Lecture 3: Linear Optimization

Zizhuo Wang

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

Sep 12, 2018

Recap: Introduction to Optimization

Three main components in optimization problems

- Decision
- Objective
- Constraints

General form of optimization problem:

$$
\begin{aligned}
\text{minimize}_x \quad & f(x) \\
\text{subject to} \quad & g_i(x) \le 0, \quad \forall i = 1, ..., s \\
& h_j(x) = 0, \quad \forall j = 1, ..., t
\end{aligned}
$$

- We have shown some examples of formulating a problem into optimization problem

Recap: Classifications

- Constrained vs Unconstrained
- Linear vs Nonlinear
- Continuous vs Discrete

Starting from today, we are going to study the most fundamental class of optimization problem — linear optimization.

Today: Linear Optimization

- Definition
- Standard Form
- Examples
- Transformations to LP

Definition

A *linear optimization* problem, or a *linear program* (LP), is an optimization problem in which the objective function and all constraint functions are linear (in the decision variables).

Example (the production planning problem):

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

- Linear optimization belongs to continuous optimization

General Formulation

A linear optimization problem can be generally written as

$$
\begin{array}{rll}
\text{minimize/maximize}_{\mathbf{x}} & \mathbf{c}^T\mathbf{x} & \\
\text{subject to} & \mathbf{a}_i^T\mathbf{x} \geq b_i & \forall i \in M_1 \\
& \mathbf{a}_i^T\mathbf{x} \leq b_i & \forall i \in M_2 \\
& \mathbf{a}_i^T\mathbf{x} = b_i & \forall i \in M_3 \\
& x_i \geq 0 & \forall i \in N_1 \\
& x_i \leq 0 & \forall i \in N_2 \\
& x_i \text{ free} & \forall i \in N_3
\end{array}
$$

where M_1, M_2, M_3 are subsets of $\{1, ..., m\}$, N_1, N_2, N_3 are subsets of $\{1, ..., n\}$.

Compact Formulation

We can write LPs in a more compact way:

$$
\begin{aligned}
\text{minimize/maximize}_{\mathbf{x}} \quad & \mathbf{c}^T\mathbf{x} \\
\text{subject to} \quad & A_1\mathbf{x} \geq \mathbf{b}_1 \\
& A_2\mathbf{x} \leq \mathbf{b}_2 \\
& A_3\mathbf{x} = \mathbf{b}_3 \\
& x_i \geq 0 \qquad \forall i \in N_1 \\
& x_i \leq 0 \qquad \forall i \in N_2 \\
& x_i \text{ free} \qquad \forall i \in N_3
\end{aligned}
$$

Here A_1, A_2 and A_3 are matrices (with dimensions $m_1 \times n$, $m_2 \times n$ and $m_3 \times n$). $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ are vectors (with dimensions $m_1 \times 1$, $m_2 \times 1$ and $m_3 \times 1$). $\mathbf{x}$ is an n dimensional column vector.

Standard Form of LP

In order to study it more systematically, we want to have a standard (and even more compact) form of LP.

An LP is said to be of *standard form* if it is of the form:

$$\begin{array}{rl} \text{minimize}_{\mathbf{x}} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array} \tag{1}$$

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

In fact, any LP can be written in the standard form, using some "tricks".

Remark: The definition of "standard form" may differ from book to book. We use (1) as the standard form in this course (consistent with Bertsimas and Tsitsiklis's book).

Transform to Standard Form

If the objective was maximization

- Use $-\mathbf{c}$ instead of $\mathbf{c}$ and change it to minimization

Eliminating inequality constraints $A\mathbf{x} \leq \mathbf{b}$ or $A\mathbf{x} \geq \mathbf{b}$

- Write it as $A\mathbf{x} + \mathbf{s} = \mathbf{b}, \mathbf{s} \geq 0$, or $A\mathbf{x} - \mathbf{s} = \mathbf{b}, \mathbf{s} \geq 0$
- We call $\mathbf{s}$ the slack variables

If one has $x_i \leq 0$

- Define $y_i = -x_i$

Eliminating "free" variables x_i (no constraint on x_i)

- Define $x_i = x_i^+ - x_i^-$, with $x_i^+ \geq 0$, $x_i^- \geq 0$

Example

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

Standard form

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

Support Vector Machine Problem

$$
\begin{array}{rll}
\text{minimize}_{\mathbf{a},b,\delta,\sigma} & \sum_i \delta_i + \sum_j \sigma_j & \\
\text{subject to} & \mathbf{x}_i^T \mathbf{a} + b + \delta_i \geq 1 & \forall i \\
& \mathbf{y}_j^T \mathbf{a} + b - \sigma_j \leq -1 & \forall j \\
& \delta_i \geq 0 & \forall i \\
& \sigma_j \geq 0 & \forall j
\end{array}
$$

Define $\mathbf{a} = \mathbf{a}^+ - \mathbf{a}^-$, $b = b^+ - b^-$, with $\mathbf{a}^+, \mathbf{a}^-, b^+, b^- \geq 0$.

