Advanced Seminar in Applied Mathematics

Nir Sharon

Spring 2020

1 Seminar description

The goal of the seminar is to expose students with sufficient mathematical background to advanced topics in applied and computational mathematics. In particular, we wish to mark the following targets:

- Learning essential topics that usually do not get the center of attention in standard undergraduate courses.
- Experiencing an academic study in a small group and with a research flavor.
- Presenting an advanced mathematical material in front of an "expert" audience.

2 Admin: Time, place, requirements

Course name:	*Applied Math Seminar*
Course #:	*0366-3013-01*
Time:	*Tuesday 12:00-14:00*
Place:	*Kaplon building, room 319*
Contact:	*Nir Sharon, Schreiber 121, nsharon@tauex.tau.ac.il*
Prerequisite:	Curiosity, Calculus and linear algebra.
Recommended:	Probability, graph and group theory

3 Guidelines

1. **Attendance in all classes is mandatory.**
2. Each student should **lecture for 1 hour** and **prepare a short set of questions** on the subject, including answers, to publish a week before the talk.
3. All students should **answer the questions** on the topic **and submit it** before the relevant lecture.
4. Each talk should be self-contained and include all the required background. Do not hesitate to repeat things that have been presented in other talks, if necessary.
5. Examples and demonstrations, live code examples are greatly appreciated.
6. Slides can be in English or Hebrew, your choice.

4 Schedule

Date	Main Topic	Lecturer and title
17/3	**Welcome**	Introduction to the seminar and its various topics
24/3	**20th Century's Most Important Algorithms**	• The simplex method • Fast Fourier Transform (FFT)
31/3	**NUFFT — where real-world meets FFT**	• NUFFT part 1 • NUFFT part 2
21/4	**Special meeting: Holocaust Memorial Day**	• Mathematicians in the dark[1]. • Math in WW2: From the Enigma to the Manhattan Project.
27/4	**Yom HaZikaron**	University's ceremony (no meeting)
5/5	**The sphere — a numerical perspective**	• Spherical Harmonics • Integration over the sphere
12/5	**Advanced optimization**	• Optimization on Lie groups • Stochastic Gradient Descent
19/5	**Introduction to learning**	• The fundamentals of learning • Classification and nearest neighbors algorithm
26/5	**Learning and classification**	• Classification: logistic regression, LDA, QDA • Boosting and Adaboost algorithm
2/6	**PCA: Principal components analysis**	• PCA part 1 • PCA part 2
9/6	**Dimension reduction**	• JL lemma and random projections • Laplacian eigenmaps
16/6	**Open subject***	
23/6	**Open subject***	

*Examples of other possible subjects (contingent upon approval): Neural Nets in inverse problems, synchronization over groups, EM algorithm, etc.

[1] In 1934, David Hilbert, by then a grand old man of German mathematics, was dining with Bernhard Rust, the Nazi minister of education. Rust asked, "How is mathematics at Göttingen, now that it is free from the Jewish influence?" Hilbert replied, "There is no mathematics in Göttingen anymore."

5 Topics: overview and references

1. **The simplex method**: An overview on linear programming (and a short intro to optimization in general) and the simplex method. Reference (available online at: `http://math.uchicago.edu/~mqt/math/teaching/math-195/vanderbei_linear-programming_ed-4.pdf`)
"Linear programming: Foundations and extensions" by Robert J. Vanderbei, chapters 1-2.

2. **Fast Fourier Transform** (FFT): A quick introduction to the realm of Fourier analysis and a detailed description of the FFT algorithm and its applications (e.g., polynomial multiplication, numerical integration, etc.). One reference (out of many) is *"A Wavelet Tour of Signal Processing" by Stephane Mallat, chapters 1-3.* Available in `https://www.di.ens.fr/~mallat/papiers/WaveletTourChap1-2-3.pdf`

3. **Nonuniform Fast Fourier Transform** (NUFFT): In many applications and real-world data, the required signal (function, image, surface, etc.) is sampled in a nonuniform fashion. This talk will present the algorithm with some motivational applications. Two papers to use are *"Accelerating the Nonuniform Fast Fourier Transform" by Leslie Greengard and June-Yub Lee.*, available on `https://math.nyu.edu/faculty/greengar/glee_nufft_sirev.pdf` and "A Parallel Nonuniform Fast Fourier Transform Library Based on an Exponential of Semicircle" Kernel by Barnett et. al., see `https://epubs.siam.org/doi/pdf/10.1137/18M120885X`. Also, many links, including to the original NUFFT paper and to software can be found in `https://cims.nyu.edu/cmcl/nufft/nufft.html`.

4. **Spherical Harmonics**: This family of polynomials plays the role of Fourier over the sphere. Their importance also stems from the tight relation with 3-D rotations. The presentation should survey main properties and their use in representation of functions over the sphere. There is a huge number of online resources, for example `https://www.cis.upenn.edu/~cis610/sharmonics.pdf`. Also, a good source of code can be found here `https://github.com/polarch/Spherical-Harmonic-Transform`.

5. **Integration over the sphere**: Integrating functions over the sphere, especially in 3-D, have been drawn a lot of attention since early days of numerical analysis. This problem combines numerical concept of quadrature together with combinatorial problem of finding "uniform" points over the sphere. A survey can be found here `https://arxiv.org/pdf/1611.02785.pdf` with numerous other references, e.g., the classical *"Numerical integration on the sphere" by Kendall Atkinson.* On t-design (a set of integration points) and examples see `http://neilsloane.com/sphdesigns/`.

