Lecture 5: The Geometry of Linear Optimization

Zizhuo Wang

Institute of Data and Decision Analytics (iDDA)
Chinese University of Hong Kong, Shenzhen

Sep 19, 2018

Recap

Up to now, we have learned how to

- Formulate linear optimization problems
- Transform an LP into a standard form
- Use MATLAB to solve linear optimizations

Why linear optimization?

- Versatile: Can model many real problems
- Easy to solve: Commercial software can solve LPs with tens of thousands of variables very easily. It is the easiest one among all optimization problems
- Fundamental: The LP theories lay the foundation for most optimization theories

Preview of the Coming Weeks

- Some basic structural properties for linear optimization
- Simplex method for solving linear optimization
- Linear program duality theory (sensitivity analysis)
- A brief introduction of interior point method

Starting Point: Graphical Solutions to LP

It is very helpful to study a small LP from a graphical point of view.

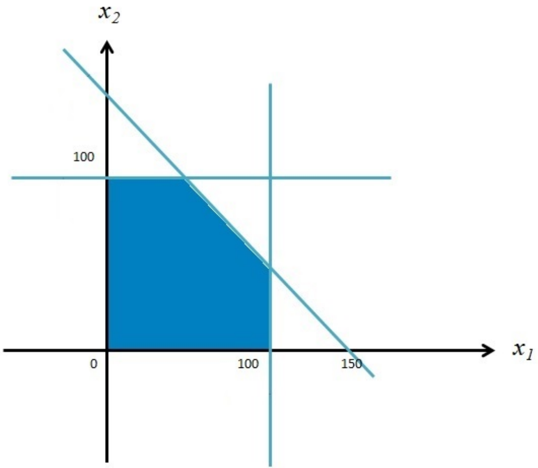
Recall the production problem:

$$\begin{array}{rlll} \text{maximize} & x_1 & +2x_2 & \\ \text{subject to} & x_1 & & \leq 100 \\ & & 2x_2 & \leq 200 \\ & x_1 & +x_2 & \leq 150 \\ & x_1, & x_2 & \geq 0 \end{array}$$

How can we solve this from a graph?

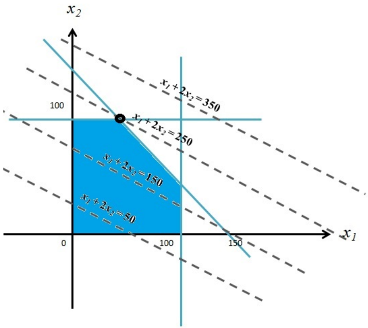
Solve LP from Graph

We first draw the feasible region.



To Maximize $x_1 + 2x_2$...

Then we draw the function $x_1 + 2x_2 = c$ for different values of c.



- The optimal solution is the highest one among these lines that touch the feasible region
- The coordinates: $(50, 100)$. Objective value: 250

Some Observations

- The feasible region of LP is a polygon
- The optimal solution tends to be at a corner of the feasible region
- Some constraints are *active* at the optimal solution ($x_2 \leq 100$, $x_1 + x_2 \leq 150$), some are not ($x_1 < 100$).

Keep these in mind as we will formalize these observations when we study the algorithms for solving LPs.

The High-Level Picture

As we see, the optimal solution to LP seems to be at the corner of the feasible sets (we will prove it). Then questions arise:

- How to find a corner point (given only the constraints)?
- How can we search for the best corner?

It is easy to see in the toy example, but when there are thousands or more variables, we need to use algebra to do it. And eventually we need an *algorithm* that is able to

- Guarantee to find the optimal solution
- Run within a certain (reasonable) amount of time

Some Definitions: Polyhedron

Definition

A *polyhedron* is a set that can be described in the form

$$\{\mathbf{x} \in \mathbb{R}^n | A\mathbf{x} \geq \mathbf{b}\}$$

where A is an $m \times n$ matrix and $\mathbf{b} \in \mathbb{R}^m$.

- Recall that in the standard form of LP, the feasible set is

$$A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0$$

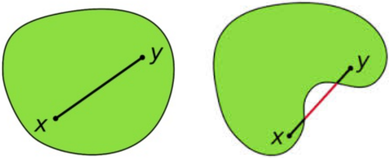
Is this a polyhedron? Why?

