Online Non-Convex Learning: 
Following Perturbed Leader is Optimal

Praneeth Netrapalli  
Microsoft Research India

Arun Sai Suggala  
CMU
Example I: Patrolling

Every night

Where do I patrol?
Example II: Portfolio selection

Every month

Where do I invest?

Stock 1

Stock 2

Stock 3
Online learning

- Time: $1, 2, \ldots, t, \ldots, T$

- At time $t$, predict $x_t \in \mathcal{X}$

- After playing $x_t$, observe loss function $\ell_t$

- **Goal**: minimize cumulative loss $\sum_{t=1}^{T} \ell_t(x_t)$
Example I: Patrolling

\( x_t = \text{indicator vector of no patrol} = [0, 1, 1, 1, 1] \)

\( c_t = \text{indicator vector of thief} = [0, 0, 0, 0, 1] \)

\( \ell_t(x_t) = \langle c_t, x_t \rangle \)
Example II: Portfolio selection

\( x_t = \text{indicator vector of investment} \)
\( = [1,0,0] \)

\( c_t = \text{negative yield of different venues} \)
\( = -[1.1,0.9,1.05] \)

\( \ell_t(x_t) = \langle c_t, x_t \rangle \)
Example III : Solving minimax problems

• Minimax problems = Zero sum games

• Widely studied in Economics

• Several recent applications in machine learning: generative adversarial networks (GANs), multi-agent reinforcement learning (RL), adversarial training etc.

• Online learning algorithms can be used to solve such problems
Online learning

• At time $t$, predict $x_t$ and observe loss function $\ell_t$  
  • $\ell_t$ fixed ahead of time

• **Goal**: minimize cumulative loss $\sum_{t=1}^{T} \ell_t(x_t)$

• **Benchmark**: $\min_{x \in \mathcal{X}} \sum_{t=1}^{T} \ell_t(x) — best \ fixed \ policy \ in \ hindsight$

• **Regret**: $\sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} \ell_t(x)$  \hspace{1cm} Minimize regret
History

• Online *linear* learning: dates back to [Brown and von Neumann 1950]

• Online *convex* learning: Heavily studied since [Zinkevich 2003]

• Regret

\[
\sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} \ell_t(x) \leq O(\sqrt{T})
\]
Online *nonconvex* learning

- Computationally intractable even if all $\ell_t(\cdot)$ are the same

What can we do?

1. Weaker notions of regret (such as stationarity in optimization)
   - [Hazan, Singh and Zhang 2017]

2. Assume access to optimization oracles (only deal with learning)
   - [Agarwal, Gonen and Hazan 2018]
Main result

Setting

• $\ell_t(\cdot)$ is Lipschitz continuous
• $x_t \in \mathcal{X}$ with bounded diameter

Our result

• Regret:
  \[
  \sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} \ell_t(x) \leq O(\sqrt{T})
  \]

• Previous best: $O(T^{2/3})$
  [Agarwal, Gonen, Hazan 2018]
Algorithm I: Follow the leader

• For any $t \leq T$ leader $\tilde{x}_t \overset{\text{def}}{=} \arg\min_{x \in \mathcal{X}} \sum_{i=1}^{t} \ell_i(x)$

• Cannot compute $\tilde{x}_t$ — do not know $\ell_t(\cdot)$

• Choose $x_t = \tilde{x}_{t-1}$
Algorithm I: Follow the leader

- Choose \( x_t \overset{\text{def}}{=} \arg\min_{x \in X} \sum_{i=1}^{t-1} \ell_i(x) \)
- Performs poorly!
- \( X = [-1,1] \)
- \( \ell_1(x) = x \)
- \( \ell_i(x) = \begin{cases} -2x, & i \text{ is even} \\ 2x, & i \text{ is odd} \end{cases} \)

Regret = \( \Omega(T) \)
Algorithm I: Follow the **perturbed** leader

- [Hannan 1957], [Kalai, Vempala 2005]

- $\sigma \sim \text{Unif}(0, \sqrt{T})$

- $x_t \overset{\text{def}}= \arg\min_{x \in X} \sum_{i=1}^{t-1} \ell_i(x) + \langle \sigma, x \rangle$

- Regret $= O(\sqrt{T})$

$$\ell_t(x)$$

$$\sum_{i=1}^{t-1} \ell_i(x) + \langle \sigma, x \rangle = \pm x + \langle \sigma, x \rangle$$
Main intuitions

