國立成功大學 高等教育 STEM 計畫

最佳化、統計與機器學習研討會

NKCU STEM Program for Higher Education
Seminar on Optimization, Statistics and Machine Learning

時間 Time	2025 年 12 月 22 日(一) 09:00 – 20:00 09:00 – 20:00, Monday, Dec. 22, 2025
地點 Venue	國立成功大學自強校區啟端館 1 樓階梯教室(96112) Lecture Hall (Room 96112), 1st Floor, Chi-Tuan Building, Department of Electrical Engineering, National Cheng Kung University (Tzu-Chiang Campus)
主辦單位 Sponsors	國立成功大學 通識教育中心 NCKU Center for General Education 理學院數學跨領域研究中心 Interdisciplinary Research Center on Mathematics 國立成功大學 電機工程學系 NCKU Department of Electrical Engineering

- 議程

時間	演講題目	演講者
09:00–10:40	Statistics Meets Optimization (I)(II) （中文）	王俊焜 助理教授 加州大學 聖地牙哥分校 電機與電腦工程學系
10:40–11:10	休息	
11:10–12:00	Making Autonomous Driving Smarter with Data Selection （中文）	馮若梅 助理教授 淡江大學 電腦科學與資訊工程學系
12:10–13:20	午餐	
13:20–14:10	Implementable Unsolvability Criteria for Systems of Two Quadratic (In)equalities （英文）	王敏齊 博士候選人 國立成功大學 數學系
14:10–15:00	Gradient methods with online scaling （英文）	朱雅琪 博士候選人 史丹佛大學 數學系
15:00–15:30	休息	
15:30–16:20	Interplay between optimization and no-regret learning （英文）	王俊焜 助理教授 加州大學 聖地牙哥分校 電機與電腦工程學系
16:20–17:10	Do Neural Networks Generalize Well? Low Norm Solutions vs. Flat Minima （英文）	Rahul Parhi 助理教授 加州大學 聖地牙哥分校 電機與電腦工程學系

註：演講大綱請見下頁

國立成功大學 高等教育 STEM 計畫

最佳化、統計與機器學習研討會

NKCU STEM Program for Higher Education
Seminar on Optimization, Statistics and Machine Learning

09:00 – 10:40 Room 96112

主持人：

演講者：王俊焜 助理教授

演講題目：Statistics Meets Optimization (I)(II) （中文）

大綱：

In this talk, I will focus on a literature overview of "Statistics Meets Optimization" from my personally biased perspective, while also covering my recent research at this intersection at a high level.

In the first part of the talk (I), I will show how the old idea of "control variate" is used to design optimization algorithms like the Stochastic Variance-Reduced Gradient method (SVRG) and the recently popular technique of Prediction Powered Inference (PPI) in statistical inference. SVRG has become a classical algorithm for reducing variance in stochastic optimization. PPI, on the other hand, concerns how to use machine learning predictions on data that mostly lack ground-truth labels to improve statistical inference such as the task of constructing a confidence interval. I will show how the same algorithmic principle has been used in these seemingly different areas. After that, I will share my recent effort in this direction.

Then, I will switch gears to the second part of the talk (II). I will first review a neat result in the literature that casts non-parametric sequential hypothesis testing as an online convex optimization problem, where an online learner tries to bet whether the null hypothesis is true or false, and a tighter regret bound suggests a faster stopping time to reject the null when the alternative is true. Then, I will show how relevant techniques can be used to design algorithms with strong statistical guarantees in online detection of LLM-generated texts. After that, I will introduce a new algorithm that overcomes the limitations of a commonly-used method for sequential hypothesis testing by betting, which potentially leads to a faster rejection time under the alternative while controlling the false positive rate.

國立成功大學 高等教育 STEM 計畫

最佳化、統計與機器學習研討會

NKCU STEM Program for Higher Education
Seminar on Optimization, Statistics and Machine Learning

11:10 – 12:00 Room 96112
主持人： 演講者：馮若梅 助理教授
演講題目：Making Autonomous Driving Smarter with Data Selection (中文)
大綱： In autonomous driving object detection, expanding data diversity while adapting models to real-world environments is crucial. Online training on every new sample is costly and impractical. We propose a data selection strategy using distance metrics and clustering. New data far from existing categories are added to training with weights inversely proportional to their class size, while overrepresented classes are ignored. Stratified sampling mixes old and new data to reduce bias. This approach improves model adaptability, efficiency, and continual learning performance in autonomous driving scenarios.

