International Conference on Statistical Optimization and Learning

30-31 December, Beijing

ICSOL 2023

The 111 Project of China

National Natural Science Foundation of China

Content

The 8th International Conference on Statistical Optimization and Learning

30-31 December 2023, Beijing

Statistical optimization and learning, as an advanced interdisciplinary field between statistics and optimization, have garnered extensive attention from scholars nationally and internationally. In order to strengthen connections and foster collaboration among scholars in related fields, as well as to further promote the integration and development of mathematics optimization and statistics, the Mathematical Programming Branch of OR Society of China, Chinese Association for Applied Statistics, the Beijing Society of Applied Statistics, Beijing Jiaotong University, Beijing University of Technology, Capital Normal University, Renmin University of China, Academy of Mathematics and Systems Science Chinese Academy of Sciences, plan to jointly host the 8th International Conference on Statistical Optimization and Learning from December 30 to 31, 2023 (ICSOL 2023).

Since 2013, ICSOL has been successfully held seven times. This academic conference will continue its commitment to fostering deep academic exchanges and advancing disciplinary development. It will focus on the current trends in statistical machine learning and big data research, addressing optimization problems within these fields. Additionally, it will facilitate academic discussions on the theoretical and practical aspects of statistical optimization and learning research within the context of artificial intelligence. Renowned domestic experts in statistical learning and big data research have been invited to present academic reports at this conference, aiming to establish a high-level, multidisciplinary platform for academic exchanges, broaden academic perspectives, and promote the profound development of the statistical optimization discipline. Scholars are warmly welcomed to participate.

1. Academic Reports

The conference extends invitations to 16 experts from both domestic and international domains to deliver presentations. The list of speakers is given as follows (in alphabetical order by last name):

Prof. Kun Chen	University of Connecticut
Prof. Ying Cui	University of California, Berkeley
Prof. Jian Huang	The Hong Kong Polytechnic University
Prof. Jiashun Jin	Carnegie Mellon University
Prof. Tracy Ke	Harvard University
Prof. Linglong Kong	University of Alberta
Prof. Runze Li	The Pennsylvania State University
Prof. Xudong Li	Fudan University
Prof. Jiawang Nie	University of California San Diego
Prof. Zhihua Su	University of Florida
Prof. Defeng Sun	The Hong Kong Polytechnic University
Prof. Ju Sun	University of Minnesota System
Prof. Falong Tan	Hunan University
Prof. Kim-chuan Toh	National University of Singapore
Dr. Jiujing Wu	Capital Normal University
Prof. Yancheng Yuan	The Hong Kong Polytechnic University

2. Short-term Courses

In order to foster the rapid integration of optimization, statistics, learning and other related fields, the conference will offer short-term courses form December 25 to 29, 2023. Course lecturers are as follows:

Prof. Zhaosong Lv	University of Minnesota System
Prof. Xinyu Zhang	Chinese Academy of Sciences
Prof. Yangjing Zhang	Chinese Academy of Sciences
Prof. Shenglong Zhou	Beijing Jiaotong University

3. Academic Committee

Director

Prof. Naihua Xiu	Beijing Jiaotong University
Prof. Defeng Sun	The Hong Kong Polytechnic University
Prof. Hui Zou	University of Minnesota

Committee Member

Prof. Hengjian Cui	Capital Normal University
Prof. Tiande Guo	University of Chinese Academy of Sciences
Prof. Houduo Qi	The Hong Kong Polytechnic University
Prof. Jie Sun	National University of Singapore & Curtin University
Prof. Dachuan Xu	Beijing University of Technology
Prof. Lingchen Kong	Beijing Jiaotong University
Prof. Liping Jing	Beijing Jiaotong University
Prof. Liping Zhu	Renmin University of China

4. Organizing Committee

Director

Prof. Lingchen Kong	Beijing Jiaotong University

Committee Member

Prof. Chao Ding	Academy of Mathematics and System Sciences
Prof. Liping Jing	Beijing Jiaotong University
Prof. Ziyan Luo	Beijing Jiaotong University
Prof. Chao Zhang	Beijing Jiaotong University
Prof. Xinyuan Zhao	Beijing University of Technology
Prof. Liping Zhu	Renmin University of China
Prof. Shenglong Zhou	Beijing Jiaotong University

Conference Affairs Team

Xiangru Xing, Mingwei Hu, Jiamei Wu, Rui Zhao

5. Contact

Xiangru Xing	Tel: 86-18437975116	Email: 21118025@bjtu.edu.cn
Mingwei Hu	Tel: 86-15801537896	Email: 21121623@bjtu.edu.cn
Jiamei Wu	Tel: 86-17779149782	Email: 23111514@bjtu.edu.cn
Rui Zhao	Tel: 86-18801080484	Email: 22121657@bjtu.edu.cn

