leq 1$, and thus the convex cone C is a convex set in $\mathbb{R}^m$, as required. ∎

Lemma 5.11

Let S be a subset of $\mathbb{R}^m$, and let C be the set of all elements of $\mathbb{R}^m$ that can be expressed as a linear combination of the form

$$s_1\mathbf{a}^{(1)} + s_2\mathbf{a}^{(2)} + \cdots + s_n\mathbf{a}^{(n)},$$

where $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ are vectors belonging to S and $s_1, s_2, \ldots, s_n$ are non-negative real numbers. Then C is a convex cone in $\mathbb{R}^m$.

Proof

Let $\mathbf{v}$ and $\mathbf{w}$ be elements of C. Then there exist finite subsets S_1 and S_2 of S such that $\mathbf{v}$ can be expressed as a linear combination of the elements of S_1 with non-negative coefficients and $\mathbf{w}$ can be expressed as a linear combination of the elements of S_2 with non-negative coefficients. Let

$$S_1 \cup S_2 = \{\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}\}.$$

Then there exist non-negative real numbers $s_1, s_2, \ldots, s_n$ and $t_1, t_2, \ldots, t_n$ such that

$$\mathbf{v} = \sum_{j=1}^{n} s_j \mathbf{a}^{(j)} \quad \text{and} \quad \mathbf{w} = \sum_{j=1}^{n} t_j \mathbf{a}^{(j)}.$$

Let λ and μ be non-negative real numbers. Then

$$\lambda \mathbf{v} + \mu \mathbf{w} = \sum_{j=1}^{n} (\lambda s_j + \mu t_j) \mathbf{a}^{(j)},$$

and $\lambda s_j + \mu t_j \geq 0$ for $j = 1, 2, \ldots, n$. It follows that $\lambda \mathbf{v} + \mu \mathbf{w} \in C$, as required. ∎

Proposition 5.12

Let m be a positive integer, let $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)} \in \mathbb{R}^m$, and let C be the subset of $\mathbb{R}^m$ defined such that

$$C = \left\{ \sum_{j=1}^{n} t_j \mathbf{a}^{(j)} : t_j \geq 0 \text{ for } j = 1, 2, \ldots, n \right\}.$$

Then C is a closed convex cone in $\mathbb{R}^m$.

Proof

It follows from Lemma 5.11 that C is a convex cone in $\mathbb{R}^m$. We must prove that this convex cone is a closed set.

The vectors $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ span a vector subspace V of $\mathbb{R}^m$ that is isomorphic as a real vector space to $\mathbb{R}^k$ for some integer k satisfying $0 \leq k \leq m$. This vector subspace V of $\mathbb{R}^m$ is a closed subset of $\mathbb{R}^m$, and therefore any subset of V that is closed in V will also be closed in $\mathbb{R}^m$. Replacing $\mathbb{R}^m$ by $\mathbb{R}^k$, if necessary, we may assume, without loss of generality that the vectors $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ span the vector space $\mathbb{R}^m$. Thus if A is the $m \times n$ matrix defined such that $(A)_{i,j} = (\mathbf{a}^{(j)})_i$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$ then the matrix A is of rank m.

Let $\mathcal{B}$ be the collection consisting of all subsets B of $\{1, 2, \ldots, n\}$ for which the members of the set $\{\mathbf{a}^{(j)} : j \in B\}$ constitute a basis of the real vector space $\mathbb{R}^m$ and, for each $B \in \mathcal{B}$, let

$$C_B = \left\{ \sum_{i=1}^{m} s_i \mathbf{a}^{(j_i)} : s_i \geq 0 \text{ for } i = 1, 2, \ldots, m \right\},$$

where $j_1, j_2, \ldots, j_m$ are distinct and are the elements of the set B. It follows from Lemma 5.7 that the set C_B is closed in $\mathbb{R}^m$ for all $B \in \mathbb{B}$.

