APPM 4720/5720 Advanced Topics in Convex Optimization Fall 2018

www.colorado.edu/amath/appm-47205720-special-topics-convex-optimization-fall-2018 and github.com/stephenbeckr/convex-optimization-class

Instructor and contacts: Stephen Becker, ECOT 231, (303) 492-0662, *stephen.becker@colorado.edu*. I will respond to most reasonable email, but (1) lengthy questions should be posed at office hours and not via email, and (2) I will not check email at odd hours or just before homework is due.

Meeting times: Monday, Wednesday and Friday, DUAN G131, 10:00 AM to 10:50 AM.

Office hours: Three hours per week, Wednesdays 3-5 PM and Thursday 3-4 PM.

Learning Goals: The course investigates landmark convex optimization algorithms and their complexity results. We aim to cover both key theory, such as polynomial time algorithms for linear programming and the optimality of Nesterov's method, while also surveying current practical state-of-the-art methods.

Prereqs: APPM 3310 "Matrix Methods" or similar; but APPM 4440 "Real Analysis" or equivalent (e.g., APPM 5440) is suggested. A first-course on optimization like CSCI 5254 "Convex Optimization" is helpful but not necessary.

Text: See next pages for details

Syllabus: See next pages for detailed syllabus

Recitations: There are no recitation session.

Exams: There are no exams (no midterms, no final)

Project: (25% of grade) The last few weeks, there will be no/less homework and instead there is a final project, which is open-ended. Students are encouraged to form groups (up to three people). It can be either theory or computation. Topics should be discussed with the instructor. The final week, groups will give brief (about 10 min) presentations on their topic.

Grades/Homeworks: (75% of grade) Your grade will be entirely based on homeworks and a project. There will be about 6 double-homeworks (due every two weeks). Each "double-homework" will consist of two parts (e.g., one part may be theory, one part computation).

1. students taking the course for graduate credit (5720 instead of 4720) may drop **two** single homeworks (e.g., one full double-homework, or two parts from different double-homeworks)
2. students taking the course for undergraduate credit (4720) may drop **four** single homeworks
3. ... however, regardless of whether you take the course for undergrad or grad credit, you must complete **all** homeworks to be eligible for a grade of "A" (if you drop a homework, then your highest possible grade is "A-").
4. Your final homework score, worth 75% of the grade, will be averaged over the homeworks you do not drop.

Late homework policy: Homework is due at the beginning of class.

Late homework will not be accepted or graded.

Course web page: (course website: https://tinyurl.com/y7sfylhz) It is your responsibility to check the web page on a regular basis. Here you will find detailed information such as homework assignments and solutions, office hours, and special announcements. In addition, it contains policies on illness, academic honesty, and special accommodations for religious holidays and documented special needs. We may also use Canvas in addition to the main course website, mainly for grades.

Dropping the course: Advice from your department advisor is recommended before dropping any course. After Sep. 12, dropping a course results in a "W" on your transcript. After November 2, dropping the course is possible only with a petition approved by the Dean's office.

Syllabus

Since there is no established optimization curriculum at CU Boulder, we will attempt to cover a lot of material in just one semester. We can divide topics into four broad categories:

Analysis Basic theory (convex analysis, optimality, duality)

Methods Standard methods, both classical and modern, and standard classifications. The idea is to give students a toolkit of standard approaches to solve actual problems.

Convergence Convergence theory (ellipsoid, Nesterov, etc.). This involves some very beautiful mathematics (and is also useful!)