6. Sponsors

The Mathematical Programming Branch of OR Society of China

Chinese Association for Applied Statistics

The Beijing Society of Applied Statistics

Academy of Fundamental and Interdisciplinary Sciences, Beijing Jiaotong University

Faculty of Science, Beijing University of Technology

Beijing National Center for Applied Mathematics, Capital Normal University

Institute of Statistics and Big Data, Renmin University of China

Academy of Mathematics and Systems Science Chinese Academy of Sciences

Program Schedule

CST:30 December 8 am -- 31 December 12 am

EST:29 December 7 pm -- 30 December 11 pm

30 December 2023 Tencent Meeting ID: 266 884 052 Passcode: 9046 Siyuan West Building 504			
Time	**Talks**	**Speaker**	**Chair**
8:00-8:30	Opening Ceremony Vice President of Beijing Jiaotong University President of the Chinese Association for Applied Statistics President of the Operations Research Society of China		Yongguang Yu
8:30-9:10	Tests for Large-Dimensional Shape Matrices via M-Estimators	Runze Li	Liqun Wang
9:10-9:50	Exploring the Training Robustness of Distributional Reinforcement Learning against Noisy State Observations	Linglong Kong	
9:50-10:00	Tea Break		
10:00-10:40	Robust Deep Learning: Where Are We?	Ju Sun	DaChuan Xu
10:40-11:20	Text Analysis and Testing of High-Dimensional Multinomials	Tracy Ke	
11:20-12:00	Some Recent Progress on The Convex Clustering Model	Yancheng Yuan	
12:00-14:00	Lunch Minghu Restaurant		

30 December 2023 Machinery Engineering Building Lecture Hall Tencent Meeting ID: 266 884 052 Passcode: 9046			
Time	**Talks**	**Speaker**	**Chair**
14:00-14:40	Statistical Deep Learning	Jian Huang	Hengjian Cui
14:40-15:20	Stochastic Bregman Gradient Methods with Application to Deep Learning	Kim-Chuan Toh	
15:20-15:30	Tea Break		
15:30-16:00	Data-Driven Minimax Optimization with Expectation Constraints	Xudong Li	Tiande Guo
16:00-16:30	Weighted Residual Empirical Processes, Martingale Transformations, and Model Specification Tests for Regressions with Diverging Number of Parameters	Falong Tan	
16:30-17:00	Model-Free Feature Screening Based on Hellinger Distance for Dltrahigh Dimensional Data	Jiujing Wu	Liping Zhu
17:00-17:40	Nonsmooth Analysis and Sparse Optimization	Defeng Sun	
18:00-20:00	Dinner Hongguoyuan Restaurant		

31 December 2023 Tencent Meeting ID: 266 884 052 Passcode: 9046 Siyuan West Building 504			
Time	**Title**	**Speaker**	**Chair**
8:20-9:00	The Statistical Triangle	Jiashun Jin	Chao Ding
9:00-9:40	Envelope-Based Partial Least Squares	Zhihua Su	
9:40-9:50	**Tea Break**		
9:50-10:30	Optimization with Superquantile Constraints: A Fast Computational Approach	Ying Cui	Wenxun Xing
10:30-11:10	Rare Feature Selection and Logic Aggregation through Convex Optimizations	Kun Chen	
11:10-11:50	The Multi-Objective Polynomial Optimization	Jiawang Nie	
11:50-12:00	**Closing Ceremony** President of the Hong Kong Mathematical Society		Lingchen Kong

Biographies, Titles and Abstracts of Talks

Runze Li is an Eberly Family Chair Professor in Statistics, the Pennsylvania State University. He served as a Co-Editor of the Annals of Statistics from 2013 to 2015, and is a Fellow of IMS, ASA and AAAS. His recent honors and awards include the Distinguished Achievement Award of International Chinese Statistical Association, 2017, Faculty Research Recognition Awards for Outstanding Collaborative Research, College of Medicine, Penn State University in 2018, and Distinguished Mentoring Award, Eberly College of Science, Penn State University in 2023. His research interests include theory and methodology in variable selection, feature screening, robust statistics, nonparametric and semiparameteric regression. His interdisciplinary research aims to promote the better use of statistics in social behavioral research, neural science research, and climate studies.

