The University of Texas at Austin
Department of Electrical and Computer Engineering

EE381K: Convex Optimization — Fall 2019

Lecture 4

Aryan Mokhtari Wednesday, September 11, 2019.

Goal: In this lecture we prove Farkas' Lemma using the Theorem of Alternatives (covered in lecture 3). Also, we study a geometric interpretation of Farkas' Lemma. [We will later use Farkas' Lemma to prove strong duality in Linear Programming.]

1 Statement and Proof of Farkas' Lemma

Recall the theorem of alternatives from the previous lecture.

Theorem 1 (Theorem of Alternatives)**.** For any given matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ and vector $\mathbf{b} \in \mathbb{R}^m$ exactly one of the following statements holds:
(a) There exists an $\mathbf{x} \in \mathbb{R}^n$ that satisfies $\mathbf{A}\mathbf{x} \leq \mathbf{b}$.

(b) There exists a $\mathbf{z} \in \mathbb{R}^m$ that satisfies $\mathbf{z} \geq \mathbf{0}$, $\mathbf{A}^\top \mathbf{z} = \mathbf{0}$, $\mathbf{b}^\top \mathbf{z} < 0$.

We will use Theorem 1 to prove the following result which is also known as Farkas' Lemma.

Lemma 1 (Farkas' Lemma)**.** For any given matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ and vector $\mathbf{b} \in \mathbb{R}^m$ exactly one of the following statements holds:
(I) There exists an $\mathbf{x} \in \mathbb{R}^n$ that satisfies $\mathbf{A}\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$.

(II) There exists a $\mathbf{y} \in \mathbb{R}^m$ that satisfies $\mathbf{A}^\top \mathbf{y} \geq \mathbf{0}$ and $\mathbf{b}^\top \mathbf{y} < 0$.

Proof. Note that the conditions in (I) can be written as

$$\mathbf{A}\mathbf{x} \leq \mathbf{b}, \quad \mathbf{A}\mathbf{x} \geq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0} \tag{1}$$

which are equivalent to

$$\mathbf{A}\mathbf{x} \leq \mathbf{b}, \quad -\mathbf{A}\mathbf{x} \leq -\mathbf{b}, \quad -\mathbf{x} \leq \mathbf{0} \tag{2}$$

We combine these inequalities and write them as

$$\begin{bmatrix} \mathbf{A} \\ -\mathbf{A} \\ -\mathbf{I} \end{bmatrix} \mathbf{x} \leq \begin{bmatrix} \mathbf{b} \\ -\mathbf{b} \\ -\mathbf{0} \end{bmatrix} \tag{3}$$

Now if we define the matrix $\hat{\mathbf{A}} := \begin{bmatrix} \mathbf{A} \\ -\mathbf{A} \\ -\mathbf{I} \end{bmatrix} \in \mathbb{R}^{(2m+n)\times n}$ and the vector $\hat{\mathbf{b}} := \begin{bmatrix} \mathbf{b} \\ -\mathbf{b} \\ -\mathbf{0} \end{bmatrix} \in \mathbb{R}^{(2m+n)}$, then we obtain that the conditions in (1) are equivalent to

$$\hat{\mathbf{A}}\mathbf{x} \leq \hat{\mathbf{b}}. \tag{4}$$

Hence, we can conclude that the statement in (I) is equivalent to
(I') There exists an $\mathbf{x} \in \mathbb{R}^n$ that satisfies $\hat{\mathbf{A}}\mathbf{x} \leq \hat{\mathbf{b}}$.

Now according to the Theorem of Alternatives, if (I') does not hold, then
(II') There exists a $\mathbf{z} \in \mathbb{R}^{2m+n}$ that satisfies $\mathbf{z} \geq \mathbf{0}$, $\hat{\mathbf{A}}^\top \mathbf{z} = \mathbf{0}$, $\hat{\mathbf{b}}^\top \mathbf{z} < 0$.

