Lecture 6. Farkas' lemma

Paul Delatte
delatte@usc.edu
University of Southern California

Last updated: 31 December, 2022

To prove this theorem, we use a geometric variation of the Hahn-Banach extension theorem which was proved in Lecture 5.

Theorem 1 (**Farkas' Lemma**). *Let* $A \in \mathbb{R}^{m\times n}$ *and* $b \in \mathbb{R}^m$. *Then exactly one the following propositions is true, but not both.*

1. *There exists* $x \in \mathbb{R}^n$ *such that* $Ax = b$ *and* $x \geq 0_n$.
2. *There exists* $y \in \mathbb{R}^m$ *such that* $A^T y \geq 0_m$ *and* $b^T y < 0$.

Proof. We first prove that both propositions cannot be true simultaneously. Suppose they are. Since $b^T y < 0$, we have $y \neq 0_m$ and $b \neq 0_m$. Since $Ax = b$, we have $x \neq 0_n$. Thus $b^T y = (Ax)^T y = x^T A^T y \geq 0$, which contradict $b^T y < 0$.

Suppose now that (1.) is false. Then $b \notin S = \{Ax : x \geq 0_n\}$. The set S is a finitely generated cone which is easily seen to be convex. By Lemma 2, we also have that S is closed. The conditions of the Hahn–Banach separation theorem in LCTVS are satisfied for $\{b\}$ compact and S closed, hence there exists $y \neq 0_m$ and $\alpha \in \mathbb{R}$ such that $b^T y < \alpha$ and $z^T y \geq \alpha$ for all $z \in S$. Since $0_m \in S$, we have $\alpha \leq 0$, and so $b^T y < 0$. By definition of S, we also have that $p^T A^T y = (Ap)^T y \geq \alpha$ for all $p \geq 0_n$. If there is $p_0 \geq 0_n$ such that $p_0^T A^T y < 0$, then this contradicts $p^T A^T y \geq \alpha$ for all $p \geq 0_n$, since for $\lambda > 0$, $\lambda p_0 \geq 0_n$, $\lambda p_0 \in S$, and $\lim_{\lambda\to+\infty}(\lambda p_0)^T A^T y = \lim_{\lambda\to+\infty} \lambda p_0^T A^T y = -\infty$. Therefore, $p^T A^T y \geq 0$ for all $p \geq 0_n$, which implies that $A^T y \geq 0_m$. Thus (2.) is true. □

Remark. There exist at least three ways of proving Farkas' lemma (see C.G. Broyden (1998)): algorithmic proofs based for instance on the dual simplex method (as originally done by Farkas); algebraic proofs (either inductive or not); geometric proofs based on the Hahn–Banach separation theorem. In the first two proofs, there is a risk of incompleteness for not handling properly the possibility of degenerate solutions (originating from redundant constraints). Geometric proofs have the advantage of conciseness and bypass the problem of degeneracy by relating "each theorem of the alternative [...] to a pair of dual problems: A primal steepest-descent problem and a dual least-norm problem" (as quoted in C.G. Broyden (2008)). A common shortcoming of the geometric proof as presented in textbooks is that the proposition that any finitely generated conical hull $S = \{Ax : x \geq 0_n\}$ is closed is admitted, while it is the nontrivial part of the proof.

Remark. In the geometric proof, the hyperplane can also be directly constructed without invoking the Hahn–Banach theorem but by taking the hyperplane with normal $y = z - b$ where z is the nearest point in S from b (whose existence need to be proved) – see S.6.5. Matousek&Gärtner Understanding and Using Linear Programming p.95. A similar idea can be used to directly proving separating hyperplane theorems in $\mathbb{R}^n$ (here for a point and a cone, but the same idea generalize to two convex sets). This elementary way of doing does not require the axiom of choice (see remark in Rockafellar CA S.11). But we already knew it from the proof of the Hahn–Banach theorem which works by transfinite induction on the dimension, hence the consequence for finite dimensions.

Reference. We take the geometric route using the Hahn–Banach separation theorem for LCTVS. We follow T.3.30. in Andreasson&Evgrafov&Patriksson ICO p.57 or L.10.2.-3. in Beck INO p.192-193. A slightly different formulation of the result with a proof in the same spirit can be found in Berge (1963) p.164. See also T.4.3.4 in Hiriart-Urruty&Lemarechal FCA p.60 and C.5.84.-85. in Aliprantis&Border (2006) p.209 for more general formulations (in different directions) than ours with the same geometric proof.

Remark. For equivalent results to Farkas' lemma, see "All linear theorems of the alternative have a common father. An addendum to a paper of C.T. Perng" by G. Giorgi (2020). For generalizations of Farkas' lemma, see "Farkas Lemma: Generalizations" by V. Jeyakumar (2009) in the Encyclopedia of Optimization.

