A NICE PROOF OF FARKAS LEMMA

DANIEL VICTOR TAUSK

Abstract. The goal of this short note is to present a nice proof of Farkas Lemma which states that if $C$ is the convex cone spanned by a finite set and if $x$ is not in $C$ then there exists a linear functional $\alpha$ that is nonnegative on $C$ with $\alpha(x) < 0$. This implies in particular that the convex cone $C$ is closed.

1. Introduction

Let $V$ be a real vector space. Recall that a subset $C$ of $V$ is said to be convex if for all $x,y \in C$ and all $t \in [0,1]$ we have $(1-t)x + ty \in C$; a simple induction argument shows that $C$ is convex if and only if $\sum_{i=1}^{k} t_i x_i \in C$, for every integer $k \geq 1$, all $x_1,\ldots,x_k \in C$ and all $t_1,\ldots,t_k \geq 0$ with $\sum_{i=1}^{k} t_i = 1$. The set $C$ is said to be a cone if for all $x \in C$, $t \geq 0$ we have $tx \in C$. It is easy to show that $C$ is a convex cone if and only if $\sum_{i=1}^{k} t_i x_i \in C$, for every integer $k \geq 1$, all $x_1,\ldots,x_k \in C$ and all $t_1,\ldots,t_k \geq 0$.

If $\{v_1,\ldots,v_k\}$ is a finite nonempty subset of $C$ then the set:

(1.1) $\left\{ \sum_{i=1}^{k} t_i v_i : t_1,\ldots,t_k \geq 0 \right\}$

is easily seen to be the smallest convex cone containing $\{v_1,\ldots,v_k\}$; we call it the convex cone spanned by $\{v_1,\ldots,v_k\}$.

We have the following:

1.1. Theorem (Farkas lemma). Let $V$ be a real vector space and $\{v_1,\ldots,v_k\}$ a finite nonempty subset of $V$. If $x \in V$ is not on the convex cone (1.1) spanned by $\{v_1,\ldots,v_k\}$ then there exists a linear functional $\alpha \in V^*$ such that $\alpha(x) < 0$ and $\alpha(v_i) \geq 0$, for all $i = 1,\ldots,k$.

Clearly, if a linear functional $\alpha$ is nonnegative on $v_1,\ldots,v_k$ then $\alpha$ is nonnegative on the entire convex cone (1.1); in particular, the condition of $x$ not being in (1.1) is also necessary for the existence of the linear functional $\alpha$ with the properties stated in Theorem 1.1.

The following is a simple consequence of Theorem 1.1:

1.2. Corollary. The convex cone spanned by a finite subset of $\mathbb{R}^n$ is closed in $\mathbb{R}^n$.

Proof. If $C$ is the convex cone spanned by a finite subset of $\mathbb{R}^n$ and if $x \in \mathbb{R}^n$ is not in $C$ then there exists a linear functional $\alpha \in \mathbb{R}^{n^*}$ with $\alpha(x) < 0$ which is nonnegative on $C$. Hence the open half-hyperplane $\{y \in \mathbb{R}^n : \alpha(y) < 0\}$ is an open neighborhood of $x$ contained in the complement of $C$. \qed
If one doesn’t use Farkas Lemma, the thesis of Corollary 1.2 apparently has no immediate proof for it, although it may seem to be a fairly intuitive result. The following example shows that one should be careful with intuition in this matter.

1.3. Example. If \( C \) is an arbitrary compact convex subset of \( \mathbb{R}^n \) then the cone:
\[
(1.2) \quad \{ tx : x \in C, \ t \geq 0 \}
\]
is not necessarily closed in \( \mathbb{R}^n \). For instance, let \( C \) be the closed disc in \( \mathbb{R}^2 \) of radius 1 centered at the point (1,0). In this case, the cone (1.2) is the union of the open half-plane \( \{(x,y) \in \mathbb{R}^2 : x > 0\} \) with the origin. Hence it is not closed in \( \mathbb{R}^2 \).

1.4. Remark. Corollary 1.2 actually holds when \( \mathbb{R}^n \) is replaced by an arbitrary Hausdorff topological vector space \( V \). Namely, the convex cone spanned by a finite subset of \( V \) is contained in a finite-dimensional subspace \( S \) of \( V \), which is complete and thus closed in \( V \). Moreover, \( S \) is linearly homeomorphic to \( \mathbb{R}^n \) and thus Corollary 1.2 implies that the convex cone is closed in \( S \).

