An Elementary Approach to Farkas' Lemma and its Relation to Hyperplane Separation

Karim El-Sharkawy and Darshini Rajamani
Professor Thomas Sinclair

June 2024

1 Introduction

Farkas' Lemma: Given a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $b \in \mathbb{R}^m$, exactly one of the following two statements is true:

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

Farkas' Lemma is a fundamental result in linear programming and optimization. It simply states that for a given matrix A and vector b, exactly one of the following statements is true:

If $b \in cone(A)$, then there's an x such that, $Ax = b$. If $b \notin cone(A)$, then there exists a vector y such that $A^T y = b$.

We aim to establish Farkas' Lemma using basic linear algebra. Existing proofs utilizing analysis, linear programming, and Fourier-Motzkin Elimination are complex and challenging for undergraduates to grasp, thereby limiting accessibility for students conducting research using optimization or linear programming.

1.1 Definitions

Cone: A set $K \subseteq \mathbb{R}^n$ is a cone if $x \in K$ implies $\alpha x \in K$ for any scalar $\alpha \geq 0$.

Conic Hull: Given a set S, the conic hull of S, denoted by cone(S), is the set of all conic combinations of the points in S, that is,

$$\text{cone}(S) = \left\{ \sum_{i=1}^{n} \alpha_i x_i \mid \alpha_i \geq 0,\ x_i \in S \right\}.$$

Hyperplanes: A hyperplane $\partial\mathcal{H}$ of a vector space V is a subspace of V of $dim(n-1)$. They can be described as the intersection of half-spaces:

$$\partial\mathcal{H} = \mathcal{H}_- \cap \mathcal{H}_+ = \{y \mid a^T y \leq b,\, a^T y \geq b\} = \{y \mid a^T y = b\}$$

Where $\mathcal{H}_-$ and $\mathcal{H}_+$ are halfspaces

A hyperplane is the solution set of a single linear equation of the form $\mathbf{a}\cdot\mathbf{x} = b$, where $a \in V$ is a nonzero, normal vector to $\partial\mathcal{H}$, $\mathbf{x}$ is a vector variable in V, and b is a constant.

2 Separating Hyperplane Theorem

Theorem 1: Let C and D be two convex sets in $\mathbb{R}^n$ that do not intersect ($i.e., C \cap D = \emptyset$). Then, there exists $a \in \mathbb{R}^n$, $a \neq 0$, $b \in$ R, such that:

$$a^T x \leq b \quad \forall x \in C,$$
$$a^T x \geq b \quad \forall x \in D.$$

Note: The hyperplane $a^T x = b$ with normal vector a separates the sets C and D.

We note that neither inequality in the conclusion of Theorem 1 can be made strict. Strict separation may not always be possible, even when both C and D are closed. (An example could be added here to illustrate this point.) However, if both sets are closed and at least one of them is compact, then the separation can be strict, as stated in the following theorem:

Theorem 2: Let C and D be two closed convex sets in $\mathbb{R}^n$ with at least one of them bounded, and assume $C \cap D = \emptyset$. Then there exist $a \in \mathbb{R}^n$, $a \neq 0$, and $b \in \mathbb{R}$ such that:

$$a^T x > b \quad \forall x \in D,$$
$$a^T x < b \quad \forall x \in C.$$

(Do we need to proof this?)

Corollary 1: Let $C \subseteq \mathbb{R}^n$ be a closed convex set and $x \in \mathbb{R}^n$ a point not in C. Then x and C can be strictly separated by a hyperplane.

Corollary 1 is a direct consequence of Theorem 2. If we consider the set C and the singleton set $\{x\}$, where $x \notin C$, the conditions of Theorem 2 are satisfied because $\{x\}$ is bounded and disjoint from C. Thus, there exists a hyperplane that strictly separates x and C.

Proof
Consider the set $D = \{x\}$, which is trivially a closed convex set and is bounded. Since $x \notin C$, we have $C \cap D = \emptyset$. By Theorem 2, there exist $a \in \mathbb{R}^n$, $a \neq 0$, and $b \in \mathbb{R}$ such that:

$$a^T y > b \quad \forall y \in D,$$
$$a^T z < b \quad \forall z \in C.$$

Since $D = \{x\}$, the first condition simplifies to $a^T x > b$. Hence, the hyperplane $a^T z = b$ strictly separates the point x from the set C.

3 Farkas Lemma

3.1 Farkas' Lemma Statement

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Then exactly one of the following statements is true:

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

3.2 Proof

3.2.1 1 → 2

This will be a proof by contradiction.

Suppose there exist $x \in \mathbb{R}^n$ such that $Ax = b$ and $x \geq 0$. Also, suppose there exist $y \in \mathbb{R}^m$ such that $A^T y \geq 0$ and $b^T y < 0$.

Since $Ax = b$,

$$(Ax)^T = b^T$$
$$x^T A^T = b^T$$
$$x^T A^T y = b^T y$$

Since $x \geq 0$ and $A^T y \geq 0$, we know $x^T A^T y \geq 0$.

However, we assumed $b^T y < 0$, leading to a contradiction.

Therefore, if there exists $x \in \mathbb{R}^n$ such that $Ax = b$ and $x \geq 0$, then there cannot exist $y \in \mathbb{R}^m$ such that $A^T y \geq 0$ and $b^T y < 0$.

3.2.2 2 → 1

Suppose there exists $y \in \mathbb{R}^m$ such that $A^T y \geq 0$ and $b^T y < 0$.

Let $a_1, \ldots, a_n$ be the columns of matrix A. Define the set $S = \{a_1, \ldots, a_n\}$. We need to show that the conic hull of S, denoted by cone(S), is convex and closed.

