Recap

Result (Existence of a vertex)
If \( F = \{ x : Ax \geq b \} \) has at least one point and \( \text{rank}(A) = n \), then \( F \) has at least one vertex.

Step 1: Start at \( x_k \) with \( A_k \). Identify inactive and independent constraint to move towards

Step 2: Compute feasible direction \( p_k \) such that
\[
\begin{pmatrix}
A_k \\
a_j^T
\end{pmatrix}
\begin{pmatrix}
p_k \\
-1
\end{pmatrix}
= \begin{pmatrix}
0 \\
1
\end{pmatrix}
\]

Step 3: Compute maximum feasible step \( \sigma_t \). Update our point to
\[
x_{k+1} = x_k + \sigma_t p_k
\]
and add blocking constraint(s) we bumped into moving along \( p_k \).
Repeat until we get to a vertex.
Theorem

Consider minimizing $c^T x$ subject to $Ax \geq b$.

(A) If no points satisfy $Ax \geq b$, the LP has no solution;

(B) If there exists a point $x^*$ satisfying the three conditions

$$Ax^* \geq b, \quad c = A_a^T \lambda_a^* \quad \text{and} \quad \lambda_a^* \geq 0,$$

where $A_a$ is the active constraint matrix at $x^*$, then $\ell^* = c^T x^*$ is the minimum value of $c^T x$ in the feasible region, and $x^*$ is a minimizer;

(C) If the constraints $Ax \geq b$ are feasible, then $\ell(x)$ is unbounded below in the feasible region if and only if the two conditions $c = A_a^T \lambda_a^*$ and $\lambda_a^* \geq 0$, are not satisfied at any feasible point.

Geometric significance of Farkas’ Lemma

Definition (The dual cone)

The **dual cone** $C_N$ at a feasible point $x$ is the set:

$$C_N(x) = \{ y : y = A_a^T \lambda \text{ for } \lambda \geq 0 \}$$

where $A_a = A_a(x)$.

Recall that

$$\text{range}(A_a^T) = \{ y : y = A_a^T \lambda \}$$

which implies that

$C_N(x)$ is a subset of $\text{range}(A_a^T)$.

Result (Farkas’ Lemma)

(A) $c^T p \geq 0$ for all $p$ such that $A_a p \geq 0$ if and only if

(B) $c = A_a^T \lambda_a^*$ for some $\lambda_a^* \geq 0$

Farkas’ Lemma is known as a **theorem of the alternative**.
Result

$C_N$ is a convex subset of $\mathbb{R}^n$.

Proof: Assume that $x, y \in C_N$, i.e.,

$x \in C_N \implies x = A_T^T \lambda_x$, for some $\lambda_x \geq 0$
$y \in C_N \implies y = A_T^T \lambda_y$, for some $\lambda_y \geq 0$

If $\theta \in [0, 1]$ then $z = (1 - \theta)x + \theta y$ satisfies

$z = (1 - \theta)x + \theta y = (1 - \theta)A_T^T \lambda_x + \theta A_T^T \lambda_y$

$= A_T^T((1 - \theta)\lambda_x + \theta \lambda_y)_{\lambda_z}$

$= A_T^T \lambda_z$ for $\lambda_z \geq 0$

Thus, $z = A_T^T \lambda_z$, for $\lambda_z \geq 0$ and $z \in C_N$. $lacksquare$

Test for optimality

does $c$ lie in $C_N(x)$?

e.g.,
can we write $c = A_T^T \lambda$ for $\lambda \geq 0$?

