LECTURE 3: GEOMETRY OF LP

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

Terminologies

- Baseline model: Min $\mathbf{c}^T \mathbf{x}$ (LP) s. t. Ax = b
- Feasible domain

$$P = \{ \mathbf{x} \in \mathbf{R}^n | \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \ge \mathbf{0} \}$$

 $\mathbf{x} \ge 0$

Feasible solution

x is a *feasible solution* if $\mathbf{x} \in P$.

Consistency

When $P \neq \phi$, LP is <u>consistent</u>.

Terminologies

Bounded feasible domain:

P is <u>bounded</u> if

 $\exists M > 0 \text{ such that } \|\mathbf{x}\| \leq M, \ \forall \mathbf{x} \in P.$

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

Bounded LP:

LP is <u>bounded</u> if

 $\exists M \in R \text{ such that } \mathbf{c}^T \mathbf{x} \geq M \ \forall \mathbf{x} \in P.$

• Question: LP has a bounded feasible domain.

 $\Downarrow \uparrow$? LP is bounded.

Terminologies

• Optimal solution:

x* is an optimal solution if

$$\mathbf{x}^* \in P$$
 and $\mathbf{c}^T \mathbf{x}^* = \underset{x \in P}{\min} \mathbf{c}^T \mathbf{x}$

Optimal solution set

$$P^* = \{ \mathbf{x}^* \mid \mathbf{x}^* \text{ is optimal} \}$$

• We say

$$\mathbf{x}^*$$
 solves LP, if $\mathbf{x}^* \in P^*$.

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 $\mathbf{a} \in \mathbf{R}^n, \mathbf{a} \neq 0$, and a scalar $\beta \in \mathbf{R}$, define

 $H = \{ \mathbf{x} \in \mathbf{R}^n | \mathbf{a}^T \mathbf{x} = \beta \} \underline{hyperplane}$

Hyperplane

Geometric representation



Properties of hyperplanes

 Property 1: The normal vector a is orthogonal to all vectors in the hyperplane H.



Properties of hyperplane

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

• Proof:

For any $\mathbf{z} \in H$, $\mathbf{w} \in H_L^i$, $\mathbf{a}^T (\mathbf{w} - \mathbf{z}) = \mathbf{a}^T \mathbf{w} - \mathbf{a}^T \mathbf{z}$ $< \beta - \beta = 0.$



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

$$P = \{\mathbf{x} \in \mathbf{R}^n | \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \ge \mathbf{0}\}$$

is a polyhedral set.

Properties of optimal solutions



Example

Give the following LP

Minimize $-x_1 - 2x_2$ s. t. $x_1 + x_2 \le 40$ $\begin{array}{ccc} 2x_1 & + & x_2 & \leq \mathbf{00} \\ x_1, & x_2, & \geq \mathbf{0} \\ \bullet \text{ Covert to standard form } & \mathbf{c} = \begin{pmatrix} -1 \\ -2 \\ 0 \\ \end{array} \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 40 \\ 60 \end{pmatrix}$ Minimize $-x_1 - 2x_2$ s. t. $x_1 + x_2 + x_3 = 40$ $2x_1 + x_2 + x_4 = 60$ $x_1, \qquad x_2, \qquad x_3, \qquad x_4 \ge 0$

Graphic Solution



Graphic Method

- Step 1: Draw the feasible domain P. (If $P = \emptyset$, STOP! No solution.)
- Step 2: Use $-\mathbf{c}$ as normal vector at each vertex to see if $P \in H_L := {\mathbf{x} \in \mathbf{R}^n | -\mathbf{c}^T \mathbf{x} \leq \beta}$ for some $\beta \in \mathbf{R}$.
 - 1. If the answer is "YES", we find an optimal solution.
 - 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 $\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^p \in \mathbf{R}^n$, $\lambda_1, \lambda_2, \dots, \lambda_p \in \mathbf{R}$, and

$$\mathbf{x} = \sum_{i=1}^{p} \lambda_i \mathbf{x}^i = \lambda_1 \mathbf{x}^1 + \lambda_2 \mathbf{x}^2 + \dots + \lambda_p \mathbf{x}^p$$

we say **x** is a linear combination of $\{\mathbf{x}^1, \dots, \mathbf{x}^p\}$.

- If $\sum_{i=1}^{p} \lambda_i = 1$, we say x is an affine combination of $\{x^1, \dots, x^p\}$.
- If $\lambda_i \ge 0$; we say x is a conic combination of $\{x^1, \dots, x^p\}$.
- If $\sum_{i=1}^{p} \lambda_i = 1, \lambda_i \ge 0$, we say x is a convex combination of $\{x^1, \dots, x^p\}$.

Sets generated by different combinations of two points



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 $\lambda \mathbf{x} \in S$ for all $\mathbf{x} \in S$ and $\lambda \ge 0$, then S is a cone.

Example

Which one is convex? Which one is affine?



