LECTURE 3: GEOMETRY OF LP

1. Terminologies
2. Background knowledge
3. Graphic method
4. Fundamental theorem of LP
Terminologies

- **Baseline model:**

\[
\begin{align*}
\text{Min } & \quad c^T x \\
\text{(LP) } & \quad \text{s. t. } A x = b \\
& \quad x \geq 0
\end{align*}
\]

- **Feasible domain**

\[P = \{ x \in \mathbb{R}^n \mid A x = b, x \geq 0 \}\]

- **Feasible solution**

\[x \text{ is a \underline{feasible solution} if } x \in P.\]

- **Consistency**

\[\text{When } P \neq \emptyset, \text{ LP is \underline{consistent}.}\]
Terminologies

• Bounded feasible domain:

\[ P \text{ is bounded if} \]
\[ \exists \ M > 0 \text{ such that } \|x\| \leq M, \forall \ x \in P. \]

In this case, we say “LP has bounded feasible domain.”

• Bounded LP:

\[ \text{LP is bounded if} \]
\[ \exists \ M \in \mathbb{R} \text{ such that } c^T x \geq M \ \forall \ x \in P. \]

• Question: LP has a bounded feasible domain.

\[ \downarrow \uparrow \]

LP is bounded.
Terminologies

• Optimal solution:
  \( x^* \) is an optimal solution if
  \[
  x^* \in P \quad \text{and} \quad c^T x^* = \min_{x \in P} c^T x
  \]

• Optimal solution set
  \[
  P^* = \{x^* \mid x^* \text{ is optimal}\}
  \]

• We say
  \( x^* \) solves LP, if \( x^* \in P^* \).
Background knowledge

- Observation 1: each equality constraint in the standard form LP is a “hyperplane” in the solution space.

- What does the equation \( x_1 - 2x_2 = 30 \) represent in the 2-d Euclidean space?

Definition:

For a vector \( a \in \mathbb{R}^n, a \neq 0 \), and a scalar \( \beta \in \mathbb{R} \), define

\[
H = \{ x \in \mathbb{R}^n | a^T x = \beta \} \text{ hyperplane}
\]
Hyperplane

- Geometric representation
Properties of hyperplanes

• Property 1: The normal vector \( \mathbf{a} \) is orthogonal to all vectors in the hyperplane \( H \).

• Proof:

\[
\forall y, z \in H, \quad a^T(y - z) = a^T y - a^T z = \beta - \beta = 0.
\]
Properties of hyperplane

• Property 2: The normal vector is directed toward the upper half space.

• Proof:

For any \( z \in H, w \in H_L^i \),

\[
\begin{align*}
\mathbf{a}^T(w - z) &= \mathbf{a}^T w - \mathbf{a}^T z \\
&< \beta - \beta = 0.
\end{align*}
\]
Properties of feasible solution set

- **Definition:**
  A polyhedral set or polyhedron is a set formed by the intersection of a finite number of a closed half spaces. If it is nonempty and bounded, it is a polytope.

- **Property 3:**
  The feasible domain of a standard form LP
  \[ P = \{ x \in \mathbb{R}^n | Ax = b, x \geq 0 \} \]
  is a polyhedral set.
Properties of optimal solutions

• Property 4:

If $P \neq \emptyset$ and $\exists \beta \in \mathbb{R}$ such that

$$P \subset H_L := \{x \in \mathbb{R}^n \mid -c^T x \leq \beta\},$$

then $\min_{x \in P} c^T x \geq -\beta$

Moreover, if $x^* \in P \cap H$ then $x^* \in P^*$. 
Example

• Give the following LP

Minimize \(-x_1 - 2x_2\)

s. t. \begin{align*}
x_1 + x_2 & \leq 40 \\
2x_1 + x_2 & \leq 60 \\
x_1, x_2, & \geq 0
\end{align*}

• Covert to standard form

Minimize \(-x_1 - 2x_2\)

s. t. \begin{align*}
x_1 + x_2 + x_3 & = 40 \\
2x_1 + x_2 + x_4 & = 60 \\
x_1, x_2, x_3, x_4 & \geq 0
\end{align*}

c = \begin{pmatrix} -1 \\ -2 \\ 0 \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 40 \\ 60 \end{pmatrix}
Since $\min_{x \in P} c^T x \geq -80$
also $-x_1 - 2x_2 = -80$ at $(0, 40)$

Hence $\begin{cases} x_1 = 0 \\ x_2 = 40 \end{cases}$ is an optimal solution.
Graphic Method

Step 1: Draw the feasible domain $P$.
   (If $P = \emptyset$, STOP! No solution.)

