Learning Minimax Estimators via Online Learning

Praneeth Netrapalli
Microsoft Research India

Joint work with

Arun Sai Suggala, Kartik Gupta, Adarsh Prasad and Pradeep Ravikumar
Carnegie Mellon University
Minimax estimation

• Given $X_1, X_2, \ldots, X_n \sim \mathcal{P}_{\theta^*} \in \{\mathcal{P}_\theta : \theta \in \Theta\}$, estimate $\theta^*$

• Goal: $\min_{\theta} \mathbb{E}_{X_1, \ldots, X_n} \left[ \left\| \hat{\theta}(X_1, X_2, \ldots, X_n) - \theta^* \right\|^2 \right]$

• Really, we do not know $\theta^*$; we would like to do

$$\min_{\hat{\theta}} \max_{\theta} \mathbb{E}_{X_1, \ldots, X_n \sim \mathcal{P}_\theta} \left[ \left\| \hat{\theta}(X_1, X_2, \ldots, X_n) - \theta \right\|^2 \right]$$

• Widely studied topic, see [Berger 1985] and [Tsybakov 2008]
Outline

• Background

• **Part I**: Nonconvex online learning

• **Part II**: Minimax estimation via online learning

• **Part III**: Example – minimax estimator for Gaussian mean
Background
Convex-concave minimax optimization

\[
\min_{\hat{\theta}} \max_{\theta} \ell(\hat{\theta}, \theta)
\]

• If \( \ell(\hat{\theta}, \theta) \) is convex in \( \hat{\theta} \) and concave in \( \theta \) then [Sion 1958]

\[
\min_{\hat{\theta}} \max_{\theta} \ell(\hat{\theta}, \theta) = \max_{\theta} \min_{\hat{\theta}} \ell(\hat{\theta}, \theta)
\]

• The optimal solution is called Nash equilibrium

• Several efficient algorithms known: gradient descent ascent, extra gradient methods, fictitious play, algorithms based on online learning
Non (convex-concave)

- $\mathbb{E}_{X_1, \ldots, X_n \sim \mathcal{N}(\theta, \mathbb{I})} \left[ \| \hat{\theta}(X_1, X_2, \ldots, X_n) - \theta \|^2 \right]$ is not convex in $\theta$!

- Minimax theorem does not hold

- Instead, $\min_{\mathcal{P}(\hat{\theta})} \max_{\mathcal{Q}(\theta)} \mathbb{E}_{\mathcal{P}, \mathcal{Q}} [\ell(\hat{\theta}, \theta)] - \mathcal{P}(\hat{\theta})$ and $\mathcal{Q}(\theta)$ are probability distributions

- This is bilinear (and so convex-concave)!

- Minimax theorem holds; leads to mixed Nash equilibrium
Can we directly apply standard convex-concave minimax algorithms?

• Not all, gradients and points become infinite dimensional

• Stochastic methods also unclear

• One feasible approach via online learning

• While convex-concave involves convex online learning, this involves nonconvex online learning
Part I
Nonconvex online learning
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: $1, 2, \cdots, t, \cdots, T$

• 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

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

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

\[ \ell_t(x_t) = \langle c_t, x_t \rangle \]
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**: $\min_{x \in X} \sum_{t=1}^{T} \ell_t(x)$ — best fixed policy in hindsight

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

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

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

• Regret

\[
\sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in X} \sum_{t=1}^{T} \ell_t(x) \leq O(\sqrt{T})
\]
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
• $x_t \in X$ with bounded diameter

Our result

• Regret:
  $$\sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in X} \sum_{t=1}^{T} \ell_t(x) \leq O(\sqrt{T})$$
• Previous best: $O(T^{2/3})$
  [Agarwal, Gonen, Hazan 2018]
Algorithm I: Follow the leader

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

• 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 = $O(\sqrt{T})$

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

• Recall: adversary fixes choices ahead of time

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

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

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

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

• Agarwal, Gonen, Hazan 2018

  $E[\|x_t - x_{t+1}\|] \leq O(T^{-1/3})$
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:** \( x_{t,i}(\sigma + ce_i) \leq x_{t,i}(\sigma) \ \forall \ c \geq 0 \)
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 \( \sigma' = \sigma + 100Lde_i \),