Let $\mathbf{b} \in C$. The definition of C then ensures that there exists some $\mathbf{x} \in \mathbb{R}^n$ that satisfies $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$. Thus the problem of determining $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$ has a feasible solution. It follows from Theorem 4.2 that there exists a basic feasible solution to this problem, and thus there exist distinct integers $j_1, j_2, \ldots, j_m$ between 1 and n and non-negative real numbers $s_1, s_2, \ldots, s_m$ such that $\mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}, \ldots, \mathbf{a}^{(j_m)}$ are linearly independent and

$$\mathbf{b} = \sum_{i=1}^{m} s_i \mathbf{a}^{(j_i)}.$$

Therefore $\mathbf{b} \in C_B$ where

$$B = \{j_1, j_2, \ldots, j_m\}.$$

We have thus shown that, given any element $\mathbf{b}$ of C, there exists a subset B of $\{1, 2, \ldots, n\}$ belonging to $\mathcal{B}$ for which $\mathbf{b} \in C_B$. It follows from this that the subset C of $\mathbb{R}^m$ is the union of the closed sets C_B taken over all elements B of the finite set $\mathcal{B}$. Thus C is a finite union of closed subsets of $\mathbb{R}^m$, and is thus itself a closed subset of $\mathbb{R}^m$, as required. ∎

5.7. Farkas' Lemma

Proposition 5.13

Let C be a closed convex cone in $\mathbb{R}^m$ and let $\mathbf{b}$ be a vector in $\mathbb{R}^m$. Suppose that $\mathbf{b} \notin C$. Then there exists a linear functional $\varphi\colon \mathbb{R}^m \to \mathbb{R}$ such that $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in C$ and $\varphi(\mathbf{b}) < 0$.

Proof

Suppose that $\mathbf{b} \notin C$. The cone C is a closed convex set. It follows from Theorem 5.9 that there exists a linear functional $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ and a real number c such that $\varphi(\mathbf{v}) > c$ for all $\mathbf{v} \in C$ and $\varphi(\mathbf{b}) < c$.

Now $\mathbf{0} \in C$, and $\varphi(\mathbf{0}) = 0$. It follows that $c < 0$, and therefore $\varphi(\mathbf{b}) \leq c < 0$.

Let $\mathbf{v} \in C$. Then $\lambda\mathbf{v} \in C$ for all real numbers λ satisfying $\lambda > 0$. It follows that $\lambda\varphi(\mathbf{v}) = \varphi(\lambda\mathbf{v}) > c$ and thus $\varphi(\mathbf{v}) > \dfrac{c}{\lambda}$ for all real numbers λ satisfying $\lambda > 0$, and therefore

$$\varphi(\mathbf{v}) \geq \lim_{\lambda \to +\infty} \frac{c}{\lambda} = 0.$$

We conclude that $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in C$.

Thus $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in C$ and $\varphi(\mathbf{b}) < 0$, as required. ∎

Lemma 5.14

(Farkas' Lemma) *Let* A *be a* $m \times n$ *matrix with real coefficients, and let* $\mathbf{b} \in \mathbb{R}^m$ *be an* m*-dimensional real vector. Then exactly one of the following two statements is true:—*

(i) *there exists* $\mathbf{x} \in \mathbb{R}^n$ *such that* $A\mathbf{x} = \mathbf{b}$ *and* $\mathbf{x} \geq \mathbf{0}$*;*

(ii) *there exists* $\mathbf{y} \in \mathbb{R}^m$ *such that* $\mathbf{y}^T A \geq \mathbf{0}$ *and* $\mathbf{y}^T \mathbf{b} < 0$*.*

Proof

Let $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ be the vectors in $\mathbb{R}^m$ determined by the columns of the matrix A, so that $(\mathbf{a}^{(j)})_i = (A)_{i,j}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, and let

$$C = \left\{ \sum_{j=1}^{n} x_j \mathbf{a}^{(j)} : x_j \geq 0 \text{ for } j = 1, 2, \ldots, n \right\}.$$

It follows from Proposition 5.12 that C is a closed convex cone in $\mathbb{R}^m$. Moreover

$$C = \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{x} \geq \mathbf{0}\}.$$

Thus $\mathbf{b} \in C$ if and only if there exists $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{b} = A\mathbf{x}$ and $\mathbf{x} \geq \mathbf{0}$. Therefore statement (i) in the statement of Farkas' Lemma is true if and only if $\mathbf{b} \in C$.

