Data-Dependent Algorithms for Bandit Convex Optimization

Mehryar Mohri[1] Scott Yang[2]

[1]Google, New York University

[2]New York University

NIPS Easy Data II, Dec 10, 2015

Learning Scenario and Set-Up

Bandit Convex Optimization

- Sequential optimization problem
- $\mathcal{K} \subset \mathbb{R}^n$ compact action space, f_t convex loss functions
- At time t, learner chooses action x_t and suffers loss $f_t(x_t)$
- Goal: minimize regret

$$\max_{x \in \mathcal{K}} \sum_{t=1}^{T} f_t(x_t) - f_t(x)$$

- Zero-th order convex optimization problem: learner has no access to gradient information!

Historical results

Summary of existing work:

1. Lipschitz [Flaxman et al 2005]: $\mathcal{O}(T^{3/4})$
2. Smooth and strongly convex loss [Levy et al 2014]: $\mathcal{O}(\sqrt{T})$
3. Smooth loss [Dekel et al 2015]: $\mathcal{O}(T^{5/8})$
4. Strongly convex loss [Agarwal et al 2010]: $\mathcal{O}(T^{2/3})$
5. etc.

Remarks:

1. Results are not data-dependent
2. Algorithms require a priori knowledge of loss function regularity

General framework for BCO Algorithms

Idea:

1. Use zero-th order information to estimate the gradient
2. Feed the gradient estimate into a normal convex optimization algorithm

Key part: estimating the gradient!

- Suppose we want to play x_t
- Instead, sample and play point y_t on ellipse E_t around x_t.
- $\nabla f_t(x_t) \approx \nabla \mathbb{E}_{y \in E_t}[\tilde{f}_t(y)] \approx \nabla f_t(y_t)$

Data-dependent sampling

Remark:

- Scaling of ellipse and learning rate both factor into the regret bound
- Historically both tuned based on worst-case data
- Algorithms do not adapt to easier data

Questions:

- Can we derive algorithms that learn faster on easier data?
- Can we characterize what easier data is for BCO problems?
- Can we construct algorithms that consolidate some of the existing regret bounds?

Data-dependent sampling

Idea:

- Scale ellipse and learning rate optimally according to the actual data that we see.

Consequences:

- Data-dependent regret bound in terms of average curvature of the ellpsoid.
- Adaptively attains smooth, strongly convex, etc. regret bounds as worst-case results.

For more details, please stop by the poster. Thank you!