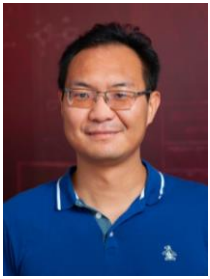
Tests for Large-Dimensional Shape Matrices via M-Estimators

Tyler's M estimator, as a robust alternative to the sample covariance matrix, has been widely applied in robust statistics. However, classical theory on Tyler's M estimator is mainly developed in the low-dimensional regime for elliptical populations. It remains largely unknown when the parameter of dimension p grows proportionally to the sample size n for general populations. By utilizing the eigenvalues of Tyler's M estimator, this article develops tests for the identity and equality of shape matrices in a large-dimensional framework where the dimension-to-sample size ratio p/n has a limit in (0,1). The proposed tests can be applied to a broad class of multivariate distributions including the family of elliptical distributions. To analyze both the null and alternative distributions of the proposed tests, we provide a unified theory on the spectrum of a large-dimensional Tyler's M estimator when the underlying population is general. Simulation results demonstrate good performance and robustness of our tests. An empirical analysis of the Fama-French 49 industrial portfolios is carried out to demonstrate the shape of the portfolios varying.

Linglong Kong is a professor in the Department of Mathematical and Statistical Sciences at the University of Alberta. He holds a Canada Research Chair in Statistical Learning, a Canada CIFAR AI Chair, and is a fellow of the Alberta Machine Intelligence Institute (AMII). His publication record includes more than 90 peer-reviewed articles in top journals such as AOS, JASA and JRSSB as well as top conferences such as NeurIPS, ICML, ICDM, AAAI, and IJCAI. Dr. Kong currently serves as associate editor of the Journal of the American Statistical Association, the Canadian Journal of Statistics, and Statistics and its Interface, as well as guest editor of Statistics and its Interface. Additionally, Dr. Kong is a member of the Executive Committee of the Western North American Region of the International Biometric Society, chair of the ASA Statistical Computing Session program, and chair of the webinar committee. He served as a guest editor of Canadian Journal of Statistics, associate editor of International Journal of Imaging Systems and Technology, guest associate editor of Frontiers of Neurosciences, chair of the ASA Statistical Imaging Session, and member of the Statistics Society of Canada's Board of Directors. He is interested in the analysis of high-dimensional and neuroimaging data, statistical machine learning, robust statistics and quantile regression, as well as artificial intelligence for smart health.

Exploring the Training Robustness of Distributional Reinforcement Learning against Noisy State Observations

In real scenarios, state observations that an agent observes may contain measurement errors or adversarial noises, misleading the agent to take suboptimal actions or even collapse while training. In this paper, we study the training robustness of distributional Reinforcement Learning (RL), a class of state-of-the-art methods that estimate the whole distribution, as opposed to only the expectation, of the total return. Firstly, we validate the contraction of distributional Bellman operators in the State-Noisy Markov Decision Process (SN-MDP), a typical tabular case that incorporates both random and adversarial state observation noises. In the noisy setting with function approximation, we then analyze the vulnerability of least squared loss in expectation-based RL with either linear or nonlinear function approximation. By contrast, we theoretically characterize the bounded gradient norm of distributional RL loss based on the categorical parameterization equipped with the Kullback–Leibler (KL) divergence. The resulting stable gradients while the optimization in distributional RL accounts for its better training robustness against state observation noises. Finally, extensive experiments on the suite of environments verified that distributional RL is less vulnerable against both random and adversarial noisy state observations compared with its expectation-based counterpart.



Ju Sun is an assistant professor at the Department of Computer Science & Engineering, the University of Minnesota at Twin Cities. His research interests span computer vision, machine learning, numerical optimization, data science, computational imaging, and healthcare. His recent efforts are focused on the foundation and computation for deep learning and applying deep learning to tackle challenging science, engineering, and medical problems. Before this, he worked as a postdoc scholar at Stanford University (2016-2019), obtained his Ph.D. degree from Electrical Engineering of Columbia University in 2016 (2011-2016), and B.Eng. in Computer Engineering (with a minor in Mathematics) from the National University of Singapore in 2008 (2004-2008). He won the best student paper award from SPARS'15, honorable mention of doctoral thesis for the New World Mathematics Awards (NWMA) 2017, and AAAI New Faculty Highlight Programs 2021.

Robust Deep Learning: Where Are We?

Deep learning (DL) models are not robust: adversarially constructed and irrelevant natural perturbations can break them abruptly. Despite intensive research in the past few years, surprisingly, there have yet to be tools for reliable robustness evaluation in the first place. I'll describe our recent efforts toward building such a reliable evaluation package. This new computational capacity leads to more concerns than hopes: we find that the current empirical robust evaluation is problematic, and adversarial training, a predominant framework toward achieving robustness, is fundamentally flawed. On the other hand, before we can obtain robust DL models, or trustworthy DL models in general, we must safeguard our models against making severe mistakes to make imperfect DL models deployable. A promising approach is to allow DL models to restrain from making predictions on uncertain samples. I'll describe our recent lightweight, universal selective classification method that performs excellently, even under distribution shifts.