And add slacks to inequality constraints

Standard Form

$$
\begin{array}{rrcl}
\text{minimize} & \sum_i \delta_i + \sum_j \sigma_j & & \\
\text{subject to} & \mathbf{x}_i^T \mathbf{a}^+ - \mathbf{x}_i^T \mathbf{a}^- + b^+ - b^- + \delta_i - s_i = \ 1 & \forall i \\
& \mathbf{y}_j^T \mathbf{a}^+ - \mathbf{y}_j^T \mathbf{a}^- + b^+ - b^- - \sigma_j + t_j = -1 & \forall j \\
& \mathbf{a}^+, \mathbf{a}^-, b^+, b^- \geq 0 & \\
& \delta_i, s_i \geq 0 & \forall i \\
& \sigma_j, t_j \geq 0 & \forall j
\end{array}
$$

Standard Form

Standard form is mainly used for analysis purposes. We don't need to write a problem in standard form unless necessary. Usually just write in a way that is easy to understand.

However, being able to transform an LP into the standard form is an important skill. It is also sometimes useful when we want to use softwares to solve LP.

In the remainder of this class, we are going to show more examples of LPs, as well as examples on how to transform a (non-LP) problem into LP.

Staffing Problem

A hospital wants to make a weekly night shift schedule for its nurses

- The demand for the night shift on day j is d_j , for $j = 1, ..., 7$
- Every nurse works 5 days in a row
- We want to minimize the total number of nurses used while meeting all demand
- Ignore the integrality constraints for now (i.e., we allow "half" nurse if necessary)

Staffing Problem

What is your choice of decision variable?

How about x_i to be the number of nurses on day i?

- Can't capture the constraints that nurses have to work 5 days in a row

A better way is to define x_i to be the number of nurses starting at day i

Our objective will be

$$\text{minimize} \quad x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7$$

Staffing Problem

The LP formulation is:

$$
\begin{array}{lccccccccc}
\min & x_1 & +x_2 & +x_3 & +x_4 & +x_5 & +x_6 & +x_7 & & \\
\text{s.t.} & x_1 & & & +x_4 & +x_5 & +x_6 & +x_7 & \geq & d_1 \\
 & x_1 & +x_2 & & & +x_5 & +x_6 & +x_7 & \geq & d_2 \\
 & x_1 & +x_2 & +x_3 & & & +x_6 & +x_7 & \geq & d_3 \\
 & x_1 & +x_2 & +x_3 & +x_4 & & & +x_7 & \geq & d_4 \\
 & x_1 & +x_2 & +x_3 & +x_4 & +x_5 & & & \geq & d_5 \\
 & & x_2 & +x_3 & +x_4 & +x_5 & +x_6 & & \geq & d_6 \\
 & & & x_3 & +x_4 & +x_5 & +x_6 & +x_7 & \geq & d_7 \\
 & x_1, & x_2, & x_3, & x_4, & x_5, & x_6, & x_7 & \geq & 0
\end{array}
$$

Matrix Representation

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} \quad \mathbf{c} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad A = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad \mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \\ d_7 \end{bmatrix}$$

- The objective $\mathbf{c}^T\mathbf{x} = x_1 + x_2 + \cdots + x_7$.
- The constraint is $A\mathbf{x} \geq \mathbf{d}$, $\mathbf{x} \geq 0$

Air Traffic Control Problem

An air traffic controller needs to control the landing time of n aircrafts

- Flights must land in the order $1, \ldots, n$
- Flight j must land in time interval $[a_j, b_j]$
- The objective is to maximize the minimum *separation time*, which is the interval between two landings

An Optimization Formulation

Decision variable

- Let t_j be the landing time of flight j

Optimization problem:

$$\begin{array}{lll} \max & \min_{j=1,\ldots,n-1}\{t_{j+1}-t_j\} & \\ \text{s.t.} & a_j \le t_j \le b_j, & j=1,\ldots,n \\ & t_j \le t_{j+1}, & j=1,\ldots,n-1 \end{array}$$

The objective function is not a linear function. We call it a maximin objective.

LP Formulation

Define

$$\Delta = \min_{j=1,\ldots,n-1} \{t_{j+1} - t_j\}$$

Therefore, $t_{j+1} - t_j \geq \Delta, \forall j$.

