



Advanced Topics in Numerical Linear Algebra

Seminar (Summer semester 2022)

Dr. Ion Victor Gosea, Dr. Sara Grundel, Dr. Jens Saak (Seminar), Prof. Dr. Peter Benner (Module Manager)

Goals and Competences:

The students learn typical implementation techniques and algorithms of numerical linear algebra (NLA), in particular of randomized numerical linear algebra (RNLA). The main emphasis will be on the construction and analysis of randomized algorithms, with application to data science. This has become more prominent with the challenge of BIG DATA, and with applications to Data Mining, Information Retrieval, Machine Learning, Bio-informatics. In the seminar, each participant should work on a specific project and present it in front of the class.

Content

The main topic of this seminar is RNLA, a fairly new sub-field of NLA. Below, we summarize some its main characteristics:

- Randomization (together with sampling) allows to design provably accurate algorithms for problems that are typically difficult, including scenarios such as:
 1. Extremely large-scale scenario (matrices too large to be stored at all).
 2. Computationally expensive or NP-hard scenario (combinatorial optimization problems).
- The main idea behind randomized algorithms is as follows:
 1. By sampling rows or columns of a matrix in a particular way, new, smaller matrices can be constructed that are close to the original matrix (with respect to some matrix norms). In most cases, this happens with high probability.
 2. By pre-processing the matrix using random projections, rows/columns can be sampled uniformly at random and still get valid bounds with high probability (hence the stochasticity of the approaches).
 3. The resulting smaller matrices are similar to the original matrices, in terms of the singular values/vectors.
 4. By analyzing such algorithms, it is of relevance to decouple the “randomized” part from the “matrix perturbation” part.

The material covered in this seminar will be based on the lectures notes of Michael W. Mahoney from University of California at Berkeley, USA, available at [1].

- Students are supposed to prepare a project based on the collection of methods and approaches included in the above-mentioned lecture notes, covering one or two sections of this collection.
- Additional valuable information can be found in the recent survey by Per-Gunnar Martinsson and Joel A. Tropp in [2].

Previous knowledge:

Introduction to Numerical Linear Algebra, Linear Algebra I and II, basic knowledge of numerical mathematics and statistics.

Availability of the module:

Compulsory elective module for the bachelor's degree and master's degree in mathematics. The seminar is particularly suitable for those majoring in mathematics or mathematics engineering.

Miscellaneous

If you are interested in the seminar, please contact Ion Victor Gosea at gosea@mpi-magdeburg.mpg.de.

Students who are interested in further studies in this field should attend both the seminar and the lecture. There, attractive topics for writing a master's thesis in numerical linear algebra will be offered.

The Max Planck Institute (CSC group) offers multiple perspectives for the master's thesis and positions for possible student assistants. You can find more details on the opportunities at <https://www.mpi-magdeburg.mpg.de/job-openings>.

Selected references

- [1] M. W. MAHONEY, *Lecture notes on randomized linear algebra*, e-print 1608.04481, arXiv, 2016. Data Structures and Algorithms (cs.DS); Machine Learning (stat.ML), available online at <https://arxiv.org/abs/1608.04481>.
- [2] P. MARTINSSON AND J. TROPP, *Randomized numerical linear algebra: foundations & algorithms*, e-print 2002.01387, arXiv, March 15th, 2021. Numerical Analysis (math.NA), available online at <https://arxiv.org/abs/2002.01387>.