Lecture 2: Formulating Optimization Problems

Zizhuo Wang

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

Sep 7, 2018

Recap: Introduction to Optimization

Three main components in optimization problems

- Decision
- Objective
- Constraints

General form of optimization problem:

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

Terminologies:

- Feasible solutions/set, optimal solutions, optimal value
- In optimization problems, we avoid dealing with strict inequality constraints

Recap: Classifications

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

By default, when we talk about an optimization problem, we assume it is continuous, unless we explicitly say that it is *discrete*

Modeling

Modeling is extremely important:

- Finding a good optimization model is at least half way in solving the problem
- In practice, we are not given a mathematical formulation — typically the problem is described verbally by a specialist within some domain. It is extremely important to be able to convert the problem into a mathematical formulation
- Modeling is *not* trivial: Sometimes, we not only want to find a formulation, but also want to find a good formulation.

In this lecture, we are going to show some examples about modeling a problem into an optimization problem.

Formulating Optimization Problem

The golden rule: Find and formulate the three components

- Decision $\longrightarrow$ Decision variables
- Objective $\longrightarrow$ Objective functions
- Constraints $\longrightarrow$ Constraint functions/inequalities

Now let's make it work.

Maximum Area Problem Revisited

You have 80 meters of fencing and want to enclose a rectangle yard as large (area) as possible. How should you do it?

- Decision variable: the length ℓ and width w of the yard
- Objective: maximize the area: ℓw
- Constraints: the total length of yard available: $2\ell + 2w \leq 80$

Therefore, the optimization problem can be written as:

$$\begin{aligned} \text{maximize}_{\ell,w} \quad & \ell w \\ \text{subject to} \quad & 2\ell + 2w \leq 80 \\ & \ell, w \geq 0 \end{aligned}$$

What category this optimization problem belongs to?

- Constrained, nonlinear, continuous.

Production Problem Revisited

Firm A needs to decide the amount of each product to produce.

	Steel	Iron	Copper	Profit
Alloy 1	1	0	1	\$1
Alloy 2	0	2	1	\$2
Resources	100	200	150	

Decision variables:

- x_1: the amount of alloy 1 to produce; x_2: the amount of alloy 2 to produce

Objective function:

- $x_1 + 2x_2$

Constraints?

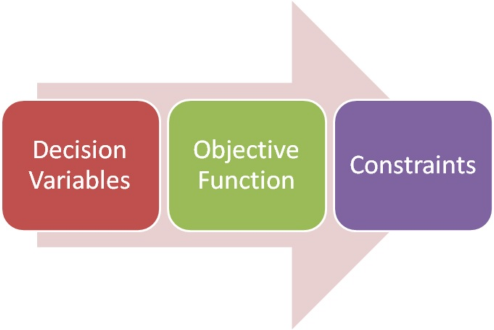
Optimization Model

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

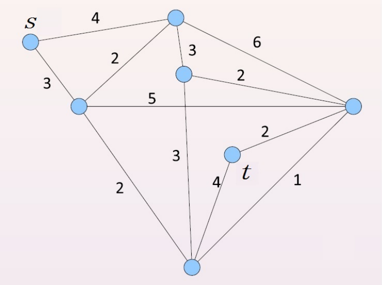
It is a linear optimization problem

We call the numbers $(1, 2, 100, 150, 200, \text{ etc.})$ the coefficients of the optimization problem.

Modeling Methods



Shortest Path Problem



What is the shortest path from s to t? How to formulate it as an optimization problem?

- This is called the shortest path problem

Some notations: We define the set of edges by E. The distance between node i and node j is w_{ij}.

Model Shortest Path Problem using Optimization

For each edge $(i, j) \in E$, define

$$x_{ij} = \begin{cases} 1 & \text{if we use edge } (i, j) \\ 0 & \text{otherwise} \end{cases}$$

An optimization model:

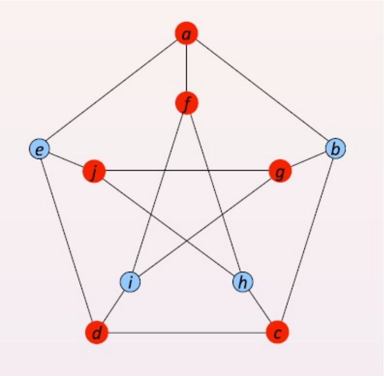
$$\begin{array}{rc} \text{minimize} & \sum_{(i,j)\in E} w_{ij} x_{ij} \\ \text{subject to} & \sum_j x_{sj} = 1 \\ & \sum_j x_{jt} = 1 \\ & \sum_j x_{ij} = \sum_j x_{ji}, \quad \forall i \neq s, t \\ & x_{ij} \in \{0, 1\}, \qquad \forall (i, j) \in E \end{array}$$

What category this optimization problem belongs to?