Tracy Ke is currently an associate professor of Statistics at Harvard University. She obtained her PhD in Operations Research and Financial Engineering from Princeton University in 2014, advised by Professor Jianqing Fan. From 2014 to 2018, she was an assistant professor of Statistics at Chicago University. She joined Harvard University in 2018. Her research interests include high-dimensional statistics, machine learning, network data analysis, and text mining. In her work on high-dimensional statistics, she is particularly interested in the optimal statistical inference when the signals are very rare and weak. In her work on network data analysis, she is particularly interested in estimating the latent community structure of a network. She is the recipient of NSF CAREER Award, IMS Peter Hall Prize and ASA Noether Young Scholar Award, and she is currently a Sloan Fellow.

Text Analysis and Testing of High-Dimensional Multinomials

Motivated by applications in text mining and discrete distribution inference, we test for equality of probability mass functions of K groups of high-dimensional multinomial distributions. Special cases of this problem include global testing for topic models, two-sample testing in authorship attribution, and closeness testing for discrete distributions. A test statistic, which is shown to have an asymptotic standard normal distribution under the null hypothesis, is proposed. This parameter-free limiting null distribution holds true without requiring identical multinomial parameters within each group or equal group sizes. The optimal detection boundary for this testing problem is established, and the proposed test is shown to achieve this optimal detection boundary across the entire parameter space of interest. The proposed method is demonstrated in simulation studies and applied to analyze two real-world datasets to examine, respectively, variation among customer reviews of Amazon movies and the diversity of statistical paper abstracts.

Yancheng Yuan is an assistant professor in the Department of Applied Mathematics at the Hong Kong Polytechnic University. His primary research interests lie in the theoretical and applied aspects of continuous optimization and machine learning. He has published papers in prestigious journals such as the "SIAM Journal on Optimization," "Journal of Machine Learning Research," and "IEEE Transactions on Neural Networks and Learning Systems", as well as presented at top academic conferences in the field of machine learning such as ICML, NeurIPS, and WWW. He has received the Best Paper Award Finalist (WWW 2021).

Some Recent Progress on The Convex Clustering Model

Convex clustering is a relaxation of the K-means model with attractive properties. In this talk, we will present some recent progress on the convex clustering model. In particular, we will discuss some dimension reduction techniques which can exploit the structure of the model in an appropriate way. We will also discuss how to adopt the deep learning technique to enhance the clustering performance of the convex clustering model. Some numerical experiment results will be shown to demonstrate the efficiency of the proposed dimension reduction techniques.

Jian Huang is a chair professor of Data Science and Analysis in the Department of Applied Mathematics at The Hong Kong Polytechnic University. He obtained his bachelor's degree in mathematics and a master's degree in statistics from Wuhan University, followed by a PhD in statistics from the University of Washington in Seattle. He once served as a professor in the Department of Statistics and Actuarial Science and the Department of Biostatistics at the University of Iowa. His research interests include machine learning, high-dimensional statistics, computational statistics, biostatistics, and bioinformatics. From 2015 to 2019, he was designated a highly cited researcher in the field of Mathematics by the Web of Science group at Clarivate and included in the list of the top 2% of the world's most cited scientists by Stanford University in 2022 and 2023. He is a fellow of American Statistical Association (ASA) and a fellow of the Institute of Mathematical Statistics (IMS).

Statistical Deep Learning

Deep learning has achieved remarkable success in a wide range of applications and has been applied to analyzing high-dimensional data in many fields of scientific research. Therefore, it would be interesting to understand why deep learning has been so successful and what its main advantages are over the traditional nonparametric methods developed over the decades, if any. In this talk, we try to explain some advantages of deep learning by considering the approximation power of deep neural networks and generalization errors of deep learning methods. Using nonparametric regression and conditional generative learning as examples, we illustrate deep learning approaches to dealing with some high-dimensional statistical problems.

Kim-Chuan Toh is the Leo Tan Professor in the Department of Mathematics at the National University of Singapore. He works extensively on convex programming, particularly large-scale matrix optimization problems such as semidefinite programming and sparse convex problems arising from machine learning and statistics. Currently he serves as a co-Editor for Mathematical Programming, an Area Editor for Mathematical Programming Computation, an Associate Editor for SIAM Journal on Optimization, and ACM Transactions on Mathematical Software. He has also served in professional committees such as SIAM Fellow Selection Committee and ICIAM's Collatz Prize Committee. He received the Farkas Prize awarded by the INFORMS Optimization Society in 2017 and the triennial Beale-Orchard Hays Prize awarded by the Mathematical Optimization Society in 2018. He received the Singapore President's Science Award in 2019. He was elected as a Fellow of SIAM (Society for Industrial and Applied Mathematics) in 2018 and a Fellow of the Singapore National Academy of Science in 2022.