- Yes, we can write it as $A\mathbf{x} \geq \mathbf{b}$, $A\mathbf{x} \leq \mathbf{b}$, $\mathbf{I}\mathbf{x} \geq 0$ where $\mathbf{I}$ is the identity matrix.

Convex Sets and Convex Combinations

Definition

A set $S \subseteq \mathbb{R}^n$ is *convex* if for any $\mathbf{x}$, $\mathbf{y} \in S$, and any $\lambda \in [0, 1]$, $\lambda\mathbf{x} + (1 - \lambda)\mathbf{y} \in S$.



(a) Convex set (b) Non-convex set

Definition

For any $\mathbf{x}_1, \ldots, \mathbf{x}_n$ and $\lambda_1, \ldots, \lambda_n \geq 0$ satisfying $\lambda_1 + \cdots + \lambda_n = 1$, we call $\sum_{i=1}^{n} \lambda_i \mathbf{x}_i$ a convex combination of $\mathbf{x}_1, \ldots, \mathbf{x}_n$.

Extreme Points

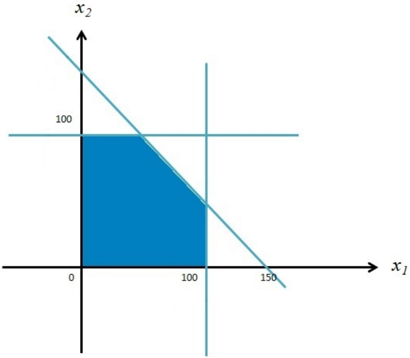
Remember we noticed that in an LP, the optimal solution tends to be on one of the corners of the feasible region. We first formalize this notion:

Definition

Let P be a polyhedron. A point $\mathbf{x} \in P$ is said to be an extreme point of P if we cannot find two vectors $\mathbf{y}, \mathbf{z} \in P$, both different from $\mathbf{x}$, and a scalar $\lambda \in [0, 1]$, such that $\mathbf{x} = \lambda \mathbf{y} + (1 - \lambda)\mathbf{z}$

- That is, $\mathbf{x}$ cannot be represented as a convex combination of other points in P.
- We sometimes call the extreme point the *vertex*, or *corner* of the polyhedron

Example



How many extreme points are there in this feasible region?

- Answer: 5

Find Extreme Points of an LP

In the following, we consider LP in its standard form:

$$\begin{array}{rl} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array}$$

where $\mathbf{x} \in \mathbb{R}^n$, A is an $m \times n$ matrix $(m < n)$ and $\mathbf{b} \in \mathbb{R}^m$.

Assumption

A has linearly independent rows (or equivalently A has full rank m)

What if it is not satisfied?

- Then it means that either there is a redundant constraint (in which case one can remove it), or the constraints are not self-consistent (in which case there is no feasible solution)

Find Extreme Points of an LP

Now we study the extreme points of an LP in its algebraic form.

Definition

We call $\mathbf{x}$ a *basic solution* of the LP if and only if

1. $A\mathbf{x} = \mathbf{b}$
2. There exist indices $B(1), ..., B(m)$ such that the columns $A_{B(1)}, ..., A_{B(m)}$ are linearly independent, and $x_i = 0$ for $i \neq B(1), ..., B(m)$

Basic solution of an LP only depends on its constraints, it has nothing to do with the objective function.

Finding a Basic Solution

Procedures to find a basic solution:

1. Choose any m independent columns of A: $A_{B(1)}, \ldots, A_{B(m)}$
2. Let $x_i = 0$ for all $i \neq B(1), \ldots, B(m)$
3. Solve the equation $A\mathbf{x} = \mathbf{b}$ for the remaining $x_{B(1)}, \ldots, x_{B(m)}$.

Since $A_{B(1)}, \ldots, A_{B(m)}$ are linearly independent, the last step must produce a unique solution.

Basic Solutions of LP

We write

$$A_B = \begin{bmatrix} | & | & & | \\ A_{B(1)} & A_{B(2)} & \cdots & A_{B(m)} \\ | & | & & | \end{bmatrix}, \quad \mathbf{x}_B = \begin{bmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{bmatrix}$$

Remember that $A_{B(i)}$'s are linearly independent, therefore A_B is invertible and $\mathbf{x}_B = A_B^{-1}\mathbf{b}$

We call $B = \{B(1), ..., B(m)\}$ the basic indices for this basic solution, $A_{B(1)}, ..., A_{B(m)}$ the basic columns, A_B the basis matrix and $x_{B(1)}, ..., x_{B(m)}$ the basic variables

- We call the remaining indices the non-basic indices, the remaining columns of A the non-basic columns and the remaining variables the non-basic variables.