Example:
Consider the dual cone associated with the active-set matrix

$A_a = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

$= \begin{pmatrix} a_1^T \\ a_2^T \end{pmatrix}$, i.e., $a_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $a_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

By definition

$C_N = \{y : y = \lambda_1 a_1 + \lambda_2 a_2 \text{ with } \lambda_1 \geq 0 \text{ and } \lambda_2 \geq 0\}$
Consider
\[ y = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 2 \end{pmatrix} \]
Is \( y \in C_4 \)?
\[ y = \begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \end{pmatrix} = \frac{1}{2} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Rightarrow \lambda = \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix} \geq 0 \]

Now consider
\[ w = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \]
for the same \( a_1 \) and \( a_2 \):
\[ a_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad a_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \]
Then
\[ w = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Rightarrow \lambda = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \not\geq 0 \]
Example:

minimize $2x_1 + x_2$ subject to the constraints:

- constraint #1: $x_1 + x_2 \geq 1$
- constraint #2: $x_2 \geq 0$
- constraint #3: $x_1 \geq 0$

Written in the form: $\min c^T x$ subject to $Ax \geq b$, we get

$$c = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

At the point $x^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, the active set is $A = \{1, 3\}$

with

$$A_a = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad b_a = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Solving for the Lagrange multipliers gives

$$A_a^T \lambda_a = c \implies \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \lambda_a = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \implies \lambda_a = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \geq 0$$

$\Rightarrow x^*$ is optimal.

At the point $\bar{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, the active set is $A = \{1, 2\}$

with

$$A_a = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad b_a = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Solving for the Lagrange multipliers gives

$$A_a^T \lambda_a = c \implies \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \lambda_a = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \implies \lambda_a = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \not\geq 0$$

$\Rightarrow \bar{x}$ is not optimal.
Proof of Farkas’ Lemma

Result (Farkas’ Lemma)

(A) \( c^T p \geq 0 \) for all \( p \) such that \( A_p p \geq 0 \)

if and only if

(B) \( c = A^T \lambda^* \) for some \( \lambda^* \geq 0 \)
Proof: (B) $\implies$ (A)

If (B) holds, then $c = A^T \lambda_a^*$ for some $\lambda_a^* \geq 0$.

Then,

$$c^T p = (A^T \lambda_a^*)^T p = (\lambda_a^*)^T (A_a p) \geq 0 \text{ for all } p \text{ with } A_a p \geq 0$$

$\implies c^T p \geq 0 \text{ for all } p \text{ with } A_a p \geq 0 \implies (A)$ holds.

Case I: $c$ does not lie in range($A^T_a$).

$c$ can be decomposed uniquely as

$$c = c_R + c_N \text{ with } c_R \in \text{range}(A^T_a) \text{ and } c_N \in \text{null}(A_a).$$

We can then show that $p = -c_N$ is a feasible descent direction

(see Lecture 9)

Case II: $c \in \text{range}(A^T_a)$, but $c = A^T_a \lambda_a^*$ with $\lambda_a^* \not\geq 0$

This is the “HARD” part of Farkas’ lemma.

However, it is EASY if we assume that the rows of $A_a$ are linearly independent.

$$A_a = \square \text{ with linearly independent rows.}$$

Proof: (A) $\implies$ (B) (by contrapositive argument)

We want to show that if there is no $\lambda_a^* \geq 0$ such that $c = A^T_a \lambda_a^*$, then $c^T p < 0$ for some $p$ such that $A_a p \geq 0$.

Two possibilities for there being no $\lambda_a^* \geq 0$ with $c = A^T_a \lambda_a^*$.

Case I: $c$ does not lie in range($A^T_a$)

Case II: $c \in \text{range}(A^T_a)$, but $c = A^T_a \lambda_a^*$ with $\lambda_a^* \not\geq 0$

In both cases we construct a $p$ such that $c^T p < 0$ and $A_a p \geq 0$. 
Case II: $c \in \text{range}(A^T_a)$, but $c = A^T_a \lambda_a^*$ with $\lambda_a^* \geq 0$

In this case, $\lambda_a^*$ exists, but $[\lambda_a^*]_s < 0$ for some index $s$.

Let $p$ be a solution of $A_a p = e_s$, where $e_s$ is the unit vector

$$e_s = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \leftarrow \text{row } s$$

The equations $A_a p = e_s$ are compatible (why?).

Moreover, $p$ is a feasible descent direction, as we now show:

$$c^T p = (A^T_a \lambda_a) T p = (\lambda^T_a A_a) p = \lambda^T_a A_a p = \lambda^T_a e_s = [\lambda_a]_s < 0$$

Hence we have constructed a direction $p$ such that $c^T p < 0$ and $A_a p = e_s \geq 0$

and we are done!

