Proceedings of Machine Learning Research vol 195:1–2, 2023 36th Annual Conference on Learning Theory

The Computational Complexity of Finding Stationary Points in Non-Convex Optimization

Alexandros Hollender ALEXANDROS.HOLLENDER@EPFL.CH
École Polytechnique Fédérale de Lausanne

Manolis Zampetakis MZAMPET@BERKELEY.EDU
University of California, Berkeley

Editors: Gergely Neu and Lorenzo Rosasco

Abstract

Finding approximate stationary points, i.e., points where the gradient is approximately zero, of non-convex but smooth objective functions f over unrestricted d-dimensional domains is one of the most fundamental problems in classical non-convex optimization. Nevertheless, the computational and query complexity of this problem are still not well understood when the dimension d of the problem is independent of the approximation error. In this paper, we show the following computational and query complexity results:

1. The problem of finding approximate stationary points over unrestricted domains is PLS-complete.
2. For $d = 2$, we provide a zero-order algorithm for finding ε-approximate stationary points that requires at most $O(1/\varepsilon)$ value queries to the objective function.
3. We show that any algorithm needs at least $\Omega(1/\varepsilon)$ queries to the objective function and/or its gradient to find ε-approximate stationary points when $d = 2$. Combined with the above, this characterizes the query complexity of this problem to be $\Theta(1/\varepsilon)$.
4. For $d = 2$, we provide a zero-order algorithm for finding ε-KKT points in constrained optimization problems that requires at most $O(1/\sqrt{\varepsilon})$ value queries to the objective function. This closes the gap between the works of Bubeck and Mikulincer (2020) and Vavasis (1993) and characterizes the query complexity of this problem to be $\Theta(1/\sqrt{\varepsilon})$.
5. Combining our results with the recent result of Fearnley et al. (2021), we show that finding approximate KKT points in constrained optimization is reducible to finding approximate stationary points in unconstrained optimization but the converse is impossible.

The full version of the paper appears in arxiv with the same title.
Keywords: non-convex optimization, PLS, CLS

Acknowledgements

AH was supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number MB22.00026. MZ was supported by the Army Research Office (ARO) under contract W911NF-17-1-0304 as part of the collaboration between US DOD, UK MOD and UK Engineering and Physical Research Council (EPSRC) under the Multidisciplinary University Research Initiative (MURI).

References

Sébastien Bubeck and Dan Mikulincer. How to trap a gradient flow. In *Proceedings of the 33rd Conference on Learning Theory (COLT)*, pages 940–960, 2020. URL http://proceedings.mlr.press/v125/bubeck20b.html.

John Fearnley, Paul W. Goldberg, Alexandros Hollender, and Rahul Savani. The complexity of gradient descent: CLS = PPAD $\cap$ PLS. In *Proceedings of the 53rd ACM Symposium on Theory of Computing (STOC)*, pages 46–59, 2021. doi: 10.1145/3406325.3451052.

Stephen A. Vavasis. Black-box complexity of local minimization. *SIAM Journal on Optimization*, 3 (1):60–80, 1993. doi: 10.1137/0803004.