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

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

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

**Upshot**

Can do nonconvex online learning with access to optimization oracles
Part II
Minimax estimation via online learning
Regret minimization vs best response

| $\hat{\theta}$ player (min) | $\theta$ player (max) |
|-----------------------------|------------------------|
| Regret minimizationalgorithm (FTPL) | Best response |
| $\mathcal{P}_0(\hat{\theta})$ | $\theta_0 = \arg\max_\theta \mathbb{E}_{\mathcal{P}_0}[\ell(\hat{\theta}, \theta)]$ |
| $\mathcal{P}_1(\hat{\theta}) = \text{FTPL}(\theta_0)$ | $\theta_1 = \arg\max_\theta \mathbb{E}_{\mathcal{P}_1}[\ell(\hat{\theta}, \theta)]$ |
| $\mathcal{P}_2(\hat{\theta}) = \text{FTPL}(\theta_0, \theta_1)$ | $\theta_2 = \arg\max_\theta \mathbb{E}_{\mathcal{P}_2}[\ell(\hat{\theta}, \theta)]$ |
Main idea

• The final output is $\frac{1}{T}\sum t \mathcal{P}_t(\hat{\theta})$

\[
\max_{\theta} \frac{1}{T}\sum_t \mathbb{E}_{\mathcal{P}_t(\bar{\theta})}[\ell(\hat{\theta}, \theta)] \\
\leq \frac{1}{T}\sum_t \mathbb{E}_{\mathcal{P}_t(\bar{\theta})}[\ell(\hat{\theta}, \theta_t)] \quad \text{(best response of } \theta) \\
\leq \min_{\mathcal{P}(\bar{\theta})} \frac{1}{T}\sum_t \mathbb{E}_{\mathcal{P}(\bar{\theta})}[\ell(\hat{\theta}, \theta_t)] + O\left(\frac{1}{\sqrt{T}}\right) \quad \text{(regret guarantee)} \\
\leq \min_{\mathcal{P}(\bar{\theta})} \max_{\theta} \mathbb{E}_{\mathcal{P}(\bar{\theta})}[\ell(\hat{\theta}, \theta)] + O\left(\frac{1}{\sqrt{T}}\right)
\]
Historical background

• Minimax estimation via online learning known from previous work [Freund and Schapire 1996]. Main new development – nonconvex online learning using nonconvex optimization oracles.

• Main challenge: Solve the associated nonconvex problems

• Contrast with other related works: guess one side of the mixed strategy [Berger 1985, Clarke and Barron 1994]
  • Results exist for very special cases only. Not clear how to extend.
Part III

Example – minimax estimator for Gaussian mean
Estimating Gaussian mean

• Given $X_1, X_2, \ldots, X_n \sim \mathcal{N} (\theta, I)$, $\theta \in \mathbb{R}^d$, $\|\theta\|_2 \leq B$, estimate $\theta$

• Goal: 
\[
\min_{\hat{\theta}} \max_{\|\theta\|_2 \leq B} \mathbb{E}_{X_1, \ldots, X_n} \left[ \|\hat{\theta} (X_1, X_2, \ldots, X_n) - \theta\|^2 \right]
\]

• For simplicity: 
\[
R(\hat{\theta}, \theta) \overset{\text{def}}{=} \mathbb{E}_{X_1, \ldots, X_n} \left[ \|\hat{\theta} (X_1, X_2, \ldots, X_n) - \theta\|^2 \right]
\]

• Several works for the case $n = 1$ but minimax estimator not known for $B \geq 1.16 \sqrt{d}$. [Bickel et al. 1981, Berry 1990, Marchand and Perron 2002]

• Our work resolves this.
Key steps

1. **Symmetry** [Berry 1990]:

\[
\min_{\hat{\theta}} \max_{\|\theta\|_2 \leq B} R(\hat{\theta}, \theta) \equiv \min_{\hat{\theta}} \max_{b \in [0,B]} \mathbb{E}_{\theta \sim \mathcal{P}_b}[R(\hat{\theta}, \theta)] \quad [\text{Berry 1990}]
\]

2. **FTPL:**

\[
b_t(\sigma) \leftarrow \arg\max_{b \in [0,B]} \sum_i \mathbb{E}_{\theta \sim \mathcal{P}_b}[R(\hat{\theta}_t, \theta)] + \sigma b
\]

Nonconvex but 1-d problem

\[
\hat{\theta}_t \leftarrow \min_{\hat{\theta}} \mathbb{E}_{b \sim \mathcal{P}_t} \left[ \mathbb{E}_{\theta \sim \mathcal{P}_b}[R(\hat{\theta}, \theta)] \right]
\]

Bayesian estimator for symmetric prior
Conclusion

• Minimax estimation a fundamental problem in statistics

• Most results obtained through problem specific approaches

• Our work:
  • General approach through nonconvex online learning
  • Efficient algorithm for nonconvex online learning based on certain optimization oracles
  • Efficiently implementing this approach for Gaussian mean estimation and some other related problems