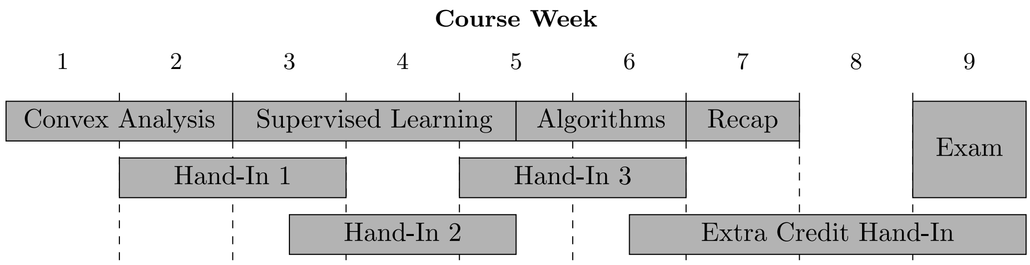
FRTN50 - Optimization for Learning

Course Program Autumn 2019

Lecturer and Course Responsible: Pontus Giselsson
TAs: Martin Morin and Mattias Fält

Overview



The course is roughly divided into three blocks, covering the fundamental convex analysis, supervised learning, and basic first order optimization algorithms. The course work consists of graded hand-ins and ungraded exercises.

Schedule

There are 14 lectures (28 hours) and 28 help sessions (56 hours) scheduled. The times and attending TAs are given below. The lectures and session halls change from week to week so we refer to `timeedit`[1].

Lectures:

Mondays	13.15–15.00
Wednesdays	13.15–15.00

Help Sessions:

Tuesdays	10.15–12.00	Martin Morin
	15.15–17.00	Mattias Fält
Thursdays	08.15–10.00	Martin Morin
	15.15–17.00	Mattias Fält

The help sessions are voluntary and open for exercise/hand-in work or general discussion. All students are free to go to any or all of the help sessions but with one TA and relatively small rooms they are dimensioned for around 20 students per session.

The graded course work have the following deadlines.

Graded Tasks:

Hand-In 1	Week 3 - Sunday	Sep 21	
Hand-In 2	Week 5 - Wednesday	Oct 2	
Hand-In 3	Week 6 - Sunday	Oct 13	
Extra-Credit	Week 9 - Sunday	Nov 3	(Voluntary)
Exam	Week 9 - Monday	Oct 28	(See `timeedit`[1] for time and place)

[1] `https://cloud.timeedit.net/lu/web/lth1/ri14565765000YQQ76Z0527007y8Y4513gQ0g5X6Y55ZQ476.html`

Examination

The examination consists of three parts: mandatory hand-ins, a written exam and a voluntary hand-in for extra credits.

Mandatory Hand-Ins There are three mandatory hand-ins. These are done in groups of 2 and are graded with a pass/fail.

The hand-ins focus on the practical aspects, each approximately corresponding to the three lecture blocks. They will require access to a computer and will be available on the course web page roughly two weeks before their deadline.

Exam The written exam is 5 hours. It will be graded with a points score between 0-20. A passing grade is 12 points.

The exam is complementary to the hand-ins and mainly cover the fundamental mathematical concepts. For this reason it will not be possible to get a 5 on the exam result alone.

Extra Credit Hand-In The voluntary extra credit hand-in is done individually. It will be graded with a points score between 0-6. These points can count towards the higher grades.

Final Grading

3 – Passing grades on the mandatory hand-ins and the exam.
4 – Criterion for 3 and at least 17 points.
5 – Criterion for 3 and at least 22 points.

The final points score is the combined score on both the exam and extra credit hand-in. Note that it is not possible to get a 5 without doing the extra credit assignment.

Hand-In Submission Submissions are sent in to `handin.frtn50@control.lth.se`. Note, this email is not usable for any other purpose.

The email should contain the following:

- Which hand-in that is being submitted.
- Names of all members of the group. (Mandatory hand-ins are done in pairs, the extra credit hand-in is done individually.)
- Attach the files specified in the hand-in manual. (pdf, source code, etc.)
- Give a rough estimate on how much time was spent on the hand-in per person.
- Answer what your primary source of information was for solving the hand-in. (Lectures, slides, hand-in manual, Google, etc.)
- If any specific `Julia` hint/knowledge helped you in your implementation, please write what that was.

The last three points are to help further development of the course.

Deadlines are set at 23:59 at the given date. However, in practice will a submission only be counted as late if we have note received it when we start the grading. We make no promises about when we start, other than it is not before the official deadline.

If the original hand-in submission fail, a maximum of 2 re-submissions are allowed. No re-submissions will be allowed on the extra-credit hand-in.

Literature and Lectures

No official course book exists and the course content is meant to be covered fully by the lectures and lecture slides. Slides and any potential supplementary material will be uploaded on the course web page.

However, *Convex Optimization* by Boyd and Vandenberghe is an excellent source, especially for a wider presentation of optimization modeling and application. It is available online for free `www.stanford.edu/~boyd/cvxbook/`. In particular; *Appendix A* is a useful review on general mathematics you are expected to know, *Chapter 1* is an introduction to convex optimization that is useful to read, *Chapters 2-3* complement our material on convex analysis.

Exercises

Preparatory exercise material for the exam and hand-ins will be available on the course web page. The exercises focus on mathematical concepts. For this reason the bulk of the exercises lie in the first and third block on convex analysis and algorithms. During the second block the focus will shift to the hand-ins and exercises will be provided as a complement.