2. The proof of Farkas lemma

Let us prove Theorem 1.1. We start by observing that there is no loss of generality in assuming that the vector space \( V \) is finite-dimensional. Namely, let \( V_0 \) be the subspace of \( V \) spanned by \( \{v_1, \ldots, v_k, x\} \); assuming the thesis valid for finite-dimensional spaces, we obtain a linear functional \( \alpha \in V_0^* \) with \( \alpha(x) < 0 \) and \( \alpha(v_i) \geq 0 \), for all \( i = 1, \ldots, k \). The conclusion is thus obtained by considering an arbitrary linear extension of \( \alpha \) to \( V \).

We also observe that we can assume that the vectors \( v_i \) are nonzero, since we can obviously delete the zero vectors from the list \( v_1, \ldots, v_k \).

To prove Theorem 1.1 we need the following:

2.1. Lemma (supporting hyperplane). Let \( V \) be a finite-dimensional nonzero real vector space, \( C \) be a convex subset of \( V \) and \( x \in V \) a point not in \( C \). Then there exists a nonzero linear functional \( \alpha \in V^* \) such that \( \alpha(y) \geq \alpha(x) \), for all \( y \in C \).

We leave the proof of Lemma 2.1 to the appendix. Geometrically, the statement of the lemma can be interpreted as follows: if the point \( x \) is not on the convex set \( C \) then there exists a hyperplane \( \{ y \in V : \alpha(y) = c \} \) passing through \( x \) such that \( C \) is contained on the closed half-space \( \{ y \in V : \alpha(y) \geq c \} \).

The following is a simple consequence of Lemma 2.1:

2.2. Corollary. Let \( C \) be a convex cone on a finite-dimensional nonzero real vector space \( V \) and let \( x \in V \) be a point not in \( C \). Then there exists a nonzero linear functional \( \alpha \in V^* \) with \( \alpha(x) \leq 0 \) and \( \alpha(y) \geq 0 \), for all \( y \in C \).

Proof. If \( C \) is empty the result is trivial, so assume \( C \) is nonempty; this implies \( 0 \in C \). Let \( \alpha \in V^* \) be as in the statement of Lemma 2.1; then \( 0 = \alpha(0) \geq \alpha(x) \), for \( 0 \in C \). Given \( y \in C \), we claim that \( \alpha(y) \geq 0 \); otherwise, \( \alpha(y) < 0 \) and \( t \alpha(y) < \alpha(x) \) for \( t > 0 \) sufficiently large. Then \( \alpha(ty) < \alpha(x) \) and since \( ty \in C \), we obtain a contradiction.

We can now begin the proof of Theorem 1.1. The proof is by induction on the dimension of \( V \). If \( V \) is one-dimensional then the nonzero vector \( v_1 \) is a basis for \( V \) and the hypothesis that \( x \) is not on the convex cone spanned by \( \{v_1\} \) implies that
independent, there exists a linear functional \( \alpha \) on the list \( v_1, \ldots, v_k \). If \( k = 1 \), we consider two possibilities: if \( v_1 \) and \( x \) are linearly independent, there exists a linear functional \( \alpha \in V^* \) with \( \alpha(v_1) = 1 \) and \( \alpha(x) = -1 \) (since the set \{\( v_1, x \)\} can be extended to a basis of \( V \)); otherwise, if \( v_1 \) and \( x \) are linearly dependent then \( x = tv_1 \) with \( t < 0 \) and thus it suffices to choose a linear functional \( \alpha \) with \( \alpha(v_1) = 1 \).

Assume now that \( k > 1 \) and that the result holds when the convex cone is spanned by less than \( k \) vectors. Let \( x \in V \) be a point outside the convex cone spanned by \{\( v_1, \ldots, v_k \)\} and let us show that there exists a linear functional \( \alpha \in V^* \) with \( \alpha(x) < 0 \) and \( \alpha(v_i) \geq 0 \), for all \( i = 1, \ldots, k \). Corollary 2.2 gives us a nonzero linear functional \( \beta \in V^* \) with \( \beta(x) \leq 0 \) and \( \beta(v_i) \geq 0 \), for all \( i = 1, \ldots, k \). If \( \alpha(x) < 0 \), we are done; thus, we only have to worry about the case where \( \alpha(x) = 0 \).

