6.215/6.255J/15.093J/IDS.200J Optimization Methods

Professor: Patrick Jaillet

TAs: Amine Bennouna, Moise Blanchard, Victor Gonzales, Tetiana Husak, Yi-Lin Liao

September 9, 2021

Course Information

Lectures: TR 1-2:30pm (32-123)
Professor: Patrick Jaillet
Office hour by appointment

Recitations: F 10-11am (32-141), or F 1-2pm (32-155)
TAs: Amine Bennouna, Moise Blanchard, Victor Gonzales,
Tetiana Husak, Yi-Lin Liao
Office hours: TBD
staff e-mail: 6.255@mit.edu (for personal issues)

Website: https://canvas.mit.edu/courses/11155
Piazza: https://piazza.com/mit/fall2021/6255
Prerequisites: Calculus, Linear algebra (18.06 or equivalent), some familiarity with computational tools (e.g., Julia) and mathematical maturity.

Required Text
The required textbook is D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific, Belmont, MA, 1997.

Structure of Class

- Linear Optimization: Lec. 1-8
- Network Flows: Lec. 9-10
- Review & Recap: Lec. 11
  (Midterm)
- Robust Optimization: Lec. 12
- Discrete Optimization: Lec 13-17
- Nonlinear Optimization: Lec. 18-21
- Convex and Semidefinite Optimization: Lec. 22-24
- Overview: Lec. 25
  (Final)

Requirements

Grading:

- Homeworks: 30% (5 psets)
- Midterm Exam: 30% (on Friday, October 15)
- Final Exam: 40% (during final week period 12/13 to 12/17)

Policy on Collaborations:

In the case of the written homework assignments, your assignment write-up may be done in pairs. If you can not find a team mate the TAs will be happy to help you along. You may also interact with other fellow students (not in your pair) when preparing your homework solutions. However, at the end, each pair must write up solutions on their own. Duplicating a solution that someone else has written (verbatim or edited), or providing solutions for another pair to copy is not acceptable.

During the midterm and the final examination, any student who either receives or knowingly gives assistance or information concerning the examination will be in violation of the policy on individual work.

Today's Lecture

Outline

- History of Optimization
- Where do Linear Optimization (LO) Problems arise?
- Examples of Formulations

History of Optimization

-Fermat, 1638; Newton, 1670

$$\min \; f(x) \text{ with } x \text{ scalar} \Rightarrow \text{necessary condition:} \frac{df(x)}{dx} = 0$$

-Euler, 1755

$$\min \; f(x_1, \ldots, x_n) \Rightarrow \nabla f(\boldsymbol{x}) = 0$$

-Lagrange, 1797

$$\begin{aligned} \min \quad & f(x_1, \ldots, x_n) \\ \text{s.t.} \quad & g_k(x_1, \ldots, x_n) = 0 \qquad k = 1, \ldots, m \end{aligned}$$

-Euler, Lagrange Problems in infinite dimensions, calculus of variations.

Optimization

The general problem

$$
\begin{aligned}
\min \quad & f(x_1, \ldots, x_n) \\
\text{s.t.} \quad & g_1(x_1, \ldots, x_n) \leq 0 \\
& \vdots \\
& g_m(x_1, \ldots, x_n) \leq 0
\end{aligned}
$$

What is Linear Optimization?

Toy illustration - a diet problem

$$
\begin{array}{rl}
\text{minimize} & 3x_1 + x_2 \\
\text{subject to} & x_1 + 2x_2 \geq 2 \\
 & 2x_1 + x_2 \geq 3 \\
 & x_1 \geq 0, x_2 \geq 0
\end{array}
$$

What is Linear Optimization?

Toy illustration - a diet problem

$$\begin{aligned} \text{minimize} \quad & 3x_1 + x_2 \\ \text{subject to} \quad & x_1 + 2x_2 \geq 2 \\ & 2x_1 + x_2 \geq 3 \\ & x_1 \geq 0, x_2 \geq 0 \end{aligned}$$

$$\boldsymbol{c} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad \boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^T \boldsymbol{x} \text{ (or } \boldsymbol{c}'\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \\ & \boldsymbol{x} \geq 0 \end{aligned}$$

History of LO

The pre-algorithmic period

-Chinese mathematicians ($\approx$ 300 BC), Newton, 1707 Linear equalities

-Fourier, 1826 Method for solving system of linear inequalities.