Contact Information

The Department offices are located in the M-building. Administrators are on the 5th floor. The course lab is on the bottom floor southwest wing. We also have facilities on floor 2, 3, and 5.

Phone and addresses

Mika Nishimura (Ladok, etc)	222 87 85	5th floor	mika.nishimura@control.lth.se
Pontus Giselsson	222 97 44	2nd floor	pontusg@control.lth.se
Mattias Fält	222 08 47	2nd floor	mattias.falt@control.lth.se
Martin Morin	222 87 60	2nd floor	martin.morin@control.lth.se

For more information about the department see `http://www.control.lth.se`

Course Program

A summary of the lecture topics and suitable exercises that covers the topics is given here. The chapter and exercise numbers refer to the exercise compendium on the web page.

w.	Lecture	Suggested Exercises
1	L1 – Course Introduction. Convex Sets. L2 – Convex Functions.	'Introduction to Julia', Ch. 1: 1-13 and 15-19.
2	L3 – Subgradients. Proximal Operators. L4 – Convex Conjugates. Fenchel Duality.	Ch. 2: 1-5, 8-9, and 14-15. Ch. 3: 1-3, 5, 7-11 and 13-16.
3	L5 – Least Squares. Regularization. L6 – Logistic Regression.	Ch. 4: 1-7.
4	L7 – Support Vector Machines. L8 – Multiclass Formulations.	Ch. 4: 8-14.
5	L9 – Neural Networks. Backpropagation. L10 – Proximal Gradient Method.	Ch. 5: 1-9
6	L11 – Coordinate Descent. Stochastic Gradient. L12 – Line Search. Acceleration.	Ch. 5: 10-12, 14, 16-19
7	L13 – Current Research (Inspiration/Bonus). L14 – Recap.	Example Exam

The remainder of the document will contain a more detailed overview of the goals of the course. The goals/content of each week will be listed as well as what could be expected to be tested on the exam.

Detailed Course Goals

Week 1

- Have Juno/Julia up and running
- Familiar with convex sets and convex functions
  - Know definitions and able to identify convexity
  - Know convexity preserving operations
  - Understand supporting and separating hyperplanes
  - Know about epigraphs, convex hull, and convex envelope
  - Understand extended-valued functions and domain
  - Know first and second order conditions for convexity
  - Understand strict-, strong convexity and smoothness
  - Know when local minima $\Rightarrow$ global minima

Week 2

- Familiar with subgradients and proximal operators
  - Able to derive formulas for subdifferentials and proximal operator

- Understand subdifferentials in terms of affine minorizers
  - Know relationship between convexity and existence of subgradient
  - Understand maximal monotonicity and Minty's theorem
  - Know strong monotonicity and relation to strong convexity
  - Understand and be able to use Fermat's rule
  - Know that prox implicitly evaluates subdifferential
  - Know subgradient as an argmax expression over conjugate
  - Know and be able to use Moreau decomposition

- Familiar with conjugate functions and Fenchel duality
  - Able to derive conjugate formulas
  - Able to prove convexity of conjugate
  - Know that biconjugate is convex envelope
  - Know when $\partial f = (\partial f^*)^{-1}$ hold
  - Know Fenchel-Young's inequality and when equality hold
  - Know that strong convexity and smoothness are dual properties
  - Formulate the Fenchel dual problem in general and for specific examples
  - Derive dual problem with primal-dual optimality conditions
  - Able to recover primal solution
  - Be aware of the inf-sup interpretation and derivation of the dual problem

Week 3

- Understand least squares and logistic regression and their purpose
- Understand the convexity of these problems
- Understand the problem of overparameterization and overfitting
- Understand the purpose and need for regularization
- Familiar with the effect of common convex regularization choices
- Understand the use and purpose of feature maps

Week 4

- Understand the support vector machine (SVM) classifier and its purpose
- Know multiclass SVM and multiclass logistic regression and their purpose
- Understand generalization and problems with overfitting to test/training data
- Understand the concept of cross-validation and how it relates to generalization
- Understand the dual SVM formulation
- Familiar with the concept of kernels and how they relate to feature maps
- Understand how SVM kernel methods rely on dual SVM problems

Week 5

- Know what a deep neural network (DNN) is
- Understand relation between DNN and classic regression/classification problems
- Understand the concept of backpropagation
- Understand regularization methods in deep learning
- Understand the difference between first and second order methods
- Understand the convergence of local algorithms on convex vs. nonconvex problems
- Know the proximal gradient method
  - Know that it is a majorization-minimization method
  - Understand its relation to the descent lemma
  - Understand the conditions for convergence and convergence proof
  - Understand what it converges to in nonconvex and convex settings
  - Able to show that the fixed-points solves the problem if convex
  - Be aware that problem scaling can impact algorithm performance

Week 6

- Know the coordinate gradient and stochastic gradient methods
  - Understand when one or the other can/should be used
  - Understand conditions for convergence
  - Know how the analyses differ and relate to the proximal gradient method one
  - Know the potential problems the algorithms might have
  - Know there is a wide range of flavors of both
    (Adam, Coordinate Minimization, ...)
- Familiar with common tricks like line search and acceleration

Week 7

- Recap and summary
- If time, bonus lecture on some current research problem. (Not part of the course.)

Exam

The written exam will cover the goals specified for week 1, 2, and 6 as well as the proximal gradient part of week 5. Most emphasis will be on the week 1 and 2.

The goals for weeks 3 and 4, the deep learning part of week 5, as well as some of the algorithms in week 5 and 6 are tested via handins.