Recap

Result (Existence of a vertex)

If $\mathbb{F} = \{x : Ax \geq b\}$ has at least one point and rank($A$) = $n$, then $\mathbb{F}$ 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

$$
\begin{pmatrix}
A_k \\
a_J^T
\end{pmatrix} p_k = \begin{pmatrix} 0 \\ -1 \end{pmatrix}
$$

Step 3: Compute maximum feasible step $\sigma$. Update our point to $x_{k+1} = x_k + \sigma p_k$ and add blocking constraint(s) we bumped into moving along $p_k$.

Repeat until we get to a vertex.

---

Back to Optimality Conditions...
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

$$Ax^* \geq b, \quad c = A^T \lambda^*_a \quad \text{and} \quad \lambda^*_a \geq 0,$$

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 $c = A^T_a \lambda^*_a$ and $\lambda^*_a \geq 0$, are not satisfied at any feasible point.

Result (Farkas’ Lemma)

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

if and only if

(B) $c = A^T_a \lambda^*_a$ for some $\lambda^*_a \geq 0$

Farkas’ Lemma is known as a theorem of the alternative
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

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

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

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

$x \in C_N \implies x = A_x^T \lambda_x$, for some $\lambda_x \geq 0$

$y \in C_N \implies y = A_y^T \lambda_y$, for some $\lambda_y \geq 0$

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

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

$= A_z^T \left((1 - \theta)\lambda_x + \theta \lambda_y\right)$

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

Thus, $z = A_z^T \lambda_z$, for $\lambda_z \geq 0$ and $z \in C_N$. \qed

Test for optimality

\begin{center}
\begin{tabular}{|c|}
\hline
\textbf{does $c$ lie in $C_n(x)$?} \\
\hline
\end{tabular}
\end{center}

\begin{center}
\begin{tabular}{|c|}
\hline
\textbf{i.e.,} \\
\hline
\textbf{can we write $c = A_x^T \lambda$ for $\lambda \geq 0$?} \\
\hline
\end{tabular}
\end{center}
Example:

Consider the dual cone associated with the active-set matrix

\[ A_a = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_1^T \\ a_2^T \end{pmatrix}, \quad \text{i.e.,} \quad a_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad a_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \]

By definition

\[ C_N = \{ y : y = \lambda_1 a_1 + \lambda_2 a_2 \quad \text{with} \quad \lambda_1 \geq 0 \quad \text{and} \quad \lambda_2 \geq 0 \} \]
Consider $y = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 2 \end{pmatrix}$.

Is $y \in C_y$?

$$y = \begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Rightarrow \lambda = \begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \end{pmatrix} \geq 0$$
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

\[
w = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \implies \lambda = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \nless 0
\]
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: $\text{min} \ c^Tx$ subject to $Ax \geq b$, we get:

\[
\begin{align*}
c &= \begin{pmatrix} 2 \\ 1 \end{pmatrix}, & A &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}, & b &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}
\end{align*}
\]

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

with

\[
A_a = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, & b_a &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}
\]

Solving for the Lagrange multipliers gives

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

$\implies x^*$ is optimal.
At the point 
\[ \bar{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{the active set is} \quad \mathcal{A} = \{1, 2\} \]

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

Solving for the Lagrange multipliers gives
\[ A_{\mathcal{A}}^T \lambda_{\mathcal{A}} = c \Rightarrow \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \lambda_{\mathcal{A}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \Rightarrow \lambda_{\mathcal{A}} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \not\geq 0 \]

\[ \Rightarrow \bar{x} \text{ is not optimal.} \]
Proof of Farkas’ Lemma

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 \( \lambda^*_a \geq 0 \)
Proof: (B) $\implies$ (A)

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

Then,

$$c^T p = (A^T \lambda^*)^T p = (\lambda^*)^T (A p \geq 0)$$

$\implies c^T p \geq 0$ for all $p$ with $A p \geq 0$

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

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

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

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

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

Case II: $c \in$ range$(A^T)$, but $c = A^T \lambda^*$ with $\lambda^* \not\geq 0$

In both cases we construct a $p$ such that $c^T p < 0$ and $A p \geq 0$. 
Case I: $c$ does not lie in range($A_a^T$).

c can be decomposed uniquely as

$$c = c_R + c_N$$

with $c_R \in \text{range}(A_a^T)$ 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: $c \in \text{range}(A_a^T)$, but $c = A_a^T \lambda^*_a$ with $\lambda^*_a \ngeq 0$

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

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

$A_a = \square$ with linearly independent rows.
Case II: $c \in \text{range}(A^T)$, but $c = A^T_\lambda \lambda^*_\lambda$ with $\lambda^*_\lambda \not\geq 0$

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

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

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

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

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

$$c^Tp = (A^T_\lambda \lambda^*_\lambda)^Tp = (\lambda^*_\lambda A_p)p = \lambda^*_\lambda (A_p)e_s = [\lambda^*_\lambda]_s < 0$$

Hence we have constructed a direction $p$ such that

$$c^Tp < 0 \quad \text{and} \quad A_p ( = e_s) \geq 0$$

and we are done! $\blacksquare$
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 $A = \{1, 2\}$ with

$$A_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

$$b_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Solving for the Lagrange multipliers gives

$$A_2^T \lambda_a = c \implies \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \lambda_a = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \implies \lambda_a = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$
\[ \lambda_a = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \leftarrow s = 2 \]

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

\[ A_a p = e_2 \Rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} p = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Rightarrow p = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \]

Observe that

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

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

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

Notes
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.,

$$A_a = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

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

$$A_a = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}, \quad b_a = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Solving the equations $A_a^T \lambda_a = c$ gives a basic solution

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} \lambda_a = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \implies \lambda_a = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} \quad \leftarrow s = 3$$

The equations $A_a p = e_3$ are

$$A_a p = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\begin{cases} p_1 + p_2 = 0 \\ p_2 = 0 \\ p_1 + 2p_2 = 1 \end{cases}$$

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

Why is this?

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

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