Anton Sugolov

Toronto, ON anton.sugolov@mail.utoronto.ca
sugolov.github.io github.com/sugolov

Education

MSc. Mathematics 09/24 – 08/25 (exp.)
University of Toronto

· Supervised by Prof. Vardan Papyan
· Coursework: Probability I & II, PDE I, Optimization, Variational Methods in Generative Neural Networks

HBSc. Applied Mathematics and Statistics 09/20 – 06/24
University of Toronto *3.84*

Experience

Faculty Affiliate Researcher 11/23 - Present
Prof. Vardan Papyan *Vector Institute, University of Toronto*

· Investigating investigating the theoretical reasons for the emergence of consistent SVD projections in Jacobians of residual blocks, based on past empirical observations by Prof. Papyan and collaborators
· Discovered coupling in SVD of transformer block Jacobians in 38+ open LLMs across depth and tokens

Research Assistant 06/20 - 06/21
Prof. Lei Sun and Dr. Andrew Paterson *University of Toronto*

· Published open-source Genome Wide Association pipeline with 1K Genomes and ERAP2 expression
· Created and led a workshop for 15 first-year level students to successfully replicate statistical tests

Publications and Preprints

1. (Preprint, arXiv) Aubry, M.[1], Meng, H.[1], **Sugolov, A.**[1], Papyan, V. Transformer Block Coupling and its Correlation with Generalization in LLMs. *Submitted to ICLR 2025. Equal contribution.*[1]

2. **Sugolov, A.**, Emmenegger, E., Paterson, A.D., Sun L. Statistical Learning of Large-Scale Genetic Data: How to Run a Genome-Wide Association Study of Gene-Expression Data Using the 1000 Genomes Project Data. *Statistics in Biosciences* (2023).

Presentations

1. Results and experiments based on '*The Emergence of Clusters in Self-Attention Dynamics*' by Geshkovski et al. Aubry, M. and **Sugolov, A.** Applied[2] Graduate Seminar. Podium. November 2023.

2. Short lectures on the Inverse Kasteleyn Matrix, and Introduction to the Ising, Potts, Percolation, and Random Cluster Models. **Sugolov, A.** Seminar in the Dimer Model and Discrete Riemann Surfaces. November 2023.

Projects

MAT1855 Course Project. An introduction to Otto calculus and some applications. (write-up)

MAT1510 Course Project. Results and experiments based on '*The Emergence of Clusters in Self-Attention Dynamics*' by Geshkovski et al. Presented at Applied[2] Graduate Seminar. (write-up, slides)

Skills

Programming: Python, Java, R
Technical: JAX (optax + eqx), PyTorch, DeepSpeed, Numpy, Slurm, Linux, WandB, HuggingFace, Git
Languages: English, Ukrainian, French (basic)

Teaching

MAT133Y: Calculus and Linear Algebra for Commerce 9/24 - Present
Teaching Assistant

Genome Wide Association Workshop 6/21
Instructor and Organizer (repo)

Honours

Vector Scholarship in AI - Masters' 2024
Vector Institute *Unable to accept*

NSERC Undergraduate Summer Research Award 2023
University of Toronto

Best Poster (1/3) 2023
Data Sciences Institute, Summer Undergraduate Research Day

Regent's Scholarship I, II, III, and Dean's List 2021-2024
University of Toronto

Volunteering

Unissued Diplomas Project 02/24 – 03/24