• Recall: adversary fixes choices ahead of time

• Be the leader lemma
  • Recall, $x_t \triangleq \arg\min_{x \in X} \sum_{i=1}^{t-1} \ell_i(x) + \langle \sigma, x \rangle$
  • $E[\sum_{t=1}^{T} \ell_t(x_{t+1})] - \min_{x \in X} \sum_{t=1}^{T} \ell_t(x) \leq O(\sqrt{T})$ since $\sigma \sim \sqrt{T}$

• Stability
  • $E[\sum_{t=1}^{T} \ell_t(x_t)] - E[\sum_{t=1}^{T} \ell_t(x_{t+1})] \leq T \cdot L \cdot E[\|x_t - x_{t+1}\|]$
Stability question

• Recall \( x_t \stackrel{\text{def}}{=} \arg\min_{x \in X} \sum_{i=1}^{t-1} \ell_i(x) + \langle \sigma, x \rangle \)

• How large can \( E[\|x_t - x_{t+1}\|] \) be?

• Agarwal, Gonen, Hazan 2018

\[
E[\|x_t - x_{t+1}\|] \leq O\left(T^{-\frac{1}{3}}\right)
\]

Our improvement

\( O(T^{-1/2}) \)
Linear case [Kalai and Vempala 2005]

• $\ell_i(x) = \langle c_i, x \rangle$; $\ell_i(\cdot)$ Lipschitz $\Rightarrow c_i$ bounded

• $\sigma \sim \text{Unif}(0, \sqrt{T})$

• $\sum_{i=1}^{t} \ell_i(x) + \langle \sigma, x \rangle = \langle \sigma + \sum_{i=1}^{t} c_i, x \rangle$

• Key idea:
  • $\sigma + \sum_{i=1}^{t-1} c_i \sim \sigma + \sum_{i=1}^{t} c_i$
  • $x_t \sim x_{t+1}$
The general nonconvex case

- $x_t(\sigma) \overset{\text{def}}{=} \arg\min_{x \in \mathcal{X}} \sum_{i=1}^{t-1} \ell_i(x) + \langle \sigma, x \rangle$

- **Weak monotonicity property:**
  
  $x_{t,i}(\sigma + ce_i) \leq x_{t,i}(\sigma) \ \forall \ c \geq 0$

![Diagram](attachment:image.png)

- $0 \leq \sigma_i \leq \sqrt{T}$

- $x_{t,i}(\sigma)$
Strong monotonicity property

• Suppose $\|x_t(\sigma) - x_{t+1}(\sigma)\|_1 \leq 10d \cdot |x_{t,i}(\sigma) - x_{t+1,i}(\sigma)|$

• Then for $\sigma' = \sigma + 100Lde_i$,

$$\max \left( x_{t,i}(\sigma'), x_{t+1,i}(\sigma') \right) \leq \max \left( x_{t,i}(\sigma), x_{t+1,i}(\sigma) \right) - \frac{9}{10} |x_{t,i}(\sigma) - x_{t+1,i}(\sigma)|$$
Strong monotonicity property

• Suppose $\|x_t(\sigma) - x_{t+1}(\sigma)\|_1 \leq 10d \cdot |x_{t,i}(\sigma) - x_{t+1,i}(\sigma)|$

• Then for $\sigma' = \sigma + 100Lde_i$,

$$\max \left( x_{t,i}(\sigma'), x_{t+1,i}(\sigma') \right) \leq \max \left( x_{t,i}(\sigma), x_{t+1,i}(\sigma) \right) - \frac{9}{10} |x_{t,i}(\sigma) - x_{t+1,i}(\sigma)|$$

$$E[|x_{t,i}(\sigma) - x_{t+1,i}(\sigma)|] \leq \frac{1}{10d} E[\|x_t(\sigma) - x_{t+1}(\sigma)\|_1] + O(d \|X\|_\infty)$$

$$E[\|x_t(\sigma) - x_{t+1}(\sigma)\|_1] \leq \frac{1}{10} E[\|x_t(\sigma) - x_{t+1}(\sigma)\|_1] + O(d^2 \|X\|_\infty)$$
Recap

• Follow the perturbed leader

• Be the leader lemma: playing $x_{t+1}$ at time $t$ is very good

• Stability: With perturbations, $\|x_t - x_{t+1}\|$ very small

• Key technical results: Tight monotonicity lemmas
Summary

• Online learning a powerful and interesting framework

• Most work so far in convex setting

• Nonconvex setting quite important in practice – work has just begun

• Perturbation based strategies are useful in this context as well