SIAM Student Chapter @NUS

13th Symposium on Applied and Computational Mathematics, 2024

SIAM Student Chapter @NUS

- Homepage: https://siamnus.github.io/website/

Committee Members

- Prof. Bao Weizhu (matbaowz@nus.edu.sg, Faculty Advisor)
- Mr. Tang Tianyun (ttang@u.nus.edu , President)
- Mr. Liu Shuigen (shuigen@u.nus.edu, Vice President)
- Mr. Li Jingyang (li_jingyang@u.nus.edu, Secretary)

Date

- 28th May 2024 Tuesday

Sponsors

- Society of Industrial and Applied Mathematics (SIAM)
- National University of Singapore (NUS)

Plenary Speakers

- Prof. Nguyen Hung Minh Tan, Department of Mathematics
- Prof. Soh Yong Sheng, Department of Mathematics
- Prof. Tan Yan Fu, Vincent, Department of Mathematics

Student Speakers

- Guo Yue, Department of Mathematics
- Hou Yunlong, Department of Mathematics
- Li Jingyang, Department of Mathematics
- Li Zhixuan, Department of Mathematics
- Dr. Wang Shiwei, Institute of Operations Research and Analytics
- Dr. Xiao Nachuan, Institute of Operations Research and Analytics
- Zhang Yijiong, Department of Mathematics
- Zhao Jiaxi, Department of Mathematics

All are welcome!

Workshop program

All talks take place at S17 (Department of Mathematics) #04-06.

May 28 (Tuesday): 13th SIAM Student Chapter		
Time	Speaker	Title
8:55-9:00am	Opening remarks	
	Session chair: Tang Tianyun	
9:00-9:45am	Prof. Nguyen Minh Tan **(Plenary)**	Principled Frameworks for Designing Deep Learning Models: Efficiency, Robustness, and Expressivity
9:45-10:30am	Prof. Soh Yong Sheng **(Plenary)**	Learning Optimal Regularizers
10:30-11am	Coffee break	
11:00-11:25am	Dr Xiao Nachuan	SGD-type Methods with Guaranteed Global Stability in Nonsmooth Nonconvex Optimization
11:25-11:50am	Li Jingyang	Towards Understanding Why FixMatch Generalizes Better Than Supervised Learning

12pm-2pm	Lunch reception	
	Session chair: Tang Tianyun	
2pm-2:45pm	Prof. Tan Yan Fu, Vincent **(Plenary)**	Optimal Clustering with Bandit Feedback
2:45-3:10pm	Hou Yunlong	Probably Anytime-Safe Stochastic Combinatorial Semi-Bandits
3:10-3:35pm	Dr Wang Shiwei	Strong Variational Sufficiency for Nonlinear Semidefinite Programming and its Implications
3:35-4:00pm	Coffee break	
4:00-4:25pm	Zhao Jiaxi	Mitigating distribution shift in machine learning-augmented hybrid simulation
4:25-4:50pm	Guo Yue	Learning Parametric Koopman Decompositions for Prediction and Control
4:50-5:15pm	Zhang Yijiong	Optimal Market Making under Model Uncertainty: A Reinforcement Learning approach
5:15-5:40pm	Li Zhixuan	Capillary folding of thin elastic sheets

Abstracts

Prof. Nguyen Minh Tan

Principled Frameworks for Designing Deep Learning Models: Efficiency, Robustness, and Expressivity

Designing deep learning models for practical applications, including those in computer vision, natural language processing, and mathematical modeling, is an art that often involves an expensive search over candidate architectures. In this talk, I present novel frameworks to facilitate the process of designing efficient and robust deep learning models with better expressivity via three principled approaches: optimization, differential equation, and statistical modeling.

From an optimization viewpoint, I leverage the continuous limit of the classical momentum accelerated gradient descent to improve Neural ODEs training and inference. The resulting Momentum Neural ODEs accelerate both forward and backward ODE solvers, as well as alleviate the vanishing gradient problem (Efficiency).

From a differential equation approach, I present a random walk interpretation of graph neural networks (GNNs), revealing a potentially inevitable over-smoothing phenomenon. Based on this random walk viewpoint of GNNs, I then propose the graph neural diffusion with a source term (GRAND++) that overcomes the over-smoothing issue and achieves better accuracy in low-labeling rate regimes (Robustness).

Using statistical modeling as a tool, I show that the attention in transformer models can be derived from solving a nonparametric kernel regression problem. I then propose the FourierFormer, a new class of transformers in which the softmax kernels are replaced by the novel generalized Fourier integral kernels. The generalized Fourier integral kernels can automatically capture the dependency of the features of data and remove the need to tune the covariance matrix (Expressivity).

