Anthony Man-Cho So, Prateek Jain, Wing-Kin Ma, and Gesualdo Scutari

Nonconvex Optimization for Signal Processing and Machine Learning

Optimization is now widely recognized as an indispensable tool in signal processing (SP) and machine learning (ML). Indeed, many of the advances in these fields rely crucially on the formulation of suitable optimization models and deployment of efficient numerical optimization algorithms. In the early 2000s, there was a heavy focus on the use of convex optimization techniques to tackle SP and ML applications. This is largely due to the fact that convex optimization problems often possess favorable theoretical and computational properties and that many problems of practical interest have been shown to admit convex formulations or good convex approximations.

In 2010—exactly a decade ago—*IEEE Signal Processing Magazine (SPM)* published a special issue on convex optimization and signal processing in which many successful applications of convex optimization techniques in SP were showcased. At that time, it was a common belief that nonconvex optimization problems were intractable and lacked strong theoretical properties. Nevertheless, not long after the publication of the aforementioned special issue, works have started to emerge showing that various SP and nondeep ML applications give rise to well-structured nonconvex formulations, which often exhibit properties akin to those of convex optimization problems and can be solved to optimality more efficiently than their convex reformulations or approximations. This line of study has since blossomed into a highly active research area and led to new insights into the structures and tractability of nonconvex optimization problems.

In view of the significant impact of nonconvex optimization techniques on both our theoretical understanding of and algorithmic capability to handle contemporary SP and ML applications, we have put together this special issue to introduce the essential elements of nonconvex optimization to the broader SP and ML communities, provide insights into how structures of the nonconvex formulations of various practical problems can be exploited in algorithm design, showcase some notable successes in this line of study, and identify important research issues that are motivated by existing or emerging applications. Through this special issue, we aim to promote cross-fertilization among SP, ML, and optimization, where advanced optimization tools enable innovations in SP and ML, while frontier applications in SP and ML drive the development of new optimization techniques.

We are extremely grateful for the enthusiastic response to our initial call for papers. We received more than 50 white papers, with coverage ranging from the fundamentals of nonconvex optimization to algorithm design and analysis to applications in SP and ML. Many of the white papers were of high quality. This created a problem where we strived to provide comprehensive coverage while obeying the magazine's page length constraints. Ultimately, we were able to accommodate only nine articles. We thank all those who have shown interest in this special issue.

As alluded to previously, a key to the success in handling nonconvex optimization problems that arise in applications is to elucidate their structures and develop algorithms that can take advantage of those structures. Each of the articles in this special issue covers one or both of these aspects. Motivated by the prevalence of nonsmooth optimization problems in applications, the article by Li et al. introduces the basic elements of nonsmooth optimization. In particular, the article takes the reader through different constructions of the subdifferential for several important classes of nonsmooth functions. These constructions give rise to different stationarity concepts, and the article explains their subtleties and the relationship among them. The article also points out the algorithmic challenges in computing the stationary points defined by these concepts.

The next three articles cover algorithmic techniques for dealing with nonconvex optimization problems that arise from SP and ML applications. The article by Curtis and Scheinberg discusses the design and analysis of adaptive optimization methods, which aim to avoid the expensive tuning of algorithmic parameters that is common in existing stochastic methods.

Digital Object Identifier 10.1109/MSP.2020.3004217
Date of current version: 2 September 2020

The authors present a unified algorithmic and analytic framework for adaptive deterministic and stochastic optimization, which sets the stage for the development of adaptive methods that are not only practical but also possess strong theoretical guarantees.

The next article, by Liu et al., surveys the recent developments in zeroth-order optimization, which is gaining more attention due to its relevance in SP and deep learning. The authors review the basics of zeroth-order optimization and present a unified framework that captures a wide range of zeroth-order optimization algorithms. They then showcase a number of applications—including adversarial example generation for deep learning models, automated ML, and policy search in reinforcement learning—for which zeroth-order optimization techniques have shown to be promising.

Another topic that has attracted much interest lately is min-max optimization. The article by Razaviyayn et al. first explains how min-max problems manifest themselves in contemporary applications such as adversarial ML, fair beamforming, and generative adversarial network training. A key feature of these min-max problems is that the objective function in question is not necessarily convex (respectively, concave) with respect to the minimization (respectively, maximization). This begs the question: what does a "stationary point" of such a min-max problem look like? The article reviews some existing stationarity notions for (non-)convex (non-)concave min-max optimization problems and provides an overview of recent theoretical and algorithmic advances in the study of such problems.

The three articles that follow provide an in-depth treatment of the structures of nonconvex optimization problems arising from concrete applications and demonstrate how these structures can be exploited in computation. The article by Vaswani focuses on the classic phase retrieval problem and surveys recent nonconvex optimization approaches that can exploit structural assumptions on the signal to achieve fast recovery with low sample complexity. It also presents experimental results to address the question of which structural assumptions are suitable for what types of real data sets.

Fu et al. provide, in their article, an overview of recent optimization approaches for large-scale structured matrix and tensor decomposition, which has become an indispensable tool for various SP and ML applications. In particular, the article presents a unified nonconvex formulation of the structured matrix/tensor decomposition problem, in which various structural requirements, such as nonnegativity, smoothness, and sparsity, can be incorporated. It then details various algorithmic strategies for handling those matrix/tensor structures and shows how they lead to provably convergent and practically efficient methods for solving the decomposition problem.

