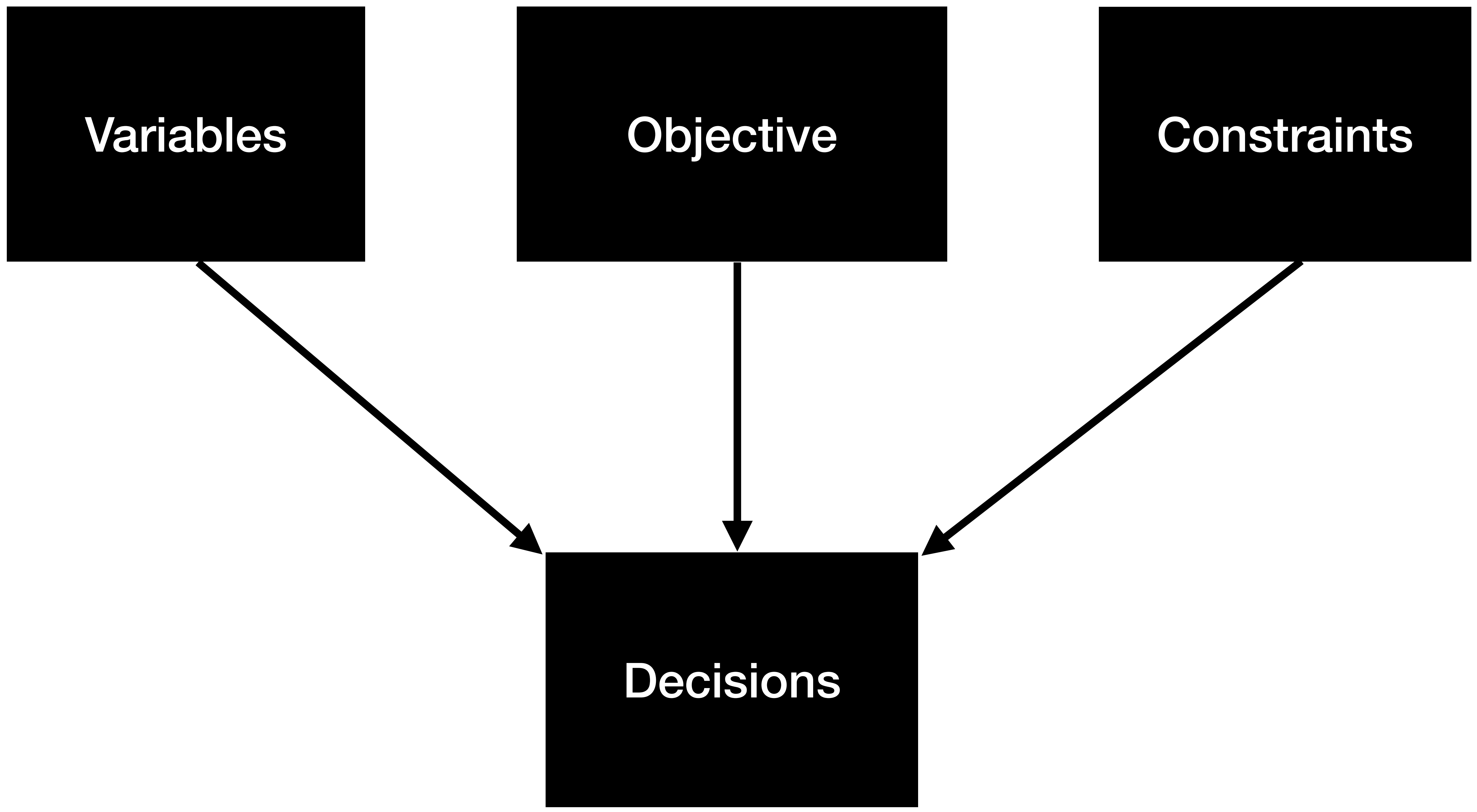
ORF522 – Linear and Nonlinear Optimization

1. Introduction

Bartolomeo Stellato - Fall 2024

What is this course about?

The mathematics behind making optimal decisions



Finance

Variables
Amounts invested in each asset

Constraints
Budget, investment per asset,
minimum return, etc.