Stochastic Bregman Gradient Methods with Application to Deep Learning

The widely used stochastic gradient methods for minimizing nonconvex composite objective functions require the Lipschitz smoothness of the differentiable part. But the requirement does not hold true for problem classes including quadratic inverse problems and training neural networks. To address this issue, we investigate a family of stochastic Bregman proximal gradient (SBPG) methods, which only require smooth adaptivity of the differentiable part. SBPG replaces the upper quadratic approximation used in SGD with the Bregman proximity measure, resulting in a better approximation model that captures the non-Lipschitz gradients of the nonconvex objective. We formulate the vanilla SBPG and establish its convergence properties under nonconvex setting without finite-sum structure. Experimental results on quadratic inverse problems testify the robustness of SBPG. Moreover, we propose a momentum-based version of SBPG (MSBPG) and prove it has improved convergence properties. We apply MSBPG to the training of deep neural networks with a polynomial kernel function, which ensures the smooth adaptivity of the loss function. Experimental results on representative benchmarks demonstrate the effectiveness and robustness of MSBPG in training neural networks.

Xudong Li is a tenure-track associate professor at the School of Data Science, Fudan University. Before joining Fudan, he was a postdoctoral researcher at Princeton University and the National University of Singapore. His research focuses on the theory, algorithms and applications of large-scale optimization problems in data science. He has won Young Researcher Prize in Continuous Optimization of the Mathematical Optimization Society, the ICML Outstanding Paper Award, and the Young Researcher Prize of the Operations Research Society of China. He is currently serving as an associate editor for Mathematical Programming Computation.

Data-Driven Minimax Optimization with Expectation Constraints

Attention to data-driven optimization approaches has grown significantly over recent decades, but data-driven constraints have rarely been studied. In this talk, we focus on the non-smooth convex-concave stochastic minimax regime and formulate the data-driven constraints as expectation constraints. Then, we propose a class of efficient primal-dual algorithms to tackle the minimax optimization with expectation constraints, and show that our algorithms converge at the optimal rate of $O(\frac{1}{\sqrt{N}})$, where N is the number of iterations. We also verify the practical efficiency of our algorithms by conducting numerical experiments on large-scale real-world applications.

Falong Tan is an associate professor in the department of Statistics at Hunan University. Prior to joining Hunan University in December 2017, he obtained his Ph.D. from Hong Kong Baptist University in 2017, under the supervision of Prof. Lixing Zhu.

Weighted Residual Empirical Processes, Martingale Transformations, And Model Specification Tests for Regressions with Diverging Number of Parameters

We propose a new methodology for testing the parametric forms of the mean and variance functions based on weighted residual empirical processes and their martingale transformations in regression models. The dimensions of the parameter vectors can be divergent as the sample size goes to infinity. We study the convergence of weighted residual empirical processes and their martingale transformation under the null and alternative hypotheses in diverging dimension settings. The proposed tests based on weighted residual empirical processes can detect local alternatives distinct from the null at the fastest possible rate of order $n^{-1/2}$ but are not asymptotically distribution-free. While tests based on martingale transformed weighted residual empirical processes can be asymptotically distribution-free, yet, unexpectedly, can only detect the local alternatives converging to the null at a much slower rate of order $n^{-1/4}$, which is somewhat different from existing asymptotically distribution-free tests based on martingale transformations. As the tests based on the residual empirical process are not distribution-free, we propose a smooth residual bootstrap and verify the validity of its approximation in diverging dimension settings. Simulation studies and a real data example are conducted to illustrate the effectiveness of our tests.

Jiujing Wu is a doctoral student at the School of Mathematical Sciences Capital Normal University. She mainly focuses on high-dimensional data analysis and nonparametric statistics.

Model-Free Feature Screening Based on Hellinger Distance for Ultrahigh Dimensional Data

With the explosive development of data acquisition and processing technology, feature dimensions increase exponentially with sample size, posing significant challenges for data analysis. It is crucial to accurately identify useful features from thousands available. In this paper, we develop an omnibus model-free feature screening procedure based on the Hellinger distance, offering several appealing merits. First, we define the Hellinger distance index for discrete response variables in discriminant analysis. Second, this procedure consistently works for continuous response variables, where the responses are discretized using a slice-and-fused technique. Third, it is robust against potential outliers and model misspecification. Theoretically, the procedure for both discrete and continuous response variables exhibit sure screening and ranking consistency properties under mild conditions. Numerical studies show that this procedure is highly competitive in heavy-tailed and skewed data, as well as maintaining comparability with existing approaches for light-tailed data, indicating robust performance across various data types. The real data sets, one with discrete and the other with continuous response variables, demonstrate the effectiveness of the proposed method.

