AFRL-AFOSR-VA-TR-2021-0049

Analysis of Nonlinear Decision-Making and Optimization Problems: Theory of Global Optimally and Linear-Time Algorithms

Sojoudi, Somayeh
REGENTS OF THE UNIVERSITY OF CALIFORNIA
2150 SHATTUCK AVE RM 313
BERKELEY, CA,
US

06/16/2021
Final Technical Report

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory
Air Force Office of Scientific Research
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
16-06-2021

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
01 Jan 2019 - 31 Dec 2020

4. TITLE AND SUBTITLE
Analysis of Nonlinear Decision-Making and Optimization Problems: Theory of Global Optimally and Linear-Time Algorithms

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-19-1-0055

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Somayeh Sojoudi

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
REGENTS OF THE UNIVERSITY OF CALIFORNIA
2150 SHATTUCK AVE RM 313
BERKELEY, CA
US

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AF Office of Scientific Research
875 N. Randolph St. Room 3112
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/AFOSR RTA2

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-AFOSR-VA-TR-2021-0049

12. DISTRIBUTION/AVAILABILITY STATEMENT
A Distribution Unlimited: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The PI's team has worked on several problems on optimization, numerical algorithms and machine learning in the past year. Some of the main results developed in this period are as follows:
Sparsity and Structure in Optimization Landscape of Non-convex Matrix Sensing:
Even under the ideal condition of no noise and zero approximation error, many highly efficient machine learning techniques involve solving potentially hard or intractable computational problems while learning from data. In practice, they are tackled by heuristic optimization algorithms, based on relaxations or greedy principals. The lack of guarantees on their performance limits their use in applications with significant cost of an error, impacting our ability to implement progressive data analysis techniques in crucial social and economic systems, such as healthcare, transportation, and energy production and distribution. Commonly, non-convexity is the main obstacle for a guaranteed learning of continuous parameters. In this work, we studied the optimization landscape of the nonconvex matrix sensing problem that is known to have many local minima in the worst case.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	UU	4	WARREN ADAMS
					19b. TELEPHONE NUMBER *(Include area code)* 00000000

Standard Form 298 (Rev.8/98)
Prescribed by ANSI Std. Z39.18

The PI's team has worked on several problems on optimization, numerical algorithms and machine learning in the past year. Some of the main results developed in this period are as follows:

Sparsity and Structure in Optimization Landscape of Non-convex Matrix Sensing:

Even under the ideal condition of no noise and zero approximation error, many highly efficient machine learning techniques involve solving potentially hard or intractable computational problems while learning from data. In practice, they are tackled by heuristic optimization algorithms, based on relaxations or greedy principals. The lack of guarantees on their performance limits their use in applications with significant cost of an error, impacting our ability to implement progressive data analysis techniques in crucial social and economic systems, such as healthcare, transportation, and energy production and distribution. Commonly, non-convexity is the main obstacle for a guaranteed learning of continuous parameters. In this work, we studied the optimization landscape of the non-convex matrix sensing problem that is known to have many local minima in the worst case. Since the existing results are related to the notion of restricted isometry property (RIP) that cannot directly capture the underlying structure of a given problem, they can hardly be applied to real-world problems where the amount of data is not exorbitantly high. To address this issue, we developed the notion of kernel structure property to obtain necessary and sufficient conditions for the inexistence of spurious local solution of any class of matrix sensing problems over a given search space. This notion precisely captures the underlying sparsity and structure of the problem, based on tools in conic optimization. We simplified the conditions for a certain class of problems to show their satisfaction and applied them to data analytics for power systems

Learning Distributed Linear-Quadratic Controllers:

Encouraged by the success of machine learning techniques applied to complex decision-making problems such as image classification, video and board games, and robotics, the use of ML for the control of autonomous systems interacting with physical environments has been an active area of research in recent years. While there is an increasing body of work studying the theoretical and practical aspects of deploying learning-enabled control policies in individual systems (e.g., self-driving cars, agile robots), there has been little work studying the use of these techniques on distributed systems, that is to say systems composed of interconnected and often spatially distributed subsystems. Examples of such distributed systems include intelligent transportation systems and cities, smart grids, and distributed sensor networks. Even when the individual components are well modeled, controlled, and understood, integrating them into a large-scale, interconnected, and heterogeneous system can make modeling and control of the full system challenging, strongly motivating the use of machine-learning-based techniques. In this work, we proposed a robust approach to design distributed controllers for unknown but-sparse linear and time-invariant systems. By leveraging modern techniques in distributed controller synthesis and structured linear inverse problems as applied to system identification, we showed that near-optimal distributed controllers can be learned with sub-linear sample complexity and computed with near-linear time complexity, both measured with respect to the dimension of the system. In particular, we provided sharp end-to-end guarantees on the stability and the performance of the designed distributed controller and proved that for sparse systems, the number of samples needed to guarantee robust and near-optimal performance of the designed controller can be significantly smaller than the dimension of the system. Finally, we showed that the proposed optimization problem can be solved to global optimality with near-linear time complexity by iteratively solving a series of small quadratic programs.