If \( \alpha(v_i) = 0 \) for all \( i = 1, \ldots, k \) then \( x \) and \( v_1, \ldots, v_k \) are all in the kernel of \( \alpha \), whose dimension is less than \( \dim(V) \). But we are assuming the result to hold for spaces of dimension less than \( \dim(V) \), so that there exists a linear functional \( \beta \) defined on the kernel of \( \alpha \) with \( \beta(x) < 0 \) and \( \beta(v_i) \geq 0 \), for \( i = 1, \ldots, k \). The linear functional \( \beta \) has a linear extension to the entire space \( V \), and thus the thesis is obtained in this case.

Let us assume that \( \alpha(v_i) > 0 \) for at least one index \( i \); then the set:

\[
\{ v_i : \alpha(v_i) = 0 \}
\]

has cardinality less than \( k \) and thus there exists a linear functional \( \beta \in V^* \) with \( \beta(x) < 0 \) and such that \( \beta(v_i) \geq 0 \) for all those \( i \) with \( \alpha(v_i) = 0 \) (we know nothing about \( \beta(v_i) \) for those \( i \) with \( \alpha(v_i) > 0 \)). Given \( \varepsilon > 0 \), consider the linear functional \( \tilde{\alpha} = \alpha + \varepsilon \beta \). We have \( \tilde{\alpha}(x) = \varepsilon \beta(x) < 0 \) and \( \tilde{\alpha}(v_i) = \varepsilon \beta(v_i) \geq 0 \) for those \( i \) with \( \alpha(v_i) = 0 \). But by choosing \( \varepsilon > 0 \) sufficiently small, we can assure that \( \tilde{\alpha}(v_i) = \alpha(v_i) + \varepsilon \beta(v_i) \) is positive for those \( i \) with \( \alpha(v_i) > 0 \). Thus the linear functional \( \tilde{\alpha} \) has the desired properties and the proof is complete.

**Appendix A. The proof of the supporting hyperplane lemma**

In this appendix we prove Lemma 2.1. We observe that this lemma is just a consequence of a version of the Hahn–Banach’s Theorem stated in many Functional Analysis books. We start by proving such version of Hahn–Banach’s Theorem.

**A.1. Theorem (Hahn–Banach).** Let \( V \) be a real vector space and let \( p : V \to \mathbb{R} \) be a nonnegative map satisfying the following two conditions:

1. \( p(x + y) \leq p(x) + p(y) \), for all \( x, y \in V \) (subadditivity);
2. \( p(cx) = cp(x) \), for all \( x \in V \) and all \( c \geq 0 \) (positive homogeneity).

Let \( \alpha \in S^* \) be a linear functional defined in a subspace \( S \) of \( V \). If \( \alpha(x) \leq p(x) \) for all \( x \in S \) then \( \alpha \) extends to a linear functional \( \tilde{\alpha} \in V^* \) such that \( \tilde{\alpha}(x) \leq p(x) \) for all \( x \in V \).

**A.2. Remark.** If \( V \) is a real vector space then a nonnegative map \( p : V \to \mathbb{R} \) satisfying the conditions \( p(x + y) \leq p(x) + p(y) \) and \( p(cx) = cp(x) \), for all \( x, y \in V \), \( c \in \mathbb{R} \) is usually called a semi-norm; if, in addition, \( p(x) > 0 \) for all \( x \neq 0 \), then \( p \) is
called a norm. Note that conditions (a) and (b) on the statement of Theorem A.1
do not imply that \( p \) be a semi-norm, since condition (b) only requires \( p(cx) = cp(x) \)
for nonnegative \( c \). A nonnegative map \( p : V \to \mathbb{R} \) is a semi-norm if and only if
it satisfies (a), (b) and the additional condition \( p(-x) = p(x) \), for all \( x \in V \). We
observe that if \( p \) is a semi-norm and \( \alpha \in V^* \) then the inequality \( \alpha(x) \leq p(x) \)
holds for all \( x \in V \) if and only if the inequality \(|\alpha(x)| \leq p(x)\) does. Many books state
the Hahn-Banach theorem only for semi-norms, but the proof in this more general
context is exactly the same.

