Optimization

Duong Thi Kim Huyen - Ta Thuy Anh

Faculty of Computer Science
PHENIKAA UNIVERSITY

Ngày 5 tháng 1 năm 2023

Course information

Teachers: Duong Thi Kim Huyen

Ta Thuy Anh

Time:

Location:

Course Description: This course will cover the basic notions of optimization, mathematical models, as well as approaches and mindsets for solving optimization problems.

Prerequisites

- Linear algebra, Analysis, Discrete Maths

Course grade

- Attendance (10%)
- Two midterm exams (40%)
- Final exam (50%)

Materials

Textbook

1. Steven Boyd and Lieven Vandenberghe. Convex optimization. Cambridge University Press (2004). ISBN: 0521833787.
2. Nguyễn Đức Nghĩa, Tối ưu hóa, NXB. Giáo dục, 2002.
3. Trần Vũ Thiệu, Nguyễn Thị Thu Thủy. Giáo trình tối ưu phi tuyến. Nxb ĐHQGHN, 2011.

References

1. Bùi Thế Tâm, Trần Vũ Thiệu, Các phương pháp tối ưu hóa, Nxb. Giao thông vận tải, 1998.
2. Bùi Minh Trí, Quy hoạch toán học, Nxb. Khoa học và Kỹ thuật, 1999.
3. Phan Quốc Khánh, Trần Huệ Nương, Quy hoạch tuyến tính, Nxb. Giáo dục, 2003
4. Cormen, Thomas, Charles Leiserson, Ronald Rivest, and Clifford Stein. Introduction to Algorithms. MIT Press, 2009.

Course Schedule

1. Introduction
   - Mathematical model
   - Types of optimization problems
   - Complexity
2. Continuous optimization
   - Convex optimization
   - Linear program
3. Discrete optimization
   - Integer linear program
   - Optimization problems on graphs
4. Applications in Machine Learning and Big Data
5. Software for solving optimization problems

Linear Algebra

- Matrix and basic operations
- Determinant, Rank, invertible matrix, etc.
- Solving a system of linear equations?
- Diagonalize a square matrix
- Vector space

Analysis

- Continuous function
- Differential of a function and Differentiation rules

What is Optimization?

Optimization problem

Maximizing or minimizing some function relative to some set, often representing a range of choices available in a certain situation. The function allows comparison of the different choices for determining which might be “best.”

Question

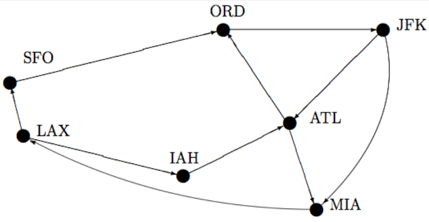
- How can one mathematically model an optimization problem?
- How can one solve an optimization problem?
- How can one evaluate the quality of solutions?
- ......

Timetabling

The Curriculum-based timetabling problem consists of the weekly scheduling of the lectures for several university courses within a given number of rooms and time periods, where conflicts between courses are set according to the curricula published by the University and not on the basis of enrolment data.

International Timetabling Competition - http://www.cs.qub.ac.uk/itc2007/index.htm

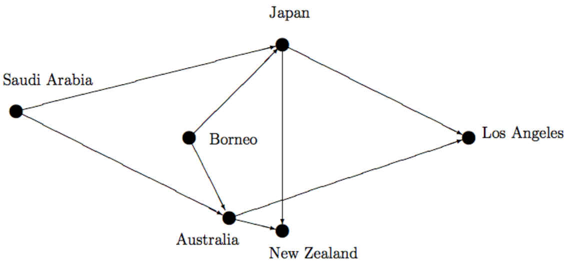
Crew Scheduling



An airline has to assign crews to its flights.

- Make sure that each flight is covered.
- Meet regulations, eg, each pilot can only fly a certain amount each day.
- Minimize costs, eg: accommodation for crews staying overnight out of town, crews dead

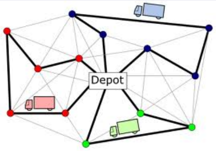
Production Planning



An oil company has oil fields in Saudi Arabia and Borneo, refineries in Japan and Australia, and customers in the US and New Zealand. The fields produce different qualities of oil, which is refined and combined into different grades of gasoline.

1. What raw oil should be shipped to which refinery?
2. How much of each type of gasoline should each refinery produce?

Vehicle routing



1. What is the optimal set of routes for a fleet of vehicles to traverse in order to deliver to a given set of customers?

