The Farkas-Minkowski Theorem

The results presented below, the first of which appeared in 1902, are concerned with the existence of non-negative solutions of the linear system

$$A\boldsymbol{x} = \boldsymbol{b}\,, \tag{1.1}$$
$$\boldsymbol{x} \geq 0\,, \tag{1.2}$$

where A is an $m \times n$ matrix with real entries, $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{b} \in \mathbb{R}^m$. Here is a basic statement of the theorem which is due to Farkas and Minkowski.

Theorem 1.1 *A necessary and sufficient condition that (1.1-1.2) has a solution is that, for all $\boldsymbol{y} \in \mathbb{R}^m$ with the property that $A^\top \boldsymbol{y} \geq 0$, we have $\langle \boldsymbol{b}, \boldsymbol{y} \rangle \geq 0$.*

This theorem may be reformulated as an alternative theorem, in other words, as a theorem asserting that one set of equalities and/or inequalities has a solution if and only if another set does not. It is easy to see that the following statement is equivalent to the first:

Theorem 1.2 *The system (1.1-1.2) has a solution if and only if the system*

$$\begin{aligned} A^\top \boldsymbol{y} &\geq 0\,, \\ \langle \boldsymbol{b}, \boldsymbol{y} \rangle &< 0\,, \end{aligned}$$

has no solution.

There are a number of ways to prove Theorem 1.1. One way is to use the duality theorem of linear programming. Since the Farkas-Minkowski Theorem is used in some discussions of linear programming, it is useful to have an independent proof even if it may be less elementary in the sense that it uses a separation theorem. This is the proof which we will present below. Once established, it can then be used to prove the Duality Theorem of Linear Programming.

Before starting, it is useful to consider some facts about cones in $\mathbb{R}^n$. We begin with a definition.

Definition 1.3 *A set $K \subset \mathbb{R}^n$ is called a* cone *if $\boldsymbol{x} \in K$ implies $\alpha \boldsymbol{x} \in K$ for all $\alpha > 0$.*

It is obvious that the entire space $\mathbb{R}^n$ as well as any subspace of $\mathbb{R}^n$ constitutes a cone. The positive orthant of $\mathbb{R}^n$ is another example. Likewise, given any $m \times n$ matrix A, the sets $\{\boldsymbol{x} \in \mathbb{R}^n | A\boldsymbol{x} \leq 0\}$ and $\{\boldsymbol{x} \in \mathbb{R}^n | A\boldsymbol{x} \geq 0\}$ are also cones. Note that these examples suggest that one should not think of a cone as having a "point".

The definition shows that a cone is a collection of half-lines emanating from the origin. The origin itself may, or may not be in the cone. It is an elemantary fact, which we leave as an exercise, that an arbitrary intersection of cones is again a cone.

Now, let $\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(k)}$ be any k elements of $\mathbb{R}^n$. We look at all vectors of the form $\sum\limits_{i=1}^{k} \mu_i \boldsymbol{x}^{(i)}$ where for each i, $\mu_i > 0$. This set is clearly a cone (in fact it is even convex) and is called the cone generated by the vectors $\boldsymbol{x}^{(i)}, i = 1, \ldots, k$. It is useful to recast the definition of this particular cone as follows. Take A to be an $n \times k$ matrix whose columns are the vectors $\boldsymbol{x}^{(i)}$. Then the cone generated by these vectors is the set $\{\boldsymbol{z} \in \mathbb{R}^n | \boldsymbol{z} = A\boldsymbol{\mu}, \boldsymbol{\mu} \geq 0\}$.

Before proving the theorem, we need to prove an auxiliary result regarding this particular cone.

Lemma 1.4 *Let A be an $m \times n$ matrix. Then the set $\mathcal{R} = \{\boldsymbol{z} \in \mathbb{R}^m \,|\, \boldsymbol{z} = A\boldsymbol{x}, \boldsymbol{x} \geq 0\}$ is a closed subset of $\mathbb{R}^m$.*

Proof: We first recognize that it is possible to write elements of the set $\mathcal{R}$ as a positive linear combination of n vectors, namely the columns $\boldsymbol{a}_j$ of the matrix A. That is

$$\mathcal{R} := \left\{ \boldsymbol{z} \in \mathbb{R}^m \,\middle|\, \boldsymbol{z} = \sum_{j=1}^{n} \mu_j \, \boldsymbol{a}_j \,, \; \mu_j \geq 0 \right\} .$$

Notice that the set $\mathcal{R}$ is a cone. We will show that it is closed by an induction argument based on the number of vectors $\boldsymbol{a}_j$, $j = 1, \ldots, k$.