Quiz

How many non-zeros could one have in a basic solution (assuming there are m constraints)?

- No more than m
- Could be anything between 0 to m, but typically it is m

How many basic solutions can one have for a linear program with m constraints and n variables?

- At most $C(n, m) = \frac{n!}{m!(n-m)!}$ (Combination number)
- Therefore for a finite number of linear constraints, there can only be a finite number of basic solutions

Definition

If a basic solution $\mathbf{x}$ also satisfies that $\mathbf{x} \geq 0$, then we call it a *basic feasible solution* (BFS).

To find a BFS

- First find a basic solution $\mathbf{x}$
- Check if $\mathbf{x} \geq 0$

Theorem

For the standard LP polyhedron $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$, *the followings are equivalent:*

1. $\mathbf{x}$ *is an extreme point*
2. $\mathbf{x}$ *is a basic feasible solution*

Example

Recall the production problem:

$$
\begin{array}{rrrl}
\text{maximize} & x_1 & +2x_2 & \\
\text{subject to} & x_1 & & \leq 100 \\
& & 2x_2 & \leq 200 \\
& x_1 & +x_2 & \leq 150 \\
& x_1, & x_2 & \geq 0
\end{array}
$$

The standard form:

$$
\begin{array}{rrrrrrl}
\text{minimize} & -x_1 & -2x_2 & & & & \\
\text{subject to} & x_1 & & +s_1 & & & = 100 \\
& & 2x_2 & & +s_2 & & = 200 \\
& x_1 & +x_2 & & & +s_3 & = 150 \\
& x_1, & x_2, & s_1, & s_2, & s_3 & \geq 0
\end{array}
$$

Example Continued

We can write the feasible set by $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$. where

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 100 \\ 200 \\ 150 \end{bmatrix}$$

Choose three independent columns of A, e.g., the first three (i.e., $B = \{1, 2, 3\}$), we get the corresponding basic solution is

$$\mathbf{x}_B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 100 \\ 200 \\ 150 \end{bmatrix} = \begin{bmatrix} 50 \\ 100 \\ 50 \end{bmatrix}$$

That is $x_1 = 50, x_2 = 100, s_1 = 50$. Therefore $(50, 100, 50, 0, 0)$ is a basic feasible solution. One can find other basic feasible solutions by choosing other sets of columns.

Example Continued

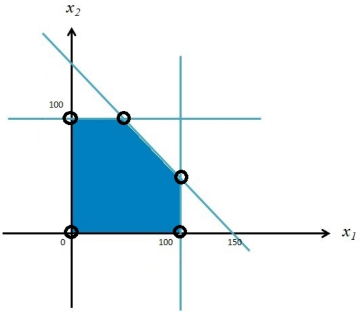
We can list the basic (feasible) solutions

Indices	$\{1,2,3\}$	$\{1,2,4\}$	$\{1,2,5\}$	$\{1,3,4\}$
Solution	(50, 100, 50, 0, 0)	(100, 50, 0, 100, 0)	(100, 100, 0, 0, -50)	(150, 0, -50, 200, 0)
Status	BFS	BFS	Basic but not feasible	Basic but not feasible
Indices	$\{1,4,5\}$	$\{2,3,4\}$	$\{2,3,5\}$	$\{3,4,5\}$
Solution	(100, 0, 0, 200, 50)	(0, 150, 100, -100, 0)	(0, 100, 100, 0, 50)	(0, 0, 100, 200, 150)
Status	BFS	Basic but not feasible	BFS	BFS

The other two choices $\{1,3,5\}$ and $\{2,4,5\}$ lead to dependent basic columns (therefore no basic solutions can be obtained)

Verify..

They indeed correspond to all the corners of the feasible sets.



Proof Sketch

We show how to prove 1 $\Rightarrow$ 2 (the other part can be found in Chapter 2 of Bertsimas's book). First, if $\mathbf{x}$ is an extreme point, then it must be a feasible solution. In the following, we prove by contradiction that $\mathbf{x}$ must be a basic solution as well.

Let $B = \{B(1), ..., B(k)\}$ be the set of indices such that $x_i > 0$. If $\mathbf{x}$ is not a basic solution, then it must be that (why?)