Applications Often in signal processing and machine learning

Details

Section 1: Background, duality, motivation

1. Intro
2. Convex Sets (ch 2 in [BV2004])
3. Convex Functions (ch 3 in [BV2004])
4. Convex Analysis, subdifferentials ([BauschkeCombettes])
5. Convex Optim. Problems (ch 4 in [BV2004])
6. Duality, optimality (ch 5 in [BV2004])
7. Some applications, ERM and machine learning, sparse recovery, image denoising, low-rank matrix recovery
8. `CVX`, `CVXPY`, `convex.jl` tutorials

Section 2: Conic programs & relaxations

1. Linear programming
   a) standard LP theory, duality ([WrightIPM])
   b) simplex method, complexity, smoothed analysis ([Vanderbei_LP], [Nemirovski_LP])
   c) IPM ([WrightIPM], [BV2004])
   d) ellipsoid method (TBD, [Nemirovski_LP])
2. Integer Linear programming ([Vanderbei_LP])
3. SDP ([BV2004], Pataki?)
4. QP, convex and non-convex (S-procedure, relaxations) ([Nemirovski_LP],[BN_Modern], [BV2004])
5. MAXCUT relaxation ([Nemirovski_LP],[BN_Modern])
6. Polynomial optimization, SOS ([Nemirovski_LP],[BN_Modern])

Section 3: Algorithms

1. Gradient Descent ([V_course])
2. Accelerated and projected/proximal gradient descent, optimality ([V_course])
3. Classical methods ([NW05], Bertsekas, [Nesterov2004])
   a) Subgradient descent
   b) Newtons method, self-concordancy
   c) Non-linear conjugate gradient
   d) quasi-Newton
   e) Levenberg-Marquardt, Gauss-Newton

f) Augmented Lagrangian
g) Sequential Quadratic Programming (SQP)
h) Step-size (Barzilai-Borwein), line search

4. Recent methods
   a) Douglas-Rachford and ADMM
   b) Primal-dual methods (e.g., Condat)
   c) Conditional Gradient and Frank-Wolfe
   d) (Randomized) Coordinate Descent
   e) Stochastic gradient descent
   f) State-of-the-art variants (SDCA, SAGA, etc.)

Possible additional topics

1. Benders decomposition
2. Geometric Programming, linear-fractional
3. Online Convex Optimization
4. Mirror Descent
5. Bregman distances
6. Robust optimization
7. Randomized Kaczmarz method; Algebraic Reconstruction Technique (ART) for tomography
8. POCS (projection onto convex sets), best approximation, Dykstra's algorithm

Texts and references

We will not follow a single textbook. The following resources (mostly available online) will be the primary references:

[BV2004] S. Boyd and L. Vandenberghe, "Convex Optimization" (Cambridge U. Press, 2004). Free electronic version at `www.stanford.edu/~boyd/cvxbook/`. This is a standard text that gives some convex analysis background and focuses on interior-point methods and engineering applications.

[B_course]S. Boyd's course notes for EE364a, which follow his book.

[V_course]L. Vandenberghe's course notes for ee236b follow the book, while his course notes for ee236c contain more details on proofs and advanced topics.

[NW05] J. Nocedal and S. Wright, "Numerical Optimization" (Springer, 2005). We have free electronic access at CU via SpringerLink. This is the go-to reference for implementing a standard method.

[Beck17] Amir Beck, "First-Order Methods in Optimization" (SIAM, 2017). A more advanced text than Beck's 2014 text; free on campus at SIAM eBOoks

Some advanced topics will be pulled from these resources:

[BubeckPDF] Sebastien Bubeck, "Convex Optimization: Algorithms and Complexity" (2015 monograph, free at `http://research.microsoft.com/en-us/um/people/sebubeck/Bubeck15.pdf`). This covers a very modern perspective

[BubeckBLOG] Sebastien Bubeck's blog for his ORF523 course (similar to his monograph): `https://blogs.princeton.edu/imabandit/orf523-the-complexities-of-optimization/`

[Beck14] Amir Beck, "Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB" (SIAM, 2014, $89). This covers classical results but by a modern researcher aware of current research

[BN_Modern] Ben-Tal and Nemirovski, Lectures on Modern Convex Optimization (the PDF is 2013; they have a SIAM book from 2001 as well)

[Nemirovski_LP] A. Nemirovski, Introduction to Linear Optimization. More lectures notes and monographs of A. Nemirovski available at his website.