Lemma 2. *Let $\{a_1, \ldots, a_n\} \subseteq \mathbb{R}^m$. Denote $A \in \mathbb{R}^{m\times n}$ the matrix with columns $a_1, \ldots, a_n$. Then the conical hull (finitely) generated by $\{a_1, \ldots, a_n\}$*

$$\operatorname{cone}(\{a_1, \ldots, a_n\}) := \left\{ \sum_{i=1}^{n} x_i a_i : x = (x_1, \ldots, x_n) \geq 0_n \right\} = \{Ax : x \geq 0_n\}$$

is closed.

Proof. Suppose first that the a_i's are linearly independent. Then the convergence of $(y^k)_{k\in\mathbb{N}} = (\sum_{i=1}^n x_i^k a_i)_{k\in\mathbb{N}}$ is equivalent to the convergence of each $(x_i^k)_{k\in\mathbb{N}}$ to some x_i, which must be nonnegative if each x_i^k in the sequence is.

Suppose on the contrary that $\sum_{i=1}^n x_i a_i = 0$ has a nonzero solution $z \in \mathbb{R}^n$. Suppose $z_i < 0$ for some i (change z to $-z$ if necessary). Each $y \in \operatorname{cone}(\{a_1, \ldots, a_n\})$ can be written as

$$y = \sum_{i=1}^{n} x_i a_i = \sum_{i=1}^{n} (x_i + t^*(y) z_i) a_i = \sum_{i \neq j(y)} x_i' a_i,$$

where

$$j(y) \in \arg\min_{z_i<0} \frac{-x_i}{z_i}, \quad \text{and} \quad t^*(y) = \frac{-x_{j(y)}}{z_{j(y)}},$$

so that each $x_i' = x_i + t^*(y) z_i$ is nonnegative. By making y vary in $\operatorname{cone}(\{a_1, \ldots, a_n\})$, we can construct a decomposition $\operatorname{cone}(\{a_1, \ldots, a_n\}) = \bigcup_{j=1}^n S_j$ where S_j is the conical hull of the $n-1$ generators a_j, $j \neq i$.

If there is some j such that the generators of S_j are linearly dependent, then the argument can be repeated to obtain a further decomposition of S_j. After finitely many such operations (why finite?), we obtain a decomposition of $\operatorname{cone}(\{a_1, \ldots, a_n\})$ as a finite union of conical hull (finitely) generated by linearly independent elements. By the first part of the proof, each of these conical hull is closed. Thus $\operatorname{cone}(\{a_1, \ldots, a_n\})$ is closed. □

Reference. The proof is standard: a representation result for conical hull is proved in a similar way as in Carathéodory's theorem's standard proof. (Carathéodory's theorem's gives a representation result for convex hulls). We follow L.4.3.3. in Hiriart-Urruty&Lemarechal FCA P.59-60. (See also L.6.32. in Beck INO p.111.) Alternative proofs exist (see alternative proofs of Carathéodory's theorem).

Lemma 3 (Hahn-Banach Separation Theorem in LCTVS)**.** *Let X be a locally convex topological vector space over $\mathbb{R}$. Let A and B be subsets of X. If A and B are disjoint nonempty convex and if A is compact and B is closed, then then there exists a continuous linear functional $f: X \to \mathbb{R}$ and $\alpha \in \mathbb{R}$ such that $f(x) < \alpha < f(y)$ for all $x \in A$ and all $y \in B$.*

Proof. See Lecture 5. □

References

Aliprantis, C. D., and K. C. Border (2006): *Infinite dimensional analysis*. Springer.

Andréasson, N., A. Evgrafov, and M. Patriksson (2005): *An introduction to continuous optimization: Foundations and fundamental algorithms*. Chalmers University of Technology Press.

Beck, A. (2014): *Introduction to nonlinear optimization: Theory, algorithms, and applications with MATLAB*. SIAM.

Berge, C. (1963): *Topological spaces*. Oliver & Boyd.

Broyden, C. (1998): "A simple algebraic proof of Farkas's lemma and related theorems," *Optimization methods and software*, 8(3-4), 185–199.

Giorgi, G. (2020): "All linear theorems of the alternative have a common father. An addendum to a paper of CT Perng," *Fundamental Journal of Mathematics and Mathematical Sciences*, 13, 43–78.

Hiriart-Urruty, J.-B., and C. Lemaréchal (2004): *Fundamentals of convex analysis*. Springer.

Jeyakumar, V. (2009): "Farkas Lemma: Generalizations.," *Encyclopedia of optimization*, 2, 87–91.

Matouşek, J., and B. Gärtner (2007): *Understanding and using linear programming*, vol. 1. Springer.

Rockafellar, R. T. (1970): *Convex Analysis*, vol. 28. Princeton University Press.