 $H = \{ \mathbf{x} \in \mathbf{R}^n | \mathbf{a}^T \mathbf{x} = \beta \}$ $H_L = \{ \mathbf{x} \in \mathbf{R}^n | \mathbf{a}^T \mathbf{x} \le \beta \}$

$$\{\mathbf{x} \in \mathbf{R}^n | \mathbf{A}\mathbf{x} = \mathbf{b}\}$$
$$P = \{\mathbf{x} \in \mathbf{R}^n | \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \ge 0\}$$

C A

C A

Example

• What's the geometric meaning of the feasible domain ?

 $P = \{ \mathbf{x} \in \mathbf{R}^n | \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \ge \mathbf{0} \}$

- 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 \ge 0" means that the rhs vector b falls in the cone generated by the columns of constraint matrix A.

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 | \mathbf{A}_2 | \cdots | \mathbf{A}_n \end{pmatrix}$$
$$A_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \qquad \mathbf{A}_k = \begin{pmatrix} \mathbf{A}_1 | \mathbf{A}_2 | \cdots | \mathbf{A}_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ w_n \end{pmatrix} = \sum_{j=1}^n x_j \mathbf{A}_j \in \mathbf{R}^{m_k}$$

Example - continue

5. Actually, the set
$$\mathbf{A}_c = \{\mathbf{y} \in \mathbf{R}^m | \mathbf{y} = \mathbf{A}\mathbf{x}, \mathbf{x} \in \mathbf{R}^n, \mathbf{x} \ge 0\}$$

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 ⊂ Rⁿ, a point
 x ∈ S is an interior point of S, if
 ∃ ε > 0 such that the ball B = {y ∈ Rⁿ | ||x − y|| ≤ ε} ⊂ S.

Otherwise, **x** is a boundary point of S.

We denote that

 $int(S) = \{ \mathbf{x} \text{ is an interior point of } S \}$ $bdry(S) = \{ \mathbf{x} \text{ is an boundary point of } S \}$

Boundary points of convex sets

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

 $S \subset \mathbf{R}^n$ is convex, then $\forall \mathbf{x} \in bdry(S), \exists$ a hyperplane H such that $\mathbf{x} \in H$ and either $S \subseteq H_L$ 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

$$H = \{\mathbf{x} \in \mathbf{R}^n | - \mathbf{c}^T \mathbf{x} = \beta\}$$

- 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 $F = P \bigcap H$ defines a <u>face</u> 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, $\mathbf{x} \in P$ is a vertex if and only if \mathbf{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?

 \mathbf{x} is an extreme point of P, then \mathbf{x} is of course a feasible solution of

$$\begin{cases} \mathbf{A}\mathbf{x} = \mathbf{b} \\ \mathbf{x} \ge \mathbf{0} \end{cases}$$

But what's special of being an extreme point? (in terms of feasible solution).

Learning from example



What's special?

• Vertices

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

• Edge Interior

$$v^{5} = \begin{pmatrix} 20 \\ 0 \\ 20 \\ 20 \end{pmatrix} \leftarrow \text{ one zero } x_{i} \qquad v^{6} = \begin{pmatrix} 15 \\ 15 \\ 10 \\ 15 \end{pmatrix} \leftarrow \text{ no zero } x_{i}$$

 $n = 4, \ m = 2, \ n - m = 2$

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

$$\begin{cases} x_1 + x_2 + x_3 = 40\\ 2x_1 + x_2 + x_3 = 60\\ x_1, x_2, x_3, x_4 \ge 0. \end{cases}$$

• 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 $\mathbf{x} \in P = {\mathbf{x} \in \mathbf{R}^n | \mathbf{A}\mathbf{x} = b, \mathbf{x} \ge 0}$ 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.,

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

also denote the first p columns of \mathbf{A} by $\mathbf{\bar{A}}$, then $\mathbf{A}\mathbf{x} = \mathbf{\bar{A}}\mathbf{\bar{x}} = \mathbf{b}$.

Proof - continue

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

Hence

$$\mathbf{y}^{1} = \begin{pmatrix} \bar{\mathbf{x}} + \epsilon \bar{\mathbf{w}} \\ \mathbf{0} \end{pmatrix} \in P$$
$$\mathbf{y}^{2} = \begin{pmatrix} \bar{\mathbf{x}} - \epsilon \bar{\mathbf{w}} \\ \mathbf{0} \end{pmatrix} \in P$$

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

Thus, **x** is an exterme point \Rightarrow columns of $\overline{\mathbf{A}}$ are linearly independent.