Defeng Sun is currently a chair professor of Applied Optimization and Operations Research at the Hong Kong Polytechnic University and the President of the Hong Kong Mathematical Society. He mainly publishes in non-convex continuous optimization and machine learning. Together with Professor Kim-Chuan Toh and Dr Liuqin Yang, he was awarded the triennial 2018 Beale-Orchard-Hays Prize for Excellence in Computational Mathematical Programming by the Mathematical Optimization Society. He served as editor-in-chief of Asia-Pacific Journal of Operational Research from 2011 to 2013 and now serves as associate editor of Mathematical Programming, SIAM Journal on Optimization, Journal of Optimization Theory and Applications, Journal of the Operations Research Society of China, Journal of Computational Mathematics, and Science China: Mathematics. In 2020, he was elected as a Fellow of CSIAM and a Fellow of SIAM and in 2021 he has received the Distinguished Collaborator Award from both the Hong Kong Research Center and Huawei Noah's Ark Lab for the contributions on developing efficient and robust techniques for solving huge scale linear programming models. He was also a recipient of the 2022 RGC Senior Research Fellow Scheme award.

Nonsmooth Analysis and Sparse Optimization

In this talk, we shall explain why nonsmooth analysis plays a critical role in solving large scale sparse optimization problems. We start with introducing some basic concepts such as Rademacher' theorem and the Moreau-Yosida regularization for convex functions. Then we talk about semismooth analysis including inverse and implicit functions theorems to demonstrate why nonsmooth systems are indispensable for solving constrained optimization problems and why smooth systems inevitably lead to singularity. Finally, we shall illustrate how we can employ nonsmooth analysis to design highly efficient sparse nonsmooth Newton methods and level set secant methods for solving several important machine learning models including convex clustering, lasso, and exclusive lasso of sparse solutions.

Jiashun Jin is Professor in Statistics & Data Science and Affiliated Professor in Machine Learning at Carnegie Mellon University. He received his Ph.D in Statistics from Stanford University in 2003. His earlier work was on large-scale multiple testing, focusing on the development of (Tukey's) Higher Criticism and practical False Discovery Rate (FDR) controlling methods. His more recent interest is on the analysis of social networks and text documents, focusing on the development of the SCORE normalization and on cycle count methods. He has also led a team collecting and analyzing a high-quality large-scale data set on the publications of statisticians. Jin is an elected IMS fellow and an elected ASA fellow, and he has delivered the highly selective IMS Medallion Lecture in 2015 and IMS AoAS (Annals of Applied Statistics) Lecture in 2016. Jin has served as Associate Editor for several statistical journals and he is current serving as the IMS treasurer.

The Statistical Triangle

In his Fisher's Lecture in 1996, Efron suggested that there is a philosophical triangle in statistics with "Bayesian", "Fisherian", and "Frequentist" being the three vertices, and many representative statistical methods can be viewed as a convex linear combination of the three philosophies. We collected and cleaned a data set consisting of the citation and bibtex (e.g., title, abstract, author information) data of 83,331 papers published in 36 journals in statistics and related fields, spanning 41 years. Using the data set, we constructed 21 co-citation networks, each for a time window between 1990 and 2015. We propose a dynamic Degree-Corrected Mixed- Membership (dynamic-DCMM) model, where we model the research interests of an author by a low-dimensional weight vector (called the network memberships) that evolves slowly over time. We propose dynamic-SCORE as a new approach to estimating the memberships. We discover a triangle in the spectral domain which we call the Statistical Triangle, and use it to visualize the research trajectories of individual authors. We interpret the three vertices of the triangle as the three primary research areas in statistics: "Bayes", "Biostatistics" and "Nonparametrics". The Statistical Triangle further splits into 15 sub- regions, which we interpret as the 15 representative sub-areas in statistics. These results provide useful insights over the research trend and behavior of statisticians.

Zhihua Su is associate professor in Department of Statistics at the University of Florida. She received her bachelor's degree from Fudan University in 2006 and got PhD in statistics at the University of Minnesota in 2012. Her research interest includes dimension reduction, variable selection, Bayesian analysis, and nonparametric statistics. Her primary research interests fall in a new area at the intersection of multivariate analysis and dimension reduction, called envelopes, which was introduced by Cook, Li and Chiaromonte (2010) under the framework of multivariate linear regression.

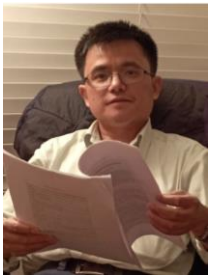
Envelope-Based Partial Least Squares

Partial least squares (PLS) is widely used in applied sciences an alternative method to ordinary least squares (OLS) for estimating the regression coefficients. It is known that PLS often has a better prediction performance compared to OLS, and the PLS algorithms can be adapted directly to the $n < p$ case. Despite its popularity, the theoretical properties of the PLS estimator are largely unknown. As a result, it is hard to determine when PLS is better than OLS, what are the limitations for PLS and how to improve PLS. Cook et al. (2013) built a connection between PLS with a dimension reduction method called the envelope model. They showed that at the population level, PLS and the envelope model have the same target parameter, but they use different algorithms for estimation. This connection allows PLS to be studied in a traditional likelihood framework and facilitates model developments. We will address three issues of PLS in this context: variable selection, categorical predictors and scale invariance.