When $k = 1$, $\mathcal{R}$ is either $\{0\}$ if $\mu_1 = 0$ or is a half-line and is therefore, in either case, is a closed set. Now, suppose that for some $k \geq 0$ all sets of the form

$$\mathcal{R}_{k-1} := \left\{ \boldsymbol{z} \,\middle|\, \boldsymbol{z} = \sum_{j=1}^{k-1} \mu_j \, \boldsymbol{a}_j \,, \; \mu_j \geq 0 \right\} ,$$

are closed. Then we consider a set of the form

$$\mathcal{R}_k := \left\{ \boldsymbol{z} \,\middle|\, \boldsymbol{z} = \sum_{j=1}^{k} \mu_j \, \boldsymbol{a}_j \,, \; \mu_j \geq 0 \right\} .$$

We show that this latter set is also closed. There are two cases. First, suppose that $\mathcal{R}_k$ contains the vectors $-\boldsymbol{a}_1, -\boldsymbol{a}_2, \cdots, -\boldsymbol{a}_k$. Then $\mathcal{R}_k$ is a subspace of dimension not exceeding k so it is closed.

In the second case, suppose that $\mathcal{R}_k$ does not contain one of the $-\boldsymbol{a}_i$. Without loss of generality, we may suppose that it does not contain $-\boldsymbol{a}_k$ (renumber if necessary). Then, every $\boldsymbol{y} \in \mathcal{R}_k$ has the form $\boldsymbol{y} = \boldsymbol{y}_{k-1} + \alpha \boldsymbol{a}_k$. To show that $\mathcal{R}_k$ is closed, suppose that $\boldsymbol{z}^{(o)}$ is a limit point. Then there exists a sequence $\{\boldsymbol{z}^{(n)}\}_{n=1}^{\infty} \subset \mathcal{R}_k$ such that $\boldsymbol{z}^{(n)} \to \boldsymbol{z}^{(o)}$ as $n \to \infty$ where the $\boldsymbol{z}^{(n)}$ have the form

$$\boldsymbol{z}^{(n)} = \boldsymbol{y}_{k-1}^{(n)} + \alpha_n \boldsymbol{a}_k \, , \; \alpha_n \geq 0 \, .$$

Let us suppose, for the moment, that the sequence $\{\alpha_n\}_{n=1}^{\infty}$ is bounded. Then, without loss of generality, we may assume that the sequence converges to a limit α as $n \to \infty$. Then, $\boldsymbol{z}^{(n)} - \alpha_n \boldsymbol{a}_k \in \mathcal{R}_{k-1}$ and this latter set is closed. Therefore

$$\boldsymbol{z} - \alpha \boldsymbol{a}_k \; = \; \lim_{n \to \infty} (\boldsymbol{z}^{(n)} - \alpha_n \boldsymbol{a}_k) \; = \; \lim_{n \to \infty} \boldsymbol{y}_{k-1}^{(n)} := \overline{\boldsymbol{y}} \in \mathcal{R}_{k-1} \, .$$

We may conclude that

$$\boldsymbol{z} \; = \; \overline{\boldsymbol{y}} + \alpha \boldsymbol{a}_k \in \mathcal{R}_k \, .$$

Hence this latter set is closed.