If $\mathbf{b} \notin C$ then it follows from Proposition 5.13 that there exists a linear functional $\varphi\colon \mathbb{R}^m \to \mathbb{R}$ such that $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in C$ and $\varphi(\mathbf{b}) < 0$. Then there exists $\mathbf{y} \in \mathbb{R}^m$ with the property that $\varphi(\mathbf{v}) = \mathbf{y}^T\mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^m$. Now $A\mathbf{x} \in C$ for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $\mathbf{x} \geq \mathbf{0}$. It follows that $\mathbf{y}^T A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $\mathbf{x} \geq \mathbf{0}$. In particular $(\mathbf{y}^T A)_i = \mathbf{y}^T A\mathbf{e}^{(i)} \geq 0$ for $i = 1, 2, \ldots, m$, where $\mathbf{e}^{(i)}$ is the vector in $\mathbb{R}^m$ whose ith component is equal to 1 and whose other components are zero. Thus if $\mathbf{b} \notin C$ then there exists $\mathbf{y} \in \mathbb{R}^m$ for which $\mathbf{y}^T A \geq \mathbf{0}$ and $\mathbf{y}^T\mathbf{b} < 0$.

Conversely suppose that there exists $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T A \geq 0$ and $\mathbf{y}^T \mathbf{b} < 0$. Then $\mathbf{y}^T A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $\mathbf{x} \geq \mathbf{0}$, and therefore $\mathbf{y}^T \mathbf{v} \geq 0$ for all $\mathbf{v} \in C$. But $\mathbf{y}^T \mathbf{b} < 0$. It follows that $\mathbf{b} \notin C$. Thus statement (ii) in the statement of Farkas's Lemma is true if and only if $\mathbf{b} \notin C$. The result follows. $\blacksquare$

Corollary 5.15

Let A be a $m \times n$ matrix with real coefficients, and let $\mathbf{c} \in \mathbb{R}^n$ be an n-dimensional real vector. Then exactly one of the following two statements is true:—

(i) *there exists $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T A = \mathbf{c}^T$ and $\mathbf{y} \geq \mathbf{0}$;*

(ii) *there exists $\mathbf{v} \in \mathbb{R}^n$ such that $A\mathbf{v} \geq \mathbf{0}$ and $\mathbf{c}^T \mathbf{v} < 0$.*

Proof

It follows on applying Farkas's Lemma to the transpose of the matrix A that exactly one of the following statements is true:—

(i) there exists $\mathbf{y} \in \mathbb{R}^m$ such that $A^T\mathbf{y} = \mathbf{c}$ and $\mathbf{y} \geq \mathbf{0}$;

(ii) there exists $\mathbf{v} \in \mathbb{R}^m$ such that $\mathbf{v}^T A^T \geq \mathbf{0}$ and $\mathbf{v}^T\mathbf{c} < 0$.

But $\mathbf{v}^T\mathbf{c} = \mathbf{c}^T\mathbf{v}$. Also $A^T\mathbf{y} = \mathbf{c}$ if and only if $\mathbf{y}^T A = \mathbf{c}^T$, and $\mathbf{v}^T A^T \geq \mathbf{0}$ if and only if $A\mathbf{v} \geq \mathbf{0}$. The result follows. ∎

Corollary 5.16

Let A be a $m \times n$ matrix with real coefficients, and let $\mathbf{c} \in \mathbb{R}^n$ be an n-dimensional real vector. Suppose that $\mathbf{c}^T\mathbf{v} \geq 0$ for all $\mathbf{v} \in \mathbb{R}^n$ satisfying $A\mathbf{v} \geq \mathbf{0}$. Then there exists some there exists $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T A = \mathbf{c}^T$ and $\mathbf{y} \geq \mathbf{0}$.

Proof

Statement (ii) in the statement of Corollary 5.15 is false, by assumption, and therefore statement (i) in the statement of that corollary must be true. The result follows. ∎