Recent Progress in Zeroth Order Optimization and Its Applications to Adversarial Robustness in Data Mining and Machine Learning

Pin-Yu Chen*
pin-yu.chen@ibm.com
IBM Research

Sijia Liu*
sijia.liu@ibm.com
IBM Research

ABSTRACT

Zeroth-order (ZO) optimization is increasingly embraced for solving big data and machine learning problems when explicit expressions of the gradients are difficult or infeasible to obtain. It achieves gradient-free optimization by approximating the full gradient via efficient gradient estimators. Some recent important applications include: a) generation of prediction-evasive, black-box adversarial attacks on deep neural networks, b) online network management with limited computation capacity, c) parameter inference of black-box/complex systems, and d) bandit optimization in which a player receives partial feedback in terms of loss function values revealed by her adversary. This tutorial aims to provide a comprehensive introduction to recent advances in ZO optimization methods in both theory and applications. On the theory side, we will cover convergence rate and iteration complexity analysis of ZO algorithms and make comparisons to their first-order counterparts. On the application side, we will highlight one appealing application of ZO optimization to studying the robustness of deep neural networks - practical and efficient adversarial attacks that generate adversarial examples from a black-box machine learning model. We will also summarize potential research directions regarding ZO optimization, big data challenges and some open-ended data mining and machine learning problems.

CCS CONCEPTS

• **Security and privacy**; • **Computing methodologies** → **Artificial intelligence**; **Machine learning approaches**;

KEYWORDS

Adversarial machine learning, adversarial robustness, gradient-free optimization, zeroth order optimization

ACM Reference Format:

Pin-Yu Chen and Sijia Liu. 2019. Recent Progress in Zeroth Order Optimization and Its Applications to Adversarial Robustness in Data Mining and Machine Learning. In *The 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'19), June 22–24, 2019, Anchorage, AK, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3292500.3332288

*Alphabetical order. Both authors contributed equally to this tutorial. Pin-Yu Chen and Sijia Liu ackowledge the support from MIT-IBM Watson AI Lab.

KDD '19, August 4–8, 2019, Anchorage, AK, USA

ACM ISBN 978-1-4503-6201-6/19/08.
https://doi.org/10.1145/3292500.3332288

1 TUTORIAL LINK

The materials and tutors' bios can be found in the folloing link: https://sites.google.com/view/adv-robustness-zoopt

2 TUTORIAL STRUCTURE

This tutorial will be divided into two parts.

2.1 First Part

First part: introduction to zeroth order (gradient-free optimization): motivation and background. The last decade has seen significant progress for the development of zeroth order (ZO) algorithms for both convex and nonconvex optimization. The existing studies suggested that ZO algorithms typically agree with the iteration complexity of first-order algorithms up to a small-degree polynomial of the problem size d.

- ZO convex optimization: Stochastic gradient descent (SGD) becomes the most widely used first-order algorithm in solving big data and machine learning tasks. Its ZO counterpart (i.e., ZO-SGD) has recently found, which yields $O(\sqrt{d}/\sqrt{T})$ convergence rate to the globally optimal solution. Here d is the number of optimization variables, and T is the number of iterations. The same rate has also been obtained by ZO mirror descent, ZO bandit convex optimization, and ZO online alternating direction method of multipliers.
- ZO nonconvex optimization: Different from convex optimization, some kind of stationary conditions are used to measure the convergence of nonconvex methods. For stochastic nonconvex optimization, ZO-SGD achieves the rate of $O(\sqrt{d}/\sqrt{T})$. Moreover, a recently proposed ZO sign-based stochastic gradient descent (ZO-signSGD) yields a graceful tradeoff between convergence accuracy and convergence speed. Furthermore, using variance reduction techniques, ZO-SGD-based algorithms can be further accelerated towards the iteration complexity bound as ZO gradient descent (ZO-GD).

2.2 Second Part

Some applications to data mining, machine learning and adversarial robustness in deep learning

- Application demands for ZO optimization grows extremely fast nowadays. On machine learning side, recent examples

have shown zeroth-order (ZO) based generation of prediction-evasive, black-box adversarial attacks on deep neural networks (DNNs) to be as effective as state-of-the-art white-box attacks, despite leveraging only the inputs and outputs of the targeted DNN. It has also been shown that ZO algorithms can be used to infer a music recommendation system based only on limited information on userâĂŹs ratings. On signal processing side, ZO algorithms have been used for network resource management to avoid involved calculation (e.g., matrix inversion) at non-expensive sensors with limited computation capacity. Also, ZO algorithm can be used for fog computing in the Internet-of-Things (IoT), where the explicit form of the loss function and constraints, are unknown to the network operator in the dynamic environment. It can also be used for hyperparameter optimziation.
- In this tutorial, we will emphasize applications of ZO algorithms to studying the robustness of deep neural networks against adversarial perturbations. In particular, we will illustrate how to formulate black-box adversarial attacks as a ZO optimization problem and how adversarial attacks can benefit from advanced ZO optimization techniques, such as providing query-efficient approaches to generating adversarial examples from ML systems with limited access, e.g., a black-box image classifier.