The proof of Theorem A.1 uses the following:

**A.3. Lemma** (simple extension). Under the hypothesis of Theorem A.1, given a
vector \( v \in V \) not in \( S \) and denoting by \( S' \) the subspace spanned by \( S \) and \( v \), then \( \alpha \)
extends to a linear functional \( \alpha' \) on \( S' \) such that \( \alpha'(x) \leq p(x) \), for all \( x \in S' \).

**Proof.** A linear extension \( \alpha' \) of \( \alpha \) to \( S' \) is uniquely determined by its value on the
vector \( v \); more explicitly, every vector of \( S' \) can be uniquely written in the form
\( x + tv \), with \( x \in S \), \( t \in \mathbb{R} \) and setting \( \alpha'(v) = c \) for some real number \( c \) we obtain:

\[
\alpha'(x + tv) = \alpha(x) + tc,
\]

for all \( x \in S \), \( t \in \mathbb{R} \). We wish to determine a value for \( c \in \mathbb{R} \) that makes the inequality:

\[
\alpha(x) + tc \leq p(x + tv)
\]

true, for all \( x \in S \), \( t \in \mathbb{R} \). For \( t = 0 \), inequality (A.1) holds. For \( t > 0 \), (A.1)
is equivalent to:

\[
\alpha(\frac{x}{t}) + c \leq p\left(\frac{x}{t} + v\right)
\]

and for \( t < 0 \), (A.1) is equivalent to:

\[
\alpha(\frac{x}{t}) + c \geq -p\left(-\frac{x}{t} - v\right).
\]

Clearly (A.2) holds for all \( x \in S \) and all \( t > 0 \) if and only if \( \alpha(y) + c \leq p(y + v) \), for
all \( y \in S \); this is equivalent to:

\[
c \leq p(y + v) - \alpha(y),
\]

for all \( y \in S \). Similarly, (A.3) holds for all \( x \in S \) and all \( t < 0 \) if and only if:

\[
c \geq -p(z - v) + \alpha(z),
\]

for all \( z \in S \). Thus, the proof will be concluded if we are able to find a number
\( c \in \mathbb{R} \) which is an upper bound for the set \( \{-p(z - v) + \alpha(z) : z \in S \} \) and a lower
bound for the set \( \{p(y + v) - \alpha(y) : y \in S \} \). But this is possible if and only if the
inequality:

\[
-p(z - v) + \alpha(z) \leq p(y + v) - \alpha(y)
\]

holds, for all \( y, z \in S \). Inequality (A.4) is equivalent to:

\[
\alpha(z + y) \leq p(y + v) + p(z - v),
\]

and (A.5) follows from the subadditivity of \( p \), as we show below:

\[
\alpha(z + y) \leq p(z + y) = p((y + v) + (z - v)) \leq p(y + v) + p(z - v).
\]

\( \square \)
A.7. Remark. and the convexity of \( C \) the proof. \( \square \)

C \infimum. Set \( p \) infimum. Set \( p \in \mathbb{R} \) finite number of times, extending \( \alpha \) successively until we get the desired linear functional \( \tilde{\alpha} \) on \( V \). For the general case, one uses Zorn’s lemma to obtain a maximal extension of \( \alpha \) to a linear functional \( \tilde{\alpha} : \tilde{S} \to \mathbb{R} \) satisfying \( \tilde{\alpha}(x) \leq p(x) \), for all \( x \in \tilde{S} \); then Lemma A.3 implies that \( \tilde{S} = V \), otherwise we would be able to find a proper extension of \( \tilde{\alpha} \). \( \square \)

A.4. Definition. A subset \( C \) of a real vector space \( V \) is called absorbent in \( V \) if for all \( x \in V \) there exists \( \varepsilon > 0 \) such that \( tx \in C \) for \( 0 \leq t \leq \varepsilon \).

Obviously absorbent sets must contain the origin.

A.5. Example. If \( C \) is a neighborhood of the origin in \( \mathbb{R}^n \) then \( C \) is absorbent; namely, for \( x \in \mathbb{R}^n \), the continuous map \( \mathbb{R} \ni t \mapsto tx \in \mathbb{R}^n \) sends \( t = 0 \) to the origin and thus we must have \( tx \in C \) for \( |t| \) sufficiently small.