- Constrained, linear, integer
- In fact, it can be transformed to a linear optimization

Another Graph Problem: Vertex Cover

Given a graph consists of nodes V and edges E, find the smallest set of vertices that touch every edge of the graph



Optimization Model for Vertex Cover Problem

Define the following decision variables:

$$x_i = \begin{cases} 1 & \text{if we choose vertex } i \\ 0 & \text{otherwise} \end{cases}$$

An optimization model can be written as:

$$\begin{array}{rll} \text{minimize} & \sum_i x_i & \\ \text{subject to} & x_i + x_j \geq 1, & \forall (i,j) \in E \\ & x_i \in \{0,1\} & \forall i \in V \end{array}$$

It is an integer (linear) optimization problem.

- Many graph/network problem can be modeled as optimization problems.

Vector Notations

In this course, we use bold font to denote vectors:

$$\mathbf{x} = (x_1, ..., x_n).$$

By default, all vectors are column vectors.

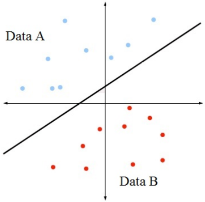
We use $\mathbf{x}^T$ to denote the transpose of a vector.

We use $\mathbf{a}^T\mathbf{x}$ to denote the inner product of $\mathbf{a}$ and $\mathbf{x}$, i.e.

$$\mathbf{a}^T\mathbf{x} = a_1x_1 + \cdots + a_nx_n = \sum_{i=1}^{n} a_ix_i$$

Support Vector Machine Problem

Given two groups of data points in $\mathbb{R}^d$, $A = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ and $B = \{\mathbf{y}_1, ..., \mathbf{y}_m\}$. We want to find a plane that separates them.



It is used in pattern recognition, machine learning, etc.

Decision: A plane $\mathbf{a}^T\mathbf{x} + b = 0$ defined by $(\mathbf{a}, b)$ that separates the points such that

$$\mathbf{a}^T\mathbf{x}_i + b > 0, \quad \forall i = 1, ..., n$$
$$\mathbf{a}^T\mathbf{y}_j + b < 0, \quad \forall j = 1, ..., m$$

Support Vector Machine Problem

This is equivalent as finding $\mathbf{a}$ and b such that:

$$\begin{aligned} \mathbf{a}^T\mathbf{x}_i + b \geq 1, &\quad \forall i = 1, \dots, n \\ \mathbf{a}^T\mathbf{y}_j + b \leq -1, &\quad \forall j = 1, \dots, m. \end{aligned}$$

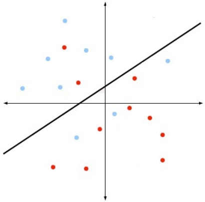
This can be written as an optimization problem:

$$\begin{array}{rll} \text{minimize}_{\mathbf{a},b} & 0 & \\ \text{subject to} & \mathbf{a}^T\mathbf{x}_i + b \geq 1, & \forall i = 1, \dots, n \\ & \mathbf{a}^T\mathbf{y}_j + b \leq -1, & \forall j = 1, \dots, m \end{array}$$

We call such problems feasibility problems: feasibility problem is a special kind of optimization problem.

Support Vector Machine Problem Continued

Sometimes a total separation is not possible:



Then we want to find a plane such that the total "error" is minimized: (we write $(w)^+$ for $\max\{w, 0\}$)

- For points in A, the error is $(1 - \mathbf{a}^T\mathbf{x}_i - b)^+$
- For points in B, the error is $(\mathbf{a}^T\mathbf{y}_i + b + 1)^+$

Support Vector Machine Problem Continued

Therefore we can write the support vector machine problem as:

$$\text{minimize}_{\mathbf{a},b} \quad \sum_i (1 - \mathbf{a}^T \mathbf{x}_i - b)^+ + \sum_j (\mathbf{a}^T \mathbf{y}_j + b + 1)^+$$

This is an unconstrained, nonlinear, continuous optimization problem.

- Next, we further show how to equivalently write it as a linear optimization problem.

A Linear Optimization Formulation

Define $\delta_i = (1 - \mathbf{a}^T\mathbf{x}_i - b)^+$ and $\sigma_j = (\mathbf{a}^T\mathbf{y}_j + b + 1)^+$.

We can first write this as

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

We claim we can relax "$=$" to "$\geq$" (why?):

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

Support Vector Machine Problem

Furthermore, $\delta_i \geq (1 - \mathbf{a}^T\mathbf{x}_i - b)^+$ is equivalent to

$$\delta_i \geq 1 - \mathbf{a}^T\mathbf{x}_i - b, \qquad \delta_i \geq 0$$

Similarly $\sigma_j \geq (\mathbf{a}^T\mathbf{y}_j + b + 1)^+$ is equivalent to

$$\sigma_j \geq \mathbf{a}^T\mathbf{y}_j + b + 1, \qquad \sigma_j \geq 0$$

Therefore the optimization problem can be transformed to

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

This is a linear optimization with decision variables $\mathbf{a}$, b, δ, σ.