Objective
Maximize profit, minus risk

Optimal control

Variables
Inputs: thrust, flaps, etc.

Constraints
System limitations, obstacles, etc.

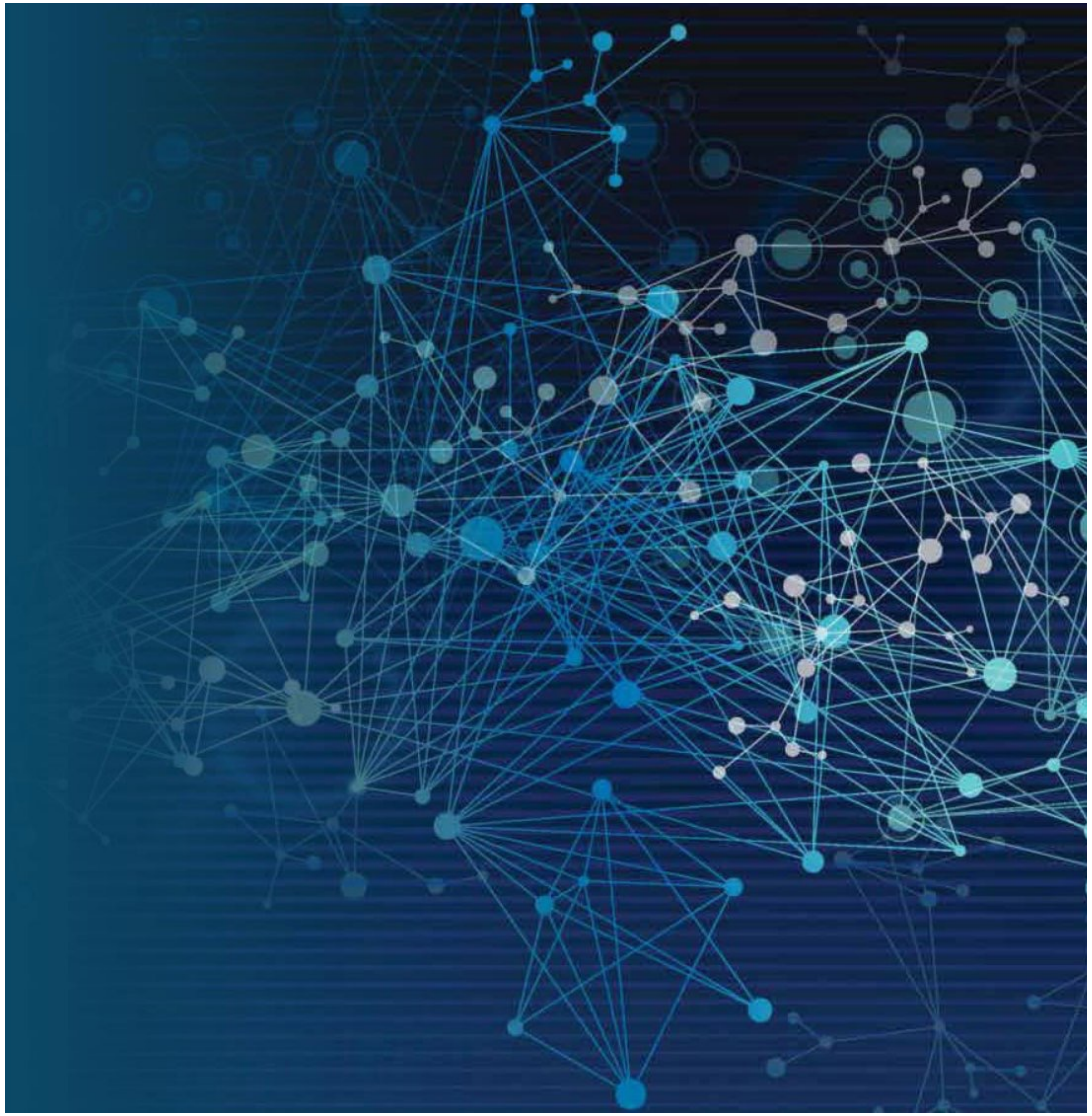
Objective
Minimize distance to target and fuel consumption