Homotopy Method for Optimal Power Flow:

The goal of optimal power flow (OPF) is to find a minimum cost production of committed generating units while satisfying technical constraints of the power system. To ensure robustness of the network, the system must be able to find new operating points within the technical limits in the event of component failures such as line and generator outages. However, finding an optimal, or even a feasible, preventive/corrective action may be difficult due to the innate nonconvexity of the problem. With the goal of finding a global solution to the post-contingency OPF problem of a stressed network, e.g. a network with a line outage, we applied a homotopy method to the problem. By parametrizing the constraint set, we defined a series of optimization problems to represent a gradual outage and iteratively solve these problems using local search. Under the condition that the global minimum of the OPF problem for the base-case is attainable, we found theoretical guarantees to ensure that the OPF problem for the contingency scenario will also converge to its global minimum. We showed that this convergence is dependent on the geometry of the homotopy path. The effectiveness of the proposed approach is demonstrated on Polish networks.

Absence of Spurious Local Trajectories in Time-Varying Optimization:

In this project, we studied the landscape of an optimization problem whose input data vary over time. This time-varying problem consists of infinitely-many individual optimization problems, whose solution is a trajectory over time rather than a single point. To understand when it is possible to find a global solution of a time-varying non-convex optimization problem, we introduced the notion of spurious (i.e., non-global) local trajectory as a generalization to the notion of spurious local solution in nonconvex (time-invariant) optimization. We developed an ordinary differential equation (ODE) which, at limit, characterizes the spurious local solutions of the time-varying optimization problem. By building upon this connection, we proved that the absence of spurious local trajectory is closely related to the transient behavior of the proposed ODE. In particular, we showed that: (1) if the problem is time-invariant, the spurious local trajectories are ubiquitous since any strict local minimum is a locally stable equilibrium point of the ODE, and (2) if the ODE is time-varying, the data variation may force all ODE trajectories initialized at arbitrary local minima at the initial time to gradually converge to the global solution trajectory. This implies that the natural data variation in the problem may automatically trigger escaping local minima over time.

Convex model to evaluate worst-case performance of local search in Optimal Power Flow:

To find an optimal solution to OPF, local search techniques such as interior point methods are typically used. However, due to the non-convex nature of the problem, these methods are likely to result in a sub-optimal solution. The goal of this project was to characterize the worst-case performance of local search on the OPF problem. To accomplish this, we formulated the OPF problem as a canonical quadratically-constrained quadratic program (QCQP). Then, we studied the problem of finding the worst-case local minimum of this QCQP, which is non-convex and hard to solve in general. We found a relaxation of this problem into a semidefinite program (SDP) and show that it is exact for certain cases. Using some test cases which are known to have multiple local minima, we demonstrated the effectiveness of the proposed relaxation to bound the worst-case local minimum. We compared the obtained upper bound on local minima to the lower bound provided by the standard SDP relaxation of the OPF problem to understand how much SDP outperforms local search for a given problem.

Tightened Convex Relaxations for Neural Network Robustness Certification:

In this project, we considered the problem of certifying the robustness of neural networks to perturbed and adversarial input data. Such certification is imperative for the application of neural networks in safety-critical decision-making and control systems. Certification techniques using convex optimization have been proposed, but they often suffer from relaxation errors that void the certificate. Our work exploits the structure of ReLU networks to improve relaxation errors through a novel partition-based certification procedure. The proposed method is proven to tighten existing linear programming relaxations, and asymptotically achieves zero relaxation error as the partition is made finer. We developed a finite partition that attains zero relaxation error and use the result to derive a tractable partitioning scheme that minimizes the worst-case relaxation error. Experiments using real data show that the partitioning procedure is able to issue robustness certificates in cases where prior methods fail. Consequently, partition-based certification procedures are found to provide an intuitive, effective, and theoretically justified method for tightening existing convex relaxation techniques.

Boundary Defense Against Cyber Threat for Power System State Estimation:

The operation of power grids is becoming increasingly data-centric. While the abundance of data could improve system efficiency, it poses major reliability challenges. In particular, state estimation aims to find the operating state of a network from the telemetered data, but an undetected attack on the data could lead to making wrong operational decisions for the system and trigger a large-scale blackout. Nevertheless, understanding the vulnerability of state estimation with regards to cyberattacks, which is a special instance of graph-structured quadratic sensing problem, has been hindered by the lack of tools for studying the topological and data-analytic aspects of networks. Algorithmic robustness is critical in extracting reliable information from abundant but untrusted grid data. For a large-scale power grid, we quantified, analyzed, and visualized the regions of the network that are not robust to cyberattacks in the sense that there exists a data manipulation strategy for each of those local regions that misleads the operator at the global scale and yields a wrong estimation of the state of the network at almost all buses. We also proposed an optimization-based graphical boundary defense mechanism to identify the border of the geographical area in which data have been manipulated. The proposed method does not allow a local attack to have a global effect on the data analysis of the entire network, which enhances the situational awareness of the grid, especially in the face of adversity. The developed mathematical framework reveals key geometric and algebraic factors that can affect algorithmic robustness and is used to study the vulnerability of the U.S. power grid in this work.