Online Non-Convex Learning

Lian Yu

May 2019

1 Introduction

In class we analyzed the online learning problem under the convex setting. We learned that the average regret of the project gradient descent algorithm under this setting is bounded by $\mathcal{O}(T^{-1/2})$, where T is the number of rounds. It is $\mathcal{O}(log(T))$ if the loss function is strongly-convex. In this paper, we analyze the average regret of the follow-the-perturbed-leader algorithm under non-convex setting. We can show that in the 1-D case the average regret is $O(T^{-1/2})$. In the multi-dimensional case, the average regret is $O(T^{-1/3})$.

2 Problem Setup

Let $\mathcal{F} \subseteq \mathbb{R}^d$ be the bounded set of possible actions available to the player. Let D denote the diameter of $\mathcal{F}$ in terms of the l_∞ norm. Let $\mathcal{Z}$ denote the possible actions of the adversary. $l : \mathcal{F} \times \mathcal{Z} \to \mathbb{R}_+$ is the loss function. It is assumed that $l(\cdot, z)$ is L-Lipschitz for all $z \in \mathcal{Z}$. We define $l_t(f_t) = l(f_t, z_t)$. At each time step t, the player chooses $f_t \in \mathcal{F}$ and the adversary chooses $z_t \in \mathcal{Z}$. Then, z_t and the loss $l(f_t, z_t)$ is revealed to the player. At time T, the regret of the strategy (f_t) is $R_T((f_t)) = (\sum_{t=1}^T l_t(f_t)) - \inf\limits_{f^* \in \mathcal{F}} (\sum_{t=1}^T l_t(f^*))$.

It is assumed that the player has access to two offline oracles: the value oracle and the optimization oracle. The value oracle takes the player action f and the loss function $l(\cdot, z)$ as inputs, and outputs the loss $l(f, z)$. The optimization oracle takes a series of loss functions $(l_1, l_2, ..., l_k)$ and a vector $\sigma \in \mathbb{R}^d$. It outputs the optimal fixed strategy $f^* = argmin_{f \in \mathcal{F}} \{\sum_{i=1}^k l_i(f) - \sigma^\intercal f\}$.

3 Algorithm

Algorithm 1: The player fixes a parameter $\eta > 0$. At each time step $1 \le t \le T$, the player draws an i.i.d random variable $\sigma_t \sim (Exp(\eta))^d$, and picks the action

$$f_t = argmin_{f \in \mathcal{F}} \{\sum_{i=1}^{t-1} l_i(f) - \sigma_t^\intercal f\}$$

This is a form of the follow-the-perturbed-leader algorithm[2]. The intuition of adding the randomization is that deterministic algorithms can be predicted by the adversary, who can always choose the least favorable cost function for the player. Since f_t is determined by σ_t, we use the notation $f_t(\sigma_t)$ to indicate this relation.

4 Bounding the Regret

4.1 Reduction to the Oblivious Adversary

In this problem, the adversary is non-oblivious in the sense that it can react to the actions of the player. Since the proposed algorithm only depends on the loss of previous actions, the regret bound in this case is equivalent to the bound when the adversary is oblivious [3][4]. Thus, we can assume that the loss function $(l_t)_{t=1}^T$ is chosen in advance, and that the vector $\sigma \sim (Exp(\eta))^d$ is drawn only once.

4.2 Bound by Stability

The follow-the-regularized-leader algorithm is

$$f_t = argmin_{f \in \mathcal{F}} \{ \sum_{i=1}^{t-1} l_i(f) + R(f) \}$$

where $R(f)$ is a regularization term.
Lemma 1: For the follow-the-regularized-leader algorithm,

$$\mathbb{E}[\mathcal{R}_T] \leq \mathbb{E}[R(f^*) - R(f_1)] + \sum_{i=1}^{T} \mathbb{E}[l_t(f_t) - l_t(f_{t+1})]$$