6. **Optimization on Lie groups**: Manifold became a very popular object to model data. This topic discusses the adaptation of classical optimization algorithms and approaches to manifold data. We will use the online available book `https://press.princeton.edu/absil`. Code of most algorithms appears as a part of this wonderful package: `https://www.manopt.org/index.html`.

7. **Stochastic Gradient Descent**: SGD is an efficient algorithm that can be implemented in a few lines of code, yet still enjoys many advantages of other gradient based methods. The simplicity of SGD allows us to use it in situations when it is not possible to apply other deterministic methods. Many resources are available online, for example see the book:
"Understanding Machine Learning From Theory to Algorithms" by Shai Shalev-Shwartz and Shai Ben-David, chapter 14. It is freely available at `https://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning`.

8. **The fundamentals of learning**: An introduction to machine learning which should include the basic terms and notions of statistical learning: fitting, training, testing, overfitting, (maximum) likelihood, maximum likelihood estimation, Bayesian estimation, model selection, curse of dimensionality, decision theory. Reference:

"Pattern Recognition and Machine Learning," Christopher M. Bishop, chapter 1.
"Understanding Machine Learning: From Theory to Algorithms," Shai Shalev-Shwartz and Shai Ben-David (available online), Chapters 1-2. See `https://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning`

9. **Classification and nearest neighbors algorithm**: A fundamental task in learning and one of the main approaches for solving it. References:
"The Elements of Statistical Learning: Data Mining, Inference, and Prediction," by Hastie, Tibshirani, Friedman, chapters 13.3-13.4. The book is available online at `https://statweb.stanford.edu/~tibs/ElemStatLearn/`. In addition, the paper
"Randomized approximate nearest neighbors algorithm," by P. Jones, A. Osipov, and V. Rokhlin, PNAS 108(38):15679-15686, 2011. Available at `http://www.cs.yale.edu/publications/techreports/tr1434.pdf`.

10. **Classification: logistic regression, LDA, QDA**: More on classification and several conventional methods for solving it. References:
"An Introduction to Statistical Learning, with Applications in R," Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Chapter 4.
"The Elements of Statistical Learning: Data Mining, Inference, and Prediction," by Hastie, Tibshirani, Friedman, chapter 18. The book is available online at `https://statweb.stanford.edu/~tibs/ElemStatLearn/`.
"Understanding Machine Learning: From Theory to Algorithms," Shai Shalev-Shwartz and Shai Ben-David (available online), chapter 22. See `https://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning`

11. **Boosting and Adaboost algorithm**: Boosting is an algorithmic paradigm that grew out of a theoretical question and became a very practical machine learning tool. Some references:
"Pattern Recognition and Machine Learning," Christopher M. Bishop, chapter 14.3-14.4. (also available online, look for it).
"The Elements of Statistical Learning: Data Mining, Inference, and Prediction," by Hastie, Tibshirani, Friedman, sections 10.1-10.5. The book is available online at `https://statweb.stanford.edu/~tibs/ElemStatLearn/`. Also, look at
"The Top Ten Algorithms in Data Mining," by Xindong Wu and Vipin Kumar, chapter 7.

12. **PCA: Principal components analysis**. One of the most fundamental tools in data science. We will use online material, for example:
"Pattern Recognition and Machine Learning," Christopher M. Bishop, chapter 12.
"Principal Component Analysis," I. T. Jolliffe, Springer, 2nd edition, 2002, chapters 1-3. Available online here `http://cda.psych.uiuc.edu/statistical_learning_course/`.
"Understanding Machine Learning: From Theory to Algorithms," Shai Shalev-Shwartz and Shai Ben-David (available online), chapter 23.1. See `https://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning`

13. **Dimension reduction: JL lemma and random projections**. The Johnson-Lindenstrauss lemma exploits the concentration of measure to show that one can dramatically decrease the dimensionality of Euclidean points with only a slight distortion on the associated metric. This remarkable fact led to many constructions of such maps and numerous studies. References:
New and Improved Johnson–Lindenstrauss Embeddings via the Restricted Isometry Property," by Felix Krahmer and Rachel Ward. And:
"Understanding Machine Learning: From Theory to Algorithms," Shai Shalev-Shwartz and Shai Ben-David (available online), Chapters 23.1. See `https://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning`

14. **Laplacian eigenmaps**: Nonlinear method for dimension reduction of data. One reference:
"Laplacian Eigenmaps for Dimensionality Reduction and Data Representation," by M. Belkin and P. Niyogi, Neural Computation 15(6):1373-1396, 2003, available at `http://jupiter.math.`

nctu.edu.tw/~weng/courses/2010_topic_discrete/Spectrum/Laplacian.pdf. And, *"Diffusion maps," by R. R. Coifman and S. Lafon, Applied and Computational Harmonic Analysis, 21: 5-30, 2006,* here https://www.sciencedirect.com/science/article/pii/S1063520306000546

Supplement: Special meeting on the Holocaust Memorial Day

This semester, the seminar meeting of April 21 takes place during the Holocaust Memorial Day. This meeting is an opportunity to deviate from regular practice and devote our effort to learn something from history. This session will combine a mix of history and math, where we will try to understand some of the vast changed underwent by individual mathematicians as well as the field in large.

Some references are:

- *"The mathematical sciences and World War II." by Mina Rees.*
- *"Mathematics and Mathematicians in World War II," by J. Barkley Rosser.*
- *"Refugee Mathematicians in the United States of America, 1933-1941: Reception and Reaction, " by Nathan Reingold.*

Also, popular science blogs and websites are good source of knowledge. For example, /culture.pl and nsa.gov