Postgraduate Course on Optimization Techniques for Big Data Analysis (MSc)

Instructor Information

Santiago Zazo

E-mail: santiago.zazo@ upm.es
Work Phone: +34 667451802

Course Information

Course Description

This course deepens in the basic concepts of optimization presented in the *'Optimization Fundamentals'* course now dealing with those specific aspects appearing in Big Data analysis. We will review the main problems in Machine Learning as regression and classification and will present the more popular distributed / parallel implementations as coordinate descent, consensus, diffusion and ADMM.

Prerequisites

Fundamentals of constrained and unconstrained optimization.

Fundamentals of Machine Learning.

A working knowledge of MATLAB and CVX.

Course Goal

The aim of the course is provide a perspective of different algorithms applicable to Big Data Analysis discussing their main characteristics as convergence rate and complexity emphasizing their capabilities of parallel and distributed implementations. Most of them will be implemented by the students in Matlab in order to make sure that the students apprehend all details.

Summary of intended course outcomes

At the end of the course a student will have used different optimization algorithms for Big Data Analysis, and will be able to:

- Reduce the dimensionality of a certain problem.
- Implement local algorithms using deterministic / stochastic optimization procedures. Low complexity and iterative techniques are emphasized.
- Implement parallelized and distributed algorithms analyzing convergence rate.

Syllabus

I Fundamentals of Large Scale Optimization

1. Introduction

2. Review of fundamentals of convex optimization
2.1. Review of convex sets and convex functions
2.2. Review of smooth / non smooth functions
2.3. Review of iterative / adaptive algorithms

3. Machine Learning Contextualization
3.1. Linear regression (Ridge regression)
3.2. Linear regression with variable selection (LASSO)
3.3. Classification
3.4. Basis pursuit

4. Data Dimensionality Reduction
4.1. Principal Component Analysis (PCA)
4.2. Randomized projections. Johnson-Lindenstrauss lemma

5. Overview of optimization procedures

II Large Scale Optimization. Local processing in parallel settings.

6. Gradient descent
6.1. Gradient descent. Smooth functions
6.2. Accelerated First Order Methods
6.3. Gradient descent. Non smooth functions

7. Second Order Methods. Quasi-Newton implementations
7.1. Conjugate Gradient & Solving Linear Systems
7.2. Quasi-Newton methods

8. Augmented Lagrangian Methods
8.1. Precursor 1. Dual Ascent
8.2. Dual decomposition
8.3. Augmented Lagrangian and the Method of Multipliers
8.4. Alternating Direction Method of Multipliers (ADMM)

III Parallelized / Distributed Large Scale Convex Optimization.

9. ADMM: a Distributed implementation using Consensus principles.

9.1. General formulation
9.2. Distributed implementations using Consensus principles. Regularization

10. In-Network optimization
10.1. Single agents stochastic gradient optimization
10.2. Multiple agents stochastic gradient optimization

11. Efficient parallelized optimization for convex / non convex function
11.1. Coordinate descent. Smooth functions
11.2. Basic model to manage non convexities
11.3. Study case. Dictionary learning

Suggested readings

1. Boyd, S., L. Vanderberghe. Convex optimization. Cambridge University Press 2004.
2. Boyd, S., N. Parikh, E. Chu, B. Peleato, J. Eckstein. Distributed Optimization and Statistical Learning Via the Alternating Direction Method of Multipliers. Foundations and Trends in Machine Learning, Vol. 3, No1, 2010.
3. Parikh,N., S. Boyd: Proximal Algorithms. Foundations and Trends in Machine Learning, Vol. 1, No3, 2013.
4. Sra, S., S. Nowocin, S.J. Wright. Optimization for Machine Learning. The MIT Press, 2012.
5. Sayed, A.H. Adaptation, Learning and Optimization over Networks. Foundations and Trends in Machine Learning, Vol. 7, No4-5, 2014.
6. Several papers to be recommended along the course

Student Assessment Criteria

A series of short individual projects will be assigned throughout the semester involving the development of Matlab computer code to simulate and optimize problems. Two final projects, one in the middle of the course and other after finishing it will include some short questions and programming a certain problem similar as those developed along the course.

Along the course Matlab Code Assignments	40%
Intermediate project	30%
Final project	30%