Ying Cui is currently an assistant professor in the Department of Industrial Engineering and Operations Research at the University of California Berkeley. Prior to that appointment, she was an assistant professor at the University of Minnesota. She worked as postdoc research associate in the Daniel J. Epstein Department of Industrial and Systems Engineering at the University of Southern California working with Professor Jong-Shi Pang. Cui completed her PhD in Mathematics at the National University of Singapore. Her research focuses on the mathematical foundation of data science with emphasis on optimization techniques for operations research, machine learning and statistical estimations. She is particularly interested in leveraging nonsmoothness to design efficient algorithms for large scale nonlinear optimization problems. She is the co-author of the recently published monograph "Modern Nonconvex Nondifferenable Optimization".

Optimization with Superquantile Constraints: A Fast Computational Approach

We present an efficient and scalable second-order computational framework for solving large-scale optimization problems with superquantile constraints. Unlike empirical risk models, superquantile models have non-separable constraints that make typical first-order algorithms difficult to scale. We address the challenge by adopting a hybrid of the second-order semismooth Newton method and the augmented Lagrangian method, which takes advantage of the structured sparsity brought by the risk sensitive measures. The key to make the proposed computational framework scalable in terms of the number of training data is that the matrix-vector multiplication in solving the resulting Newton system can be computed in a reduced space due to the aforementioned sparsity. The computational cost per iteration for the Newton method is similar or even smaller than that of the first-order method. Our developed solver is expected to help improve the reliability and accuracy of statistical estimation and prediction, as well as control the risk of adverse events for safety-critical applications.



Jiawang Nie is currently a professor of Mathematics in University of California, San Diego. He works in the broad area of applied and computational mathematics. His research areas are optimization, convex algebraic geometry, tensor computation, and their applications in data sciences. He received the Tucker Prize Finalist (2009), NSF Career Award (2009), Informs Optimization Prize for Young Researchers (2014), Chang-Jiang Scholar (2017), SIAM SIAG/Linear Algebra Best Paper Prize (2018), and Feng Kang Prize 2021. He serves in the editor boards for the journals: Journal of Operations Research Society of China, Computational Optimization and Applications, Mathematics of Operations Research, and SIAM Journal on Matrix Analysis and Applications.

The Multi-Objective Polynomial Optimization

The multi-objective optimization is to optimize several objective functions over a common feasible set. Because the objectives usually do not share a common optimizer, people often consider (weakly) Pareto points. This paper studies multi-objective optimization problems that are given by polynomial functions. First, we study the geometry for (weakly) Pareto values and represent Pareto front as the boundary of a convex set. Linear scalarization problems (LSPs) and Chebyshev scalarization problems (CSPs) are typical approaches for getting (weakly) Pareto points. For LSPs, we show how to use tight relaxations to solve them and how to detect existence or nonexistence of proper weights. For CSPs, we show how to solve them by moment relaxations. Moreover, we show how to check whether a given point is a (weakly) Pareto point or not and how to detect existence or nonexistence of (weakly) Pareto points. We also study how to detect unboundedness of polynomial optimization, which is used to detect nonexistence of proper weights or (weakly) Pareto points.

School of Mathematics and Statistics

The School of Mathematics and Statistics of Beijing Jiaotong University was founded in March 2022. Its predecessor was the Department of Applied Science of Northern Jiaotong University in February 1960, the Department of Mathematics was established in October 1986, and the College of Arts and Science was merged in July 1996. In the past 20 years, mathematics has made outstanding achievements in discipline development, talent training and scientific research. It was for this aim that the School of Science was founded in August 1998, consisting of five second Discipline Departments: Department of Mathematics, Department of Physics, Department of Chemistry, Institute of Optoelectronic Technology, and the College of Life Sciences and Bioengineering.