where $f^* = argmin_{f \in \mathcal{F}} \{ \sum_{i=1}^{T} l_i(f) \}$.
Lemma 2: Under Algorithm 1, $\mathbb{E}[\|f_t(\sigma) - f_{t+1}(\sigma)\|_1] = \mathcal{O}(\frac{poly(d)\eta}{\delta} + d\delta)$. In the one-dimensional case, $\mathbb{E}[|f_t(\sigma) - f_{t+1}(\sigma)|] = \mathcal{O}(\eta)$.
proof for 1-D case:

$$\because f_t(\sigma) = argmin_{f \in \mathcal{F}} \{ \sum_{i=1}^{t-1} l_i(f) - \sigma^\intercal f \}$$

$$\therefore \sum_{i=1}^{t-1} l_i(f_t(\sigma)) - \sigma^\intercal f_t(\sigma) \leq \sum_{i=1}^{t-1} l_i(f_{t+1}(\sigma')) - \sigma^\intercal f_{t+1}(\sigma')$$

$$\sum_{i=1}^{t} l_i(f_{t+1}(\sigma')) - \sigma'^\intercal f_{t+1}(\sigma') \leq \sum_{i=1}^{t} l_i(f_t(\sigma)) - \sigma'^\intercal f_t(\sigma)$$

$$\therefore l_t(f_{t+1}(\sigma') - l_t(f_t(\sigma) \leq (\sigma' - \sigma)(f_{t+1}(\sigma') - f_t(\sigma))$$

Letting $\sigma' = \sigma + 2L$ and the using the fact that l_t is L-Lipschitz:

$$-L|f_{t+1}(\sigma') - f_t(\sigma)| \leq l_t(f_{t+1}(\sigma') - l_t(f_t(\sigma) \leq 2L(f_{t+1}(\sigma') - f_t(\sigma))$$

$$f_t(\sigma) \leq f_{t+1}(\sigma')$$

Similarly, $f_{t_1}(\sigma) \leq f_t(\sigma')$. Let $f_{min}(\sigma) = min\{f_t(\sigma), f_{t+1}(\sigma)\}$, $f_{max}(\sigma) = max\{f_t(\sigma), f_{t+1}(\sigma)\}$. We have $f_{max}(\sigma) \leq f_{min}(\sigma')$.

$$\begin{aligned}
\mathbb{E}[f_{min}(\sigma)] &= \int_{\sigma=0}^{2L} \eta e^{-\eta\sigma} f_{min}(\sigma) d\sigma + \int_{\sigma=2L}^{\infty} \eta e^{-\eta\sigma} f_{min}(\sigma) d\sigma \\
&= \int_{\sigma=0}^{2L} \eta e^{-\eta\sigma} f_{min}(\sigma) d\sigma + \int_{\sigma=0}^{\infty} \eta e^{-\eta(\sigma+2L)} f_{min}(\sigma') d\sigma \\
&\because \text{D is the diameter of } \mathcal{F}, f_{max}(\sigma) \leq f_{min}(\sigma') \\
\therefore \mathbb{E}[f_{min}(\sigma)] &\geq \int_{\sigma=0}^{2L} \eta e^{-\eta\sigma} (\mathbb{E}[f_{max}(\sigma)] - D) d\sigma + \int_{\sigma=0}^{\infty} \eta e^{-\eta(\sigma+2L)} f_{max}(\sigma) d\sigma \\
&= (1 - e^{-2L\eta})(\mathbb{E}[f_{max}](\sigma) - D) + e^{-2L\eta} \mathbb{E}[f_{max}(\sigma)] \\
&= \mathbb{E}[f_{max}(\sigma)] - (1 - e^{-2L\eta}) D \\
&\because e^x \geq 1 + x \\
&\therefore 1 - e^{-2L\eta} \leq 2L\eta \\
&\mathbb{E}[|f_t(\sigma) - f_{t+1}(\sigma)|] = \mathbb{E}[f_{max}(\sigma) - f_{min}(\sigma)] \leq 2LD\eta
\end{aligned}$$

proof for multi-dimensional case:

$$\because f_t(\sigma) = argmin_{f \in \mathcal{F}} \{\sum_{i=1}^{t-1} l_i(f) - \sigma^\intercal f\}$$