Objective

- Minimize the global transportation cost based on the global distance travelled as well as the fixed costs associated with the used vehicles and drivers
- Minimize the number of vehicles needed to serve all customers
- Least variation in travel time and vehicle load
- Minimize penalties for low quality service
- Maximize a collected profit/score

Other applications

- Machine learning
- Data science
- Artificial intelligence (AI)

Problem 1.
Find the maximum value of $f(x) = x^2 - 3x + 2$ over the domain $[-2; 3]$?

$$\begin{aligned}
\text{maximize} \quad & f(x) = x^2 - 3x + 2 && (1)\\
\text{subject to} \quad & x \geq -2 && (2)\\
& x \leq 3 && (3)
\end{aligned}$$

Problem 1.
Find the maximum value of $f(x) = x^2 - 3x + 2$ over the domain $[-2; 3]$?

$$\begin{aligned} \text{maximize} \quad & f(x) = x^2 - 3x + 2 && (1) \\ \text{subject to} \quad & x \geq -2 && (2) \\ & x \leq 3 && (3) \end{aligned}$$

Problem 2.

$$\begin{aligned} \text{maximize} \quad & g(x, y) = 2x + y && (4) \\ \text{subject to} \quad & 3x + y \leq 6 && (5) \\ & x + y \leq 4 && (6) \\ & x \geq 0 && (7) \\ & y \geq 0 && (8) \end{aligned}$$

Optimization problem
Find the maximum value of $f_0(x_1, \ldots, x_n)$ over the domain D, where

$$D = \{\mathrm{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n \mid f_i(\mathrm{x}) \le b_i, i = 1, \ldots, m\}$$

Optimization problem
Find the maximum value of $f_0(x_1, \ldots, x_n)$ over the domain D, where

$$D = \{\mathrm{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n \mid f_i(\mathrm{x}) \leq b_i, i = 1, \ldots, m\}$$

Mathematical model

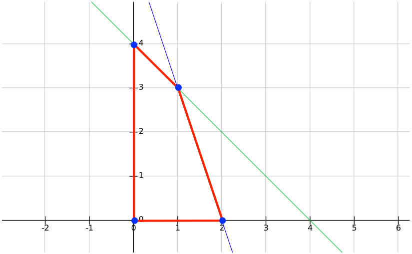
$$\begin{aligned} &\text{maximize} && f_0(\mathrm{x}) && (9)\\ &\text{subject to} && \mathrm{x} \in D && (10) \end{aligned}$$

$$\begin{aligned} &\text{maximize} && f_0(\mathrm{x}) && (11)\\ &\text{subject to} && f_1(\mathrm{x}) \leq b_1 && (12)\\ &&& f_2(\mathrm{x}) \leq b_2 && (13)\\ &&& \ldots\ldots && \\ &&& f_m(\mathrm{x}) \leq b_m && (14) \end{aligned}$$

- $\mathrm{x} = (x_1, \ldots, x_n)$ is optimization variable
- $f_i : \mathbb{R}^n \to \mathbb{R}$ is constraint function, for $i = 1, \ldots, m$
- b_i is the limit or bound for the constraint, for $i = 1, \ldots, m$
- $\mathrm{x} \in D$: is a feasible solution
- **optimal** solution is a feasible solution x^* such that $f_0(\mathrm{x}^*) \geq f_0(\mathrm{x})$, $\forall \mathrm{x} \in D$

Example

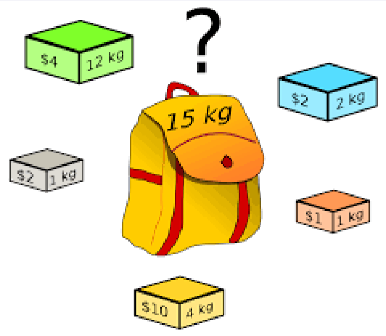
$$\begin{aligned} \text{maximize} \quad & g(x,y) = 2x + y \\ \text{subject to} \quad & 3x + y \leq 6 \\ & x + y \leq 4 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$



- Continuous optimization
- Discrete optimization (or Combinatorial optimization)
- Constrained optimization
- Unconstrained optimization
- Convex (continuous) optimization
- Non-Convex (continuous) optimization
- Multi-objective optimization