國立成功大學 高等教育 STEM 計畫

最佳化、統計與機器學習研討會

NKCU STEM Program for Higher Education
Seminar on Optimization, Statistics and Machine Learning

13:20 – 14:10 Room 96112
主持人： 演講者：王敏齊 博士候選人
演講題目：Implementable Unsolvability Criteria for Systems of Two Quadratic (In)equalities（英文）
大綱： Let $f(x)$ and $g(x)$ be two real quadratic functions defined on $\mathbb{R}^n$. The decision problem of whether there exists a common solution to the systems of (in)equalities $[f(x) = 0, g(x) = 0]$ (Calabi Theorem) and $[f(x) = 0, g(x) \leq 0$f] (Strict Finsler Lemma), respectively, has rarely been studied in the literature. In this talk, we equip non-homogeneous variants of the Calabi Theorem as well as the strict Finsler Lemma with necessary and sufficient conditions. To implement these criteria, we reformulate key conditions using matrix pencils and provide an efficient procedure for computing, enabling fast numerical feasibility tests. Finally, we benchmark our method against existing approaches, demonstrating clear gains in both computational speed and reliability.

國立成功大學 高等教育 STEM 計畫

最佳化、統計與機器學習研討會

NKCU STEM Program for Higher Education
Seminar on Optimization, Statistics and Machine Learning

14:10 – 15:00 Room 96112
主持人：林敏雄 主任 國立成功大學 數學系 演講者：朱雅琪 博士候選人
演講題目：Gradient methods with online scaling（英文）
大綱： Matrix stepsizes, commonly referred to as preconditioners, play a crucial role to accelerate modern first-order optimization methods by adapting to optimization landscapes with highly heterogeneous curvature across dimensions. It remains challenging to determine optimal matrix stepsizes in practice, which typically requires problem-specific expertise or computationally expensive pre-processing. In this talk, we present a novel family of algorithms, online scaled gradient methods (OSGM), that employs online learning to learn the matrix stepsizes and provably accelerate first-order methods. Convergence rate of OSGM is asymptotically no worse than the rate achieved by the optimal matrix stepsize. On smooth convex problems, OSGM provides a new trajectory-dependent global convergence guarantees; on strongly convex problems, OSGM constitutes a new family of first-order methods with nonasymptotic superlinear convergence, joining the celebrated quasi-Newton methods. Our experiments show that OSGM substantially outperforms existing adaptive first-order methods and frequently matches the performance of L-BFGS, an efficient quasi-Newton method, while utilizing less memory and requiring less computational effort per iteration.

國立成功大學 高等教育 STEM 計畫

最佳化、統計與機器學習研討會

NKCU STEM Program for Higher Education
Seminar on Optimization, Statistics and Machine Learning

15:30 – 16:20 Room 96112
主持人：彭勇寧 主任 國立成功大學 數學跨領域研究中心 演講者：王俊焜 助理教授
演講題目：Interplay between optimization and no-regret learning（英文）
大綱： In this talk, I will show how to design and analyze first-order convex optimization algorithms by playing a two-player zero-sum game. In particular, I will highlight a strong connection between online learning (a.k.a. no-regret learning) and optimization. It turns out that several classical optimization updates can be generated from the game dynamics by pitting pairs of online learners against each other. These include Nesterov's methods, accelerated proximal method (Bech and Teboulle 2009), Frank-Wolfe Method, and more. This fresh perspective of "optimization as a game" also gives rise to new accelerated Frank–Wolfe methods over certain types of constraint sets. Furthermore, by summing the weighted average regrets of both players in the game, one obtains the convergence rate of the resulting optimization algorithm, which provides a simple and modular technique to analyze these optimization methods.

國立成功大學 高等教育 STEM 計畫

最佳化、統計與機器學習研討會

NKCU STEM Program for Higher Education
Seminar on Optimization, Statistics and Machine Learning

16:20 – 17:00 Room 96112

主持人：

演講者：Rahul Parhi 助理教授

演講題目：Do Neural Networks Generalize Well? Low Norm Solutions vs. Flat Minima（英文）

大綱：

This talk investigates the fundamental differences between low-norm and flat solutions of shallow ReLU networks training problems, particularly in high-dimensional settings. We show that global minima with small weight norms exhibit strong generalization guarantees that are dimension-independent. In contrast, local minima that are “flat” can generalize poorly as the input dimension increases. We attribute this gap to a phenomenon we call neural shattering, where neurons specialize to extremely sparse input regions, resulting in activations that are nearly disjoint across data points. This forces the network to rely on large weight magnitudes, leading to poor generalization. Our theoretical analysis establishes an exponential separation between flat and low-norm minima. In particular, while flatness does imply some degree of generalization, we show that the corresponding convergence rates necessarily deteriorate exponentially with input dimension. These findings suggest that flatness alone does not fully explain the generalization performance of neural networks.

講者簡介：

Rahul Parhi is an Assistant Professor in the Department of Electrical and Computer Engineering at the University of California, San Diego. Prior to joining UCSD, he was a Postdoctoral Researcher at the École Polytechnique Fédérale de Lausanne (EPFL), where he worked from 2022 to 2024. He completed his PhD in Electrical Engineering at the University of Wisconsin-Madison in 2022. His research interests lie at the interface between functional and harmonic analysis and data science.