SF2526 Numerical algorithms for data-intensive science
aka Numerics for data science
7.5 ECTS

Course data

- Study period 3. Exercise sessions: None
- Homeworks (3): corresponding to 3.5 ECTS
- Exam: corresponding to 4 ECTS
- Number of students (canvas): 22
- Three blocks
  - Algorithms for low-rank data
  - Algorithmic spectral graph clustering
  - Data with Fourier and Toeplitz structure
- Learning activities:
  - Lectures (via zoom, with slides and ipad and live programming). The lectures contained google form quizzes with break-out rooms.
  - Homeworks
  - Mandatory CANVAS-quizzes
  - Course training area (moderated wiki) collected into exam prep problems

Aim

The course aims to provide an introduction to numerical algorithms used in data science. Applications are show to arise from various fields such as image/video analysis, classification of data and audio analysis and signal processing.

Analysis

The course lecture material are based on notes (block1.pdf, block2.pdf and block3.pdf) written by me, which are partially self-contained or refer to specific pages in other material. The three homeworks are quite application oriented, with substantial programming and handling of data. The students could choose between MATLAB and Julia, and 15 % of the submissions were in Julia.

The course was given completely digitally due to the pandemic with live zoom lectures. Several (but not all) lectures were recorded, anonymized and made available online. In order to lessen the weight on the quite heavy zoom-lectures, I also recorded some short videos explaining for some material, which I called X.5 lectures, intended to be watched asynchrounosly. This left more space in the lectures to have short google form quizzes which were discussed by the students in break-out rooms, which were appreciated.

The course uses (and has used since it started) the active collaborative learning by wiki, which worked even better this year since students were more prepared to work online. Some creative problems were submitted, e.g., problems from computer games (super mario).

Despite the restrictions, I am happy with the learning I observed during the course and the exam. There were no critical remarks in the course evaluation. One student requested information about where one could learn more about these topic.

Changes from last year:

- Since so many things had to be changed due to the pandemic, I added a course survey in the middle of the course and did change the style more to asynchronous videos as requested.
- Lectures via zoom. See above.
- Scheduled Q&A session where students can ask questions
- Last year comment: The lecture note were very useful, but it would be nicer with more self-contained material.
- Change: I continued to work on the lecture notes (5 new pages) and did not receive similar comments this year.
- Last year comment: Use more data sets from the wild that combines different topics, e.g. gapminder
- Change: I added new material in the homework from Fashion MNIST. It seemed well perceived.

Selection of student comments / comments for next year students:

The homeworks are a great and fun way to learn the course content, start early so you have time asking questions and do them thoroughly.

Clustering was very interesting and seemed to have more applications in real life

Very interesting course, this is the first course since the pandemic were the teacher is really trying to keep things interactive. Also I felt all the tools to succeed were provided in this course.

In the first block, I really liked the content and the homework. Working with the videos and images was a lot of fun.

I think all parts were interesting in their own way, but maybe the first block was the best one. It made it easier to get an overview of the course because of the clear division

Thank you for starting to upload the notes from the lectures! It felt good that you really looked at the "middle" survey, and cared about our opinions.

I think Elias was a good lecturer and look forward in having him in SF2524!

This topic is really great and would love to be pointed in some general direction or if there are some great courses at KTH that builds on this knowledge. Further reading is also a possibility.

The homeworks are a great and fun way to learn the course content, start early so you have time asking questions and do them thoroughly.

Do take the time to read the referenced literature, while the lecture notes are great the literature can be helpful in understanding some of the more difficult concepts/algorithms and give a deeper understanding.