[Myths] Harvey Greenberg (UC Denver, emeritus), Myths and counterexamples in mathematical programming. Entertaining and informative

[Duals] A. Benjamin, Sensible rules for remembering duals. SIAM Review (37)1 1995.

[Vanderbei_LP] R. Vanderbei, "Linear Programming." (2008) Standard LP theory, including brief vignettes into integer programming. Free download at Vanderbei's website and also via SpringerLink

[Wright_IPM] Stephen Wright, "Primal-Dual Interior-Point Methods" (SIAM 1997), available electronically from SIAM. Good summary of LP theory, and more practical/thorough IPM discussion than in [BV2004] (e.g., includes Mehotra's predictor-corrector method), though restricted to LP of course.

[Rockafellar1970] R. Tyrrell Rockafellar, "Convex Analysis", Princeton 1970 (reprinted 1997). You can download via John Datorro's website. Terry Rockafellar wrote the first, and definitive, book on the subject.

[BauschkeCombettes] H. Bauschke and P. Combettes, "Convex Analysis and Monotone Operator Theory in Hilbert Spaces", Springer 2011, available electronically via SpringerLink. More advanced, recommended as a reference to lookup results.

[Nesterov2004] Y. Nesterov, "Introductory Lectures on Convex Optimization", Springer 2004. Concise book, presenting the material in a unique way, from one of the best researchers. Free access via SpringerLink

[BertsekasNedicOzdaglar] D. Bertsekas, A. Nedic, A.E. Ozdaglar, "Convex Analysis and Optimization" (Athena Scientific).

[Bertsekas] D. Bertsekas, "Convex Optimization Theory" (Athena Scientific). [Bertsekas has a lot of other books with useful information as well]

[BorweinLewis] J. M. Borwein and A. S. Lewis, "Convex Analysis and Nonlinear Optimization" (Springer). A bit more mathematical

[Hiriart-UrrutyLemarechal] J.B. Hiriart-Urruty and C. Lemarechal, "Convex Analysis and Minimization Algorithms" (Springer). More on analysis, less on optimization. Technical and a useful reference.

[LuenbergerYe] D. Luenberger and Y. Ye, "Linear and Nonlinear Programming" (Springer). A bit more mathematical than [BV2004], slightly different topics.

Research articles

Especially for the state-of-the-art topics, we will use journal articles are references. Here are a few we may use:

Splitting and Proximal methods

- "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers" by S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein (2010). Readable and popular survey on ADMM.
- "Proximal Algorithms" by N. Parikh and S. Boyd (2013). Popular survey article.
- "A Primer on Monotone Operator Methods" by E. Ryu and S. Boyd (2016). Easier introduction to advanced topic.
- P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," 2011. Good survey
- Laurent Condat "A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms", 2011 (J. Optim. Theory and Appl. 2013)

Coordinate descent and stochastic methods

- "Coordinate Descent Algorithms" by S. Wright (2015). Survey article, easier to follow, good starting point.
- SDCA: "Stochastic Dual Coordinate Ascent Methods for Regularized Loss Minimization" by S. Shalev-Shwartz and T. Zhang (2012). Well-known algorithm, improved to mini-batch (Takac, Richtarik, Srebro)and proximal versions (Shalev-Shwartz and Zhang 2014) in later years
- Lin Xiao, "Dual Averaging Methods for Regularized Stochastic Learning and Online Optimization", 2010 JMLR. Not coordinate descent, more like stochastic gradient descent (SGD) with averaging. Analyzes in an *online learning* framework.
- Defazio, Bach and Lacoste-Julien, "SAGA: A Fast Incremental Gradient Method With Support for Non-Strongly Convex Composite Objectives", NIPS 2014. Recent variance-reduction technique (variant of SGD) for solving very large problems.