Now we proceed to show that the statement in (II') is equivalent to the one in (II). Since $\mathbf{z} \in \mathbb{R}^{2m+n}$ we can write it as $\hat{\mathbf{z}} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{w} \end{bmatrix}$, where $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{v} \in \mathbb{R}^m$, and $\mathbf{w} \in \mathbb{R}^n$. Note that the conditions in statement (II') which are $\mathbf{z} \geq \mathbf{0}$, $\hat{\mathbf{A}}^\top \mathbf{z} = \mathbf{0}$, $\hat{\mathbf{b}}^\top \mathbf{z} < 0$ can be written as

$$\begin{aligned} \mathbf{z} \geq \mathbf{0} \quad &\Rightarrow \quad \mathbf{u} \geq \mathbf{0}, \quad \mathbf{v} \geq \mathbf{0}, \quad \mathbf{w} \geq \mathbf{0} \\ \hat{\mathbf{A}}^\top \mathbf{z} = \mathbf{0} \quad &\Rightarrow \quad \mathbf{A}^\top \mathbf{u} - \mathbf{A}^\top \mathbf{v} - \mathbf{w} = \mathbf{0} \\ \hat{\mathbf{b}}^\top \mathbf{z} < 0 \quad &\Rightarrow \quad \mathbf{b}^\top \mathbf{u} - \mathbf{b}^\top \mathbf{v} < 0 \end{aligned}$$

If we regroup the terms, these conditions can be written as

$$\mathbf{u} \geq \mathbf{0}, \quad \mathbf{v} \geq \mathbf{0}, \quad \mathbf{w} \geq \mathbf{0}, \quad \mathbf{A}^\top (\mathbf{u} - \mathbf{v}) = \mathbf{w}, \quad \mathbf{b}^\top (\mathbf{u} - \mathbf{v}) < 0. \tag{5}$$

We can eliminate the vector $\mathbf{w}$ and combine the conditions $\mathbf{w} \geq \mathbf{0}$ and $\mathbf{A}^\top (\mathbf{u} - \mathbf{v}) = \mathbf{w}$ to write $\mathbf{A}^\top (\mathbf{u} - \mathbf{v}) \geq 0$. Hence, we can rewrite the conditions in (5) as

$$\mathbf{u} \geq \mathbf{0}, \quad \mathbf{v} \geq \mathbf{0}, \quad \mathbf{A}^\top (\mathbf{u} - \mathbf{v}) \geq 0, \quad \mathbf{b}^\top (\mathbf{u} - \mathbf{v}) < 0. \tag{6}$$

Now note that both variables $\mathbf{u}$ and $\mathbf{v}$ are non-negative, therefore if we define $\mathbf{y} := \mathbf{u} - \mathbf{v}$, then the variable $\mathbf{y}$ can take any value in $\mathbb{R}^m$. Hence, we can rewrite the conditions in (6) as

$$\mathbf{A}^\top \mathbf{y} \geq 0, \quad \mathbf{b}^\top \mathbf{y} < 0. \tag{7}$$

Hence, we showed that the statement in (II') is equivalent to the following statement:
There exists a vector $\mathbf{y} \in \mathbb{R}^m$ that satisfies $\mathbf{A}^\top \mathbf{y} \geq 0, \mathbf{b}^\top \mathbf{y} < 0$,

which is identical to the statement in (II).

To summarize, we showed that according to the theorem of alternatives only one of the statements in (I') and (II') hold, and since (I') is equivalent to (I) and (II') is equivalent to (II), then we can conclude that only one of the statements in (I) and (II) can be true. The proof is complete. □

2 Geometric Interpretation of Farkas' Lemma

Farkas' Lemma is essential for proving strong duality of Linear Programming. But, it also has a nice geometric interpretation. To explain its geometric interpretation we first need to define a "convex cone".

Definition 1 (Convex cone)**.** A nonempty set $\mathcal{S}$ is a convex cone if it has the following property

$$\text{For any} \quad \mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathcal{S} \quad \text{and} \quad \theta_1 \geq 0, \ldots, \theta_k \geq 0 \quad \Rightarrow \quad \theta_1 \mathbf{x}_1 + \cdots + \theta_k \mathbf{x}_k \in \mathcal{S} \tag{8}$$

Example 1. *The solution of a finite system of homogeneous linear inequalities is a convex cone, i.e., the set* $\mathcal{S} := \{\mathbf{x} \mid \mathbf{A}\mathbf{x} \leq \mathbf{0}, \mathbf{C}\mathbf{x} = \mathbf{0}\}$ *is a convex cone.*

(We'll recap the definition of convex cone later in the class when we talk about convex optimization)

In Farkas' lemma we are given a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ (with columns $\mathbf{a}_i$) and a vector $\mathbf{b} \in \mathbb{R}^m$.