Step 2: Use $-c$ as normal vector at each vertex to see
   if $P \in H_L := \{ x \in \mathbb{R}^n | -c^T x \leq \beta \}$ for some
   $\beta \in \mathbb{R}$.

1. If the answer is “YES”, we find an optimal
   solution.

2. If all answers are “NO”, the problem is
   unbounded below.
Pros and Cons

• Advantages:
  - Geometrically simple.

• Disadvantages
  - Algebraically difficult
    How many vertices are there?
    How to identify each vertex?
Any better way?

- Simplex method

A way to *generate and manage the vertices* of the feasible solution set, which is a polyhedral set.
Background knowledge

• Definition: Let \( x^1, x^2, \ldots, x^p \in \mathbb{R}^n, \lambda_1, \lambda_2, \ldots, \lambda_p \in \mathbb{R}, \) and

\[
x = \sum_{i=1}^{p} \lambda_i x^i = \lambda_1 x^1 + \lambda_2 x^2 + \cdots + \lambda_p x^p
\]

we say \( x \) is a linear combination of \( \{x^1, \ldots, x^p\} \).

• If \( \sum_{i=1}^{p} \lambda_i = 1 \), we say \( x \) is an affine combination of \( \{x^1, \ldots, x^p\} \).

• If \( \lambda_i \geq 0 \), we say \( x \) is a conic combination of \( \{x^1, \ldots, x^p\} \).

• If \( \sum_{i=1}^{p} \lambda_i = 1, \lambda_i \geq 0 \), we say \( x \) is a convex combination of \( \{x^1, \ldots, x^p\} \).
Sets generated by different combinations of two points

Affine combination

Convex combination

Conical combination
Affine set, convex set, and cone

- Definition: Let $S$ be a subset of $\mathbb{R}^n$.

  If the affine combination of any two points of $S$ falls in $S$, then $S$ is an **affine set**.

  If the convex combination of any two points of $S$ falls in $S$, then $S$ is a **convex set**.

  If $\lambda x \in S$ for all $x \in S$ and $\lambda \geq 0$, then $S$ is a **cone**.
Example

- Which one is convex? Which one is affine?

\[ H = \{ x \in \mathbb{R}^n | a^T x = \beta \} \]
\[ H_L = \{ x \in \mathbb{R}^n | a^T x \leq \beta \} \]
\[ P = \{ x \in \mathbb{R}^n | Ax = b, x \geq 0 \} \]
Example

What’s the geometric meaning of the feasible domain?

\[ P = \{ x \in \mathbb{R}^n | Ax = b, x \geq 0 \} \]

1. P is a **polyhedral** set.
2. P is a **convex** set.
3. P is the intersection of *m* hyperplanes and the cone of the first orthant.
4. “Ax = b and x ≥ 0” means that the rhs vector b falls in the cone generated by the columns of constraint matrix A.
Example - continue

5. Actually, the set

$$A_c = \{ y \in \mathbb{R}^m | y = Ax, x \in \mathbb{R}^n, x \geq 0 \}$$

is a convex cone generated by the columns of matrix A.
Interior and boundary points

• Given a set, what’s the difference between an interior point and a boundary point?

• Definition: Given a set $S \subseteq \mathbb{R}^n$, a point $x \in S$ is an interior point of $S$, if
  \[ \exists \epsilon > 0 \text{ such that } \{ y \in \mathbb{R}^n \mid \| x - y \| \leq \epsilon \} \subseteq S. \]
  Otherwise, $x$ is a boundary point of $S$.

