EPFL

EE-556   **Mathematics of data: from theory to computation**

Cevher Volkan

Cursus	Sem.	Type
Computational science and Engineering	MA1, MA3	Opt.
Data Science	MA1	Opt.
Electrical and Electronical Engineering	MA1, MA3	Obl.
Managmt, tech et entr.	MA1, MA3	Opt.

Language of teaching	English
Credits	4
Session	Winter
Semester	Fall
Exam	During the semester
Workload	120h
Weeks	14
Hours	**4 weekly**
Courses	2 weekly
Exercises	2 weekly
Number of positions	

Summary

This course reviews recent advances in convex optimization and statistical analysis in the wake of Big Data. We provide an overview of the emerging convex formulations and their guarantees, describe scalable solution techniques, and illustrate the role of parallel and distributed computation.

Content

The course consists of the following topics
Lecture 1: "Objects in Space": Definitions of norms, inner products, and metrics for vector, matrix and tensor objects. Basics of complexity theory.
Lecture 2: Maximum likelihood principle as a motivation for convex optimization. Fundamental structures in convex analysis, such as cones, smoothness, and conjugation.
Lecture 3: Unconstrained, smooth minimization techniques. Gradient methods. Variable metric algorithms. Time-data tradeoffs in ML estimation.
Lecture 4: Convex geometry of linear inverse problems. Structured data models (e.g., sparse and low-rank) and convex gauge functions and formulations that encourage these structures. Computational aspects of gauge functions.
Lecture 5: Composite convex minimization. Regularized M-estimators. Time-data tradeoffs in linear inverse problems.
Lecture 6: Convex demixing. Statistical dimension. Phase transitions in convex minimization. Smoothing approaches for non-smooth convex minimization.
Lecture 7: Constrained convex minimization-I. Introduction to convex duality. Classical solution methods (the augmented Lagrangian method, alternating minimization algorithm, alternating direction method of multipliers, and the Frank-Wolfe method) and their deficiencies
Lecture 8: Constrained convex minimization-II. Variational gap characterizations and dual smoothing. Scalable, black-box optimization techniques. Time data-tradeoffs for linear inverse problems.
Lecture 9: Classical black-box convex optimization techniques. Linear programming, semidefinite programming, and the interior point method (IPM). Hierarchies of classical formulations. Time and space complexity of the IPM.
Lecture 10: Time-data tradeoffs in machine learning.
Lecture 11: Convex methods for Big Data I: Randomized coordinate descent methods. The Page Rank problem and Nesterov's solution. Composite formulations.
Lecture 12: Convex methods for Big Data II: Stochastic gradient descent methods. Least squares: conjugate gradients vs. a simple stochastic gradient method. Dual and gradient averaging schemes. Stochastic mirror descent.
Lecture 13: Randomized linear algebra routines for convex optimization. Probabilistic algorithms for constructing approximate low-rank matrix decompositions. Subset selection approaches. Theoretical approximation guarantees.
Lecture 14: Role of parallel and distributed computing. How to avoid communication bottlenecks and synchronization. Consensus methods. Memory lock-free, decentralized, and asynchronous algorithms.

Learning Prerequisites

 Important concepts to start the course

Previous coursework in calculus, linear algebra, and probability is required.
Familiarity with optimization is useful.

Learning Outcomes

By the end of the course, the student must be able to:

- Choose an appropriate convex formulation for a data analytics problem at hand
- Estimate the underlying data size requirements for the correctness of its solution
- Implement an appropriate convex optimization algorithm based on the available computational platform
- Decide on a meaningful level of optimization accuracy for stopping the algorithm
- Characterize the time required for their algorithm to obtain a numerical solution with the chosen accuracy

Assessment methods

Homework assignments. (Continuous control)