Math Models of OR: Theorems of the Alternative

John E. Mitchell

Department of Mathematical Sciences
RPI, Troy, NY 12180 USA

October 2018
The Theorem (Farkas Lemma) states:

Let $A$ be an $m \times n$ matrix and let $b \in \mathbb{R}^m$. Exactly one of the following two systems is consistent:

- There exists an $x \in \mathbb{R}^n$ with $Ax = b$ and $x \geq 0$.
- There exists a $y \in \mathbb{R}^m$ with $A^T y \leq 0$ and $b^T y > 0$. 

This theorem is also known as the Farkas Lemma.
Theorem (Farkas Lemma)

Let $A$ be an $m \times n$ matrix and let $b \in \mathbb{R}^m$. Exactly one of the following two systems is consistent:

I. There exists an $x \in \mathbb{R}^n$ with $Ax = b$ and $x \geq 0$.

II. There exists a $y \in \mathbb{R}^m$ with $A^T y \leq 0$ and $b^T y > 0$. 
Theorem (Farkas Lemma)

Let $A$ be an $m \times n$ matrix and let $b \in \mathbb{R}^m$. Exactly one of the following two systems is consistent:

I. There exists an $x \in \mathbb{R}^n$ with $Ax = b$ and $x \geq 0$.

II. There exists a $y \in \mathbb{R}^m$ with $A^T y \leq 0$ and $b^T y > 0$. 
Example 1

Let

\[ A = \begin{bmatrix} 2 & 3 & -1 \\ 4 & 1 & 3 \end{bmatrix}, \quad b = \begin{bmatrix} 7 \\ 19 \end{bmatrix} \]

System I is consistent: we can take \( \bar{x} = (3, 1, 2)^T \). It can be checked that \( A\bar{x} = b \).
Example 1: System II is inconsistent

Farkas then says that there does not exist $y \in \mathbb{R}^2$ satisfying

$$A^T y = \begin{bmatrix} 2 & 4 \\ 3 & 1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and $b^T y = 7y_1 + 19y_2 > 0$.

Adding together the three equations for $A^T y$ with weights equal to $\bar{x} = (3, 1, 2)^T$ gives:

$$3(2y_1 + 4y_2) \leq 0$$
$$+ \quad 1(3y_1 + y_2) \leq 0$$
$$+ \quad 2(-y_1 + 3y_2) \leq 0$$

$$\implies \quad 7y_1 + 19y_2 \leq 0,$$

contradicting $b^T y > 0$. 
Example 2

\[ A^T \bar{y} = \begin{bmatrix} -2 & 3 & 0 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ -4 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \]

Let

\[ A = \begin{bmatrix} -2 & 3 & 0 \\ 1 & -1 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \]

The vector \( \bar{y} = (-1, -2)^T \) satisfies both \( A^T y \leq 0 \) and \( b^T y > 0 \), so System II is consistent.

For \( Ax = b \), need both

\[
\begin{align*}
-1 \left( -2x_1 + 3x_2 \right) &= 1 \\
-2 \left( x_1 - x_2 + 2x_3 \right) &= -1 \\
0x_1 - x_2 - 4x_3 &= 1
\end{align*}
\]

Adding these two equalities with weights \(-1\) and \(-2\) respectively implies \( x \) must satisfy \( 0x_1 - x_2 - 4x_3 = 1 \), which is impossible for \( x \geq 0 \), so System I is inconsistent.
The Farkas Lemma

Proof of Farkas Lemma

We first show that at most one of Systems I and II has a solution:

Assume $\bar{x}$ satisfies System I and $\bar{y}$ satisfies System II.

So $A\bar{x} = b$, $\bar{x} \geq 0$, $A^T\bar{y} \leq 0$, and $b^T\bar{y} > 0$. Then

$$\bar{y}^T(A\bar{x}) = \bar{y}^Tb \quad \text{since} \quad A\bar{x} = b$$
$$\implies (A^T\bar{y})^T\bar{x} > 0 \quad \text{since} \quad b^T\bar{y} > 0,$$
and matrix multiplication is associative

But $A^T\bar{y} \leq 0$ and $\bar{x} \geq 0$ so $(A^T\bar{y})^T\bar{x} \leq 0$, which is a contradiction.

So at most one of the Systems has a solution.
Proof of Farkas, part 2

We now show that at least one of Systems I and II has a solution:

Consider the primal-dual pair of linear optimization problems:

\[
\begin{align*}
\text{min}_{x \in \mathbb{R}^n} & \quad 0^T x \\
\text{subject to} & \quad Ax = b & (P) \\
& \quad x \geq 0
\end{align*}
\]

I: \( Ax = b, x \geq 0 \)

\[
\begin{align*}
\text{max}_{y \in \mathbb{R}^m} & \quad b^T y \\
\text{subject to} & \quad A^T y \leq 0 & (D) \\
& \quad y \text{ free}
\end{align*}
\]

II: \( A^T y \leq 0, b^T y > 0 \)

Assume I does not have a solution, show II must have a solution.

Since System I is inconsistent, the LOP (P) is infeasible.

