Proceedings of Machine Learning Research vol 291:1–2, 2025 38th Annual Conference on Learning Theory

Alternating Regret for Online Convex Optimization

Soumita Hait* HAIT@USC.EDU
University of Southern California

Ping Li* PINGLEE@STU.SUFE.EDU.CN
Shanghai University of Finance and Economics

Haipeng Luo* HAIPENGL@USC.EDU
University of Southern California

Mengxiao Zhang* MENGXIAO-ZHANG@UIOWA.EDU
University of Iowa

Editors: Nika Haghtalab and Ankur Moitra

Abstract

Motivated by alternating learning dynamics in two-player games, a recent work by Cevher et al. (2024) shows that $o(\sqrt{T})$ alternating regret is possible for any T-round adversarial Online Linear Optimization (OLO) problem, and left as an open question whether the same is true for general Online Convex Optimization (OCO). We answer this question in the affirmative by showing that the continuous Hedge algorithm achieves $\widetilde{\mathcal{O}}(d^{\frac{2}{3}}T^{\frac{1}{3}})$ alternating regret for any adversarial d-dimensional OCO problems. We show that this implies an alternating learning dynamic that finds a Nash equilibrium for any convex-concave zero-sum games or a coarse correlated equilibrium for any convex two-player general-sum games at a rate of $\widetilde{\mathcal{O}}(d^{\frac{2}{3}}/T^{\frac{2}{3}})$.

To further improve the time complexity and/or the dimension dependence, we propose another simple algorithm, Follow-the-Regularized-Leader with a regularizer whose convex conjugate is 3rd-order smooth, for OCO with smooth and self-concordant loss functions (such as linear or quadratic losses). We instantiate our algorithm with different regularizers and show that, for example, when the decision set is the ℓ_2 ball, our algorithm achieves $\widetilde{\mathcal{O}}(T^{\frac{2}{5}})$ alternating regret with no dimension dependence. In addition, for quadratic losses, we achieve an even better $\widetilde{\mathcal{O}}(T^{\frac{1}{3}})$ bound.

While our work significantly advances our understanding on alternating regret upper bounds, there are no existing alternating regret lower bounds at all. Towards closing this gap, we make an initial attempt by considering the special case of the expert problem and providing several algorithm-specific lower bounds. Specifically, we first show that worst-case alternating regret of Hedge is $\Omega(T^{\frac{1}{3}})$, thereby giving a tight characterization of Hedge's alternating regret with a fixed learning rate. Somewhat surprisingly, we also show that two optimistic algorithms — Predictive Regret Matching+ (PRM+) (Farina et al., 2021) and Optimistic Online Gradient Descent (OOGD) (Rakhlin and Sridharan, 2013) — exhibit an even worse lower bound of $\Omega(\sqrt{T})$ for alternating regret, despite the former being widely used in alternating learning dynamics.

Keywords: Online Convex Optimization, alternating regret, alternating learning dynamics

* Authors are listed in alphabetical order.

† **Extended abstract. Full version appears as arXiv:2502.12529v2.**

‡ HL thanks Yang Cai, Gabriele Farina, Julien Grand-Clément, Christian Kroer, Chung-Wei Lee, and Weiqiang Zheng for discussions related to regret matching with alternation. He is supported by NSF award IIS-1943607.

References

Volkan Cevher, Ashok Cutkosky, Ali Kavis, Georgios Piliouras, Stratis Skoulakis, and Luca Viano. Alternation makes the adversary weaker in two-player games. *Advances in Neural Information Processing Systems*, 36, 2024.

Gabriele Farina, Christian Kroer, and Tuomas Sandholm. Faster game solving via predictive blackwell approachability: Connecting regret matching and mirror descent. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 5363–5371, 2021.

Sasha Rakhlin and Karthik Sridharan. Optimization, learning, and games with predictable sequences. *Advances in Neural Information Processing Systems*, 2013.