3 OUTLINE

The outline of the tutorial is as follows:

First Part: an introduction to zeroth order (gradient-free optimization): overview of recent advances in zeroth order optimization.

- ZO algorithms: iteration complexity versus query complexity for both convex and nonconvex optimization
  - ZO-GD, ZO-SGD, and ZO-signSGD
  - Variance reduced ZO algorithms
  - ZO operator splitting method for smooth + nonsmooth composite optimization
  - ZO adaptive momentum methods
  - ZO distributed optimization
- Applications in machine learning, data mining and signal processing
  - Recommendation system
  - Network resource management
  - Sensor network
  - Hyperparameter optimization

Second part: ZO optimization for adversarial robustness in deep learning.

- Brief introduction to adversarial machine learning and robustness
  - What is adversarial example?
  - White-box vs black-box adversarial attacks and defenses
- ZO optimization and black-box attacks to deep neural networks
  - Connecting ZO algorithm to black-box adversarial attacks
  - Score-based black-box attacks: ZOO and AutoZOOM
  - Decision-based black-box attacks
  - Universal perturbation attacks

REFERENCES

[1] S. Liu, J. Chen, P.-Y. Chen, and A. Hero, "Zeroth-Order Online Alternating Direction Method of Multipliers: Convergence Analysis and Applications," AISTATS, 2018
[2] P.-Y. Chen, H. Zhang, Y. Sharma, J. Yi, and C.-J. Hsieh, "ZOO: Zeroth Order Optimization based Black-box Attacks to Deep Neural Networks without Training Substitute Models," ACM CCS Workshop on AI-Security, 2017
[3] S. Liu, B. Kailkhura, P.-Y. Chen, P. Ting, S. Chang, and L. Amini, "Zeroth-Order Stochastic Variance Reduction for Nonconvex Optimization," NeurIPS, 2018
[4] Liu, S., Li, X., Chen, P.-Y., Vinzamuri, B., Haupt, J., andAmini, L. Zeroth-order stochastic projected gradientdescent for non-convex optimization. In GlobalSIP. IEEE, 2018.
[5] D. Hajinezhad, M. Hong, A. Garcia, "Zeroth Order Nonconvex Multi-Agent Optimization over Networks", accepted, IEEE Transactions on Automatic Control, Jan. 2018
[6] S. Liu, P.-Y. Chen, X. Chen, M. Hong, "signSGD via Zeroth-Order Oracle", Proc. International Conference on Learning Representation (ICLR) 2019
[7] M. Cheng, T. Le, P.-Y. Chen, J. Yi, H. Zhang, and C.-J. Hsieh, "Query-Efficient Hard-label Black-box Attack: An Optimization-based Approach", Proc. International Conference on Learning Representation (ICLR) 2019
[8] Tu, C.-C., Ting, P., Chen, P.-Y., Liu, S., Zhang, H.,Yi, J., Hsieh, C.-J., and Cheng, S.-M. Autozoom:Autoencoder-based zeroth order optimization method for attacking black-box neural networks. AAAI, 2019
[9] X. Chen, S. Liu, R. Sun and M. Hong, "On the convergence of a class of adam-type algorithms for non-convex optimization". Proc. International Conference on Learning Representation (ICLR) 2019
[10] Duchi, J. C., Jordan, M. I., Wainwright, M. J., and Wibisono,A. Optimal rates for zero-order convex optimization: The power of two function evaluations.IEEE Transactions on Information Theory, 61(5):2788âĂŞ2806, 2015
[11] Gao, X., Jiang, B., and Zhang, S. On the information-adaptive variants of the ADMM: an iteration complexity perspective.Optimization Online, 12, 2014
[12] Ghadimi, S. and Lan, G. Stochastic first-and zeroth-order methods for nonconvex stochastic programming.SIAM Journal on Optimization, 23(4):2341âĂŞ2368, 2013.
[13] Lian, X., Zhang, H., Hsieh, C.-J., Huang, Y., and Liu, J. A comprehensive linear speedup analysis for asynchronous stochastic parallel optimization from zeroth-order to first-order. In Advances in Neural Information Processing Systems, pp. 3054âĂŞ3062, 2016
[14] Reddi, S. J., Sra, S., Poczos, B., and Smola, A. J. Proximal stochastic methods for nonsmooth nonconvex finite-sum optimization. In Advances in Neural Information Processing Systems, pp. 1145âĂŞ1153, 2016.
[15] Bernstein, J., Wang, Y., Azizzadenesheli, K., and Anandkumar, A. SignSGD: compressed optimisation for non-convex problems, ICML, 2018