- Columns $A_{B(1)}, ..., A_{B(k)}$ are linearly dependent.

Say $\alpha_1 A_{B(1)} + \cdots + \alpha_k A_{B(k)} = 0$, where $(\alpha_1, ..., \alpha_k) \neq 0$.

Now since $x_i > 0$, $\forall i \in B$, we can choose a small ϵ such that both

$$\bar{\mathbf{x}} = \begin{cases} x_i + \epsilon\alpha_j & i = B(j) \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \tilde{\mathbf{x}} = \begin{cases} x_i - \epsilon\alpha_j & i = B(j) \\ 0 & \text{otherwise} \end{cases}$$

are still feasible solutions.

Then we have $\mathbf{x} = 0.5\bar{\mathbf{x}} + 0.5\tilde{\mathbf{x}}$, which contradicts with that $\mathbf{x}$ is an extreme point. $\square$

Basic Feasible Solutions

Theorem

(LP fundamental theorem) Given a linear optimization problem where A has full row rank m

1. *If there is a feasible solution, there is a basic feasible solution;*
2. *If there is an optimal solution, there is an optimal solution that is a basic feasible solution.*

Corollary

In order to find an optimal solution, we only need to look among basic feasible solutions.

Corollary

If an LP with m constraints (in the standard form) has an optimal solution, then there must be an optimal solution such that there is no more than m positive entries.

Proof Sketch

We first prove part (i). Assume there is a feasible solution $\mathbf{x}$. Let B denote the index set for which $\mathbf{x}$ is positive, and let k be the size of B. Then we show that if $\mathbf{x}$ is not a basic solution, we can always find a solution that has an index set with size less than k. By repeatedly doing this we must be able to find a basic feasible solution.

To show that, if $\mathbf{x}$ is not a basic solution, then it must be that A_i, $i \in B$ are linearly dependent. Suppose $\sum_{i \in B} \alpha_i A_i = 0$ where at least one α_i is positive (we can always achieve that). And we define a vector $\boldsymbol{\alpha}$ by

$$\boldsymbol{\alpha} = \begin{cases} \alpha_i & i \in B \\ 0 & \text{otherwise} \end{cases}.$$

Then we consider $\mathbf{x} - \epsilon\boldsymbol{\alpha}$. There must exist an ϵ such that it is still feasible and that there is one more 0 entry in $\mathbf{x} - \epsilon\boldsymbol{\alpha}$. Thus part (i) is proved.

Proof Sketch Continued

Now we consider part (ii). We continue to use the arguments for part (i). Let $\mathbf{x}$ be an optimal solution with k positive entries. If $\mathbf{x}$ is not a basic solution, then we show that we can always find an optimal solution with fewer than k positive entries (or reach a contradiction). By repeatedly doing this, we must be able to find an optimal basic feasible solution.

We continue to define the same $\boldsymbol{\alpha}$. We claim that it must be that $\mathbf{c}^T\boldsymbol{\alpha} = 0$, otherwise, one can construct a better solution than $\mathbf{x}$ by considering $\mathbf{x} + \epsilon\boldsymbol{\alpha}$ and $\mathbf{x} - \epsilon\boldsymbol{\alpha}$ for some small α, which contradicts with the optimality of $\mathbf{x}$.

If $\mathbf{c}^T\boldsymbol{\alpha} = 0$, then $\mathbf{x} - \epsilon\boldsymbol{\alpha}$ are still optimal solutions (why?). And by previous arguments, we can find one optimal solution with fewer positive entries. $\square$

Basic Feasible Solutions

Now we know that we only need to search among basic feasible solutions for the optimal solution.

How to search among the basic feasible solutions?

- One may suggest to list all the basic feasible solutions and compare their objective values. However, there are too many of them.
- For a linear optimization with m constraints and n variables, how many basic feasible solutions it may have?
- $C(n, m)$.. If $n = 1000$, $m = 100$, then the value is 10^{143}..

Simplex Method

Therefore we need a smarter way to find the optimal solution.

- Simplex method

The simplex method proceeds from one BFS (a corner point of the feasible region) to a neighboring one, in such a way as to continuously improve the value of the objective function until reaching optimality.

- We need to define what it means by *adjacent* or *neighboring* solution
- We need to design efficient way to find (and move to) the neighboring BFS (e.g., we should try to avoid taking matrix inversions every time)
- We need to design a valid stopping criterion

We will talk about these steps in the following lectures