SN Operations Research Forum (2020) 1: 6
https://doi.org/10.1007/s43069-020-0006-9

BOOK REVIEWS

Ioannis C. Demetriou and Panos M. Pardalos (eds): Approximation and Optimization: Algorithms, Complexity and Applications

Oleg Burdakov[1]

Published online: 5 March 2020

Springer Optimization and Its Applications,
ISBN 978-3-030-12766-4, ISBN 978-3-030-12767-1 (eBook),
https://doi.org/10.1007/978-3-030-12767-1,
Springer Nature Switzerland AG 2019, 237+x pp

Approximation and Optimization: Algorithms, Complexity and Applications, 2019 is an edited and reviewed collection of invited papers that were presented at a conference held in June, 2017 in National and Kapodistrian University of Athens, Greece.

The collection of individual papers is a mix of state of the art research articles on Approximation and Optimization, surveys of new work, and relevant applications in science, engineering, and social sciences. The book consists of eleven chapters. A brief survey to the papers is presented in the first chapter. A short review of the later chapters is given in the following paragraphs.

The second chapter ("Evaluation Complexity Bounds for Smooth Constrained Nonlinear Optimization using Scaled KKT Conditions and High-order Models" by C. Cartis, N.I.M. Gould, and P.L. Toint) considers evaluation complexity for solving constrained optimization. This paper presents theoretical results on a subject where the authors have made important contributions. It starts by considering the case with convex constraints and shows that the complexity bound of $O(\epsilon^{-3/2})$ derived from previous work of the authors for computing an ϵ-approximate first-order critical point can be obtained under significantly weaker assumptions. This result is generalized to the case where pth-order derivatives are available resulting in a bound of $O(\epsilon^{-(p+1)/p})$ evaluations. It is also shown that the bound of

✉ Oleg Burdakov
oleg.burdakov@liu.se

1 Department of Mathematics, Linköping University, Linköping, Sweden

$O(\epsilon_P^{-1/2}\epsilon_D^{-3/2})$ evaluations for the general nonconvex case subject to both equality and inequality constraints can be generalized to yield a bound of $O(\epsilon_P^{-1/p}\epsilon_D^{-(p+1)/p})$ evaluations under similarly weaker assumptions, where ϵ_P and ϵ_D are the primal and dual accuracy thresholds.

Data-dependent approximation has attracted a lot of research especially in the area of social computing. The third chapter ("Data-Dependent Approximation in Social Computing" by Weili Wu, Yi Li, P.M. Pardalos, and Ding-Zhu Du) presents some theoretical results on nonsubmodular optimization problems and discusses on related problems open to research.

The fourth chapter ("Multi-objective evolutionary optimization algorithms for machine learning: A recent survey" by S.-A.N. Alexandropoulos, C.K. Aridas, S.B. Kotsiantis, and M.N. Vrahatis) presents recent multi-objective evolutionary approaches for four major data mining and machine learning tasks, namely data preprocessing, classification, clustering, and association rules. The chapter provides a survey of the basic ideas, rather than a simple list of all the research papers that have discussed or used these ideas. The references cover the most important theoretical issues, give guidelines to the main branches of literature regarding such techniques and schemes, and lead the reader to up-to-date research directions. The paper points out connections not noticed before and provides a well-guided tour through this research area with efficient information flow.

Researchers in optimization rate the effectiveness of their algorithms by comparing them against their competitors on classes of selected optimization problems. Analogously, comparison methods of supervised learning algorithms when dealing with issues such as generalization, early-stopping and cross validation of classification, and other machine learning problems rely on statistical analysis of the results obtained subject to interpretation, which frequently provides controversial conclusions. Based on solid mathematical constructs, D.H. Wolpert addressed these matters formulating a theory known as No Free Lunch Theorems, which proves that it is meaningless to claim that an algorithm without re-sampling performs better than its competitors in all optimization problems. The fifth chapter ("No free lunch theorem: A review" by S.P. Adam, S.-A.N. Alexandropoulos, P.M. Pardalos, and M.N. Vrahatis), with respect to the original work by Wolpert, delineates the essential concepts of the theory presenting a detailed discussion of the theoretical constructs of No Free Lunch Theorems both for optimization and supervised learning. It devotes distinct sections to scientific area of interest such as Optimization and Search, Optimization and Evolutionary Algorithms including Meta-Heuristic Techniques, and Supervised Learning issues with emphasis on Real-World Machine Learning. In each section, the most relevant theoretical issues are presented with respect to No Free Lunch Theorems and supported by a list of the most significant research efforts reported in the literature. The wideness of the scientific subjects addressed together with the clarity and the consistency of the presentation and, finally, the application issues raised are the constituents making this paper a very helpful guide for any researcher interested in the No Free Lunch Theorems.

Chapter 6 ("Piecewise Convex-Concave Approximation in the Minimax Norm" by M.P. Cullinan) addresses the problem of making the least maximum change to n noisy data subject to q sign changes in the second-order consecutive-divided differences of the smoothed values. An algorithm is developed that obtains the optimal solution in only $O(qn \log n)$ computer operations, attention is paid to complexity analysis, and diagrams illustrate the proofs. Some numerical results show that the calculation is particularly efficient for large densely packed data, even when the errors are quite large. The construction approach is as follows. When $q = 0$, the best convex approximation is obtained by adding a constant to the lower convex hull of the data. When $q = 1$, the best convex-concave approximation is found from the best approximation when $q = 0$. The best approximations with larger values of q are found from the best approximation when q is reduced by 2. The optimization techniques obtain the solution by adjustments that depend on local information. In summary, the method has definite practical importance and the paper makes a useful contribution to the subject of data approximation.

