On the Necessity of Adaptive Regularisation: Optimal Anytime Online Learning on ℓ_p-Balls

NeurIPS 2025

Emmeran Johnson (Imperial College London)
David Martínez-Rubio (Carlos III University of Madrid)
Ciara Pike-Burke (Imperial College London)
Patrick Rebeschini (University of Oxford)

Online Convex Optimisation (OCO)

- A sequential game over T rounds where in each round $t = 1, ..., T$:
  - A learner picks a point x_t in a convex set $V \subset \mathbb{R}^d$.
  - An adversary chooses a convex loss function $\ell_t : V \to [-1, 1]$, it is revealed to the learner who suffers a loss of $\ell_t(x_t)$.
- The goal is to minimize cumulative regret:

$$R_T = \sum_{t=1}^{T} \ell_t(x_t) - \inf_{u \in V} \sum_{t=1}^{T} \ell_t(u).$$

OCO on ℓ_p-balls

- $V = \mathcal{B}_p = \{x \in \mathbb{R}^d : \|x\|_p \leq 1\}$, $p > 2$.
- Follow The Regularised Leader (FTRL) algorithm:

$$x_{t+1} = \arg\min_{x \in V} \Big\{ \psi(x) + \eta_t \sum_{s=1}^{t} \ell_s(x) \Big\}.$$

- $\psi(x) = \frac{1}{2}\|x\|_2^2 \implies R_T = O(\sqrt{Td^{1-2/p}})$.
  - Optimal for $d \leq T$ (low-dimensional setting).
- $\psi(x) = \frac{1}{p}\|x\|_p^p \implies R_T = O(T^{1-1/p})$.
  - Optimal for $d > T$ (high-dimensional setting).

Optimal Regret with Adaptive Regularisation

- If T is unknown, how to choose ψ ?
- FTRL using $\frac{1}{p}\|x\|_p^p$ until $t_0 \approx d$ and then switching to $\frac{1}{2}\|x\|_2^2$ guarantees anytime optimality.
- Anytime optimal = optimal without knowledge of T.
- Adaptive Regularisation = switching regulariser.

Necessity of Adaptive Regularisation

- FTRL with $\frac{1}{2}\|x\|_2^2$ and any decreasing step-size η_t suffers

$$R_T = \Omega\Big(\min\big(T, \sqrt{Td^{1-2/p}}\big)\Big).$$

- FTRL with $\frac{1}{p}\|x\|_p^p$ and any decreasing step-size η_t suffers

$$R_T = \Omega\big(T^{1-1/p}\big).$$

- **Main result (separable regularisers):** consider $\psi(x) = \sum_{i=1}^d g(x_i)$, where g is a 1-dimensional regulariser. FTRL with ψ and any sequence of decreasing η_t cannot be optimal across all dimensions.

Linear Bandit Setting

- Bandit setting: learner only observes $\ell_t(x_t)$, not ℓ_t.
- High-dimensional linear bandit problem is not learnable:

Fix $p \geq 1$. For d large enough and any OCO algorithm with bandit feedback on $V = \mathcal{B}_p$, there exists a sequence of random linear losses such that $\mathbb{E}\left[\bar{R}_T\right] = \Omega(T)$.

Conclusion

Thank you!