Prof. Soh Yong Sheng

Learning Optimal Regularizers

In optimization-based approaches to inverse problems and to statistical estimation, it is common to augment criteria that enforce data fidelity with a regularizer that promotes desired structural properties in the solution. The choice of a suitable regularizer is typically driven by a combination of prior domain information and computational considerations. Convex regularizers are attractive computationally but they are limited in the types of structure they can promote. On the other hand, nonconvex regularizers are more flexible in the forms of structure they can

promote and they have showcased strong empirical performance in some applications, but they come with the computational challenge of solving the associated optimization problems.
In this paper, we seek a systematic understanding of the power and the limitations of convex regularization by investigating the following questions: Given a distribution, what is the optimal regularizer for data drawn from the distribution? What properties of a data source govern whether the optimal regularizer is convex? We address these questions for the class of regulariz- ers specified by functionals that are continuous, positively homogeneous, and positive away from the origin. We say that a regularizer is optimal for a data distribution if the Gibbs density with energy given by the regularizer maximizes the population likelihood (or equivalently, minimizes cross-entropy loss) over all regularizer-induced Gibbs densities. As the regularizers we consider are in one-to-one correspondence with star bodies, we leverage dual Brunn-Minkowski theory to show that a radial function derived from a data distribution is akin to a “computational sufficient statistic” as it is the key quantity for identifying optimal regularizers and for assessing the amenability of a data source to convex regularization. Using tools such as Γ-convergence from variational analysis, we show that our results are robust in the sense that the optimal regularizers for a sample drawn from a distribution converge to their population counterparts as the sample size grows large. Finally, we give generalization guarantees for various families of star bodies that recover previous results for polyhedral regularizers (i.e., dictionary learning) and lead to new ones for a variety of classes of star bodies.

Dr Xiao Nachuan

SGD-type Methods with Guaranteed Global Stability in Nonsmooth Nonconvex Optimization

In this talk, we focus on providing convergence guarantees for variants of the stochastic subgradient descent (SGD) method in minimizing nonsmooth nonconvex functions. We first develop a general framework to establish global stability for general stochastic subgradient methods, where the corresponding differential inclusion admits a coercive Lyapunov function. We prove that, with sufficiently small stepsizes and controlled noises, the iterates asymptotically stabilize around the stable set of its corresponding differential inclusion. Then we introduce a scheme for developing SGD-type methods with regularized update directions for the primal variables. Based on our developed framework, we prove the global stability of our proposed scheme under mild conditions. We further illustrate that our scheme yields variants of SGD-type methods, which enjoy guaranteed convergence in training nonsmooth neural networks. In particular, by employing the sign map to regularize the update directions, we propose a novel subgradient method named the Sign-map Regularized SGD method (SRSGD). Preliminary numerical experiments exhibit the high efficiency of SRSGD in training deep neural networks.

Li Jingyang

Towards Understanding Why FixMatch Generalizes Better Than Supervised Learning

Semi-supervised learning (SSL), exemplified by FixMatch, has shown significant generalization advantages over supervised learning (SL), particularly in the context of deep neural networks (DNNs). However, it is still unclear, from a theoretical standpoint, why FixMatch-like SSL algorithms generalize better than SL on DNNs. In this work, we present the first theoretical justification for the enhanced test accuracy observed in FixMatch-like SSL applied to DNNs by taking convolutional neural networks (CNNs) on classification tasks as an example. Our theoretical analysis reveals that the semantic feature learning processes in FixMatch and SL are rather different. In particular, FixMatch learns all the discriminative features of each semantic class, while SL only randomly captures a subset of features due to the well-known lottery ticket hypothesis. Furthermore, we show that our analysis framework can be applied to other FixMatch-like SSL methods, e.g., FlexMatch, FreeMatch, Dash, and SoftMatch. Inspired by our theoretical analysis, we develop an improved variant of FixMatch, termed Semantic-Aware FixMatch (SA-FixMatch). Experimental results corroborate our theoretical findings and the enhanced generalization capability of SA-FixMatch.

Prof. Tan Yan Fu, Vincent

Optimal Clustering with Bandit Feedback

This work considers the problem of online clustering with bandit feedback. A set of arms (or items) can be partitioned into various groups that are unknown. Within each group, the observations associated to each of the arms follow the same distribution with the same mean vector. At each time step, the agent queries or pulls an arm and obtains an independent observation from the distribution it is associated to. Subsequent pulls depend on previous ones as well as the previously obtained samples. The agent's task is to uncover the underlying partition of the arms with the least number of arm pulls and with a probability of error not exceeding a prescribed constant δ. The problem proposed finds numerous applications from clustering of variants of viruses to online market segmentation. We present an instance-dependent information-theoretic lower bound on the expected sample complexity for this task, and design a computationally efficient and asymptotically optimal algorithm, namely Bandit Online Clustering (BOC). The algorithm includes a novel stopping rule for adaptive sequential testing that circumvents the need to exactly solve any NP-hard weighted clustering problem as its subroutines. We show through extensive simulations on synthetic and real-world datasets that BOC's performance matches the lower bound asymptotically, and significantly outperforms a non-adaptive baseline algorithm.

