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, MA3	Opt.
Electrical Engineering		Opt.
Electrical and Electronical Engineering	MA1, MA3	Opt.
MNIS	MA3	Opt.
Managmt, tech et entr.	MA1, MA3	Opt.

Language of teaching	English
Credits	6
Session	Winter
Semester	Fall
Exam	Written
Workload	180h
Weeks	14
Hours	**6 weekly**
Courses	3 weekly
TP	3 weekly
Number of positions	

Summary

This course reviews recent advances in continuous optimization and statistical analysis along with models. We provide an overview of the emerging learning 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: Introduction. The role of models and data. Maximum-likelihood formulations. Sample complexity bounds for estimation and prediction.
Lecture 2: The role of computation. Challenges to optimization algorithms. Optimality measures. Structures in optimization. Gradient descent. Convergence rate of gradient descent.
Lecture 3: Optimality of convergence rates. Accelerated gradient descent. Concept of total complexity. Stochastic gradient descent.
Lecture 4: Concise signal models. Compressive sensing. Sample complexity bounds for estimation and prediction. Challenges to optimization algorithms for non-smooth optimization.
Lecture 5: Introduction to proximal-operators. Proximal gradient methods. Linear minimization oracles. Conditional gradient method for constrained optimization.
Lecture 6: Time-data trade-offs. Variance reduction for improving trade-offs.
Lecture 7: A mathematical introduction to deep learning. Double descent curves and over-parameterization. Implicit regularization.
Lecture 8: Structures in non-convex optimization. Optimality measures. Escaping saddle points. Adaptive gradient methods.
Lecture 9: Adversarial machine learning and generative adversarial networks (GANs). Wasserstein GAN. Difficulty of minimax optimization.
Lecture 10: Primal-dual optimization-I: Fundamentals of minimax problems. Pittfalls of gradient descent-ascent approach.
Lecture 11: Primal-dual optimization-II: Extra gradient method. Chambolle-Pock algorithm. Stochastic primal-dual methods.
Lecture 12: Primal-dual III: Lagrangian gradient methods. Lagrangian conditional gradient methods.
Recitation 1: Generalized linear models. Logistic regression.
Recitation 2: Computation of Gradients. Reading convergence plots. Helpful definitions on linear algebra.
Recitation 3: Activation functions in neural networks. Backpropogation. Introduction to pytorch.

Keywords

Machine Learning. Signal Processing. Optimization. Statististical Analysis. Linear and non-linear models. Algorithms. Data and computational trade-offs.

Learning Prerequisites

Required courses

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

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)