Example of discrete optimization problem

Knapsack problem



Mathematical model

$$\begin{aligned} \text{maximize} \quad & a_1x_1 + \cdots + a_nx_n \\ \text{subject to} \quad & b_1x_1 + \cdots + b_nx_n \leq c, \\ & x_i \in \{0, 1\}. \end{aligned}$$

- a_i is the profit of item i
- b_i is the weight of item i
- c is the capacity of the Knapsack
- x_i is discrete variables
- $x_i = 1$ iff item i is chosen
- $x_i = 0$ iff item i is not chosen

Production Planning

Given: A farmer has a piece of farm land, say L km2, to be planted with either wheat or barley or some combination of the two. He has a limited amount of fertilizer, F kilograms, and pesticide, P kilograms. Every square kilometer of wheat requires F_1 kilograms of fertilizer and P_1 kilograms of pesticide, while every square kilometer of barley requires F_2 kilograms of fertilizer and P_2 kilograms of pesticide. Let S_1 be the selling price of wheat per square kilometer, and S_2 be the selling price of barley.

Task: Find the area of land planted with wheat and barley so that the profit is maximized.

Mathematical model

$$
\begin{aligned}
\text{maximize} \quad & S_1 \cdot x_1 + S_2 \cdot x_2 && \text{(maximize the revenue)} \\
\text{subject to} \quad & x_1 + x_2 \leq L, && \text{(limit on total area)} \\
& F_1 \cdot x_1 + F_2 \cdot x_2 \leq F, && \text{(limit on fertilizer} \\
& P_1 \cdot x_1 + P_2 \cdot x_2 \leq P, && \text{(limit on pesticide)} \\
& x_1; x_2 \geq 0 && \text{(cannot plant a negative area)}
\end{aligned}
$$

Machine scheduling

Given: There are n jobs needed to be processed on m identical machines. Each job j has a processing time (or size) p_j, and can be processed on exactly one machine. At every time, each machine can only processes not more than one job.

Task: Find a schedule of jobs to machines such that the maximum completion time of all jobs is minimized.

- Describing the problem;
- Formulating the OR model;
- Solving the OR model;
- Performing some analysis of the solution;
- Presenting the solution and analysis.

Machine scheduling

Given: There are n jobs needed to be processed on m identical machines. Each job j has a processing time (or size) p_j, and can be processed on exactly one machine. At every time, each machine can only processes not more than one job.

Task: Find a schedule of jobs to machines such that the maximum completion time of all jobs is minimized.

$$
\begin{aligned}
\text{minimize} \quad & \max_{i=1}^{m} \left\{ \sum_{j=1}^{n} p_j \cdot x_{ij} \right\} \\
\text{subject to} \quad & \sum_{i=1}^{m} x_{ij} = 1, \quad \text{for} \quad j = 1, \ldots, n \\
& x_{ij} \in \{0, 1\}
\end{aligned}
$$

The problem of finding the shortest path

Given: There is an undirected graph $G = (V, E)$ of n vertices, each edge (i, j) has a weight w_{ij}. Let s and t be two vertices of G.

Task: Find a path of minimum weight from s to t.

The problem of finding the shortest path

Given: There is an undirected graph $G = (V, E)$ of n vertices, each edge (i, j) has a weight w_{ij}. Let s and t be two vertices of G.

Task: Find a path of minimum weight from s to t.

The variable

$$x_{ij} = \begin{cases} 1 & \text{if the shotest path contains } i \to j \\ 0 & \textit{otherwise}. \end{cases}$$

There is a formula for calculating the amount of flows at each single node i.

$$\begin{aligned} b(i) &= \text{Amount of outgoing flow from } i - \text{Amount of incoming flow to } i \\ &= \sum_j x_{ij} - \sum_k x_{ki} \end{aligned}$$

We have

$$b_i = \begin{cases} 1 & \text{if } i \text{ is the satarting node} \\ -1 & \text{if } i \text{ is the ending node} \\ 0 & \textit{otherwise}. \end{cases}$$

Mathematical model

$$
\begin{aligned}
\text{minimize} \quad & \sum_{(i,j)} w_{ij} \cdot x_{ij} \\
\text{subject to} \quad & \sum_{j} x_{sj} - \sum_{k} x_{ks} = 1, \\
& \sum_{j} x_{tj} - \sum_{k} x_{kt} = -1, \\
& \sum_{j} x_{ij} - \sum_{k} x_{ki} = 0, \quad i \neq s, t \\
& x_{ij} \in \{0, 1\}
\end{aligned}
$$