Therefore, by strong duality, the dual problem (D) is either infeasible or unbounded. Now, (D) is always feasible: just take \( y = 0 \in \mathbb{R}^m \).

Therefore, (D) has an unbounded optimal value. Thus, it has a feasible solution \( \bar{y} \) with positive objective function value.

Then \( \bar{y} \) is a feasible solution to System II.
Outline

1. The Farkas Lemma
2. Visualizing the Farkas Lemma
3. Other Theorems of the Alternative
The columns of $A$

Let

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix}. $$

For what vectors $b \in \mathbb{R}^2$ is System I consistent?

For what $b \in \mathbb{R}^2$ is System II consistent?
Consistency of System I

For System I, we require $b = Ax$ for some $x \geq 0$.

We can express $Ax$ as follows:

$$
Ax = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 3x_2 + 2x_3 \\ 4x_1 + 2x_2 + x_3 \end{bmatrix}
$$

$$
= x_1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 1 \end{bmatrix}
$$

Thus, System I is consistent if and only if $b$ is a nonnegative linear combination of the columns of $A$. This is a cone in $\mathbb{R}^m$. 

Mitchell Theorems of the Alternative

11 / 20
Nonnegative combinations of the columns of $A$

1: $Ax = b$, $x \geq 0$

$$A = \begin{bmatrix}
1 & 3 & 2 \\
4 & 2 & 1
\end{bmatrix}$$

Mitchell
Nonnegative combinations of the columns of $A$

\[ A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix} \]

\[ x = \begin{bmatrix} 0.5 \\ 0 \\ 0 \end{bmatrix} \]

\[ Ax = \begin{bmatrix} 0.5 \\ 2 \end{bmatrix} \]

I: $Ax = b$, $x \geq 0$
Nonnegative combinations of the columns of $A$

I: $Ax = b, \ x \geq 0$

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix}$$

$$x = \begin{bmatrix} t \\ 0 \\ 0 \end{bmatrix}$$

$$Ax = \begin{bmatrix} t \\ 4t \\ t \geq 0 \end{bmatrix}$$
Nonnegative combinations of the columns of $A$

\[ I: Ax = b, \ x \geq 0 \]

\[ A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix} \]

\[ x = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \]

\[ Ax = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \]
Nonnegative combinations of the columns of $A$

I: $Ax = b, \ x \geq 0$

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix}$$

$$x = \begin{bmatrix} 0 \\ 0 \\ s \end{bmatrix}$$

$$Ax = \begin{bmatrix} 2s \\ s \end{bmatrix}$$

$$s \geq 0$$
Nonnegative combinations of the columns of $A$

\[ I: \ Ax = b, \ x \geq 0 \]

\[ A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix} \]

\[ x = \begin{bmatrix} 0.5 \\ 1.5 \\ 2 \end{bmatrix} \]

\[ Ax = \begin{bmatrix} 9 \\ 7 \end{bmatrix} \]
Nonnegative combinations of the columns of $A$

System I consistent for $b$ here

$A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix}$

$I: Ax = b, \ x \geq 0$
A polar cone

System II considers the $y \in \mathbb{R}^m$ satisfying $A^T y \leq 0$. We have:

$$\{ y \in \mathbb{R}^m : A^T y \leq 0 \} = \left\{ y \in \mathbb{R}^2 : \begin{bmatrix} 1 & 4 \\ 3 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

$$= \left\{ y \in \mathbb{R}^2 : \begin{bmatrix} 1 \\ 4 \end{bmatrix}^T \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \leq 0, \begin{bmatrix} 3 \\ 2 \end{bmatrix}^T \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \leq 0, \begin{bmatrix} 2 \\ 1 \end{bmatrix}^T \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \leq 0 \right\}$$

so we consider the set of vectors $y$ that make an angle of at least $\pi/2$ with each column of $A$. This is another cone in $\mathbb{R}^m$.

If $b$ is a nonzero vector in this cone then one solution to System II is to take $y = b$, since then $b^T y = b^T b > 0$. 

Mitchell
Theorems of the Alternative
Picturing the polar cone

II: $A^T y \leq 0$, $b^T y > 0$

$A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix}$

$y$: non-positive inner product with each column

$g$: non-positive inner product with each column
Picturing the polar cone

II: $A^T y \leq 0, \ b^T y > 0$

$A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix}$

$y_1 + 4y_2 \leq 0$
Picturing the polar cone

II: $A^T y \leq 0, b^T y > 0$

\[ A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix} \]

$2y_1 + y_2 \leq 0$

\[ (1,4), (3,2), (2,1) \]
Picturing the polar cone

II: $A^T y \leq 0, \ b^T y > 0$

$A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix}$

$3y_1 + 2y_2 \leq 0$

$A^T y \leq 0$

$2y_1 + y_2 = 0$

$(1, 4)$

$(3, 2)$

$(2, 1)$
Picturing the polar cone

II: $A^T y \leq 0, \ b^T y > 0$

$A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix}$

$y_1 + 4y_2 \leq 0$

$3y_1 + 2y_2 \leq 0$

$2y_1 + y_2 \leq 0$
Picturing the polar cone

II: $A^T y \leq 0$, $b^T y > 0$

$A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix}$

$y_1 + 4y_2 \leq 0$
$3y_1 + 2y_2 \leq 0$
$2y_1 + y_2 \leq 0$

System II consistent: take $\bar{y} = b$
Other choices of $b$

If $b$ is in neither cone then there is a point $y$ on the boundary of the second cone that makes positive inner product with $b$. This $y$ satisfies the conditions of System II.

