ESE 415: Optimization
Making “best” decisions in a high-dimensional universe

Prof. Ulugbek Kamilov
Computational Imaging Group (CIG)
Today we will talk about

- Motivating example from CIG
  Optimization as a pillar of machine learning
- Brief overview of optimization
  What will we learn during this semester?
- Class structure
  Assignments, grades, assistant instructors, etc.
Today we will talk about

- Motivating example from CIG
  Optimization as a pillar of machine learning

- Brief overview of optimization
  What will we learn during this semester?

- Class structure
  Assignments, grades, assistant instructors, etc.
What are some examples of imaging systems?

MRI

Phone

Optical microscope

Depth sensor

Electron microscope

CT
Imaging systems are going through a paradigm shift with computation at its core

**Past:** Solely rely on hardware for image formation

**Present:** Use digital processing for improved performance

**Future:** Advanced inference for retrieving “hidden” information
Medical image quality directly depends on the amount of collected data

MRI in an ideal world where patients do not move

\[ b = A \times x \]

full sampling = as many measurements as unknown image pixels
Medical image quality directly depends on the amount of collected data

MRI in an ideal world where patients do not move

high-quality image
Medical image quality directly depends on the amount of collected data

\[ b = A x \]

undersampling = less measurements than unknown image pixels

MRI in a real world where patients breath and move
Medical image quality directly depends on the amount of collected data.

MRI in an ideal world where patients do not move:

- High-quality image

MRI in a real world where patients breath and move:

- Noisy image
Medical image quality directly depends on the amount of collected data

MRI in an ideal world where patients do not move

MRI in a real world where patients breath and move

Question: How to compensate for the missing data?
A modern machine learning solution: 
Train a deep neural net on a dataset

Assume that we have a dataset consisting of noisy images with their clean counterparts

\[ \{y_i\}_{i=1,\ldots,n} \]

\[ x_i = f_\theta(y_i) \]

\[ \{x_i\}_{i=1,\ldots,n} \]

Question: How to “train” the deep neural net?
A modern machine learning solution: Train a deep neural net on a dataset

Assume that we have a dataset consisting of noisy images with their clean counterparts

\[ \{y_i\}_{i=1,...,n} \]

\[ x_i = f_\theta(y_i) \]

Training is an optimization of the weights in the neural net using some fast and reliable algorithm

\[ \theta^* = \arg \min_{\theta} \left\{ \frac{1}{n} \sum_{i=1}^{n} \|x_i - f_\theta(y_i)\|^2 \right\} \]

find weights that produce the smallest average prediction error
A modern machine learning solution:
Train a deep neural net on a dataset
A modern machine learning solution:
Train a deep neural net on a dataset

evolution of the cost across iterations
Today we will talk about

- Motivating example from CIG
  Optimization as a pillar of machine learning

- Brief overview of optimization
  What will we learn during this semester?

- Class structure
  Assignments, grades, assistant instructors, etc.
The goal of optimization is to find the smallest value of a function under constraints.

Mathematical optimization problem

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad x \in \mathcal{X}
\end{align*}
\]

- Optimization variable: \( x = (x_1, \ldots, x_n) \)
- Objective function: \( f : \mathbb{R}^n \rightarrow \mathbb{R} \)
- Constraint set: \( \mathcal{X} = \{x \mid h(x) = 0\} \cap \{x \mid g(x) \leq 0\} \)
- Optimal solution \( x^* \) has smallest value of \( f \) among all vectors that satisfy the constraints
Sample applications of optimization: finance, engineering, and computer science

**portfolio optimization**
- variables: amounts invested in different assets
- constraints: budget, minimum return
- objective: overall risk or return variance

**device sizing in electronic circuits**
- variables: device width and lengths
- constraints: manufacturing limits, timing requirements
- objective: power consumption

**machine learning**
- variables: model parameters
- constraints: prior information, parameter limits
- objective: prediction error
Brief history of optimization

Antiquity: Greek mathematicians solve some optimization problems that are related to their geometrical studies

- Euclid (300bc): minimal distance between a point and a line
- Heron (100 bc): light travels between two points through the path with shortest length when reflecting from a mirror

17th-18th: before the invention of calculus of variations only some separate optimization problems are being investigated

- Kepler (1615): Optimal dimensions of wine barrel. Formulated the secretary problem when searching for a new wife
- Fermat: derivative of a function vanishes at the extreme point the (1636). Shows that light travels between two points in minimal time (1657)

Newton (1660s) and Leibniz (1670s) create mathematical analysis that forms the basis of calculus of variations

http://www.mitrikitti.fi/opthist.html
Brief history of optimization

19th century: the first optimization algorithms are presented
- Legendre (1806): least square method, which also Gauss claims to have invented
- Cauchy (1847): presents the gradient method
- Gibbs shows that chemical equilibrium is an energy minimum

20th century: the field of algorithmic research expands as electronic calculation develops
- von Neuman and Morgenstern (1944): dynamic programming for solving sequential decision problems
- Dantzig (1947): simplex method for solving LP-problems
- Kuhn and Tucker (1951): reinvent optimality conditions for nonlinear problems. Similar conditions in 1939 by Karush.
- Modern era: nonsmooth analysis, stochastic optimization...

http://www.mitrikitti.fi/opthist.html
Today we will talk about

- Motivating example from CIG
  Optimization as a pillar of machine learning
- Brief overview of optimization
  What will we learn during this semester?
- Class structure
  Assignments, grades, assistant instructors, etc.
Practical information

Course website: cigroup.wustl.edu/teaching/ese415-2019/

Schedule:
Lectures: Tue and Thu at 1:00-2:30 pm
Additional tutorials: TBD

Optional reading to complement lecture notes:
By the end of the semester, hopefully, you will be able to:
• recognize and formulate problems as optimization
• characterize and understand optimal solutions
• develop code for optimization algorithms

Topics:
• optimality conditions
• convex sets and functions
• constrained and unconstrained optimization
• optimization algorithms
• analysis of optimization algorithms
• examples and applications
Practical information

Grading policy:
• homework (40%) consists of 6 assignments
• midterm (30%) is on Thursday, 7 March 2019
• final (30%): is on Tuesday, 7 May 2019

Grading convention:
• 93-100 A
• 90-92 A-
• 88-89 B+
• 83-87 B
• 80-82 B-
• 70-79 C
• 60-69 D
• below 60 F
Homework assignments will be collected only via Gradescope (use code: M2P3R8)

Assignment #0 is out and due this Thursday!

It is mandatory, but you will not get graded on it

Academic integrity policy:
You are encouraged to discuss with colleagues, but you must submit your own individual homework. No copying from colleagues will be tolerated in this class.
Practical information

Head assistant instructors:
• Xiaojian Xu (xiaojianxu@wustl.edu)
• Hesam Mazidi (hmazidi@wustl.edu)
• Yu Sun (sun.yu@wustl.edu)

The rest of the team:
• Yiran Sun
• Weiran Wang
• Jiaming Liu
• Joseph Han
• Yifan Wang
• Fa Long
• Guancheng Jiang
To conclude

- Optimization is extensively used in almost all engineering applications
- The goal of ESE 415 is to build solid basics that you can rely on for the rest of your career
- Optimization is still an active research area with many open questions
Computational Imaging Group (CIG) at Washington University in St. Louis (WashU)

**CONTACT INFO**

Prof. Ulugbek Kamilov  
Email: kamilov@wustl.edu  
Personal Twitter: @ukmlv

Web: http://cigroup.wustl.edu  
Group Twitter: @wustlcig

If you are interested in doing research at the intersection of Computational Imaging, Machine Learning, and Optimization contact me or my PhD students