

Optimization for big Data CM

Course title - Intitulé du cours	Optimization for big Data CM
Level / Semester - Niveau /semestre	M1 / S2
School - Composante	Ecole d'Economie de Toulouse
Teacher - Enseignant responsable	GADAT Sebastien
Lecture Hours - Volume Horaire CM	15
TA Hours - Volume horaire TD	0
TP Hours - Volume horaire TP	12
Course Language - Langue du cours	Anglais
TA and/or TP Language - Langue des TD et/ou TP	Anglais

Teaching staff contacts - Coordonnées de l'équipe pédagogique :

sebastien.gadat@math.univ-toulouse.fr

Office: T586

Office hours: monday afternoon - Send an email first for an eventual appointment!

Course's Objectives - Objectifs du cours :

Objectives:

The "Optimisation for big data" lecture is an introductory course that describes some modern algorithms useful in statistics and optimisation. These algorithms will pay specific attention to the high dimensional framework involved by big data and the associated computation time.

In the nowadays big data era, old fashioned statistical or econometrics methods become useless for possibly different reasons: high dimensional settings, on-line learning or difficult non euclidean structure. It induces serious troubles with simple regression methods (among other linear model and logistic regression) and to bypass these difficulties, some efforts are needed both on the algorithmic side and on the statistical side. The goal of this course is to present some recent advances on optimization that helps for solving numerical problems derived from statistical modeling (among other L1 penalized mean square or logistic loss, on-line learning).

1) Deterministic optimization

- 1-a) Reminders on convexity
- 1-b) Strong convexity algorithms (beyond MSE)
- 1-c) Duality (beyond classification)
- 1-d) Non smooth optimisation (beyond linear model)

1-e) What goes wrong with old stats/ecox methods? Do old pipes give the sweetest smoke everytime? Beyond old fashioned methods... Lasso and SVM

2) On-line optimization

2-a) Stochastic optimization (Sequential approaches)

2-b) Optimization under uncertainty (Bandit problems)

3) Optimisation with structural data: Graphs

Prerequisites - Pré requis :

Requirements:

Mathematical statistics (1&2)

Convex analysis

Practical information about the sessions - Modalités pratiques de gestion du cours :

Laptops are accepted

Students are expected to ask questions during the lecture and the practical sessions!

Several homeworks in groups of 2 or 3 students

Grading system - Modalités d'évaluation :

final mark: aggregation of several homeworks.

Bibliography/references - Bibliographie/références :

References:

S. Bubeck, 2015, Convex Optimization: Algorithms and Complexity, Foundations and Trends in Machine Learning, Vol. 8, No. 3-4 (2015) 231–357

Y. Nesterov. Introductory lectures on convex optimization: A basic course. Kluwer Academic Publishers, 2004a.

R. Rockafellar. Convex Analysis. Princeton University Press, 1970.

Lecture notes available on moodle

Practical sessions available on moodle

Distance learning – Enseignement à distance :

En cas de nécessité, un enseignement à distance sera assuré en mobilisant des séances de cours en visio par zoom aux horaires habituels tels que définis dans l'emploi du temps. Nous effectuerons également des tds et tps à distance mis en ligne sur moodle.