$$\therefore \sum_{i=1}^{t-1} l_i(f_t(\sigma)) - \sigma^\intercal f_t(\sigma) \leq \sum_{i=1}^{t-1} l_i(f_{t+1}(\sigma')) - \sigma^\intercal f_{t+1}(\sigma')$$

$$\sum_{i=1}^{t} l_i(f_{t+1}(\sigma')) - \sigma'^\intercal f_{t+1}(\sigma') \leq \sum_{i=1}^{t} l_i(f_t(\sigma)) - \sigma'^\intercal f_t(\sigma)$$

$$\therefore l_t(f_{t+1}(\sigma') - l_t(f_t(\sigma) \leq (\sigma' - \sigma)(f_{t+1}(\sigma') - f_t(\sigma))$$

Letting $\sigma' = \sigma + B\delta^{-1} \cdot e^k$, where e^k is the k^{th} element of the standard basis of $\mathbb{R}^d$. Assuming that the range of l_t is [0,B]:

$$-B \leq l_t(f_{t+1}(\sigma') - l_t(f_t(\sigma) \leq B\delta^{-1}(f_{t+1}(\sigma') - f_t(\sigma)) \cdot e^k$$

$$f_{t,k}(\sigma) - \delta \leq f_{t+1,k}(\sigma')$$

where $f_{t,k}$ is the k^{th} coordinate of f_t. Similarly, $f_{t,k}(\sigma) - \delta \leq f_{t+1,k}(\sigma')$. Let $f_{k,min}(\sigma) = min\{f_{t,k}(\sigma), f_{t+1,k}(\sigma)\}$, $f_{k,max}(\sigma) = max\{f_{t,k}(\sigma), f_{t+1,k}(\sigma)\}$. We have $f_{k,max}(\sigma) - \delta \leq f_{k,min}(\sigma')$. For a fixed k, we denote by $\mathbb{E}_{-k}$ the conditional mean given the noise σ except for its k^{th} coordinate σ_k.

$$\begin{aligned}
\mathbb{E}_{-k}[f_{min}(\sigma_k)] &= \int_{\sigma=0}^{B\delta^{-1}} \eta e^{-\eta\sigma_k} f_{k,min}(\sigma_k) d\sigma_k + \int_{\sigma_k=B\delta^{-1}}^{\infty} \eta e^{-\eta\sigma_k} f_{k,min}(\sigma_k) d\sigma_k \\
&= \int_{\sigma_k=0}^{B\delta^{-1}} \eta e^{-\eta\sigma_k} f_{k,min}(\sigma_k) d\sigma_k + \int_{\sigma_k=0}^{\infty} \eta e^{-\eta(\sigma_k+B\delta^{-1})} f_{k,min}(\sigma_k + B\delta^{-1}) d\sigma_k \\
&\because \text{D is the } l_\infty \text{ diameter of } \mathcal{F}, f_{k,max}(\sigma_k) - \delta \leq f_{k,min}(\sigma_k + B\delta^{-1}) \\
\therefore \mathbb{E}_{-k}[f_{k,min}(\sigma_k)] &\geq \int_{\sigma_k=0}^{B\delta^{-1}} \eta e^{-\eta\sigma_k} (\mathbb{E}_{-k}[f_{k,max}(\sigma_k)] - D) d\sigma_k \\
&\quad + \int_{\sigma_k=0}^{\infty} \eta e^{-\eta(\sigma_k+B\delta^{-1})} (f_{k,max}(\sigma_k) - \delta) d\sigma_k \\
&= (1 - e^{-B\delta^{-1}\eta})(\mathbb{E}_{-k}[f_{k,max}](\sigma_k) - D) + e^{-B\delta^{-1}\eta}(\mathbb{E}_{-k}[f_{k,max}(\sigma_k)] - \delta) \\
&= \mathbb{E}_{-k}[f_{k,max}(\sigma)] - (1 - e^{-B\delta^{-1}\eta})D - e^{-B\delta^{-1}\eta}\delta \\
&\geq \mathbb{E}_{-k}[f_{k,max}(\sigma)] - (1 - e^{-B\delta^{-1}\eta})D - \delta \\
&\because e^x \geq 1 + x \\
&\therefore 1 - e^{-B\delta^{-1}\eta} \leq B\delta^{-1}\eta \\
&\mathbb{E}_{-k}[f_{k,max}(\sigma_k) - f_{k,min}(\sigma_k)] \leq B\delta^{-1}D\eta + \delta
\end{aligned}$$