It remains to prove that the sequence $\{\alpha_n\}_{n=1}^{\infty}$ is a bounded sequence. Assume the contrary, namely that $\alpha_n \to \infty$ as $n \to \infty$. Then since the $\boldsymbol{z}^{(n)}$ converge, they form a bounded sequence. Hence $(1/\alpha_n)\boldsymbol{z}^{(n)} \to 0$ as $n \to \infty$. It follows that $(1/\alpha_n)\boldsymbol{y}_{k-1}^{(n)} + \boldsymbol{a}_k \to 0$ as $n \to \infty$ Therefore $\lim_{n \to \infty}(1/\alpha_n)\boldsymbol{y}_{k-1}^{(n)} = -\boldsymbol{a}_k$. But since $\mathcal{R}_{k-1}$ is closed, this means that $-\boldsymbol{a}_k \in \mathcal{R}_{k-1}$ which is a contradiction. □

Having the results of this lemma in hand, we may turn to the proof of Theorem 1.1.

Proof: First, it is easy to see that the condition is necessary. Indeed, if the system (1.1-1.2) has a non-negative solution $\overline{\boldsymbol{x}} \geq 0$, then, for all $\boldsymbol{y} \in \mathbb{R}^m$ such that $A^\top \boldsymbol{y} \geq 0$, we have

$$\langle \boldsymbol{y}, \boldsymbol{b} \rangle \; = \; \langle \boldsymbol{y}, A\overline{\boldsymbol{x}} \rangle \; = \; \langle A^\top \boldsymbol{y}, \overline{\boldsymbol{x}} \rangle \geq 0 \, ,$$

since all terms in the inner product are products of non-negative real numbers.

To see that the condition is sufficient we assume that the system (1.1-1.2) has no solution and show that there is some vector $\boldsymbol{y}$ such that $A^\top \boldsymbol{y} \geq 0$ and $< \boldsymbol{b}, \boldsymbol{y} > < 0$.
In order to do this, we will apply the basic separation theorem. Consider the set

$$\mathcal{R} := \{\boldsymbol{z} \in \mathbb{R}^m \,|\, \boldsymbol{z} = A\boldsymbol{x}, \boldsymbol{x} \geq 0\}\,.$$

Clearly this set is convex and, by the preceeding lemma, it is closed. To say that the system (1.1-1.2) has no solution says that $\boldsymbol{b} \notin \mathcal{R}$. Observe that the set $\{\boldsymbol{b}\}$ is closed, bounded and convex. Hence, by the strict separation theorem, there exists a vector $\boldsymbol{a} \in \mathbb{R}^m, \boldsymbol{a} \neq 0$ and a scalar α, such that

$$\langle \boldsymbol{a}, \boldsymbol{y} \rangle < \alpha \leq \langle \boldsymbol{a}, \boldsymbol{b} \rangle\,, \ \text{ for all } \ \boldsymbol{y} \in \mathcal{R}\,.$$

Since $0 \in \mathcal{R}$ we must have $\alpha > 0$. Hence $\langle \boldsymbol{a}, \boldsymbol{b} \rangle > 0$. Likewise, $\langle \boldsymbol{a}, A\boldsymbol{x} \rangle \leq \alpha$ for all $\boldsymbol{x} \geq 0$. From this it follows that $A^\top \boldsymbol{a} \leq 0$. Indeed, if the vector $\boldsymbol{w} = A^\top \boldsymbol{a}$ were to have a positive component, say w_j then we can take $\hat{\boldsymbol{x}} = (0, 0, \ldots, 0, M, 0, \ldots, 0)^\top$ where $M > 0$ appears in the jth position. Then certainly $\hat{\boldsymbol{x}} \geq 0$ and

$$\langle A^\top \boldsymbol{a}, \hat{\boldsymbol{x}} \rangle \; = \; w_j \, M\,,$$

which can be made as large as desired by choosing M sufficiently large. In particular, if we choose $M > \alpha / w_j$ then the bound $\langle \boldsymbol{a}, A\boldsymbol{x} \rangle \leq \alpha$ is violated. This shows that $A^\top \boldsymbol{a} \leq 0$ and completes the proof. Indeed, we simply set $\boldsymbol{y} = -\boldsymbol{a}$ to get the required result. □

There are a number of variants which can be reduced to the basic theorem.

(a) The system

$$\begin{aligned} A\,\boldsymbol{x} \;\; &\leq \;\; \boldsymbol{b}\,, \\ \boldsymbol{x} \;\; &\geq \;\; 0\,, \end{aligned}$$

has a solution if and only if, for all $\boldsymbol{y} \geq 0$ such that $A^\top \boldsymbol{y} \geq 0$ we have $\langle \boldsymbol{y}, \boldsymbol{b} \rangle \geq 0$.