-de la Vallée Poussin, 1910s Simplex-like method for objective function with absolute values.

-Kantorovich, Koopmans, 1930s Formulations and solution method.

-von Neumann, 1928 Game theory, duality.

-Farkas, Minkowski, Carathéodory, 1870-1930 Foundations.

History of LO

The modern period

- 1947 George Dantzig, Simplex method.
- 1950s Applications.
- 1960s Large scale optimization. Stochastic optimization.
- 1970s Complexity theory.
- 1979 The ellipsoid algorithm.
- 1980s Interior point algorithms.
- 1990s Semidefinite and conic optimization.
- 2000s Robust optimization, online optimization.
- 2010s Huge scale optimization.
- 2020s ...

Where do LO problems arise?

Wide applicability

- Transportation
- Telecommunications
- Manufacturing
- Medicine
- Engineering
- Typesetting (TEX, LATEX)
- Finance
- Machine learning ...

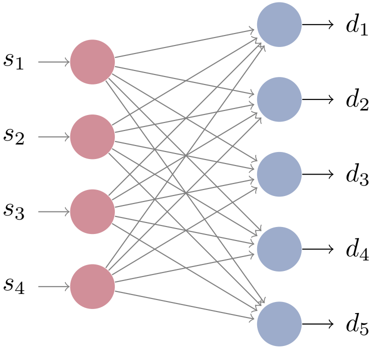
"If one would take statistics about which mathematical problem is using up most of the computer time in the world, then the answer would probably be linear optimization." Lászlo Lovász, Microsoft Research, 1980

Transportation Problem

Data

Description:

- m plants, n warehouses
- s_i supply of ith plant, $i = 1, \ldots, m$
- d_j demand of jth warehouse, $j = 1, \ldots, n$
- c_{ij}: cost of transportation $i \to j$



Problem: Satisfy the demand of all warehouses with minimal transportation cost.

Transportation Problem

Decision Variables, Formulation

x_{ij} = number of units to send $i \rightarrow j$

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} & \\
\text{s.t.} \quad & \sum_{i=1}^{m} x_{ij} \geq d_j & j = 1 \ldots n \\
& \sum_{j=1}^{n} x_{ij} \leq s_i & i = 1 \ldots m \\
& x_{ij} \geq 0 &
\end{aligned}
$$

Sorting using LO

- Given n numbers $c_1, c_2, \ldots, c_n$;
- The order statistic $c_{(1)}, c_{(2)}, \ldots, c_{(n)}$: $c_{(1)} \leq c_{(2)} \leq \ldots \leq c_{(n)}$;
- Use LO to find $\sum_{i=1}^{k} c_{(i)}$.

Sorting using LO

- Given n numbers $c_1, c_2, \ldots, c_n$;
- The order statistic $c_{(1)}, c_{(2)}, \ldots, c_{(n)}$: $c_{(1)} \leq c_{(2)} \leq \ldots \leq c_{(n)}$;
- Use LO to find $\sum_{i=1}^{k} c_{(i)}$.

$$\begin{aligned} \min \quad & \sum_{i=1}^{n} c_i x_i \\ \text{s.t.} \quad & \sum_{i=1}^{n} x_i = k \\ & 0 \leq x_i \leq 1 \qquad i = 1, \ldots, n \end{aligned}$$

Investment Problem

- Five investment choices A, B, C, D, E.
- A, C, and D are available in 2020.
- B is available in 2021.
- E is available in 2022.
- Cash earns 6% per year.
- $1,000,000 in 2020.

Investment Problem

Cash Flow per Dollar Invested

Year:	A	B	C	D	E
2020	-1.00	0	-1.00	-1.00	0
2021	+0.30	-1.00	+1.10	0	0
2022	+1.00	+0.30	0	0	-1.00
2023	0	+1.00	0	+1.75	+1.40
LIMIT	$500,000	NONE	$500,000	NONE	$750,000

Problem: How much to invest in each option in order to maximize profit.