• We denote that
  \[
  \text{int}(S) = \{ x \text{ is an interior point of } S \} \\
  \text{bdry}(S) = \{ x \text{ is an boundary point of } S \}
  \]
Boundary points of convex sets

- What’s special about boundary points of a convex set?
- Separation Theorem:

\[ S \subseteq \mathbb{R}^n \text{ is convex, then } \forall x \in \text{bdry}(S), \exists \text{ a hyperplane } H \text{ such that } x \in H \text{ and either } S \subseteq H_L \text{ or } S \subseteq H_U. \]
Question

• Can you now see that if an LP (in two or three dimensions) has a finite optimal solution, then one vertex of P is optimal?

• Hint: Consider the supporting hyperplane

\[ H = \{ x \in \mathbb{R}^n | -c^T x = \beta \} \]

• How about higher dimensional case?
  - This leads to the Fundamental Theorem of LP.
Are all boundary points the same?

- Some sits on the shoulders of others, and some don’t.
- Definition: *x* is an extreme point of a convex set *S* if *x* cannot be expressed as a convex combination of other points in *S*. 
Geometrical meaning of extreme points

**Definition:**

Let $P$ be a convex polyhedron and $H$ be a supporting hyperplane of $P$, then $F = P \cap H$ defines a **face** of $P$.

- When $\dim(F) = 0$, it is a **vertex**
- $\dim(F) = 1$, it is an **edge**
- $\dim(F) = \dim(P) - 1$, a **facet**

**Theorem:**

Let $P$ be a convex polyhedron, $x \in P$ is a vertex if and only if $x$ is an extreme point of $P$. 
Representation of extreme points

- For the feasible domain $P$ of an LP, its **vertices are** the **extreme points**. How can we take this advantage to generate and manage all vertices?

$x$ is an extreme point of $P$, then $x$ is of course a feasible solution of

\[
\begin{cases}
    Ax = b \\
    x \geq 0
\end{cases}
\]

But what’s special of being an extreme point? (in terms of feasible solution).
Learning from example

Minimize \( x_1 - 2x_2 \)
subject to \( x_1 + x_2 + x_3 = 40 \)
\( 2x_1 + x_2 + x_4 = 60 \)
\( x_1, x_2, x_3, x_4 \geq 0. \)
What’s special?

- Vertices

\[
v^1 = \begin{pmatrix} 0 \\ 0 \\ 40 \\ 60 \end{pmatrix}, v^2 = \begin{pmatrix} 30 \\ 0 \\ 10 \\ 0 \end{pmatrix}, v^3 = \begin{pmatrix} 20 \\ 20 \\ 0 \\ 0 \end{pmatrix}, v^4 = \begin{pmatrix} 0 \\ 40 \\ 0 \\ 20 \end{pmatrix}.
\]

- Edge

\[
v^5 = \begin{pmatrix} 20 \\ 0 \\ 20 \\ 20 \end{pmatrix} \quad \text{← one zero } x_i
\]

- Interior

\[
v^6 = \begin{pmatrix} 15 \\ 15 \\ 10 \\ 15 \end{pmatrix} \quad \text{← no zero } x_i
\]

\[n = 4, \ m = 2, \ n - m = 2\]
Observations

• Ax = b has $n$ variables in $m$ linear equations.

• When $n > m$, we only need to consider $m$ variables in $m$ equations for solving a system of linear equations.

• An extreme point of P is obtained by setting $n - m$ variables to be zero and solving the remaining $m$ variables in $m$ equations.

• the columns of A corresponding to the non-zero (positive) variables better be linear independent!
Example

- System of equations

\[
\begin{align*}
    x_1 + x_2 + x_3 &= 40 \\
    2x_1 + x_2 + x_4 &= 60 \\
    x_1, x_2, x_3, x_4 &\geq 0.
\end{align*}
\]

- Linear independence of the columns

\[
\begin{pmatrix}
    1 \\ 2
\end{pmatrix} x_1 +
\begin{pmatrix}
    1 \\ 1
\end{pmatrix} x_2 +
\begin{pmatrix}
    1 \\ 0
\end{pmatrix} x_3 +
\begin{pmatrix}
    0 \\ 1
\end{pmatrix} x_4 =
\begin{pmatrix}
    40 \\ 60
\end{pmatrix}
\]
Finding extreme points

• Theorem:
A point $x \in P = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$ is an extreme point of $P$ if and only if the columns of $A$ corresponding to the positive components of $x$ are linearly independent.