Convexity of cone(S)

To prove that cone(S) is convex, let $x, y \in$ cone(S). Then there exist non-negative scalars α_i, β_i such that:

$$x = \sum_{i=1}^{n} \alpha_i a_i, \quad y = \sum_{i=1}^{n} \beta_i a_i$$

For any $\lambda \in [0, 1]$, consider:

$$\lambda x + (1 - \lambda) y = \lambda \sum_{i=1}^{n} \alpha_i a_i + (1 - \lambda) \sum_{i=1}^{n} \beta_i a_i = \sum_{i=1}^{n} (\lambda \alpha_i + (1 - \lambda) \beta_i) a_i$$

Since $\lambda \alpha_i + (1 - \lambda) \beta_i \geq 0$ for all i, it follows that $\lambda x + (1 - \lambda) y \in$ cone(S). Thus, cone(S) is convex.

Closedness of cone(S)

To show that cone(S) is closed, consider a sequence $\{z_k\}$ in cone(S) that converges to a point $\bar{z}$. Each z_k can be written as:

$$z_k = \sum_{i=1}^{n} \alpha_{(k)} a_i \quad \text{with} \quad \alpha_{(k)} \geq 0$$

We need to show that $\bar{z} \in$ cone(S).

Consider the following linear program:

$$\begin{aligned} \min_{\alpha, z} \quad & \|z - \bar{z}\|_\infty \\ \text{s.t.} \quad & \sum_{i=1}^{n} \alpha_i a_i = z, \\ & \alpha_i \geq 0 \end{aligned}$$

The objective function $\|z - \bar{z}\|_\infty$ is an infinity norm and is always non-negative. For each z_k, there exists $\alpha_{(k)}$ such that the pair $(z_k, \alpha_{(k)})$ is feasible for the LP since $z_k \in$ cone(S).

As the sequence $\{z_k\}$ converges to $\bar{z}$, the optimal value of the LP approaches zero. Since LPs achieve their optimal values, it follows that $\bar{z} \in$ cone(S).

In conclusion, since $\bar{z} \in$ cone(S), cone(S) is closed. Thus, the set $S =$ cone$\{a_1, \ldots, a_n\}$ is a closed convex set.

Now we can use the Separating Hyperplane Theorem.

Given that $b \notin C$ by the assumption that the first condition is infeasible. By Corollary 1, the point b and the set C can be strictly separated; i.e., there exist $y \in \mathbb{R}^m$, $y \neq 0$, and $r \in \mathbb{R}$ such that:

$$y^T z \geq r \quad \forall z \in C \quad \text{and} \quad y^T b < r.$$

Since $0 \in C$, we must have $r \leq 0$. If $r < 0$, then there exists some $z \in \text{cone}(S)$ such that $y^T z < 0$.

Since $\text{cone}(S)$ is a cone, for any $\alpha \geq 0$, $\alpha z \in \text{cone}(S)$. Therefore, $y^T(\alpha z) = \alpha y^T z$. If $y^T z < 0$, then $y^T(\alpha z) = \alpha y^T z$ can be made arbitrarily negative by choosing a sufficiently large α. This contradicts the assumption that $y^T z \geq r \geq 0$ for all $z \in \text{cone}(S)$.

Thus, r cannot be less than zero, so r must be zero. The new condition becomes:

$$y^T z \geq 0 \quad \forall z \in C \quad \text{and} \quad y^T b < 0.$$

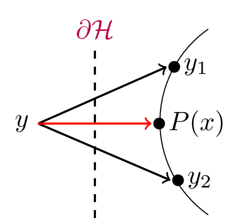
Since $\{a_1, \ldots, a_n\} \subseteq C$, we see that $A^T y \geq 0$.

Therefore, if there exists $y \in \mathbb{R}^m$ such that $A^T y \geq 0$ and $b^T y < 0$, then there cannot exist $x \in \mathbb{R}^n$ such that $Ax = b$ and $x \geq 0$.

This completes the proof of Farkas' Lemma.

3.3 Connection to Hyperplane Separation Theorem

We will now understand Farkas' lemma in the light of The Hyperplane Separation Theorem. Consider the set: $C = \{Ax \mid x \geq 0\}$, a convex cone, which we will denote by a semi-circle, as if you're looking top-down:



Where y is a point outside of C, $P(x) \in C$ is the closest point to y, and $\partial\mathcal{H}$ is the hyperplane created by $P(x)$

The existence of $P(x)$ follows from Weierstrass' Theorem, which asserts that optimization problems in Euclidean space with bounded and compact sets must attain their minimum or maximum values. To find $P(x)$, we want to minimize $\|y - x\|^2$ where $x \in C$ and the minimum is attained at $x = P(x)$.

We denote the red vector by $\eta_x = y - P(x)$ where it is outward-normal with respect to $\mathcal{H}_-$ and inward-normal with respect to $\mathcal{H}_+$. $\partial\mathcal{H}$ is the hyperplane created by η_x: $\langle \eta_x, P(x) \rangle = 0$. y lies in $\mathcal{H}_- := \langle \eta_x, P(x) \rangle \leq 0$.

The cone C is defined as follows:

$$C = \bigcap_{\mathcal{H}_+ \subseteq \mathbb{C}}^{\infty} \mathcal{H}_+$$

This implies C is composed of half-spaces. Moreover, by its definition, $\partial\mathcal{H} = \mathcal{H}_- \cap \mathcal{H}_+$. Hence, $\partial\mathcal{H}$ can only intersect C except at its border.

This concludes our comprehension of Farkas' lemma in the context of hyperplanes.