Machine learning

Variables
Model parameters

Constraints
Prior information, parameter limits

Objective
Minimize prediction error, plus regularization



Mathematical optimization

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

$x = (x_1, \dots, x_n)$ Variables

$f : \mathbf{R}^n \to \mathbf{R}$ Objective function

$g_i : \mathbf{R}^n \to \mathbf{R}$ Constraint functions

$x^\star$ Solution/Optimal point

$f(x^\star)$ Optimal value

Most optimization problems cannot be solved

Solving optimization problems

General case → **Very hard!**

Compromises

- Long computation times
- Not finding the solution
  (in practice it may not matter)

Exceptions

- Linear optimization
- Convex optimization

→ **Can be solved very efficiently and reliably**

Meet your teaching staff

Instructor

Bartolomeo Stellato

I am a Professor at ORFE. I obtained my PhD from Oxford and I was a postdoc at MIT.

email: bstellato@princeton.edu

office hours: Thu 2pm—4pm EST, Sherred 123

website: stellato.io

Assistant in instruction

Yanjun Liu

PhD student at ORFE.

email: yanjun.liu@princeton.edu

office hours: Mon 4:30pm—6:30pm EST, Sherrerd 003

Meet your classmates!

Name? Year?

Where are you from?

What is your department?

What do you want to use optimization for?

Today's agenda

- Optimization problems
- History of optimization
- Course contents and information
- A glance into modern optimization

Linear optimization

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & a_i^T x \leq b_i, \quad i = 1. \dots, m
\end{array}
$$

No analytical formula (99% of the time there will be none in this course!)

Efficient algorithms and software we can solve problems with several thousands of variables and constraints

Extensive theory (duality, degeneracy, sensitivity)

Linear optimization

Example: resource allocation

$$
\begin{array}{ll}
\text{maximize} & \sum_{i=1}^{n} c_i x_i \\
\text{subject to} & \sum_{i=1}^{n} a_{ji} x_i \leq b_j, \quad j = 1, \ldots, m \\
& x_i \geq 0, \quad i = 1, \ldots, n
\end{array}
$$

- c_i: profit per unit of product i shipped
- b_j: units of raw material j on hand
- a_{ji}: units of raw material j required to produce on unit of product i

Nonlinear optimization

$$
\begin{array}{ll}
\text{minimize} & f(x) \\
\text{subject to} & g_i(x) \leq 0, \quad i = 1, \ldots, m
\end{array}
$$

Hard to solve in general

- multiple local minima
- discrete variables $x \in \mathbf{Z}^n$
- hard to certify optimality

Convex optimization

convex functions
nonnegative (upward) curvature

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

All local minima are global!

Efficient algorithms and software

Extensive theory (convex analysis and conic optimization) [ORF523]

Used to solve non convex problems

Prehistory of optimization

Calculus of variations

Fermat/Newton

minimize $\quad f(x), \; x \in \mathbf{R}$

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = 0$$

1670

Euler

minimize $\quad f(x), \; x \in \mathbf{R}^n$

$$\nabla f(x) = 0$$

1755

Lagrange

minimize $\quad f(x)$

subject to $\quad g(x) = 0$

1797

Time

History of optimization

Age of computers

Algorithms

- 1930s — Origin of linear optimization (Kantorovich, Koopmans, von Neumann)
- 1947 — Simplex algorithm (Dantzig)
- 1984 — Interior-point methods (Karmarkar)
- 2000s — Large-scale optimization

Applications

- Operations Research, Economics
- 1990s — Engineering, Statistics
- 2000s — Machine learning, Image processing, Communication systems, Embedded intelligent systems

