Output Perturbation for Differentially Private Convex Optimization

With Improved Population Loss and Runtime Bounds, and Applications to Adversarial Training

Andrew Lowy and Meisam Razaviyayn (USC)

Background

- Differential Privacy (DP) provides a rigorous guarantee that one's data cannot be leaked. Formally, a randomized algorithm $\mathscr{A} : \mathscr{X}^n \to \mathbb{R}^d$ is said to be (ϵ, δ)-DP if $\mathbb{P}(\mathscr{A}(X) \in S) \leq \mathbb{P}(\mathscr{A}(X') \in S)e^{\epsilon} + \delta$ for all $S \subset Range(\mathscr{A})$ and all adjacent data sets $X, X' \in \mathscr{X}^n$ that differ by one observation.
  - $\delta = 0$ provides strongest privacy guarantee
- Output perturbation: $\mathscr{A}(X) = w^*(X) + z$, where $w^*(X) \in \arg\min_{w \in \mathbb{R}^d} F(w, X)$
  - Applied as conceptual method for smooth classification in ERM (Chaudhuri et al., 2011)
  - Implemented with gradient descent for smooth ERM (Zhang et al., 2017)

Related Work and Gaps

1. Most work in DP learning has focused on average loss: Empirical Risk Minimization (ERM) (e.g. Bassily et al. 2014) or Stochastic Convex Optimization (SCO)

- It is important to have DP algorithms for ***general* loss** functions that are *not of ERM/SCO form*
  - Applications: fairness, robustness, sensitivity to outliers, bias/variance tradeoff
  - Example: $F(w, X) = \frac{1}{\tau} \log\left(\frac{1}{n}\sum_{i=1}^{n} e^{\tau f(w, x_i)}\right)$ (TERM) (Li et al. 2020)

2. Another gap: tight population loss and runtime bounds for $(\epsilon, 0)$**-DP SCO**
   - Bassily et al. (2019) and Feldman et al. (2020) give tight loss bounds and fast runtimes for (ϵ, δ)-DP SCO for $\delta > 0$

3. Need for practical algorithms for **DP Adversarial Training**

Main Contributions

1. We provide a simple algorithm for **general DP convex optimization**
   - Excess risk and runtime bounds for general convex Lipschitz loss (and under smoothness and/or strong convexity)
2. We prove the tightest known excess population loss and runtime bounds for $(\epsilon, 0)$**-DP SCO**
3. We develop a practical algorithm for **DP Adversarial Training** with excess adversarial risk and runtime guarantees
4. We attain our results through **output perturbation**
   - We detail how output perturbation can be seen as a powerful practical method for transforming *any* non-private solver into a DP one

Results: General DP Optimization

1. Solving $\min_{w \in \mathbb{R}^d} F(w, X)$ for **general (non-ERM) loss**

Function Class	Excess Risk	Runtime	Assumptions
L-Lipschitz, μ-strongly convex	$O\left(\frac{L^2}{\mu}\frac{d}{\epsilon}\right)$	$O(n\epsilon)$	$d \lesssim \epsilon$
L-Lipschitz, μ-strongly convex, β-smooth	$O\left(\frac{L^2}{\mu}\left(\frac{d}{\epsilon}\right)\min\left\{\kappa\left(\frac{d}{\epsilon}\right), 1\right\}\right)$	$\widetilde{O}(dn\sqrt{\kappa})$	$d \lesssim \epsilon$
L-Lipschitz, convex	$O\left(LR\sqrt{\frac{d}{\epsilon}}\right)$	$O\left(\frac{n\epsilon^2}{d}\right)$	$d \lesssim \epsilon$
L-Lipschitz, convex, β-smooth	$O\left(\min\left\{\beta^{1/3}L^{2/3}R^{4/3}\left(\frac{d}{\epsilon}\right)^{2/3}, LR\sqrt{\frac{d}{\epsilon}}\right\}\right)$	$\widetilde{O}\left(n\max\left\{\left(\frac{\beta R}{L}\right)^{1/3}\epsilon^{1/3}d^{2/3}, \left(\frac{\beta R}{L}\right)^{1/2}\epsilon^{1/4}d^{3/4}\right\}\right)$	$\left(\frac{d}{\epsilon}\right)^2 \lesssim \frac{L}{\beta R}$
TERM	$O\left(\frac{L^2 C_\tau}{\mu}\frac{d}{\epsilon n}\right)$	$O\left(\frac{n^2 d}{C_\tau}\max\left\{\frac{d}{\epsilon}, \frac{n}{C_\tau}\right\}\right)$	f bounded on $\mathcal{W} \times \mathcal{X}$, $\frac{d}{\epsilon n} \leq 1$
Smooth TERM	$O\left(\frac{L^2 C_\tau}{\mu}\frac{d}{\epsilon n}\min\left\{1, \kappa_\tau\frac{d}{\epsilon n}\right\}\right)$	$\widetilde{O}(nd\kappa_\tau)$	f bounded on $\mathcal{W} \times \mathcal{X}$, $\frac{d}{\epsilon n} \leq 1$

Table 1: General loss (non-ERM, non-SCO), $\delta = 0$. For $\delta > 0$, d gets replaced by $\sqrt{d}\log\left(\frac{1}{\delta}\right)$ in excess risk bounds.