Investment Problem

Formulation, Decision Variables

- $A, \ldots E$: amount invested in \$ millions
- $Cash_t$: amount invested in cash in period t, $t = 1, 2, 3$

$$
\begin{array}{ll}
\max & 1.06 Cash_3 + 1.00B + 1.75D + 1.40E \\
\text{s.t.} & A + C + D + Cash_1 \leq 1 \\
 & Cash_2 + B \leq 0.3A + 1.1C + 1.06 Cash_1 \\
 & Cash_3 + 1.0E \leq 1.0A + 0.3B + 1.06 Cash_2 \\
 & A \leq 0.5, \quad C \leq 0.5, \quad E \leq 0.75 \\
 & A, \ldots, E \geq 0
\end{array}
$$

- Solution: $A = 0.5M$, $B = 0$, $C = 0$, $D = 0.5M$, $E = 0.659M$, $Cash_1 = 0$, $Cash_2 = .15M$, $Cash_3 = 0$; Objective: $1.7976M$

Manufacturing
Data

- n products, m raw materials
- c_j: profit of product j
- b_i: available units of material i.
- a_{ij}: # units of material i product j needs in order to be produced.

Manufacturing

Formulation, Decision variables

x_j = amount of product j produced.

$$
\begin{aligned}
\max \quad & \sum_{j=1}^{n} c_j x_j \\
\text{s.t.} \quad & a_{11}x_1 + \cdots + a_{1n}x_n \leq b_1 \\
& \vdots \\
& a_{m1}x_1 + \cdots + a_{mn}x_n \leq b_m \\
& x_j \geq 0, \qquad j = 1 \ldots n
\end{aligned}
$$

Data classification

Problem

Given a set of measurements ("features") of an object (e.g., color, weight, etc.), determine its type

- We have labeled samples $\{a^1, a^2, \ldots, a^n\}$ and $\{b^1, b^2, \ldots, b^m\}$, where $a^i, b^j \in \mathbb{R}^d$ (training set)
- Find a (possibly nonlinear) classifier to distinguish the sets.

For simplicity, consider the case $d = 2$, and a linear classifier.

$$c_0 + c_1 x_1 + c_2 x_2$$

- Decision variables: c_0, c_1, c_2
- A perfect classifier must satisfy the inequalities:

$$\begin{aligned} c_0 + c_1 a_1^i + c_2 a_2^i &\geq 1, & i = 1, \ldots, n \\ c_0 + c_1 b_1^j + c_2 b_2^j &\leq -1, & j = 1, \ldots, m \end{aligned}$$

Scheduling

Data & Constraints; Decision variables?

- Hospital wants to make weekly nightshift for its nurses
- d_j demand for nurses, $j = 1 \ldots 7$
- Every nurse works 5 days in a row
- Goal: hire minimum number of nurses

Scheduling

Data & Constraints; Decision variables?

- Hospital wants to make weekly nightshift for its nurses
- d_j demand for nurses, $j = 1 \dots 7$
- Every nurse works 5 days in a row
- Goal: hire minimum number of nurses

Decision Variables

x_j: # nurses starting their week on day j

Scheduling

Formulation

$$
\begin{array}{llllllllll}
\min & \sum_{j=1}^{7} x_j & & & & & & & & \\
\text{s.t.} & x_1+ & & & x_4+ & x_5+ & x_6+ & x_7 & \geq & d_1 \\
 & x_1+ & x_2 & & & x_5+ & x_6+ & x_7 & \geq & d_2 \\
 & x_1+ & x_2+ & x_3 & & & x_6+ & x_7 & \geq & d_3 \\
 & x_1+ & x_2+ & x_3+ & x_4+ & & & x_7 & \geq & d_4 \\
 & x_1+ & x_2+ & x_3+ & x_4+ & x_5 & & & \geq & d_5 \\
 & & x_2+ & x_3+ & x_4+ & x_5+ & x_6 & & \geq & d_6 \\
 & & & x_3+ & x_4+ & x_5+ & x_6+ & x_7 & \geq & d_7 \\
 & x_j \geq 0 & & & & & & & &
\end{array}
$$

1. Define your decision variables clearly.
2. Write constraints and objective function.
3. No systematic method available.

What is a good LO formulation?
A formulation with a small number of variables and constraints, and the matrix $\boldsymbol{A}$ is sparse.

(Nonlinear) Optimization

The general problem

$$
\begin{array}{ll}
\min & f(x_1, \ldots, x_n) \\
\text{s.t.} & g_1(x_1, \ldots, x_n) \leq 0 \\
 & \vdots \\
 & g_m(x_1, \ldots, x_n) \leq 0
\end{array}
$$

Convex functions

- $f : S \longrightarrow R$
- For all $x_1, x_2 \in S$

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$

- $f(\boldsymbol{x})$ concave if $-f(\boldsymbol{x})$ convex.