(b) The system

$$A\,\boldsymbol{x} \leq \, \boldsymbol{b}\,,$$

has a solution if and only if for all $\boldsymbol{y} \geq 0$ such that $A\boldsymbol{y} = 0$ we have $\langle \boldsymbol{y}, \boldsymbol{b} \rangle \geq 0$.

There are also a number of closely related results. Here is one.

Theorem 1.5 *(Gordon) Let A be an $m \times n$ real matrix, $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{y} \in \mathbb{R}^m$. Then one and only one of the following conditions holds:*

1. *There exists and $\boldsymbol{x} \in \mathbb{R}^n$ such that $A\boldsymbol{x} < 0$;*

2. *There exists a* $\boldsymbol{y} \in \mathbb{R}^m, \boldsymbol{y} \neq 0$ *such that* $A^\top \boldsymbol{y} = 0$ *and* $\boldsymbol{y} \geq 0$.

Proof: Let $\hat{\boldsymbol{e}} = (1, 1, \ldots, 1)^\top \in \mathbb{R}^m$. Then the first condition is equivalent to saying that $A\boldsymbol{x} \leq -\hat{\boldsymbol{e}}$ has a solution. By Theorem 1.1 this is equivalent to the statement that if $\boldsymbol{y} \geq 0$ and $A^\top \boldsymbol{y} = 0$ then $\langle -\boldsymbol{y}, \hat{\boldsymbol{e}} \rangle \geq 0$. Hence there is no $\boldsymbol{y} \neq 0$ such that $A^\top \boldsymbol{y} = 0$ *and* $\boldsymbol{y} \geq 0$.

Conversely, if there is a $\boldsymbol{y} \neq 0$ such that $A^\top \boldsymbol{y} = 0$ and $\boldsymbol{y} \geq 0$, then the condition of the Farkas-Minkowski Theorem does not hold and hence $A\boldsymbol{x} < 0$ has no solution. □

We now turn to a basic solvability theorem from Linear Algebra which is sometimes called the Fredholm Alternative Theorem. While the usual proofs of this result do not use the Farkas-Minkowski result, we do so here.

Theorem 1.6 *Either the system of equations* $A\boldsymbol{x} = \boldsymbol{b}$ *has a solution, or else there is a vector* $\boldsymbol{y}$ *such that* $A^\top \boldsymbol{y} = 0$ *and* $y^\top \boldsymbol{b} \neq 0$.

Proof: In order to use the Farkas-Minkowski Theorem, we need to rewrite the equation $A\boldsymbol{x} = \boldsymbol{b}$ which has no sign constraint, as an equivalent system which does have such a constraint. We do this by introducing two new variables $\boldsymbol{u} \geq 0$ and $\boldsymbol{v} \geq 0$ with $\boldsymbol{x} + \boldsymbol{v} = \boldsymbol{u}$. The equation then is $A(\boldsymbol{u} - \boldsymbol{v}) = \boldsymbol{b}$ or, in partitioned form

$$[A, -A] \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} = \boldsymbol{b},$$

with non-negative solution $(\boldsymbol{u}, \boldsymbol{v})^\top \geq 0$. The Farkas alternative is then that $\boldsymbol{y}^\top [A, -A] \geq 0$, $\boldsymbol{y}^\top \boldsymbol{b} < 0$ has a solution $\boldsymbol{y}$. These inequalities mean, in particular, that $\boldsymbol{y}^\top A \geq 0$ and $-\boldsymbol{y}^\top A \geq 0$ which together imply that $\boldsymbol{y}^\top A = 0$. So the Farkas alternative says $\boldsymbol{y}^\top A = 0, y^\top b < 0$ has a solution.

Without loss of generality, we may replace $\boldsymbol{y}$ by $-\boldsymbol{y}$ so that we now have $\boldsymbol{y}^\top A = 0$ and $\boldsymbol{y}^\top \boldsymbol{b} \neq 0$. This is exactly the Fredholm Alternative. □