Proof - continue

Suppose that x is not an extreme point, then $\mathbf{x} = \lambda \mathbf{y}^1 + (1 - \lambda)\mathbf{y}^2$ for some $\mathbf{y}^1, \mathbf{y}^2 \in P, \ \mathbf{y}^1 \neq \mathbf{y}^2$ and $0 < \lambda < 1$, Since $\mathbf{y}^1 \ge 0, \mathbf{y}^2 \ge 0$ and $0 < \lambda < 1$. the last n - p components of \mathbf{y}^1 must be zero, *i.e.*

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

Now

$$x-y^1=\left(\begin{array}{c}\bar{x}-\bar{y}^1\\0\end{array}\right)\neq 0$$

and $A(x - y^1) = Ax - Ay^1 = b - b = 0$ \Rightarrow columns of A are linearly dependent. Thus, columns of \overline{A} are linearly independent \Rightarrow x is an extreme point.

Managing extreme points algebraically

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

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix} \leftarrow \text{basic variables} & \mathbf{A} = \begin{pmatrix} \mathbf{B} & | & \mathbf{N} \end{pmatrix} \end{pmatrix}$$

 $\leftarrow \text{ non-basic variables} & \uparrow & \uparrow \end{pmatrix}$
 $\leftarrow \text{ non-basic variables} & \text{ basis non-basis}$

Definition: (basic solution and basic feasible solution)

If we set $\mathbf{x}_N = 0$ and solve \mathbf{x}_B for $\mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{x}_B = \mathbf{b}$ then \mathbf{x} is a <u>basic solution</u> (bs). Furthermore, if $\mathbf{x}_B \ge 0$, then \mathbf{x} is a basic feasible solution (bfs).

Example of basic and basic feasible solutions



$$v^{1} = \begin{pmatrix} 0 \\ 0 \\ 40 \\ 60 \end{pmatrix}, v^{2} = \begin{pmatrix} 30 \\ 0 \\ 10 \\ 0 \end{pmatrix}, v^{3} = \begin{pmatrix} 20 \\ 20 \\ 0 \\ 0 \end{pmatrix}, v^{4} = \begin{pmatrix} 0 \\ 40 \\ 0 \\ 20 \end{pmatrix} \text{bfs}$$
$$v^{7} = \begin{pmatrix} 40 \\ 0 \\ 0 \\ -20 \\ -20 \end{pmatrix}, v^{8} = \begin{pmatrix} 0 \\ 60 \\ -20 \\ 0 \end{pmatrix} \text{bs}$$
$$v^{5} = \begin{pmatrix} 20 \\ 0 \\ 20 \\ 20 \end{pmatrix}, v^{6} = \begin{pmatrix} 15 \\ 15 \\ 10 \\ 15 \end{pmatrix}$$

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. <u>Proof:</u> # of ways to choose m linearly independent columns from n columns ≤ C(n,m) = n!/(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 $P = \{\mathbf{x} \in \mathbf{R}^n | \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \ge 0\}$ is a nonempty polytope, then

any point in P can be represented x as a convex combination of the v5 v3 V v^4

 v^1

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 (≠ 0) ∈ Rⁿ is an
 extremal direction of P, if
 {x ∈ Rⁿ | x = x⁰ + λd, λ ≥ 0} ⊂ P.

for all $\mathbf{x}^0 \in P_{\dots}$



- Observations: (1) P is unbounded \Leftrightarrow P has an extremal direction.
 - (2) $\mathbf{d} \ (\neq \mathbf{0})$ is an extremal direction of $P \Leftrightarrow$ $\mathbf{A}\mathbf{d} = \mathbf{0}$ and $\mathbf{d} \ge \mathbf{0}$

Resolution theorem

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

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

where

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

and either d = 0 or

d is an extermal direction of P.

• We can also write

$$\mathbf{x} = \sum_{i \in I} \lambda_i v^i + \mathbf{s} \, \mathsf{d}$$
, 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 leas one extreme point.

Note that $\mathbf{x} = \sum_{i \in I} \lambda_i v^i + s d$ 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^Tx 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 \mathbf{d} such that $\mathbf{c}^T \mathbf{d} < \mathbf{0}$. Hence P is unbounded and $\mathbf{z} \to -\infty$ along \mathbf{d} .

Proof - continue

- Case 2: *P* does not have any extremal direction d such that $\mathbf{c}^T \mathbf{d} < 0$, then $\forall \mathbf{x} \in P$, either $\mathbf{x} = \sum_{i \in I} \lambda_i v^i$ with $\sum_{i \in I} \lambda_i = 1$, $\lambda_i \ge 0$, or $\mathbf{x} = \sum_{i \in I} \lambda_i v^i + \overline{\mathbf{d}}$ with $\mathbf{c}^T \overline{\mathbf{d}} \ge \mathbf{0}$.
- In both cases, $\mathbf{c}^T \mathbf{x} = \mathbf{c}^T [\sum_{i \in I} \lambda_i v^i] (+ \mathbf{c}^T \mathbf{d})$ $\geq \sum_{i \in I} \lambda_i (\mathbf{c}^T v^i)$ $\geq \min_{i \in I} {\mathbf{c}^T v^i} (\sum_{i \in I} \lambda_i)$ $= \min_{i \in I} {\mathbf{c}^T v^i}$ $= \mathbf{c}^T v^{\min}.$

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