Convex optimization — Project

Léonard Benedetti
leonard.benedetti@platypus.academy

March 2021

To conclude this course on convex optimization, you will be asked to carry out a project to verify your understanding of the different concepts covered during the lectures and the lab sessions, and your ability to use them in a relatively autonomous way when faced with concrete problems.

This document presents this project, which is declined into two subjects: each group must choose and carry out one or the other of these subjects.

Sections 1 and 2 define the two possible topics, and section 3 describes the instructions and the different points that will be taken into account for the evaluation.

Good luck!

1 Subject A: application of optimization techniques to a concrete machine learning problem

Build a model to solve a machine learning task (classification, regression, clustering, recommender system, etc.) on a specific dataset that you will have to choose, while evaluating this model — ideally with a comparison to already existing work — and detailing your approach.

You will focus your analysis on the optimization techniques implemented by the machine learning algorithms (SVM, SGD, neural networks, etc.) you choose to use.

1.1 Datasets

To choose the task and dataset of your project, we refer you to two resources:

- UCI Machine Learning Repository: a repository where a wide variety of interesting tasks can be found, often with references to existing work.
- Kaggle: a platform on which a very wide variety of datasets can be found, but with tasks that are not always explicit.

In any case, validate with your teacher your choice of task and dataset.

1.2 Important notes

About optimization Although the construction of accurate machine learning models and their evaluation are very interesting and can be discussed in this project, remember that an important part of it must be the analysis of the optimization methods used by those models.

About convexity It is extremely common that the functions to optimize in order to train a model (loss functions) for a given dataset are not at all convex. Take an interest in this situation, and try to explain how it may affect the models and possible ways to address this issue.

2 Subject B: study of a particular case of convex optimization

Study and develop an analysis on a particular case of convex optimization that you will have to choose: either the optimization of a particular case of convex programs (quadratic functions, equality constrained, second-order cone program, etc.) or a particular technique (conjugate gradient method, quasi-Newton methods, interior-point method, etc.) to solve convex optimization programs.

Explain the ins and outs and present solutions to the various problems raised.

3 Instructions and evaluation criteria

You will have to carry out this project in groups of four, over a period of about a month. At the end, you will present your work, in English, with a report of 10 to 20 pages and during a defense. Throughout the project, you must strive—as much as possible—to follow a scientific approach.

Regarding deliverables and deadlines, the report and the source code of your experiments must be sent (by the means indicated by your teacher) no later than Tuesday, March 30th at 11:59 pm CEST (23 h 59 Paris time). The defense will take place on Wednesday, March 31st.

Your project will be evaluated according to several criteria detailed below (roughly in descending order, elements at the top of this list are more important than those at the bottom):

- Interest, relevance and editorial quality of your report. Its length must be between 10 and 20 pages and it must detail your approach and your research or realization. In particular, an analysis of the optimization techniques used within the scope of this project is expected[1].

[1] Model construction and implementation details are also relevant, but remain secondary.

- Your performance during your defense. Try to be clear and interesting for your audience!
- Study of the state of the art. For subject A: comparison of your model with the existing ones. For subject B: in addition to studying the state of the art of the particular case chosen, compare the different alternatives.
- The quality of your source code and implementations. It is first about its cleanliness and readability, and then possibly its performance.