Technological innovations

Lots of data

easy storage
and
transmission

Massive computations

computers
are
super fast

High-level programming languages

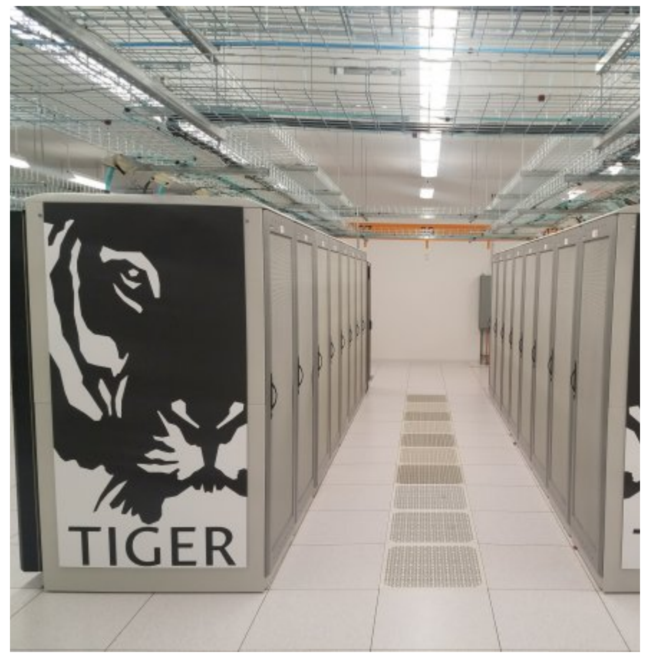
easy to
do complex
stuff

What is happening today?

Huge scale optimization

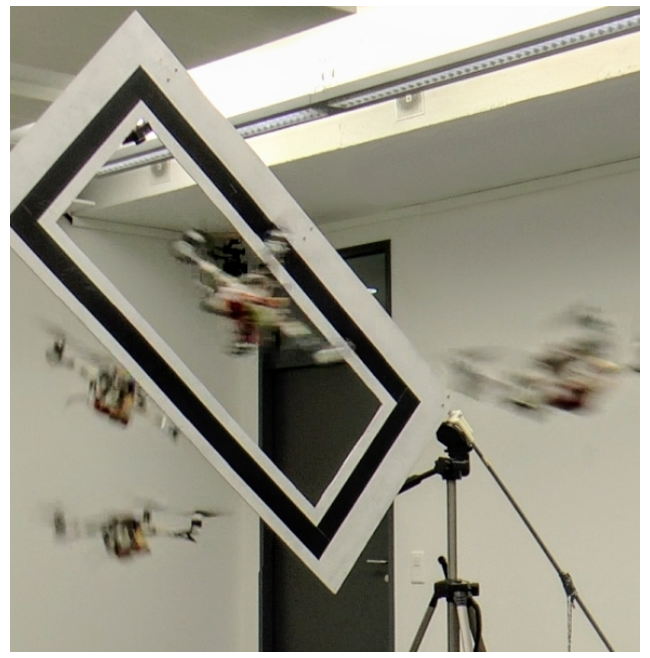
Massive data

Massive computations



Real-time optimization

Fast real-time requirements



Low-cost computing platforms

Renewed interest in old methods (70s)

- Subgradient methods
- Proximal algorithms

- Cheap iterations
- Simple implementation

Contents of this course

Linear optimization

- Modeling and applications
- Geometry
- Simplex method
- Duality and sensitivity analysis

Large-scale convex optimization

- Modeling and applications
- Optimality conditions
- First-order methods
- Operator-splitting algorithms
- Acceleration schemes
- Computer-aided analysis

Nonconvex and stochastic optimization

- Sequential convex programming
- Branch and bound algorithms
- Robust optimization
- Distributionally robust optimization

Course information

Grading

- **30% Homeworks**
  5 bi-weekly homeworks with coding component. Collaborations are encouraged!

- **30% Midterm**
  90 minutes written exam in class. No collaborations.

- **40% Final**
  Take-home assignment with coding component. No collaborations.