Write an LP:

$$\begin{array}{ll} \text{maximize} & \Delta \\ \text{subject to} & t_{j+1} - t_j - \Delta \geq 0, \quad j = 1, \ldots, n-1 \\ & a_j \leq t_j \leq b_j, \quad j = 1, \ldots, n \\ & t_j \leq t_{j+1}, \quad j = 1, \ldots, n-1 \end{array}$$

At optimal, Δ must equal the minimal separation (since we try to maximize Δ).

Minimax Objective

Similar to the previous example, sometimes we are interested in a minimax objective:

$$\begin{aligned} \text{minimize}_{\mathbf{x}} \quad & \max_{i=1,\ldots,n}\{\mathbf{c}_i^T\mathbf{x} + d_i\} \\ \text{subject to} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

We can deal it in a similar manner

- Define $y = \max_{i=1,\ldots,n}\{\mathbf{c}_i^T\mathbf{x} + d_i\}$

$$\begin{aligned} \text{minimize}_{\mathbf{x},y} \quad & y \\ \text{subject to} \quad & y \geq \mathbf{c}_i^T\mathbf{x} + d_i \quad \forall i \\ & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

Dealing with Absolute Values

Problems with absolute values might be handled as well by LP.

$$\begin{array}{rl} \text{minimize} & \sum_{i=1}^{n} |x_i| \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \end{array}$$

This can be equivalently written as (why?)

$$\begin{array}{rl} \text{minimize} & \sum_{i=1}^{n} y_i \\ \text{s.t.} & y_i \geq x_i \\ & y_i \geq -x_i \\ & A\mathbf{x} = \mathbf{b} \end{array}$$

Similar idea can be applied when there are constraints like $|\mathbf{a}^T\mathbf{x} + b| \leq c$.

Other ways to formulate it?

Other Ways to Formulate Minimizing Absolute Values

The previous problem can also be formulated as:

$$\begin{array}{rc} \text{minimize} & \sum_{i=1}^n x_i^+ + \sum_{i=1}^n x_i^- \\ \text{s.t.} & A\mathbf{x}^+ - A\mathbf{x}^- = \mathbf{b} \\ & \mathbf{x}^+, \mathbf{x}^- \geq 0 \end{array}$$

Another way (illustrated using $n = 2$):

$$\begin{array}{rc} \text{minimize} & y \\ \text{s.t.} & y \geq x_1 + x_2 \\ & y \geq x_1 - x_2 \\ & y \geq -x_1 + x_2 \\ & y \geq -x_1 - x_2 \\ & A\mathbf{x} = \mathbf{b} \end{array}$$

Absolute Values

Consider a similar problem

$$\begin{aligned} \text{maximize} \quad & \textstyle\sum_{i=1}^n |x_i| \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \end{aligned} \tag{2}$$

Can we use the similar idea and transform it into:

$$\begin{aligned} \text{maximize} \quad & \textstyle\sum_{i=1}^n y_i \\ \text{s.t.} \quad & y_i \geq x_i \\ & y_i \geq -x_i \\ & A\mathbf{x} = \mathbf{b} \end{aligned}$$

Answer: No. There is some intrinsic property of problem (2) that prevents us to formulate it as an LP (non-convexity). We will talk about it later in this course.

Consider

$$\begin{array}{ll} \text{minimize}_{\mathbf{x}} & \dfrac{\mathbf{c}^T\mathbf{x}+d}{\mathbf{e}^T\mathbf{x}+f} \\ \text{s.t.} & A\mathbf{x} \le \mathbf{b} \end{array}$$

Assume that $\mathbf{e}^T\mathbf{x}+f>0$ for any $\mathbf{x}$ that satisfies $A\mathbf{x} \le \mathbf{b}$.

Any idea to transform it to LP?

- Define

$$\mathbf{y} = \frac{\mathbf{x}}{\mathbf{e}^T\mathbf{x}+f}, \qquad z = \frac{1}{\mathbf{e}^T\mathbf{x}+f}$$

Then we can write the problem as

$$\begin{array}{rl} \text{minimize}_{\mathbf{y},z} & \mathbf{c}^T\mathbf{y} + dz \\ \text{s.t.} & A\mathbf{y} - bz \le 0 \\ & \mathbf{e}^T\mathbf{y} + fz = 1 \\ & z \ge 0 \end{array}$$

This is an LP

- Why are they equivalent?
- See Boyd and Vandenberghe (supplemental reading, page 151)

Why Linear Program?

Easiest to solve.

- Theoretically, it is polynomial time solvable, and the complexity is low among all optimization problems.
- Practically, commercial softwares can solve LPs with tens of thousands of variables very easily. Can solve LPs with millions of variables if there are some structures

Extremely versatile

- Can model many real problems, either exactly or approximately

Fundamental

- The LP theories lay the foundation for most optimization theories

In the next lecture, we will show how to use MATLAB to solve LPs.