SI231B - Matrix Computations, Spring 2020-21

Course Projects

Prof. Ziping Zhao

SI231B projects are meant to be substancial and are, in fact, the most significant part of the final grade in the course. They should contain an overview of a particular problem and some independent work.

1 Specifications

Some specifications for the project:

- The length should be at leat 8 single-column pages with 10pt font. Be brief and to the point.
- The project should be written using LaTeX (use an IEEE paper style or similar) and submitted in PDF format (do not use Word!).
- Keep a pdf copy of all the cited references (as you will have to send them to me with the final project).
- The project should contain individual work by at most two students as a group.
- The student should read around 8 research papers on the chosen topic.
- The project should start with a description of the problem (citing at least 8 related papers): this should not be a separate overview of each of the papers one after another; instead, it should contain a unified problem formulation, explaining how each of the related papers fits within the general formulation.
- The related papers should be criticized: you should have an opinion on the papers that you read, you should be able to comment on the contribution and difficulty of each paper.
- Independent research: after the problem formulation and overview of the state of the art, the student should try to propose something new that improves on the existing approaches based on the contents learnt from this course.
- No plagiarism or self-plagiarism is allowed. (The student is never allowed to reuse his/her own published papers as the final project.)

2 Structure of the written report

In order to make the evaluation of the project as objective as possible, the written report should stricly adhere to the following structure with the sections (a penalty will be applied if the report is not organised according to the guideline):

1. Introduction: 10% grade
2. Overview of existing work (with unified notation): 20% grade
3. Criticism of the existing work: 20% grade
4. New contribution (if any): 20% grade
5. Numerical results: 10% grade
6. Conclusions: 10% grade

References: 10% grade

3 Schedule and submission

Please follow the deadlines below. They are strict deadlines and there will be penalties for not respecting them. In particular, the final reports late by 1 day will be penalized with 20% of the grade, late by 2 days will be penalized with 40% of the grade, and late by 3 days is most likely a Fail.

1. *Topic:* By Apr. 19th, the student should choose a topic (either inspired on the list of topics below or not, preferably the student will come up with a topic of his/her interest) with a brief plan on what he/she wants to do and email the instructors to get a confirmation.
2. *Initial proposal:* By May 23rd 11pm, the student should submit an initial proposal (2-3 pages) containing the title, introduction, preliminary ideas, and references. Submit the proposal with filename YourName-proposal.pdf to the link:
   http://pan.shanghaitech.edu.cn/cloudservice/outerLink/decode?c3Vnb24xNjE5MDgwMzcwOTM1c3Vnb24=
3. *Final report:* By June 20th 11pm, submit your final report with all the cited references and codes (optional) with filename YourName-finalreport.zip to the link:
   http://pan.shanghaitech.edu.cn/cloudservice/outerLink/decode?c3Vnb24xNjE5MDgwMzM0Mjc3c3Vnb24=

4 List of topics

Some potential topics for the projects include:

- X. Bai, K. Scheinberg, and R. Tutuncu, *Least-Squares Approach to Risk Parity in Portfolio Selection.*
- Z. Alsulaimawi, *A Non-Negative Matrix Factorization Framework for Privacy-Preserving and Federated Learning.*
- H. Ye and T. Zhang, *DeEPCA: Decentralized Exact PCA with Linear Convergence Rate.*
- B. Xiang, A. Gang, and W. U. Bajwa, *Distributed Principal Subspace Analysis for Partitioned Big Data: Algorithms, Analysis, and Implementation.*
- O. Costa and A. Paiva, *Robust Portfolio Selection Using Linear-Matrix Inequalities.*
- B. Shi and J. M. Phillips, *A Deterministic Streaming Sketch for Ridge Regression.*
- Z. Huang, X. Lin, W. Zhang, and Y. Zhang, *Communication-Efficient Distributed Covariance Sketch, with Application to Distributed PCA.*
- J. Fan, Y. Fan, and J. Lv, *High Dimensional Covariance Matrix Estimation Using a Factor Model.*
- E. Ollila, D. P. Palomar, and F. Pascal, *Shrinking the Eigenvalues of M-Estimators of Covariance Matrix.*
- T. Bodnar, S. Dmytriv, Y. Okhrin, N. Parolya, and W. Schmid, *Statistical Inference for the Expected Utility Portfolio in High Dimensions.*
- M. Aharon, M. Elad, and A. Bruckstein, *K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation.*