Online Learning via the Differential Privacy Lens

Jacob Abernethy @ Georgia Institute of Technology
Young Hun Jung @ University of Michigan
Chansoo Lee @ Google Brain
Audra McMillan @ Boston University
Ambuj Tewari @ University of Michigan

NeurIPS 2019
Online Learning via the Differential Privacy Lens

- DP inspired stability
- Unifying analysis framework for existing online algos
- New algos with the first-order regret bounds

DP inspired stability is well-suited to analyzing OL algorithms
Adversarial Online Learning Problems

- A sequential game between Learner and Adversary
- Learner chooses its action $x_t \in X$, which can be random
- Adversary chooses a loss function $\ell_t \in Y$ (NOT random)
- Full Info.: the entire function $\ell_t$ is revealed to the learner
- Partial Info.: only the function value $\ell_t(y_t)$ is revealed
Adversarial Online Learning Problems

- The learner’s goal is to minimize the expected regret:
  \[
  \mathbb{E}[\text{Regret}_T] = \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(x_t)\right] - L^*_T, \quad \text{where } L^*_T = \min_{x \in X} \sum_{t=1}^{T} \ell_t(x).
  \]

- Zero-order bound proves \( \mathbb{E}[\text{Regret}_T] = o(T) \)
- First-order bound proves \( \mathbb{E}[\text{Regret}_T] = o(L^*_T) \)
  - The first-order bound is more desirable if \( L^*_T = o(T) \)
- OCO, OLO, expert problems, MABs, bandits with experts
Differential Privacy

Let $\mathcal{A}$ be a randomized algorithm that maps a data set $S$ to a decision rule in $\mathcal{X}$

- $\mathcal{A}(S)$ will be available to users but NOT $S$ itself
- We do NOT want the users to infer our data set $S$ from $\mathcal{A}(S)$
- Suppose $S$ and $S'$ differ only by a single entry
  $\Rightarrow$ We want $\mathcal{A}(S)$ and $\mathcal{A}(S')$ to be similar
Differential Privacy

- The $\delta$-approximate max-divergence between two distributions $P$ and $Q$ is (sup takes over all measurable sets)

$$D_\infty^\delta(P, Q) = \sup_{P(B) > \delta} \log \frac{P(B) - \delta}{Q(B)}$$

- We say $\mathcal{A}$ is $(\epsilon, \delta)$-DP if $D_\infty^\delta(\mathcal{A}(S), \mathcal{A}(S')) < \epsilon$
Main Observation
In online learning, Follow-The-Leader algorithm performs badly while F-T-Perturbed-L or F-T-Regularized-L do well.

Definition 1 (One-step differential stability)
For a divergence $D$, $A$ is called DiffStable($D$) at level $\epsilon$ iff for any $t$ and any $\ell_{1:t} \in \mathcal{Y}^t$, we have $D(A(\ell_{1:t-1}), A(\ell_{1:t})) \leq \epsilon$

Definition 2 (DiffStable, when losses are vectors)
For a norm $\| \cdot \|$, $A$ is called DiffStable($D,\| \cdot \|$) at level $\epsilon$ iff for any $t$ and any $\ell_{1:t} \in \mathcal{Y}^t$, we have $D(A(\ell_{1:t-1}), A(\ell_{1:t})) \leq \epsilon \| \ell_t \|$

Remark. $\ell_{1:t-1}$ and $\ell_{1:t}$ only differ by one item!
Suppose loss functions always belong to \([0, B]\) for some \(B\) and \(\mathcal{A}\) is DiffStable\((\mathcal{D}_\infty^\delta)\) at level \(\epsilon \leq 1\). Then the regret of \(\mathcal{A}\) satisfies

\[
\mathbb{E}[\text{Regret}(\mathcal{A})_T] \leq 2\epsilon L_T^* + 3\mathbb{E}[\text{Regret}(\mathcal{A}^+)_T] + \delta BT.
\]

- We can adopt DiffStable algorithms from DP community
- \(\mathbb{E}[\text{Regret}(\mathcal{A}^+)_T]\) is usually small (independent of \(T\))
- \(\delta\) can be set to be as small as \(1/\!BT\)
Online Convex Optimization

Algorithm 1 Online convex optimization using Obj-Pert

1: **Given** Obj-Pert solves the convex optimization while preserving DP
2: **for** $t = 1, \cdots, T$ **do**
3: Play $x_t = \text{Obj-Pert}(\ell_{1:t-1}; \epsilon, \delta, \beta, \gamma)$
4: **end for**

- Algorithm 1 is automatically DiffStable due to Obj-Pert (object perturbation) algorithm from DP literature
- When applying the Key Lemma, $\mathbb{E}[\text{Regret}(A^+)_T]$ scales as $\frac{1}{\epsilon}$
  
  $$
  \mathbb{E}[\text{Regret}(A)_T] \leq 2\epsilon L_T^* + 3\mathbb{E}[\text{Regret}(A^+)_T] + \delta BT
  $$

- Tuning $\epsilon$ and setting $\delta = 1/BT$, we get the first-order regret bound of $O(\sqrt{L_T^*})$
Other Applications

- OLO/OCO, Expert Learning, MABs, Bandits with Experts
- Zero-order and First-order regret bounds
- Provide a unifying framework to analyze OL algorithms
- Come to Poster #TBD for more details

Thanks!