Course information

Course website

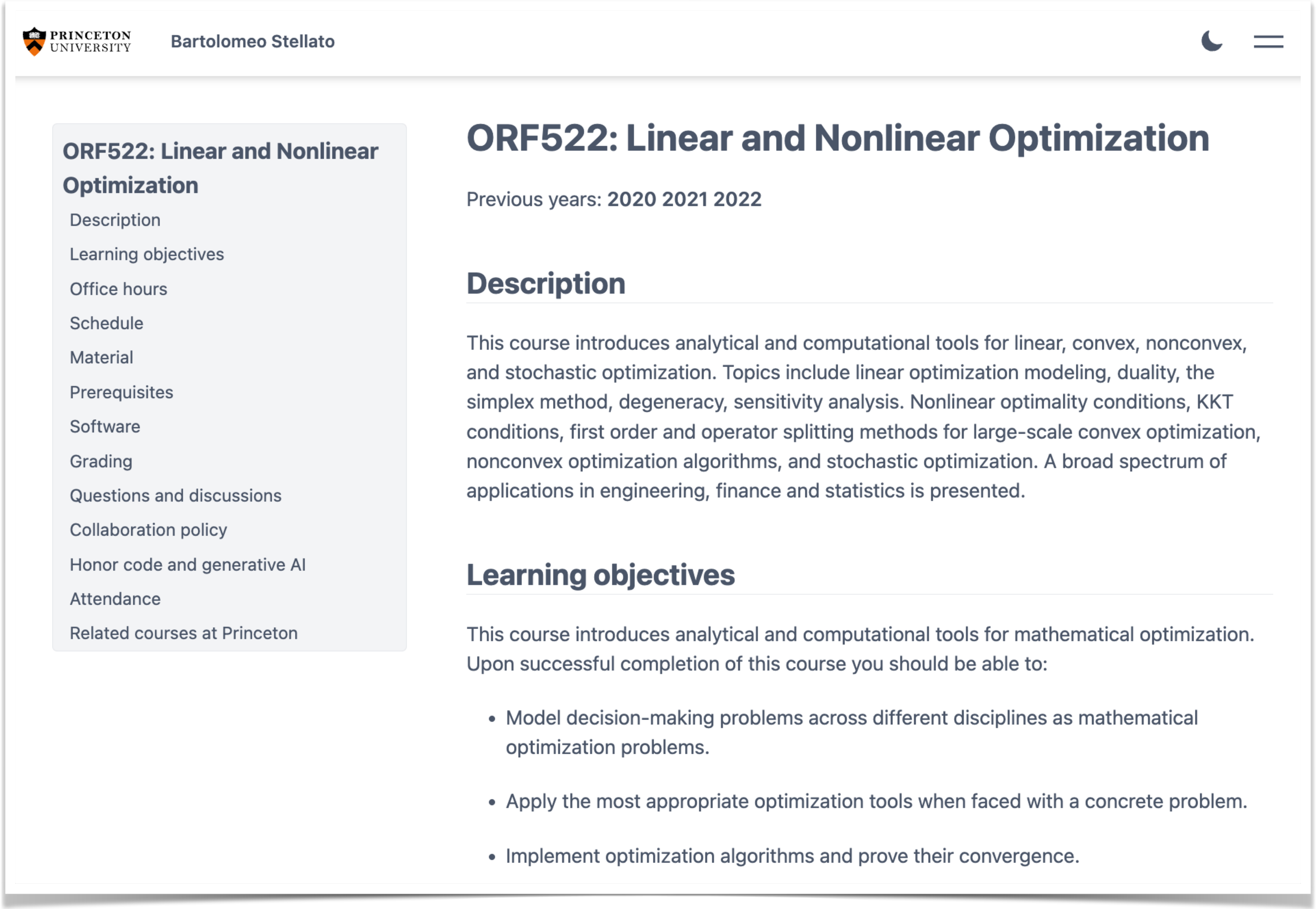
`https://stellato.io/teaching/orf522`

Prerequisites

- Good knowledge of linear algebra and calculus.

  For a refresher, read Appendices A & C of [CO] Boyd, Vandenberghe: *Convex Optimization* (available **online**).

- Familiarity with Python.


PRINCETON UNIVERSITY Bartolomeo Stellato

ORF522: Linear and Nonlinear Optimization
- Description
- Learning objectives
- Office hours
- Schedule
- Material
- Prerequisites
- Software
- Grading
- Questions and discussions
- Collaboration policy
- Honor code and generative AI
- Attendance
- Related courses at Princeton

ORF522: Linear and Nonlinear Optimization

Previous years: **2020 2021 2022**

Description

This course introduces analytical and computational tools for linear, convex, nonconvex, and stochastic optimization. Topics include linear optimization modeling, duality, the simplex method, degeneracy, sensitivity analysis. Nonlinear optimality conditions, KKT conditions, first order and operator splitting methods for large-scale convex optimization, nonconvex optimization algorithms, and stochastic optimization. A broad spectrum of applications in engineering, finance and statistics is presented.