A.6. Lemma. Let \( V \) be a real vector space and let \( C \) be an absorbent convex subset of \( V \). Then there exist a nonnegative map \( p : V \to \mathbb{R} \) satisfying conditions (a) and (b) in the statement of Theorem A.1, such that:

\[
(A.6) \quad \{ x \in V : p(x) < 1 \} \subset C \subset \{ x \in V : p(x) \leq 1 \}.
\]

Proof. For each \( x \in V \), consider the set:

\[
I_x = \{ t > 0 : \frac{x}{t} \in C \}.
\]

The fact that \( C \) is absorbent implies that \( I_x \) is nonempty; thus, it has a nonnegative infimum. Set \( p(x) = \inf I_x \), for all \( x \in V \). Observe that if \( x = 0 \) then \( I_x = [0, +\infty[ \) and thus \( p(0) = 0 \); this proves (b) for \( c = 0 \). If \( c > 0 \) then clearly \( t \in I_{cx} \) if and only if \( \frac{x}{t} \in I_x \), which show that \( I_{cx} = cI_x \); thus:

\[
p(cx) = \inf I_{cx} = c \inf I_x = cp(x).
\]

This proves (b). Let us prove (a). Let \( x, y \in V \) be fixed, set \( t = p(x), s = p(y) \) and choose \( \varepsilon > 0 \). We know that there exists \( t' \in I_x, s' \in I_y \) with \( t' \leq t + \varepsilon \) and \( s \leq s' < s + \varepsilon \). Thus \( \frac{x}{t'} \in C, \frac{y}{s'} \in C \) and the convexity of \( C \) implies that:

\[
\frac{x + y}{t' + s'} = \frac{t' x}{t' x + s' x} + \frac{s' y}{s' y + t' y}.
\]

is also in \( C \); thus \( t' + s' \in I_{x+y} \) and:

\[
p(x+y) \leq t' + s' < t+s+2\varepsilon = p(x)+p(y)+2\varepsilon.
\]

Since \( \varepsilon > 0 \) is arbitrary, we obtain (a). Finally, let us prove (A.6). Given \( x \in V \) with \( p(x) < 1 \) then there exists \( t \in I_x \) with \( p(x) \leq t < 1 \). Thus \( \frac{x}{t} \in C, 0 \in C \) and the convexity of \( C \) implies that \( x = t \frac{x}{t} + (1- t) 0 \) is in \( C \). Moreover, given \( x \in C \) then obviously \( 1 \in I_x \), so that \( p(x) \leq 1 \). This proves (A.6) and concludes the proof. \( \square \)

A.7. Remark. The map \( p \) constructed in the proof of Lemma A.6 is not in general a semi-norm, unless \( C \) is symmetric, i.e., if \( x \in C \) implies \( -x \in C \).

A.8. Corollary. Lemma 2.1 holds if we add the hypothesis that the convex set \( C \) be absorbent (in which case the hypothesis of finite-dimensionality for \( V \) is no longer necessary).
Proof. Choose \( p \) as in the statement of Lemma A.6. Since \( 0 \in C \) and \( x \notin C \), we have \( x \neq 0 \) and thus \( \{x\} \) is a basis for a one-dimensional subspace \( S \) of \( V \). Let \( \alpha \in S^* \) be the linear functional such that \( \alpha(x) = 1 \). Notice that \( p(x) \geq 1 \) for, otherwise, (A.6) would imply \( x \in C \). Thus, for \( t \geq 0 \) we have:

\[
\alpha(tx) = t \leq tp(x) = p(tx);
\]

obviously, for \( t < 0 \) we have \( \alpha(tx) = t < 0 \leq p(tx) \). We have thus shown that \( \alpha(y) \leq p(y) \) for all \( y \in S \), and now Theorem A.1 gives us a linear extension \( \tilde{\alpha} \) of \( \alpha \) to \( V \) such that \( \tilde{\alpha}(y) \leq p(y) \), for all \( y \in V \). For all \( y \in C \), we have \( p(y) \leq 1 \) and thus:

\[
\tilde{\alpha}(y) \leq p(y) \leq 1 = \tilde{\alpha}(x).
\]

Hence \( -\tilde{\alpha}(y) \geq -\tilde{\alpha}(x) \), for all \( y \in C \), so that \( -\tilde{\alpha} \) is the linear functional we are looking for. \( \square \)

A.9. Lemma. If \( V \) is a finite-dimensional real vector space and \( C \) is a nonempty convex subset of \( V \) then there exists a point \( p \in C \) and a subspace \( S \) of \( V \) such that the set \( C - p = \{x - p : x \in C\} \) is contained in \( S \) and is absorbent in \( S \).