• Proof:
Without loss of generality, we may assume that the first $p$ components of $x$ are positive and rest are zero, i.e.,

$$
x = \begin{pmatrix} \bar{x} \\ 0 \end{pmatrix} \text{ where } \bar{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} > 0
$$

also denote the first $p$ columns of $A$ by $\bar{A}$, then

$$
Ax = \bar{A}\bar{x} = b.
$$
Proof - continue

Suppose that the columns of $\bar{A}$ are not linearly independent, then $\exists \bar{w} \neq 0$ such that $\bar{A}\bar{w} = 0$. Notice that for $\epsilon$ is small enough $\bar{x} + \epsilon\bar{w} \geq 0$ and $\bar{A}(\bar{x} + \epsilon\bar{w}) = \bar{A}\bar{x} = b$

Hence

$$y^1 = \left( \begin{array}{c} \bar{x} + \epsilon\bar{w} \\ 0 \end{array} \right) \in P$$

$$y^2 = \left( \begin{array}{c} \bar{x} - \epsilon\bar{w} \\ 0 \end{array} \right) \in P$$

and $x = \frac{1}{2} y^1 + \frac{1}{2} y^2$, i.e. $x$ can not be a vertex (extreme point) of $P$.

Thus, $x$ is an extreme point $\Rightarrow$ columns of $\bar{A}$ are linearly independent.
Proof - continue

Suppose that $x$ is not an extreme point, then

$x = \lambda y^1 + (1 - \lambda)y^2$ for some $y^1, y^2 \in P$, $y^1 \neq y^2$ and $0 < \lambda < 1$.

Since $y^1 \geq 0, y^2 \geq 0$ and $0 < \lambda < 1$.

the last $n - p$ components of $y^1$ must be zero, i.e.

$$y^1 = \begin{pmatrix}
\bar{y}^1 \\
0
\end{pmatrix}$$

Now

$$x - y^1 = \begin{pmatrix}
\bar{x} - \bar{y}^1 \\
0
\end{pmatrix} \neq 0$$

and $A(x - y^1) = Ax - Ay^1 = b - b = 0$

$\Rightarrow$ columns of $A$ are linearly dependent.

Thus, columns of $\bar{A}$ are linearly independent

$\Rightarrow$ $x$ is an extreme point.
Managing extreme points algebraically

- Let $A$ be an $m$ by $n$ matrix with $m \leq n$, we say $A$ has full rank (full row rank) if $A$ has $m$ linearly independent columns.
- In this, we can rearrange

$$
\begin{align*}
x &= \begin{pmatrix} x_B \\ x_N \end{pmatrix} & \leftarrow & \text{basic variables} \\
A &= \begin{pmatrix} B & | & N \end{pmatrix} \\
& \uparrow & \uparrow \\
& \text{basis} & \text{non-basis}
\end{align*}
$$

- Definition: (basic solution and basic feasible solution)

If we set $x_N = 0$ and solve $x_B$ for $Ax = Bx_B = b$, then $x$ is a basic solution (bs).

Furthermore, if $x_B \geq 0$, then $x$ is a basic feasible solution (bfs).
Example of basic and basic feasible solutions

Minimize $x_1 - 2x_2$
subject to $x_1 + x_2 + x_3 = 40$
$2x_1 + x_2 + x_4 = 60$
$x_1, x_2, x_3, x_4 \geq 0.$

$v^1 = \begin{pmatrix} 0 \\ 0 \\ 40 \\ 60 \end{pmatrix}$, $v^2 = \begin{pmatrix} 30 \\ 10 \\ 0 \\ 0 \end{pmatrix}$, $v^3 = \begin{pmatrix} 20 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $v^4 = \begin{pmatrix} 0 \\ 0 \\ 40 \\ 20 \end{pmatrix}$, $v^5 = \begin{pmatrix} 20 \\ 0 \\ 20 \\ 0 \end{pmatrix}$, $v^6 = \begin{pmatrix} 15 \\ 0 \\ 10 \\ 15 \end{pmatrix}$

$v^7 = \begin{pmatrix} 40 \\ 0 \\ 0 \\ -20 \end{pmatrix}$, $v^8 = \begin{pmatrix} 0 \\ 60 \\ -20 \\ 0 \end{pmatrix}$
Further results

• Observation: When A does not have full rank, then either
  (1) Ax = b has no solution and hence \( P = \emptyset \), or
  (2) some constraints are redundant.

