Farkas' lemma for cones

France Dacar, Jožef Stefan Institute
France.Dacar@ijs.si

April 18, 2012

Let E be a finite-dimensional real vector space, of dimension $n > 0$.

We shall resort to two devious tricks: we shall make E into an Euclidean space, and we shall make use of the topology of E. For the first, we choose a basis — any basis — of E, define the scalar product $x \cdot y := x_1 y_1 + \cdots + x_n y_n$, where x_k and y_k are the coordinates of the vectors x resp. y in the chosen basis, and define the norm $\|x\| := \sqrt{x^2}$ (where $x^2 = x \cdot x$). The topology is the standard topology of E; here, we define it as the smallest topology that makes all linear functionals $E \to \mathbb{R}$ continuous (where $\mathbb{R}$ carries the topology of the real line). Some properties of the topological space E:

- If E' is another finite-dimensional real vector space equipped with the standard topology, then every linear (or affine) mapping $E \to E'$ is continuous.
- Every linear (or affine) isomorphism $E \to \mathbb{R}^n$ is a homeomorphism of E onto the product topological space.
- For any norm on E, the topology of the corresponding metric space is the standard topology of E.
- If E' is a vector subspace of E, then E' is a closed subset of E, and the standard topology of E' is the same as the subspace topology induced from E.
- A subset X of E is compact if and only if it is closed and bounded; boundedness of X means that the norm (no matter which one) is bounded on X.

We could go on, listing more properties, but these will more than suffice.

First, existence of separating hyperplanes.

Let C be a closed convex subset of E, and $y \in E \setminus C$. Then there exists an affine functional f on E such that $f(x) \leqslant 0$ for every $x \in C$, while $f(y) > 0$.

If C is empty, we choose for f the constant functional 1, and are done; so from now on we assume that $C \neq \varnothing$. There exists $\overline{y} \in C$ that is closest to y, that is, $\alpha := \|y - \overline{y}\| \leqslant \|y - x\|$ for all $x \in C$. (Pick $x_0 \in C$. The set C_0 of all $x \in C$ satisfying $\|x - y\| \leqslant \|x_0 - y\|$ is nonempty and compact, thus the continuous function $\|x - y\|$ of $x \in C_0$ attains a minimum α at some $\overline{y} \in C_0$, and this α is also the minimum of $\|x - y\|$ over all $x \in C$. The point $\overline{y}$ is unique: if $x_1, x_2 \in C$, and $\|y - x_1\| = \|y - x_2\| = \alpha$, then

$$\left\| y - \frac{x_1 + x_2}{2} \right\|^2 = \alpha^2 - \frac{1}{4} \|x_1 - x_2\|^2 \geqslant \alpha^2$$

implies $x_1 = x_2$.) Set $a := y - \overline{y} \neq 0$. We claim that

$$a \cdot (x - \overline{y}) \leqslant 0 \qquad \text{for every } x \in C \ . \tag{1}$$

Let $x \in C$, and let $0 < \lambda \leqslant 1$. The point $x_\lambda := \overline{y} + \lambda(x - \overline{y})$ belongs to C, therefore $(y - x_\lambda)^2 \geqslant \alpha^2 = a^2$. Since $y - x_\lambda = a - \lambda(x - \overline{y})$, we have

$$\big(a - \lambda(x - \overline{y})\big)^2 = a^2 - 2\lambda a \cdot (x - \overline{y}) + \lambda^2 (x - \overline{y})^2 \geqslant a^2 ,$$

and hence

$$a \cdot (x - \overline{y}) \leqslant \frac{\lambda}{2}(x - \overline{y})^2 .$$

In the limit $\lambda \to 0$ we obtain (1). Since $a \cdot (y - \overline{y}) = a^2 > 0$, we have the required affine functional $f(x) := a \cdot (x - \overline{y})$.

Next, convex cones and some of their properties.

A convex cone in E is a subset C of E with the following properties:

(c0) $0 \in C$,

(c1) if $x \in C$ and $\lambda \geqslant 0$, then $\lambda x \in C$,

(c2) if $x, y \in C$, then $x + y \in C$.

Evidently, every convex cone is a convex subset of E.

Since the properties of a convex cone are closure rules, the set of all convex cones in E is a closure system on $\mathscr{P}E$ (it is closed under arbitrary intersections). Each subset X of E generates the smallest convex cone $\operatorname{cone}(X)$ containing the set X. Since the closure rules (c0), (c1), (c2) are finitary (accidentally, the closure rule (ck) is k-ary), the closure operator $\operatorname{cone}(-)$ is finitary: for every subset X of E

$$\operatorname{cone}(X) = \bigcup \{\operatorname{cone}(Z) \mid Z \in \mathscr{P}_{\mathrm{f}} X\}$$

(where $\mathscr{P}_{\mathrm{f}} X$ denotes the set of all finite subsets of X). It is easy to show that $\operatorname{cone}(X)$ is the set of all linear combinations, with non-negative coefficients, of elements of X.