The first alternative. There exists a vector $\mathbf{x} \in \mathbb{R}^n$ such that

$$\mathbf{b} = \sum_{i=1}^{n} x_i \mathbf{a}_i \qquad x_i \geq 0 \quad \text{for all} \quad i = 1, \ldots, n$$

The vector $\mathbf{b}$ belongs to the convex cone created by the columns of $\mathbf{A}$

The second alternative. There exists a vector $\mathbf{y} \in \mathbb{R}^m$ such that

$$\mathbf{y}^\top \mathbf{a}_i \geq 0 \quad \text{for all} \quad i = 1, \ldots, n \qquad\qquad \mathbf{y}^\top \mathbf{b} < 0$$

The hyperplane $\mathbf{y}^\top \mathbf{z} = 0$ separates $\mathbf{b}$ from $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$

Conclusion: Either $\mathbf{b}$ belongs to the convex cone generated by the columns of $\mathbf{A}$, or it does not belong to the convex cone and therefore they can be separated by a hyperplane.

3 Other Side Effects of Theorem of Alternatives

Now we consider the case that we have mixed inequalities and equalities.

Theorem 2. For any $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{C} \in \mathbb{R}^{p \times n}$, and $\mathbf{d} \in \mathbb{R}^p$ exactly one of the following statements holds:

(a) There exists an $\mathbf{x} \in \mathbb{R}^n$ that satisfies $\mathbf{Ax} \leq \mathbf{b}$ and $\mathbf{Cx} = \mathbf{d}$.

(b) There exist $\mathbf{z} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^p$ that satisfy $\mathbf{z} \geq \mathbf{0}$, $\mathbf{A}^\top \mathbf{z} + \mathbf{C}^\top \mathbf{y} = \mathbf{0}$, $\mathbf{b}^\top \mathbf{z} + \mathbf{d}^\top \mathbf{y} < 0$.

Proof. Homework (Problem Set 2). □

4 Different forms of LP

LP in "inequality form" (only inequality constraints)

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^\top \mathbf{x} \\ \text{subject to} \quad & \mathbf{Ax} \leq \mathbf{b} \end{aligned} \tag{9}$$

LP in "standard form" (constraints are equality, variables are non-negative)

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^\top \mathbf{x} \\ \text{subject to} \quad & \mathbf{Ax} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{10}$$

Tips:
- Inequalities can be converted to equalities by adding or subtracting non-negative slack variables
Example 1: $\mathbf{Ax} \leq \mathbf{b}$ can be replaced by $\mathbf{Ax} + \mathbf{s} = \mathbf{b}$, $\mathbf{s} \geq \mathbf{0}$.
Example 2: $\mathbf{Ax} \geq \mathbf{b}$ can be replaced by $\mathbf{Ax} - \mathbf{s} = \mathbf{b}$, $\mathbf{s} \geq \mathbf{0}$.
-To write an LP in standard form we can write an unconstrained variable $\mathbf{x}$ as the difference of two positive variables: $\mathbf{x} = \mathbf{y} - \mathbf{z}$, $\mathbf{y} \geq \mathbf{0}$, $\mathbf{z} \geq \mathbf{0}$.

Example 2. Write the following LP in a standard form.

$$\begin{aligned} \text{minimize} \quad & x_1 + x_2 \\ \text{subject to} \quad & x_1 - x_2 \leq 3 \\ & x_1 \geq 0 \\ & x_2 \leq 4 \end{aligned}$$

Define a new variable $y \geq 0$ to replace the constraint $x_1 - x_2 \leq 3$ by $x_1 - x_2 + y = 3$, and define a new variable $z \geq 0$ to replace $x_2 \leq 4$ by $x_2 + z = 4$. Hence, it can be written as

$$\begin{aligned} \text{minimize} \quad & x_1 + x_2 \\ \text{subject to} \quad & x_1 - x_2 + y = 3 \\ & x_2 + z = 4 \\ & x_1 \geq 0 \\ & y \geq 0 \\ & z \geq 0 \end{aligned}$$

But now we don't have any constraints on x_2, therefore we should define two new variables $u \geq 0$ and $v \geq 0$ and replace x_2 with $u - v$ to obtain

$$\begin{aligned} \text{minimize} \quad & x_1 + u - v \\ \text{subject to} \quad & x_1 - u + v + y = 3 \\ & u - v + z = 4 \\ & x_1 \geq 0 \\ & y \geq 0 \\ & z \geq 0 \\ & u \geq 0 \\ & v \geq 0 \end{aligned}$$