---

Implications of Farkas’ Lemma

In our proof, we defined a feasible descent direction $p$ such that

$$A_a p = e_s$$

where $s$ is an index such that $[\lambda_a^*]_s < 0$

The direction $p$ satisfies:

$$a_j^T p = \begin{cases} 0 & \text{for } j \neq s, \text{ with } a_j^T \text{ the } j\text{th row of } A_a \\ 1 & \text{for } j = s, \text{ with } a_s^T \text{ the } s\text{th row of } A_a \end{cases}$$

A step $x + \alpha p$ keeps all active constraint residuals fixed at zero except the $s$th, which increases.

i.e., $x + \alpha p$ “stays on” all the active constraints and “moves off” of the constraint with the chosen negative component of $\lambda_a$.

At $\bar{x} = (1 \ 0)^T$, the active set is $\mathcal{A} = \{1, 2\}$ with

$$A_a = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

$$b_a = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Solving for the Lagrange multipliers gives

$$A_a^T \lambda_a = c \implies \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \lambda_a = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \implies \lambda_a = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$
\[ \lambda_a = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \leftarrow s = 2 \]

We can construct a feasible descent direction \( p \) such that

\[ A_a p = e_s \iff \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} p = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \iff p = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \]

Observe that

\[ a_1^T p = 0, \quad a_2^T p = 1, \quad c^T p = (2 \quad 1) \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -1 \]

\[ \Rightarrow p \text{ is a feasible descent direction.} \]

\[ \Rightarrow \text{we “move off” constraint #2, but “stay on” constraint #1.} \]

Complications of degeneracy

Unfortunately, the simplified proof using a \( p \) such that

\[ A_a p = e_s \]

will not work when the active constraints are dependent, e.g.,

\[ A_a = \]

i.e., if \( A_a \) defines a degenerate vertex (more than \( n \) active constraints).

In this case, \( A_a p = e_s \) is incompatible and no \( p \) exists.

Example:

minimize \( 2x_1 + x_2 \) subject to the constraints:

- constraint #1: \( x_1 + x_2 \geq 1 \)
- constraint #2: \( x_2 \geq 0 \)
- constraint #3: \( x_1 \geq 0 \)
- constraint #4: \( x_1 + 2x_2 \geq 1 \)

Written in the form \( \min c^T x \) subject to \( Ax \geq b \), we get

\[ c = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \]
At the point \( \bar{x} = (1, 0)^T \),

the active set is \( \mathcal{A} = \{1, 2, 4\} \) with

\[
A_\mathcal{A} = \begin{pmatrix}
1 & 1 \\
0 & 1 \\
1 & 2
\end{pmatrix},
\]

\[
b_\mathcal{A} = \begin{pmatrix}
1 \\
0 \\
1
\end{pmatrix}
\]

Solving the equations \( A_\mathcal{A}^T \lambda_\mathcal{A} = c \) gives a basic solution

\[
\begin{pmatrix}
1 & 0 & 1 \\
1 & 1 & 2
\end{pmatrix} \lambda_\mathcal{A} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \implies \lambda_\mathcal{A} = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} \iff s = 3
\]

The equations \( A_\mathcal{A} p = e_3 \) are

\[
A_\mathcal{A} p = \begin{pmatrix}
1 & 1 & 1 \\
0 & 1 & 2
\end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}
\]

\[
p_1 + p_2 = 0 \\
p_2 = 0 \\
p_1 + 2p_2 = 1
\]

incompatible equations!

Yet \( p = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \) is still a feasible descent direction.

In this example, there is no feasible direction that moves off one active constraint while staying on all the other active constraints.

i.e., \( p = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \) moves off two of the active constraints.

So how did Farkas prove his lemma?

i.e., how did Farkas construct his vector \( p \) such that

\[ c^T p < 0 \quad \text{and} \quad A_\mathcal{A} p \geq 0? \]