

Convex Optimization: Fall 2013

Machine Learning 10-725/Statistics 36-725

Instructors:

Barnabas Poczos, Dept. of Machine Learning, bapoczos@cs.cmu.edu

Ryan Tibshirani, Dept. of Statistics, ryantibs@stat.cmu.edu

Teaching assistants:

Adona Iosif, aiosif@cs.cmu.edu

Yifei Ma, yifeim@cs.cmu.edu

Aaditya Ramdas, aramdas@cs.cmu.edu

Sashank Reddi, sjakkamr@cs.cmu.edu

Lectures: Tuesdays and Thursdays 3-4:20pm, Doherty Hall 1212

Overview and objectives

Nearly every problem in machine learning and statistics can be formulated in terms of the optimization of some function, possibly under some set of constraints. As we obviously cannot solve every problem in machine learning or statistics, this means that we cannot generically solve every optimization problem (at least not efficiently). Fortunately, many problems of interest in statistics and machine learning can be posed as optimization tasks that have special properties—such as convexity, smoothness, separability, sparsity etc.—permitting standardized, efficient solution techniques.

This course is designed to give a graduate-level student a thorough grounding in these properties and their role in optimization, and a broad comprehension of algorithms tailored to exploit such properties. The main focus will be on convex optimization problems, though we will also discuss nonconvex problems at the end. We will visit and revisit important applications in statistics and machine learning. Upon completing the course, students should be able to approach an optimization problem (often derived from a statistics or machine learning context) and:

- (1) identify key properties such as convexity, smoothness, sparsity, etc., and/or possibly reformulate the problem so that it possesses such desirable properties;
- (2) select an algorithm for this optimization problem, with an understanding of the advantages and disadvantages of applying one method over another, given the problem and properties at hand;
- (3) implement this algorithm or use existing software to efficiently compute the solution.

Outline of material

Here is an rough outline of the course topics.

Core topics:

- Basics of convex analysis
- First-order methods
- Second-order and related methods
- Duality and KKT conditions
- Interior point methods
- Quadratic, semidefinite, cone programs

Advanced topics:

- Uses of duality, dual methods
- Coordinate-based methods
- Nonconvex methods
- Large-scale optimization

Logistics

Prerequisites: Students entering the class are expected to have a pre-existing working knowledge of algorithms; however, the class has been designed to allow students that do not have this background, but have strong analytical skills, to catch up and fully participate. Students are expected to be comfortable with rigorous mathematical arguments. Though not required, having taken 10-701 or an equivalent machine learning or statistics class will be helpful, since we will frequently use applications in machine learning and statistics to demonstrate the concepts we learn in class.

Class website: The class website is <http://www.stat.cmu.edu/~ryantibs/convexopt/>. The class schedule, lecture notes, homeworks, etc., will be posted there.

Attendance: Attendance at lectures is highly encouraged. We think you will learn more by coming to lectures, paying attention, and asking questions. Plus, it will be more fun.

Recitations: Recitations are Wednesdays 4:30-6pm in Gates 4307. Attendance at recitations is optional, but recitations can be a great way for you to hear a second take on the lecture material and relevant background material, in a (likely) more private setting.

Office hours: The weekly schedule for office hours is given below.

BP: Thursdays 4:30-5:30pm, Gates 8231

RT: Wednesdays 10-11am, Baker 229B

AI: Wednesdays 3-4pm, Gates 8221

YM: Mondays 4:30-5:30pm, Gates 8223

AR: Tuesdays 4:30-5:30pm, Gates 8015

SR: Mondays 2-3pm, Gates 8223

Discussion group: We will use the discussion board through blackboard.

Anonymous feedback: We are setting up a system through blackboard that will allow students to provide anonymous feedback for the course instructors and TAs at any time throughout the semester. Details to come, once this has been set up.

Textbook: Lectures are intended to be mostly self-contained. But a terrific reference for most of the material covered in class is *Convex Optimization* by Boyd and Vandenberghe, which is available online for free (<http://www.stanford.edu/~boyd/cvxbook/>). We will often draw examples or problems from this book. Alternatively, another good general reference *Nonlinear Programming* by Bertsekas. Finally, a wonderfully thorough mathematical reference for much of the underlying convex analysis we cover is *Convex Analysis* by Rockafellar.

Homework: There will be 5 homework assignments, approximately one every two weeks. The assignments will be posted on the course website, and your homeworks will be submitted electronically through blackboard. You can submit your homework anytime before 12am (midnight) on its due date. **Late homework will not be accepted.**

The homeworks are structured to give you experience in both written mathematical exercises and programming exercises. As we will often reuse problems from previous versions of the course, or problems covered by papers and webpages, we expect you not to copy, refer to, or look at solutions in preparing your answers. Also, while it is completely acceptable for you to collaborate with other students in order to solve the problems, we assume that you will be taking full responsibility in terms of writing up your own solutions and implementing your own code. You must indicate on each homework the students with whom you collaborated.

Midterm: There will be one in-class midterm, on November 5. This will cover about the first 2/3 of the course material (the core topics, before the advanced topics).

Little test: There will be one final little test, on December 3. This will consist of multiple choice and true false questions, and will cover everything learned in the course.

Final project: You have the option to complete a final project. If you choose to do the final project, you should enroll in the 12 unit option for the course; those not doing the final project must enroll in the 9 unit option. Details on the project to come.

Scribing: Students are required to scribe for at least one lecture. There is a minimum of two scribes per lecture, though three is the recommended number, and a fourth is optional—if all the lectures have all three scribe slots taken, then you can sign up as the fourth scribe for any lecture of your choice. You can sign up for scribing lectures on the class website.

Evaluation: The grading breakdown is as follows.

Homework	45%
Scribing	2%
Midterm	15%
Little test	8%
Final project	30%

(For those of you enrolled in the 9 unit option with no project, just subtract the 30% for the project and renormalize.)

TA roles: Of course you can ask any of us questions, but if you have specific questions about the following topics, it would be best to consult the corresponding TA.

Adona: discussion group, anonymous feedback system

Yifei: lecture videos

Aaditya: mastery sets, homework policies

Sashank: scribing