Chapter 7 ("A Decomposition Theorem for the Least Squares Piecewise Monotonic Data Approximation Problem" by I.C. Demetriou) considers the least squares change to n univariate data subject to the condition that the first differences of the estimated values have at most q sign changes. The situation compared with the one in the previous paragraph where the objective function is the minimax norm is quite different. Indeed, in the least squares case, any general optimization algorithm that uses local information will stop at a local minimum, since the set of local minima is composed of discrete points. These difficulties are caused by the enormous number of isolated local solutions of the optimization calculation that can occur in practice, namely $O(n^q)$. However, much less work is needed because of a decomposition property of the solution and this chapter gives necessary and sufficient conditions. Specifically, the case when $q = 0$ gives a monotonic fit to the data and when $q > 0$, the problem decomposes into at most q least squares monotonic approximation problems to disjoint sets of adjacent data. The important consequence of the decomposition is that it allows the development of a dynamic programming procedure that provides the solution in only $O(n^2 + qn \log n)$ operations, which is a previous work of the author. Some numerical results show that the calculation is quite efficient. The efficacy of the calculation is demonstrated by estimating the peaks of a nuclear magnetic resonance spectrum that consists of $110,000$ measurements when $q = 49$. This problem requires a formidable number of combinations to be tested in order to obtain optimality, but the given procedure gave the optimal solution in less than a minute on a pc. Some figures illustrate this result.

The development of modern electronic devices is closely related to many deep mathematical problems. The itself invention by W. Cauer of now ubiquitous low and high-pass elliptic electrical filters became possible only after the explicit solution of the best uniform approximation problem on two intervals of the signum function by a rational one was given by E. I. Zolotarëv. Challenges of contemporary telecommunication industry bring us to significantly more sophisticated problems. One of those is the extension of Zolotarëv problem to more than two intervals. Chapter 8 ("Recent Progress in Optimization of Multiband Electrical Filters" by A. Bogatyrëv) gives a historical review of those optimization problems and a description of several approaches to their solution, including a recently developed one which is based on the geometry of algebraic curves. Several figures illustrate these results.

Chapter 9 ("Impact of Error in Parameter Estimations on Large Scale Portfolio Optimization" by V.A. Kalyagin and S.V. Slashchinin) examines how estimation error for means and covariance matrix of stock returns may affect the results of selected portfolios. Different experiments using both real data from different stock markets and generated data were run in order to compare the out-of-sample performance of the estimators and the influence of the estimation error on the portfolio selection. It is observed that for large-scale portfolio optimization, the efficiency of the obtained optimal portfolio is biased with respect to the true optimal portfolio. Different aspects of this phenomenon and possible ways to reduce its negative effect are discussed. These results are illustrated by some figures.

Chapter 10 ("Optimal Design of Smart Composites" by G. Tairidis, G. Foutsitzi, and G. Stavroulakis) investigates optimal design problems related to smart composites. First, the mechanical properties of a smart composite are tailored to meet required specifications. Second, pointwise optimization leading to functionally graded composites or even topology optimization is applied and a cantilever beam with two materials is briefly presented. Further, the control subsystem has several parameters to be optimized: number and position of sensors and actuators, as well as the parameters of the controller. Here, some basic techniques regarding soft control based on fuzzy and neuro-fuzzy strategies are presented, along with optimization options and methods which can be used for the fine-tuning of the parameters of the system. Rich numerical results indicate both the capabilities of control in structural design and the availability of optimization techniques to built an efficient control procedure.

Chapter 11 ("Tax Evasion as an Optimal Solution to a Partially Observable Markov Decision Process" by P. Papadopoulou and D. Hristu-Varsakelis) explores tax evasion as an optimization problem from the point of view of a risk-neutral self-interested firm. Starting with a parametric model of a tax system which includes audits, tax penalties, occasional amnesties, and a statute of limitations, it is shown that the firm's expected behavior arises as the optimal solution to a partially observable Markov decision process. In this process, the firm may possibly engage in tax evasion but is uncertain as to the future effects of its actions (e.g., getting caught). Using existing point-based solution algorithms, various scenarios of practical interest are then explored, including the effects of tax parameter changes on the revenue collected by the State, the effects of information "leaks" regarding upcoming tax amnesties, and the effect of extending the statute of limitations for tax audits.

The volume contains a wealth of interesting results. A number of interesting new proposals are made, and directions of current and future research can be found in selected topics in data approximation and optimization.

For a reader who is interested in evolutionary optimization for machine learning and the no free lunch theorems, the survey articles would be useful sources of recent results. These articles inform on the latest interest of the scientific community, provide an extended bibliography, and are not too mathematical to follow.

For the researcher, the whole book would be useful to providing insightful work in topics that have been drawn from approximation to discrete noisy data in order to obtain piecewise monotonicity or convexity, complexity bounds for non-linearly constrained optimization, data-dependent approximation in social computing, filter design as an extension of a Zolotarëv problem, optimal design of smart composites, peak estimation in spectroscopy, portfolio optimization by estimating model parameters, and tax evasion as an optimal control problem.

The book by content, expertise, and application areas will be useful to academics, researchers, data science practitioners, business analysts, social sciences investigators, and graduate students.

Oleg Burdakov,

Linköping University, Sweden