PDE-Based Optimal Strategy for Unconstrained Online Learning

Zhiyu Zhang

Joint work with Ashok Cutkosky and Yannis Paschalidis

ICML 2022

BOSTON UNIVERSITY

Motivation

Practical ML models are naturally defined on high-dimensional unconstrained domains.

In theory, even with convex losses,

- Most practical optimizers are only well-studied on bounded domains.
- Without boundedness, they typically require carefully tuned learning rates.

Motivation

Practical ML models are naturally defined on high-dimensional unconstrained domains.

In theory, even with convex losses,

– Most practical optimizers are only well-studied on bounded domains.

– Without boundedness, they typically require carefully tuned learning rates.

How can an optimizer quickly find a far-away optimum, without hyperparameter tuning?

Unconstrained Online Linear Optimization (OLO)

In each round, an algorithm picks a prediction $x_t \in \mathbb{R}^d$, receives a loss gradient $g_t \in \mathbb{R}^d$, and suffers a loss $\langle g_t, x_t \rangle$.

In each round, an algorithm picks a prediction $x_t \in \mathbb{R}^d$, receives a loss gradient $g_t \in \mathbb{R}^d$, and suffers a loss $\langle g_t, x_t \rangle$.

Define regret as

$$\mathrm{Regret}_T(u) := \sum_{t=1}^{T} \langle g_t, x_t \rangle - \sum_{t=1}^{T} \langle g_t, u \rangle .$$

The goal is to achieve low regret for all time horizon T, loss gradients $g_{1:T}$, and comparator $u \in \mathbb{R}^d$.

Challenge: adapt to unbounded $u \in \mathbb{R}^d$.

Our result

1. A new comparator-adaptive regret bound. For any constant $C > 0$, with 1-Lipschitz losses,

$$\mathrm{Regret}_T(u) \leq C\sqrt{T} + \|u\| \sqrt{2T} \left[\sqrt{\log\left(1 + \frac{\|u\|}{\sqrt{2}C}\right)} + 2 \right].$$

For the first time in comparator-adaptive OL,

- The bound achieves the optimal $O(\sqrt{T})$ rate without an impractical doubling trick.
- The leading constant $\sqrt{2}$ is optimal.

Our result

1. A new comparator-adaptive regret bound. For any constant $C > 0$, with 1-Lipschitz losses,

$$\mathrm{Regret}_T(u) \leq C\sqrt{T} + \|u\| \sqrt{2T} \left[\sqrt{\log\left(1 + \frac{\|u\|}{\sqrt{2}C}\right)} + 2 \right].$$

   For the first time in comparator-adaptive OL,

   - The bound achieves the optimal $O(\sqrt{T})$ rate without an impractical doubling trick.
   - The leading constant $\sqrt{2}$ is optimal.

2. At its core is a new framework, which designs parameter-free optimizers by solving a PDE.
   – Similar ideas have appeared in *expert* problems [Drenska and Kohn, 2020, Harvey et al., 2020].
   – We show that PDEs applied to *adaptive* OL can largely reduce the amount of guessing.

General recipe

Step 1 A loss-regret duality [McMahan and Orabona, 2014] to remove u.

If the loss $\sum_{t=1}^{T} \langle g_t, x_t \rangle \leq -f(-\sum_{t=1}^{T} g_t)$, then $\mathrm{Regret}_T(u) \leq f^*(u)$, i.e., f conjugate.

General recipe

Step 1 A loss-regret duality [McMahan and Orabona, 2014] to remove u.

If the loss $\sum_{t=1}^{T} \langle g_t, x_t \rangle \leq -f(-\sum_{t=1}^{T} g_t)$, then $\mathrm{Regret}_T(u) \leq f^*(u)$, i.e., f conjugate.

Step 2 Consider value functions V defined by

$$V(t, S) = \min_{x \in \mathbb{R}^d} \max_{\|g\| \leq 1} \left[V(t+1, S-g) + \langle g, x \rangle \right].$$

Given any V, we can define $f(\cdot) := V(T, \cdot)$ in Step 1.

General recipe

Step 1 A loss-regret duality [McMahan and Orabona, 2014] to remove u.

If the loss $\sum_{t=1}^{T} \langle g_t, x_t \rangle \le -f(-\sum_{t=1}^{T} g_t)$, then $\mathrm{Regret}_T(u) \le f^*(u)$, i.e., f conjugate.

Step 2 Consider value functions V defined by

$$V(t, S) = \min_{x \in \mathbb{R}^d} \max_{\|g\| \le 1} \left[V(t+1, S-g) + \langle g, x \rangle \right].$$

Given any V, we can define $f(\cdot) := V(T, \cdot)$ in Step 1.

Step 3 Taking a continuous-time scaling limit,

$$\nabla_t V + \frac{1}{2} \max\{\lambda_{\max}(\nabla_{SS} V), 0\} = 0.$$

Solutions are treated as potential functions, which yields concrete algorithms.

Our improved regret bound translates to strong empirical performance.

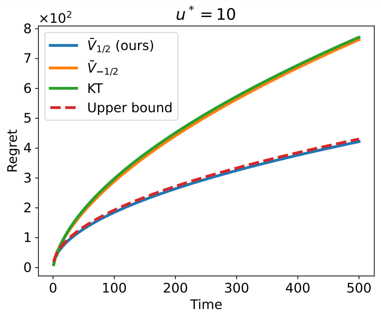

Figure 1: Regret as a function of T.

Blue: our algorithm, which guarantees

$$\mathrm{Regret}_T(u) \leq C\sqrt{T} + O\left(\|u\| \sqrt{T \log\left(C^{-1}\|u\|\right)}\right).$$

Red dashed: the regret upper bound of our algorithm.

Green and Orange: baselines that guarantee

$$\mathrm{Regret}_T(u) \leq C + O\left(\|u\| \sqrt{T \log\left(C^{-1}\|u\|\sqrt{T}\right)}\right).$$

Thank you for your attention.

Drenska, N. and Kohn, R. V. (2020).
Prediction with expert advice: A PDE perspective.
Journal of Nonlinear Science, 30(1):137–173.

Harvey, N. J., Liaw, C., Perkins, E. A., and Randhawa, S. (2020).
Optimal anytime regret for two experts.
In *IEEE 61st Annual Symposium on Foundations of Computer Science*, pages 1404–1415. IEEE.

McMahan, H. B. and Orabona, F. (2014).
Unconstrained online linear learning in hilbert spaces: Minimax algorithms and normal approximations.
In *Conference on Learning Theory*, pages 1020–1039.