- Excess risk := $\mathbb{E}_{\mathscr{A}} F(\mathscr{A}(X), X) - F(w^*(X), X)$

Results: DP SCO $(\delta = 0)$

2. Solving $\min_{w \in \mathbb{R}^d} F(w, X) = \mathbb{E}_{x \sim \mathscr{D}}\left[f(w, x)\right]$

Function Class	Excess Population Loss	Runtime	Assumptions
L-Lipschitz, μ-strongly convex	$O\left(\frac{L^2}{\mu}\left(\frac{d}{\epsilon n} + \frac{1}{n}\right)\right)$	$O\left(nd\max\left\{n, \frac{\epsilon}{d}\right\}\right)$	$\frac{d}{\epsilon n} \lesssim 1$
L-Lipschitz, μ-strongly convex, β-smooth	$O\left(\frac{L^2}{\mu}\left(\min\left\{\kappa\left(\frac{d}{\epsilon n}\right)^2, \frac{d}{\epsilon n}\right\} + \frac{1}{n}\right)\right)$	$\widetilde{O}\left(d(n + \sqrt{n\kappa})\right)$	$\frac{d}{\epsilon n} \lesssim 1$
L-Lipschitz, convex	$O\left(LR\left(\sqrt{\frac{d}{\epsilon n}} + \frac{1}{\sqrt{n}}\right)\right)$	$O\left(n^2 d\max\left\{1, (\frac{\epsilon}{d})^2\right\}\right)$	$\frac{d}{\epsilon n} \lesssim 1$
L-Lipschitz, convex, β-smooth	$O\left(\min\left\{\beta^{1/3}L^{2/3}R^{4/3}\left(\left(\frac{d}{\epsilon n}\right)^{2/3} + \frac{1}{\sqrt{n}}\right), LR(\sqrt{\frac{d}{\epsilon n}} + \frac{1}{\sqrt{n}})\right\}\right)$	$\widetilde{O}\left(nd + \max\left\{n^{5/6}d^{2/3}\epsilon^{1/3}\left(\frac{\beta R}{L}\right)^{1/3}, n^{3/4}d^{3/4}\epsilon^{1/4}\sqrt{\frac{\beta R}{L}}\right\}\right)$	$\left(\frac{d}{\epsilon n}\right)^2 \lesssim \frac{L}{\beta R}$

Table 2: Excess population Loss, $\delta = 0$. All excess risk bounds should be read as $\min\{LR, ...\}$ by taking the trivial algorithm (e.g. if $d > \epsilon n$).

- Excess population loss := $\mathbb{E}_{\mathscr{A}; x \sim \mathscr{D}} f(\mathscr{A}(X), x) - \min_{w \in \mathbb{R}^d} \mathbb{E}_{x \sim \mathscr{D}} f(w, x)$

DP Adversarial Training

3. Solving $\min_{w \in \mathbb{R}^d} \max_{v \in S^n} F(w, X + v)$: adversary perturbs feature components of data set X, while corresponding labels are fixed (not perturbed)

- Applying our framework results in the first excess adversarial risk and runtime bounds for DP adversarial training
- Excess adversarial risk := $\mathbb{E}_{\mathscr{A}} \max_{v \in S^n} F(\mathscr{A}(X), X + v) - \min_{w \in \mathbb{R}^d} \max_{v \in S^n} F(w, X + v)$

References

1. Bassily et al. (2014). Differentially Private Empirical Risk Minimization: Efficient Algorithms and Tight Error Bounds. https://arxiv.org/pdf/1405.7085.pdf
2. Bassily et al. (2019). Private Stochastic Convex Optimization with Optimal Rates. https://arxiv.org/abs/1908.09970
3. Chaudhuri et al. (2011). Differentially Private Empirical Risk Minimization. https://www.jmlr.org/papers/volume12/chaudhuri11a/chaudhuri11a.pdf
4. Feldman et al. (2020). Private Stochastic Convex Optimization: Optimal Rates in Linear Time. http://vtaly.net/papers/FKT_fast_DPSCO.pdf
5. Li et al. (2020). Tilted Empirical Risk Minimization. https://arxiv.org/abs/2007.01162
6. Zhang et al. (2017). Efficient Private ERM for Smooth Objectives. https://arxiv.org/abs/1703.09947

lowya@usc.edu