Joint work with Junwen Yang (IORA, NUS) and Zixin Zhong (University of Alberta). To appear in the Journal of Machine Learning

Hou Yunlong

Probably Anytime-Safe Stochastic Combinatorial Semi-Bandits

Motivated by concerns about making online decisions that incur undue amount of risk at each time step, in this paper, we formulate the probably anytime-safe stochastic combinatorial semi-bandits problem. In this problem, the agent is given the option to select a subset of size at most K from a set of L ground items. Each item is associated to a certain mean reward as well as a variance that represents its risk. To mitigate the risk that the agent incurs, we require that with probability at least $1-\delta$, over the entire horizon of time T, each of the choices that the agent makes should contain items whose sum of variances does not exceed a certain variance budget. We call this probably anytime-safe constraint. Under this constraint, we design and analyze an algorithm {\sc PASCombUCB} that minimizes the regret over the horizon of time T. By developing accompanying information-theoretic lower bounds, we show that under both the problem-dependent and problem-independent paradigms, {\sc PASCombUCB} is almost asymptotically optimal. Experiments are conducted to corroborate our theoretical findings. Our problem setup, the proposed {\sc PASCombUCB} algorithm, and novel analyses are applicable to domains such as recommendation systems and transportation in which an agent is allowed to choose multiple items at a single time step and wishes to control the risk over the whole time horizon.

Dr Wang Shiwei

Strong Variational Sufficiency for Nonlinear Semidefinite Programming and its Implications

Strong variational sufficiency is a newly proposed property, which turns out to be of great use in the convergence analysis of multiplier methods. However, what this property implies for non-polyhedral problems remains a puzzle. In this talk, we will introduce the equivalence between the strong variational sufficiency and the strong second order sufficient condition (SOSC) for nonlinear semidefinite programming (NLSDP), without requiring the uniqueness of multiplier or any other constraint qualifications. Based on this characterization, the local convergence property of the augmented Lagrangian method (ALM) for NLSDP can be established under strong SOSC in the absence of constraint qualifications. Moreover, under the strong SOSC, we can apply the semi-smooth Newton method to solve the ALM subproblems of NLSDP as the positive definiteness of the generalized Hessian of augmented Lagrangian function is satisfied.

Zhao Jiaxi

Mitigating distribution shift in machine learning-augmented hybrid simulation

We study the problem of distribution shift generally arising in machine-learning augmented hybrid simulation, where parts of simulation algorithms are replaced by data-driven surrogates. A mathematical framework is established to understand the structure of machine-learning augmented hybrid simulation problems and the cause and effect of the associated distribution shift. We show correlations between distribution shift and simulation error both numerically and theoretically. Then, we propose a simple methodology based on tangent-space regularized estimator to control the distribution shift, thereby improving the long-term accuracy of the simulation results. In the linear dynamics case, we provide a thorough theoretical analysis to quantify the effectiveness of the proposed method. Moreover, we conduct several numerical experiments, including simulating a partially known reaction-diffusion equation and solving Navier-Stokes equations using the projection method with a data-driven pressure solver. In all cases, we observe marked improvements in simulation accuracy under the proposed method, especially for systems with high degrees of distribution shift, such as those with relatively strong non-linear reaction mechanisms, or flows at large Reynolds numbers.

Guo Yue

Learning Parametric Koopman Decompositions for Prediction and Control

We present an approach to constructing approximate Koopman-type decompositions for dynamical systems depending on static or time-varying parameters. Our method simultaneously constructs an invariant subspace and a parametric family of projected Koopman operators acting on this subspace. We parametrize both the projected Koopman operator family and the dictionary that spans the invariant subspace by neural networks, and jointly train them with trajectory data. We show theoretically the validity of our approach, and demonstrate via numerical experiments that it exhibits significant improvements over existing methods in solving prediction problems, especially those with large state or parameter dimensions, and those possessing strongly non-linear dynamics. Moreover, our method enables data-driven solution of optimal control problems involving non-linear dynamics, with some interesting implications on controllability.

Zhang Yijiong

Optimal Market Making under Model Uncertainty: A Reinforcement Learning approach

We study the optimal market-making problem in order-driven electronic markets, focusing on developing an interpretable quoting strategy and deriving a robust solution that excels amidst model ambiguity. We address two types of ambiguity in market dynamics transitions: variability in various parameters, ranging from order arrival intensities to price dynamics, and distributional variability, defined via Wasserstein uncertainty. To overcome technical challenges, we introduce a tractable model for limit order books using Markov decision processes and develop a model-based robust reinforcement learning method to navigate the complex optimization problem. Our framework enables the incorporation of realistic settings to accurately capture the intricacies of market microstructure and ensure sustainably superior performance under ever-changing and diverse market conditions. Through simulations and real-data studies, we demonstrate the effectiveness of our approach, emphasizing the importance of accounting for model ambiguity. Additionally, our work presents a comprehensive framework that illustrates the application of advanced machine learning techniques in quantitative finance, offering a valuable blueprint for addressing similar challenges within this field.