Learning objectives

This course introduces analytical and computational tools for mathematical optimization. Upon successful completion of this course you should be able to:

- Model decision-making problems across different disciplines as mathematical optimization problems.
- Apply the most appropriate optimization tools when faced with a concrete problem.
- Implement optimization algorithms and prove their convergence.

Course information

Main books

Linear optimization

- [LP] R. J. Vanderbei: *Linear Programming: Foundations & Extensions* (available on **SpringerLink**)
- [LO] D. Bertsimas, J. Tsitsiklis: *Introduction to Linear Optimization* (available **Princeton Controlled Digital Lending**)

Large-scale convex optimization

- [NO] J. Nocedal, S. J. Wright: *Numerical Optimization* (available on **SpringerLink**)
- [CO] S. Boyd, L. Vandenberghe: *Convex Optimization* (available for **free**)
- [FMO] A. Beck: *First-order methods in optimization* (available on **SIAM**)
- [LSMO] E. K. Ryu and W. Yin: *Large-Scale Convex Optimization via Monotone Operators* (available for **free**)

Software (open-source)

Numerical computations

Numerical computations on *numpy* and *scipy*.

CVXPY

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$

```
x = cp.Variable(n)
prob = cp.Problem(
    cp.Minimize(c.T@x),
    [A @ x <= b]
)
prob.solve()
print("The optimal value is", prob.value)
print("The solution x is", x.value)
```

Learning goals

- **Model** your favorite decision-making problems as mathematical optimization problems.

- **Apply** the most appropriate optimization tools when faced with a concrete problem.

- **Implement** optimization algorithms and prove their convergence.

Glance into modern optimization

Huge scale optimization

Dataset with
billions of datapoints (x^i, y^i) $\longrightarrow$ **Goal:** Design predictor $\hat{y}^i = g_\theta(x^i)$

Optimization problem

Loss Regularizer

minimize $\underbrace{\mathcal{L}(\theta)}_{\text{Loss}} + \lambda \underbrace{r(\theta)}_{\text{Regularizer}} = \sum_{i=1}^n \ell(\hat{y}^i, y^i) + \lambda r(\theta)$

Many examples

- Support vector machines
- Regularized regression
- Neural networks

Large-scale computing

- Parallel
- Distributed

How large are the largest problems we can solve? (how many variables?)

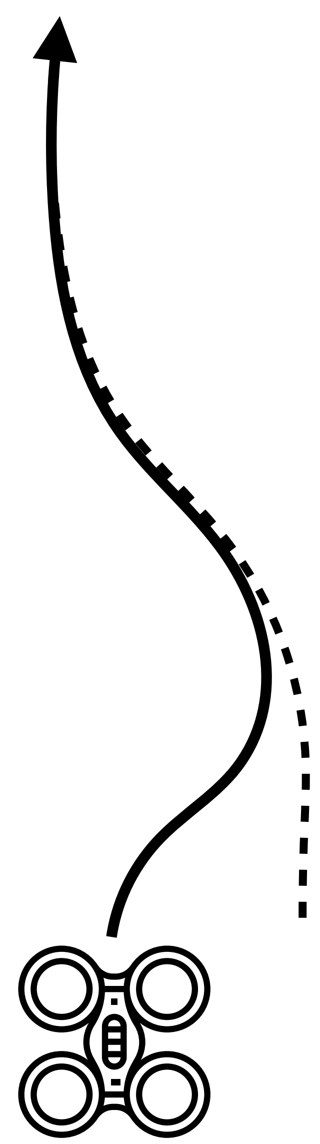
Glance into modern optimization

Real-time optimization

Dynamical system: $x_{t+1} = Ax_t + Bu_t$

$x_t \in \mathbf{R}^n$: state

$u_t \in \mathbf{R}^m$: input



Goal: track trajectory minimize $\sum_{t=0}^{T} \|x_t - x_t^{\mathrm{des}}\|$

Constraints: inputs $\|u\| \leq U$, states $a \leq x_t \leq b$

Solve and repeat…..

How fast can we solve these problems?

1-norm $\longrightarrow$???

∞-norm $\longrightarrow$???

Next lecture

Linear optimization

- Definitions
- Modelling
- Formulations
- Examples