

Advanced Topics in Convex Optimization and Learning: Oracle, Online and Stochastic Models

מספר הקורס - 236606

מרצה:	פרופ'ח נתי סרברו
מתרגל/בודק תרגילים:	
שעות הרצאה:	Thursday 13:30-15:30
דרישות קדם:	Linear Algebra, Calculus 2, Algorithms 1, Probability.
אתר הקורס: (כתובת האתר)	

תאור הקורס

The course is intended as both a mathematical introduction to Convex Optimization, which also lies at the basis of much of modern Machine Learning, and a practical tour of contemporary methods in large scale optimization, including online and stochastic optimization/learning. We will rigorously define and study the complexity of optimization problems, providing both algorithms and upper bounds, as well as lower bound analysis. We will be particularly interested in big data / large scale regimes, and accordingly focus on first order methods and stochastic methods. We will study local oracle models for convex optimization, as well as online optimization and stochastic optimization models, understanding the relationship to machine learning.

The course will complement CS236330 "Introduction to Optimization", in that it will focus more on rigorous analysis of optimization problems, and thus only convex optimization, and in that it will focus on first-order large-scale methods, and encompass also online and stochastic optimization and learning.

Taking CS236330 before this course will allow students to gain a broader view and more insight into convex optimization, but is not required--we will define and develop convex optimization from the ground up.

דרישות הקורס

Written homeworks and either project or exam.

Topics:

- What is an optimization problem and how do we measure its complexity? Oracle models, local and first order oracles.
- The importance of convexity and a brief introduction to convex analysis: convex sets and functions, separating and supporting hyperplanes, subgradients, the Fenchel conjugate.
- Optimality conditions and duality.
- Classic unconstrained and optimization: gradient methods, Newton's method, optimality of the conjugate gradient method for quadratic functions.
- Historical overview of low-dimensional optimization complexity: center of mass, ellipsoid and interior point methods.
- Alternate geometries and Mirror Descent: Potential functions, Bregman divergences and prox maps.
- "Modern" first-order large-scale methods: projected gradient descent / mirror descent and dual averaging for non-smooth or non-strongly-convex problems, Nesterov acceleration, smoothing, FISTA and variants.
- Online optimization and regret analysis.
- Stochastic optimization: machine learning as stochastic optimization, stochastic gradient descent / mirror descent, Sample Average Approximation/Empirical Risk Minimization versus Stochastic Approximation, sample complexity and stochastic oracle complexity.

ספרות

R.T. Rockafeller: "Convex Analysis"

A Ben-Tal and A Nemirovski: "Lecture on Modern Convex Optimization"

Y Nesterov: "Introductory Lectures on Convex Optimization, A Basic Course"