The School of Mathematics and Statistics consists of 4 departments and 1 virtual body research institute: Department of Information and Computing Science, Department of Mathematics and Applied Mathematics, Department of Statistics and Operations Research, Department of Data Science, and Institute of Basic and Interdisciplinary Science of Beijing Jiaotong University (Virtual body). It has two first-level doctoral programs in mathematics and statistics, covering six directions: basic mathematics, computational mathematics, probability theory and mathematical statistics, applied mathematics, operations research and cybernetics, and statistics. It also has a postdoctoral mobile station in mathematics, and has obtained the 111 innovation and intelligence Introduction base of information and transportation Operations research. Statistics and Operations Research has ranked among the world's top 160 in the "QS World University Subject Rankings" for three consecutive years, and the top 5% in ESI global subject rankings, which has made important contributions to the innovation and integration development of disciplines and the "double first-class" construction of our university. The school has research platforms such as Statistics Experimental Center, Data Science and Engineering Laboratory, and Scientific Computing Center Key Laboratory, and is actively preparing to build a comprehensive statistics experimental platform to impact the key laboratory of the Ministry of Education.

The school has 88 full-time teachers, including 1 academician of Chinese Academy of Sciences, 2 famous teaching teachers of Beijing City, 2 famous young teaching teachers of

Beijing City, 2 professors selected in the National Youth Talent Plan, 27 doctoral tutors, 59 master tutors, 29 professors and 34 associate professors. 75% of them have more than one year of study abroad experience. The proportion of teachers with doctoral degrees is 86%. There are 1 provincial and ministerial level undergraduate gold courses, 4 postgraduate professional basic courses "professional gold courses", (Foundation of Modern Analysis, Foundation of Algebra, Foundation of Topology and Geometry, Foundation of Probability Theory). In terms of professional construction, Information and computing science and Statistics were rated as the national first-class undergraduate program construction sites, and Mathematics and Applied mathematics were rated as the first-class undergraduate programs in Beijing. The school has a high level of teaching staff, gives full play to the advantages of operations research and cybernetics, applied mathematics, statistics and other disciplines, and is active in the forefront of applied research. In addition, it participated in interdisciplinary research, as well as applied combinatorics and coding theories to research practical problems such as information transmission, information security, and network, which provided scientific support for social development and industry standard formulation. In the past five years, the school has presided over 4 National Natural Fund key projects, 1 provincial and ministerial level or above scientific and technological achievement award and hosted many large domestic high-level academic conferences and international conferences.

Academy of Fundamental and Interdisciplinary Sciences

To implement the "Twelve–Five planning", develop new field of interdisciplinary subjects and achieve the goal of the 10th congress of party representatives of establishing the world's first–class universities, Beijing Jiaotong University set up Academy of Fundamental and Interdisciplinary Sciences (AFIS) on May 3, 2013. The research team consists of staffs from School of Science, Electronic Information Engineering, Computer and Information Technology, Traffic and Transportation, Economics and Management who are engaged in researches on basic and interdisciplinary subjects. Now, the academy has three research centers: the research center of stochastic structure and data science, the research center of combination and optimization, the research center of bioinformatics and interdisciplinary.

The positioning and development directions: strengthening the integration between basic science and other characteristic disciplines such as information, management, transportation and engineering science; building some high-level creative teams; for major problems in those disciplines, researchers can jointly work and obtain distinct results; culturing a group of innovative talents in interdisciplinary; elevating overall strength of basic and engineering science in our university so as to reach the leading level in China and abroad.

Research Center of Stochastic Structure and Data Science

In order to process mass data generated by information, management, transportation, biology and engineering research, the center mainly work on constructing statistic and stochastic model, applying high–speed computer to deal with all kinds of data which provide theoretical and technical supports for practical use. The center's members have got various research results in the above fields. Hundreds of articles appeared in the related top journals. With the support from the National 973 Program and Natural Science Foundation of China and other national projects, one Second Class National Prize of Natural Science has been awarded no mention to several provincial science and technology awards.

Research Center of Combination and Optimization

Currently, with the rapid development of computer technology and information technology, combination and optimization are one of the most basic and supportive disciplines. The center mainly works on some scientific problems of coding and cryptology theory, symmetry of graphs, reliability of the network and optimization. In these areas, the center's members have got various research results. Many articles appeared in the related top journals such as IEEE Transaction on Information Theory, IEEE Transaction on Computers, Transportation Research–Part B, Journal of Combinatorial Theory and Series A (B). Four achievements have been awarded by Ministry of Education. The members of the center also take several National Natural Science Foundation and other national research projects. Extensive cooperation and communication are under way with renowned universities and research institutions in Australia, Japan, Korea, Italy, Slovenia, etc.

Research Center of Bioinformatics and Interdisciplinary

The center mainly studies on some important scientific problems of biological information and biological interdisciplinary. There are many papers published in top journals of the related field, such as American Journal of Human Genetics, Bioinformatics, Journal of Bone and Mineral Research, Human Genetics, Journal of Materials Chemistry and Advanced Materials, etc. The members of the center also take several National Natural Science Foundation and other national research projects. Extensive cooperation and communication are under way with renowned American universities and research institutions in American.

Map of Beijing Jiaotong University