We can graph the problem in $\mathbb{R}^m$. The columns of $A$ are points in $\mathbb{R}^m$, and the two cones are subsets of $\mathbb{R}^m$. 
The cones

System I
consistent

System II
consistent: take $\bar{y} = b$

$y_1 + 4y_2 = 0$

$\{ y : A^T y \leq 0 \}$

$2y_1 + y_2 = 0$
The cones

System II consistent:
\[
\begin{align*}
\text{take } & \bar{y} = (-4, 1) \\
\text{so } & b^T \bar{y} > 0
\end{align*}
\]

System I consistent:
\[
\begin{align*}
(1,4) \\
(3,2) \\
(2,1)
\end{align*}
\]

\[
\begin{align*}
\{y : A^T y \leq 0\}
\end{align*}
\]

System II consistent:
\[
\begin{align*}
\text{take } & \bar{y} = b
\end{align*}
\]
The cones

System I
consistent:
take $\bar{y} = (4, 1)$
so $b^T \bar{y} > 0$

System II
consistent:
take $\bar{y} = (2, 1)$
so $b^T \bar{y} > 0$

System I
consistent:
take $\bar{y} = (1, 4)$

System II
consistent:
take $\bar{y} = (1, -2)$
so $b^T \bar{y} > 0$
Outline

1. The Farkas Lemma

2. Visualizing the Farkas Lemma

3. Other Theorems of the Alternative
Gordan’s Theorem

The Farkas Lemma is an example of a *theorem of the alternative*: either one system is consistent, or another, but not both. Many of these theorems can be proved using linear optimization duality.

**Theorem (Gordan’s Theorem)**

Let $A$ be an $m \times n$ matrix. Exactly one of the following two systems is consistent:

I. There exists a nonzero $x \in \mathbb{R}^n$ with $Ax = 0$, $x \geq 0$.

II. There exists a vector $y \in \mathbb{R}^m$ with $A^T y > 0$.

Gordan’s Theorem is used in the proof of optimality conditions for nonlinear optimization problems.
Gordan’s Theorem

The Farkas Lemma is an example of a theorem of the alternative: either one system is consistent, or another, but not both. Many of these theorems can be proved using linear optimization duality.

**Theorem (Gordan’s Theorem)**

Let $A$ be an $m \times n$ matrix. Exactly one of the following two systems is consistent:

I. There exists a nonzero $x \in \mathbb{R}^n$ with $Ax = 0, \ x \geq 0$.

II. There exists a vector $y \in \mathbb{R}^m$ with $A^T y > 0$.

Gordan’s Theorem is used in the proof of optimality conditions for nonlinear optimization problems.
Gordan’s Theorem

The Farkas Lemma is an example of a *theorem of the alternative*: either one system is consistent, or another, but not both. Many of these theorems can be proved using linear optimization duality.

**Theorem (Gordan’s Theorem)**

*Let $A$ be an $m \times n$ matrix. Exactly one of the following two systems is consistent:

1. There exists a nonzero $x \in \mathbb{R}^n$ with $Ax = 0$, $x \geq 0$.
2. There exists a vector $y \in \mathbb{R}^m$ with $A^T y > 0$.*

Gordan’s Theorem is used in the proof of optimality conditions for nonlinear optimization problems.
Handling strict inequalities

The proof of Gordan’s Theorem is a little tricky because of the strict vector inequality in System II.

One remedy is to notice that there exists a $y$ with $A^T y > 0$ if and only if there exists a $\tilde{y} \in \mathbb{R}^m$ with every component of $A^T \tilde{y}$ equal to at least 1: we can rescale.

\[ A^T y = 0 \]
A primal-dual pair of linear optimization problems

Let \( e = (1, \ldots, 1)^T \in \mathbb{R}^n \), the vector of all ones. We can then set up the primal dual pair of linear optimization problems:

\[
\begin{align*}
\text{max}_{x \in \mathbb{R}^n} & \quad e^T x \\
\text{subject to} & \quad Ax = 0 \quad (P) \quad x \geq 0 \\
\text{I: } & \quad Ax = 0, x \geq 0, x \neq 0
\end{align*}
\]

\[
\begin{align*}
\text{min}_{y \in \mathbb{R}^m} & \quad 0^T y \\
\text{subject to} & \quad A^T y \geq e \quad (D) \\
\text{II: } & \quad A^T y > 0
\end{align*}
\]

Notice that (P) is always feasible: take \( x = 0 \). Further, if (D) is feasible then it has an optimal value of 0. It can then be argued that:

- System I is consistent \( \iff \) (P) has a feasible solution with positive objective function value 
- \( \iff \) (P) has an unbounded optimal value 
- \( \iff \) (D) is infeasible 
- \( \iff \) System II is inconsistent.