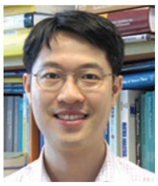
Another topic that has been the subject of intense study in recent years is understanding the landscape of the loss function associated with neural network training. This is motivated by the empirical observation that neural networks can often be trained to attain the global minimum of the loss function at hand, even though the loss function is nonconvex. The article by Sun et al. surveys recent attempts to rigorously justify such an intriguing phenomenon and elucidates the role that the neural network architecture plays in achieving good training results. Furthermore, it discusses how the loss function can be designed to achieve more favorable landscape properties.

The last two articles introduce theoretical and algorithmic tools for tackling general structured nonconvex optimization problems. The article by Manton gives an introduction to optimization over smooth manifolds. Such problems arise in a number of SP and ML applications, including dictionary learning, principal component analysis, and low-rank matrix completion, just to name a few. The article demonstrates that many unconstrained optimization techniques and methods can be generalized to tackle optimization problems with manifold constraints.

Finally, the article by Tohidi et al. gives a survey on submodular optimization and its applications in SP and ML. Submodularity is a property of set functions that is akin to convexity of real-valued functions. Although submodular optimization is a branch of discrete optimization, the nice structures of submodular set functions make it possible to optimize these functions in an efficient manner. The article takes the reader through different classes of submodular optimization problems and demonstrates how the submodular structure can be exploited in the design of scalable optimization methods with provable performance guarantees. The article also showcases several applications—including resource selection in multiple-input, multiple-output radar and feature selection for classification—from which submodular optimization problems arise.

We would like to express our deepest gratitude to the contributing authors and anonymous reviewers. We especially thank the members of the editorial board for their continual support and guidance. This special issue would simply not be possible without the contributions of all those involved. To close with a metaphor, looking back to *SPM*'s special issue a decade ago, we were climbing hills; today, we seek to explore jagged mountain ranges.

Guest Editors



Anthony Man-Cho So (manchoso@se.cuhk.edu.hk) received his B.S.E. degree in computer science from Princeton University, New Jersey, with minors in applied and computational mathematics, engineering and management systems, and German language and culture. He then received his M.Sc. and Ph.D. degrees in computer science and his Ph.D. minor in mathematics from Stanford University, California. He joined the Chinese University of Hong Kong (CUHK) in 2007, where he is now a professor in the Department of Systems Engineering and Engineering Management. His research focuses

on optimization theory and its applications in various areas of science and engineering. He was appointed as an Outstanding Fellow of the Faculty of Engineering at CUHK in 2019. He received the 2018 IEEE Signal Processing Society Best Paper Award, the 2015 IEEE Signal Processing Society Signal Processing Magazine Best Paper Award, the 2014 IEEE Communications Society Asia-Pacific Outstanding Paper Award, and the 2010 INFORMS Optimization Society Optimization Prize for Young Researchers. He is a Senior Member of IEEE.

Prateek Jain (prajain@microsoft.com) received his Ph.D. degree from the Computer Science Department at the University of Texas at Austin and his B.Tech. degree from the Indian Institute of Technology, Kanpur (IIT Kanpur). He is a senior principal researcher at Microsoft Research India and an adjunct faculty member at IIT Kanpur. He wrote a monograph on nonconvex optimization in machine learning (ML), summarizing many of his results. He regularly serves on the senior program committee of top ML conferences and is an action editor of *Journal of Machine Learning Research*. He was the recipient of the Best Student Paper Award at the International Conference on Machine Learning in 2007 and at the Conference on Computer Vision and Pattern Recognition in 2008. More recently, his work on alternating minimization was selected as the 2020 Best Paper by the IEEE Signal Processing Society.

Wing-Kin Ma (wkma@ieee.org) received his M.Phil. and Ph.D. degrees from the Chinese University of Hong Kong, where he is currently a professor with the Department of Electronic Engineering. His research focuses on signal processing, optimization, and communications. He is a corecipient of the 2015 IEEE Signal Processing Magazine Best Paper Award, the 2016 IEEE Signal Processing Letters Best Paper Award, and a 2018 IEEE Signal Processing Best Paper Award. He previously served as an associate editor and a senior area editor of *IEEE Transactions on Signal Processing* and a special issue guest editor and lead guest editor of *IEEE Signal Processing Magazine*. He is a Fellow of IEEE, was an IEEE Signal Processing Society (SPS) Distinguished Lecturer (2018–2019), and is currently an IEEE SPS regional director-at-large (Region 10).

Gesualdo Scutari (gscutari@purdue.edu) received his electrical engineering and Ph.D. degrees (both with honors) from the University of Rome "La Sapienza," Italy, in 2001 and 2005, respectively. He is the Thomas and Jane Schmidt Rising Star Associate Professor with the School of Industrial Engineering, Purdue University, West Lafayette, Indiana. His research interests include continuous and distributed optimization, equilibrium programming, and their applications. Among others, he was the recipient of the 2013 National Science Foundation CAREER Award, the 2015 Anna Maria Molteni Award for Mathematics and Physics, and the 2015 IEEE Signal Processing Society Young Author Best Paper Award. He serves on the editorial boards of *SIAM Journal on Optimization* and *IEEE Transactions on Signal Processing* (senior area editor). He is also a guest editor of this issue of *IEEE Signal Processing Magazine*. He is a Senior Member of IEEE.

SP