Since the above bound holds for all fixed σ excluding the k^{th} coordinate, the unconditional mean is bounded in the same way:

$$\mathbb{E}[f_{k,max}(\sigma_k) - f_{k,min}(\sigma_k)] \leq B\delta^{-1}D\eta + \delta$$

Thus,

$$\mathbb{E}[\|f_{k,max}(\sigma_k) - f_{k,min}(\sigma_k)\|_1] = \sum_{k=1}^{d} \mathbb{E}[f_{k,max}(\sigma_k) - f_{k,min}(\sigma_k)] \leq dB\delta^{-1}D\eta + d\delta$$

Theorem 1: The average regret of algorithm 1 is $\mathcal{O}(poly(d)T^{-1/3})$. In the 1-D case, the avergae regret is $\mathcal{O}(poly(d)T^{-1/2})$.

proof: From lemma 1

$$\begin{aligned}
\mathbb{E}[\mathcal{R}_T] \leq & \mathbb{E}[\sigma^\intercal (f^* - f_1)] + \sum_{i=1}^{T} \mathbb{E}[l_t(f_t) - l_t(f_{t+1})] \\
\leq & \mathbb{E}[\|\sigma\|_\infty] \mathbb{E}[\|f^* - f_1\|_1] + L \sum_{i=1}^{T} \mathbb{E}[f_t - f_{t+1}] \\
& \because \sigma \sim Exp(\eta)^d \\
& \therefore \mathbb{E}[\|\sigma\|_\infty] \leq \eta^{-1}(log(d) + 1) \\
\mathbb{E}[\mathcal{R}_T] \leq & \eta^{-1}(log(d) + 1)D + L \sum_{i=1}^{T} \mathbb{E}[f_t - f_{t+1}]
\end{aligned}$$

From lemma 2

$$\begin{aligned}
\mathbb{E}[\mathcal{R}_T] \leq & \eta^{-1}(log(d) + 1)D + \mathcal{O}(LT(\frac{poly(d)\eta}{\delta} + d\delta)) \\
= & \mathcal{O}(poly(d)(\eta^{-1}(log(d) + 1) + T(\frac{\eta}{\delta} + \delta)))
\end{aligned}$$

Setting $\eta = T^{-2/3}, \eta = T^{-1/3}$ we have:

$$\begin{aligned}
\mathbb{E}[\mathcal{R}_T] \leq & \mathcal{O}(poly(d) T^{2/3}) \\
\mathbb{E}[\mathcal{R}_T / T] \leq & \mathcal{O}(poly(d) T^{-1/3})
\end{aligned}$$

In the 1-D case:

$$\mathbb{E}[\mathcal{R}_T] \leq \eta^{-1} D + LT\mathcal{O}(\eta)$$

Setting $\eta = T^{-1/2}$ we have:

$$\begin{aligned}
\mathbb{E}[\mathcal{R}_T] \leq & \mathcal{O}(T^{1/2}) \\
\mathbb{E}[\mathcal{R}_T / T] \leq & \mathcal{O}(T^{-1/2})
\end{aligned}$$

References

[1] N. Agarwal, A. Gonen, and E. Hazan, *"Learning in Non-convex Games with an Optimization Oracle,"* arXiv:1810.07362 [cs, stat], Oct. 2018.

[2] A. Kalai and S. Vempala, *"Efficient algorithms for online decision problems,"* Journal of Computer and System Sciences, vol. 71, no. 3, pp. 291–307, Oct. 2005.

[3] N. Cesa-Bianchi and G. Lugosi, *"Prediction, Learning, and Games."* Cambridge University Press, 2006.

[4] E. Hazan and T. Koren, *"The Computational Power of Optimization in Online Learning,"* arXiv:1504.02089 [cs], Apr. 2015.