Proof. Choose an arbitrary point \( q \in C \) and set \( C' = C - q \), so that \( C' \) is a convex subset of \( V \) containing the origin. Now let \( \{b_1, \ldots, b_k\} \) be a maximal linearly independent subset of \( C' \); the existence of such subset follows easily from the finite-dimensionality of \( V \). The subspace \( S \) spanned by \( \{b_1, \ldots, b_k\} \) contains \( C' \); otherwise, there would exist \( x \in C' \setminus S \) and thus \( \{b_1, \ldots, b_k, x\} \) would be a linearly independent subset of \( C' \) properly containing \( \{b_1, \ldots, b_k\} \). Let \( L : \mathbb{R}^k \to S \) be the linear isomorphism that carries the canonical basis of \( \mathbb{R}^k \) to \( \{b_1, \ldots, b_k\} \). Then \( L^{-1}(C') \) is a convex subset of \( \mathbb{R}^k \) containing the origin and the canonical basis; this implies that \( L^{-1}(C') \) also contains the set:

\[
\Delta_k = \{(t_1, \ldots, t_k) \in \mathbb{R}^k : \sum_{i=1}^k t_i \leq 1, \ t_i \geq 0, \ i = 1, \ldots, k\}.
\]

The point \( u = (\frac{1}{k+1}, \frac{1}{k+1}, \ldots, \frac{1}{k+1}) \in \mathbb{R}^k \) is clearly an interior point of \( \Delta_k \), so that \( \Delta_k - u \) is a neighborhood of the origin in \( \mathbb{R}^k \). Thus \( L^{-1}(C') - u \) is also a neighborhood of the origin in \( \mathbb{R}^k \) and it is therefore an absorbent subset of \( \mathbb{R}^k \) (recall Example A.5). The image of \( L^{-1}(C') - u \) under the isomorphism \( L : \mathbb{R}^k \to S \) is equal to \( C' - L(u) \) and it is an absorbent subset of \( S \). But

\[
C' - L(u) = C - (q + L(u))
\]

and the conclusion is obtained by setting \( p = q + L(u) \). \( \square \)

Proof of Lemma 2.1. If \( C \) is empty, any nonzero linear functional \( \alpha \) will do. Assume that \( C \) is nonempty and choose \( p \) and \( S \) as in the statement of Lemma A.9. If the point \( x - p \) is not in \( S \) then there exists a linear functional \( \alpha \in V^* \) such that \( \alpha|_S = 0 \) and \( \alpha(x - p) = -1 \). Then, for all \( y \in C \) we have \( y - p \in S \) and thus:

\[
\alpha(y) - \alpha(p) = \alpha(y - p) = 0 > -1 = \alpha(x - p) = \alpha(x) - \alpha(p),
\]

which implies \( \alpha(y) > \alpha(x) \), for all \( y \in C \). On the other hand, if \( x - p \) is in \( S \) then, since \( C - p \) is convex and absorbent in \( S \), Corollary A.8 gives us a nonzero linear functional \( \alpha \) on \( S \) with \( \alpha(z) \geq \alpha(x - p) \), for all \( z \in C - p \); this implies:

\[
\alpha(y) - \alpha(p) = \alpha(y - p) \geq \alpha(x - p) = \alpha(x) - \alpha(p),
\]

and thus \( \alpha(y) \geq \alpha(x) \), for all \( y \in C \). The conclusion is obtained by considering an arbitrary linear extension of \( \alpha \) to \( V \). \( \square \)
DEPARTAMENTO DE MATEMÁTICA,
UNIVERSIDADE DE SÃO PAULO, BRAZIL
E-mail address: tausk@ime.usp.br
URL: http://www.ime.usp.br/~tausk