Software

You may program in whatever language you like. Programming in `C`, `Fortran` or `C++`, you may use standard libraries like `NAG` or `GLPK`, but these are not always state-of-the-art. Most researchers are using high-level languages; we recommend these either **MATLAB** (try CVX), **Python** (try cvxpy) or **Julia** (try Convex.jl).

The following policies are standard CU-Boulder policies

Accommodation for Disabilities If you qualify for accommodations because of a disability, **please submit to your professor a letter from Disability Services** in a timely manner (for exam accommodations provide your letter at least one week prior to the exam) so that your needs can be addressed. Disability Services determines accommodations based on documented disabilities in the academic environment. Information on requesting accommodations is located on the Disability Services website. Contact Disability Services at 303-492-8671 or *dsinfo@colorado.edu* for further assistance. If you have a temporary medical condition or injury, see Temporary Medical Conditions under the Students tab on the Disability Services website. Any student requiring exam accommodations should contact their instructor *and* the Help Room Coordinator, Rachel Cox (*rachel.cox@colorado.edu*; office: ECCR 241). Make sure to schedule arrangements with Ms. Cox 5 business days in advance of any exam requiring accommodation.

Religious Holidays Campus policy regarding religious observances requires that faculty make every effort to deal reasonably and fairly with all students who, because of religious obligations, have conflicts with scheduled exams, assignments or required attendance. In this class, there should be minimal conflict since there is no attendance grade. If you must miss the final exam due to religious reasons, talk to the professor at the beginning of the semester to make special arrangements. If the homework is due on the date of a religious holiday, you are expected to turn the homework in early. If you have a religious holiday that lasts longer than one week, and so therefore you would not receive the homework with enough time to turn it in early, talk to the professor at the beginning of the semester. See the campus policy regarding religious observances for full details.

Classroom Behavior Students and faculty each have responsibility for maintaining an appropriate learning environment. Those who fail to adhere to such behavioral standards may be subject to discipline. Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, color, national origin, sex, pregnancy, age, disability, creed, religion, sexual orientation, gender identity, gender expression, veteranstatus, political affiliation or political philosophy. Class rosters are provided to the instructor with the student's legal name. I will gladly honor your request to address you by an alternate name or gender pronoun. Please advise me of this preference early in the semester so that I may make appropriate changes to my records. For more information, see the policies at classroom behavior and the Student Code of Conduct.

Sexual Misconduct, Discrimination, Harassment and/or Related Retaliation The University of Colorado Boulder (CU Boulder) is committed to maintaining a positive learning, working, and living environment. CU Boulder will not tolerate acts of sexual misconduct (including sexual assault, exploitation, harassment, dating or domestic violence, and stalking), discrimination, and harassment by members of our community. Individuals who believe they have been subject to misconduct or retaliatory actions for reporting a concern should contact the Office of Institutional Equity and Compliance (OIEC) at 303-492-2127 or cureport@colorado.edu. Information about the OIEC, university policies, anonymous reporting, and the campus resources can be found on the OIEC website.

Please know that faculty and instructors have a responsibility to inform OIEC when made aware of incidents of sexual misconduct, discrimination, harassment and/or related retaliation, to ensure that individuals impacted receive information about options for reporting and support resources.

Honor Code All students enrolled in a University of Colorado Boulder course are responsible for knowing and adhering to the Honor Code. Violations of the policy may include: plagiarism, cheating, fabrication, lying, bribery, threat, unauthorized access to academic materials, clicker fraud, submitting the same or similar work in more than one course without permission from all course instructors involved, and aiding academic dishonesty. All incidents of academic misconduct will be reported to the Honor Code (*honor@colorado.edu*; 303-492-5550). Students who are found responsible for violating the academic integrity policy will be subject to nonacademic sanctions from the Honor Code Council as well as academic sanctions from the faculty member. Additional information regarding the academic integrity policy can be found at the Honor Code Office website. Students are encouraged to work in groups, however all work turned in must be your own, and you are responsible and accountable for all group work associated with your name.