Let $X \subseteq E$ and $y \in \operatorname{cone}(X)$. The point y is in $\operatorname{cone}(Z)$ for some finite subset Z of X, therefore there exists an inclusion-minimal finite set Z with this property.

Carathéodory's theorem for cones. *If X is a subset of E and $y \in \operatorname{cone}(X)$, while $y \notin \operatorname{cone}(X \smallsetminus \{x\})$ for every $x \in X$, then X is a set of linearly independent vectors.*

The point y lies in the convex cone generated by some finite subset of X that cannot be a proper subset of X, hence is the set X itself. It follows that $y = \sum_{x \in X} \lambda_x x$ with $\lambda_x > 0$ for every $x \in X$. Suppose that $\sum_{x \in X} \alpha_x x = 0$ for some $\alpha_x \in \mathbb{R}$ ($x \in X$), not all of them 0; we can assume that $\alpha_{x_0} > 0$ for some $x_0 \in X$ (for otherwise we reverse the signs of all α_x). Let us express y as a linear combination $y - \mu 0 = \sum_{x \in X} (\lambda_x - \mu \alpha_x) x$; for all small enough $\mu > 0$ the coefficients of this linear combination are non-negative, and there is the largest such μ, namely $\mu = \mu_{\max} = \min\{\lambda_x / \alpha_x \mid \alpha_x > 0\} \leqslant \lambda_{x_0} / \alpha_{x_0}$, for which at least one coefficient is zero, contrary to the assumptions.

We have the following useful consequence of the theorem: for every subset X of E, the convex cone generated by X is the union of convex cones generated by the linearly independent subsets of X.

If X is finite, then $\operatorname{cone}(X)$ *is a closed subset of the topological space E.*

Since X has only finitely many linearly independent subsets, and since the union of finitely many closed sets is a closed set, it suffices to prove the assertion in the special case of a linearly independent X. The set X is a basis of the vector subspace E_X of E spanned by X; let $\{x^* \mid x \in X\}$ be the dual basis. Then

$$\operatorname{cone}(X) \;=\; \bigl\{y \in E_X \bigm| x^*(y) \geqslant 0 \text{ for every } x \in X\bigr\}$$

is a closed subset of the closed subspace E_X of E, hence is a closed subset of E.

Let E^* be the dual vector space of E. The relation $f(x) \leqslant 0$ between $f \in E^*$ and $x \in E$ determines a contravariant Galois connection $(\Xi, \Phi) : \mathscr{P}E^* \rightleftharpoons (\mathscr{P}E)^{\mathrm{op}}$, where

$$\Xi(F) \;=\; \bigl\{x \in E \bigm| f(x) \leqslant 0 \text{ for every } f \in F\bigr\}\ , \qquad\qquad F \subseteq E^*\ ,$$

$$\Phi(X) \;=\; \bigl\{f \in E^* \bigm| f(x) \leqslant 0 \text{ for every } x \in X\bigr\}\ , \qquad\qquad X \subseteq E\ .$$

The composite $\Xi\Phi$ is a closure operator on $\mathscr{P}E$, while the composite $\Phi\Xi$ is a closure operator on $\mathscr{P}E^*$. For every subset F of E^*, the set $\Xi(F)$ is a convex cone in E, and for every subset X of E, the set $\Phi(X)$ is a convex cone in E^*.

If X is a finite subset of E, then $\Xi\Phi(X) = \operatorname{cone}(X)$.

If $f \in \Phi(X)$, and $\lambda_x \geqslant 0$ for every $x \in X$, then $f\bigl(\sum_{x\in X} \lambda_x x\bigr) = \sum_{x\in X} \lambda_x f(x) \leqslant 0$, whence $\Xi\Phi(X) \supseteq \operatorname{cone}(X)$.

To prove the opposite inclusion, assume that $y \in E \setminus \operatorname{cone}(X)$; we shall prove that $y \in E \setminus \Xi\Phi(X)$ by exhibiting a linear functional $f \in \Phi(X)$ such that $f(y) > 0$. Put $a := y - \overline{y}$, where $\overline{y}$ is the point of the closed convex subset $\operatorname{cone}(X)$ of E that is closest to y. We know that $a\cdot(x - \overline{y}) \leqslant 0$ for every $x \in \operatorname{cone}(X)$, and that $a\cdot(y - \overline{y}) = a^2 > 0$. We are almost there: all it remains to show is that the constant term $-a\cdot\overline{y}$ of the affine functional $f(x) := a\cdot x - a\cdot\overline{y}$ is 0. This is easy: because the origin 0 and the point $2\overline{y}$ belong to $\operatorname{cone}(X)$, we have $-a\cdot\overline{y} = f(0) \leqslant 0$ and $a\cdot\overline{y} = f(2\overline{y}) \leqslant 0$.

Exchanging the roles of E and E^*, we get

Farkas' lemma for cones. *If F is a finite subset of E^*, then* $\Phi\Xi(F) = \operatorname{cone}(F)$.