  For the second case, after removing the redundant constraints, new A has full rank.

• Corollary: A point \( x \) in \( P \) is an extreme point of \( P \) if and only if \( x \) is a bfs corresponding to some basis \( B \).

• Corollary: The polyhedron \( P \) has only a finite number of extreme points.

  \textbf{Proof:} \# of ways to choose \( m \) linearly independent columns from \( n \) columns
  \[ \leq C(n, m) = \frac{n!}{m!(n-m)!}. \]
Are there many vertices for LP?

• Yes!

\[ C(n, m) = \frac{n!}{m!(n-m)!} \]

• This is **not a small number**, when \( n \) and \( m \) become large. Please try it out by taking \( n = 100 \) and \( m = 50 \).
What do extreme points bring us?

• Observation:
  When \( P = \{ x \in \mathbb{R}^n | Ax = b, x \geq 0 \} \)
  is a nonempty polytope, then any point in \( P \) can be represented
  as a convex combination of the extreme points of \( P \).

Question: Can it be more general?
Extremal direction for unboundedness

- When \( P \) is unbounded, we need a direction leading to infinity.

- Definition:
  - A vector \( d \neq 0 \in \mathbb{R}^n \) is an extremal direction of \( P \), if
    \[
    \{ x \in \mathbb{R}^n \mid x = x^0 + \lambda d, \quad \lambda \geq 0 \} \subset P
    \]
    for all \( x^0 \in P \).

- Observations:
  1. \( P \) is unbounded \( \iff \) \( P \) has an extremal direction.
  2. \( d \neq 0 \) is an extremal direction of \( P \) \( \iff \)
    \[
    Ad = 0 \quad \text{and} \quad d \geq 0
    \]
Resolution theorem

• Theorem:

Let $V = \{v^i \in \mathbb{R}^n | i \in I\}$ be a set of all extreme points of $P$, $I$ is a finite index set, then $\forall x \in P$, we have

$$x = \sum_{i \in I} \lambda_i v^i + d$$

where

$$\sum_{i \in I} \lambda_i = 1, \lambda_i \geq 0 \ \forall i \in I.$$ 

and either $d = 0$ or $d$ is an external direction of $P$.

• We can also write

$$x = \sum_{i \in I} \lambda_i v^i + s \ d, \ \text{for some } s \geq 0.$$
Implications of resolution theorem

• Corollary:
  If P is bounded (a polytope), then any $x$ in P can be expressed as a convex combination of its extreme points.

• Corollary:
  If P is nonempty, then it has at least one extreme point.

Note that $x = \sum_{i \in I} \lambda_i v^i + s d$ implies that the objective value of $x$ is determined by the objective values of extreme points and extremal direction.
Fundamental theorem of LP

• Theorem: For a standard form LP, if its feasible domain $P$ is nonempty, then the optimal objective value of $z = c^T x$ over $P$ is either unbounded below, or it is attained at (at least) an extreme point of $P$.

• Proof:
  By the resolution theorem, there are two cases:
  Case 1:
  
  $P$ has an extremal direction $d$ such that $c^T d < 0$. Hence $P$ is unbounded and $z \to -\infty$ along $d$. 
Proof - continue

Case 2: \( P \) does not have any extremal direction \( d \) such that \( c^T d < 0 \), then \( \forall x \in P \), either
\[
x = \sum_{i \in I} \lambda_i v^i \quad \text{with} \quad \sum_{i \in I} \lambda_i = 1, \ \lambda_i \geq 0, \quad \text{or}
\]
\[
x = \sum_{i \in I} \lambda_i v^i + \bar{d} \quad \text{with} \quad c^T \bar{d} \geq 0.
\]

In both cases,
\[
c^T x = c^T \left[ \sum_{i \in I} \lambda_i v^i \right] (+c^T \bar{d})
\geq \sum_{i \in I} \lambda_i (c^T v^i)
\geq \min_{i \in I} \{c^T v^i\} \left( \sum_{i \in I} \lambda_i \right)
= \min_{i \in I} \{c^T v^i\}
= c^T v^{\text{min}}.
\]

Hence the minimum of \( z \) is attained at one extreme point!