CS291A Introduction to Differential Privacy **Fall 2021**

Lecture 4: Private Multiplicative Weights (Oct. 6)

Lecturer: Yu-Xiang Wang *Scribes: Xi Gong*

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

4.1 Private multiplicative weights

In the previous lecture, we have seen that the Laplace mechanism yields normalized error of $\frac{|Q|\log(|Q|/\delta)}{n\epsilon}$ for query release, and normalized error of $\frac{\sqrt{|X|\log(\frac{1}{\delta})}}{n\epsilon}$ for data release. Yet, this leaves the question of whether it is possible for the error to depend polylogarithmically on both $|Q|$ and $|X|$. This question is answered positively by the private multiplicative weights algorithm which we will introduce in this lecture.

We first state a non-private version of the algorithm.

Algorithm: Online query release without privacy

- True data $p = \frac{x}{n}$, initialize synthetic dataset $\tilde{p}_i = \frac{1}{|X|}$, initialize accuracy parameter α.
- Adversary selects an online sequence of queries
- If $|q^\top \tilde{p}_t - q^\top p| \geq \alpha$:
  1. Privately release $y = q^\top p$.
  2. Set the loss vector to be $\ell_t := \text{sign}(q^\top \tilde{p}_t - q^\top p) \times q$.
  3. Update $\tilde{p}_{t+1} = \text{Normalize}(\tilde{p}_t e^{-\eta \ell_t})$.
  4. Increment t, i.e., $t = t + 1$. Break if $t > N$.
- Else: release $q^\top \tilde{p}$.

While this algorithm is clearly non-private, its utility analysis will provide useful insight when we extend it to the private version. Before analysing its utility, we need to take a detour and state a classical result on the multiplicative weights algorithm. From this result, one can conclude that even when facing fully adaptive queries that aim to maximize the number of rounds for $|q^\top \tilde{p}_t - q^\top p| \geq \alpha$, the number of times we are above threshold grows only logarithmically in $|X|$ (which is good!). Later we will see that this translates to a small number of times we will need to spend from our privacy budget.

Algorithm: Multiplicative weights

1: Initialize: $\forall i \in [N], W_1(i) = 1$
2: for $t = 1$ to T do
3: Pick $i_t \sim_R W_t$, i.e., $i_t = i$ with probability $\mathbf{x}_t(i) = \frac{W_t(i)}{\sum_j W_t(j)}$
4: Incur loss $\ell_t(i_t)$
5: Update weights $W_{t+1}(i) = W_t(i) e^{-\varepsilon \ell_t(i)}$
6: end for

An important theorem on the Multiplicative Weights algorithm is the following:

Theorem 4.1. *Let ℓ_t^2 denote the N-dimensional vector of square losses, i.e., $\ell_t^2(i) = \ell_t(i)^2$, let $\varepsilon = \sqrt{\frac{\log(N)}{T}}$, and assume $\ell_t(i) \in [-1, 1]$. The Multiplicative Weights algorithm satisfies for any expert $i^\star \in [N]$:*

$$\sum_{t=1}^{T} \mathbf{x}_t^\top \ell_t \leq \sum_{t=1}^{T} \ell_t\left(i^\star\right) + 2\sqrt{T \log N}.$$

Instead of competing against a fixed best expert, we can easily generalize this theorem to be against arbitrary distribution over the experts. This results in the following corollary.

Corollary 4.2. *Under the same setting as the previous theorem, the Multiplicative weights algorithm satisfies for arbitrary $p \in \Delta_N$:*

$$\sum_{t=1}^{T} \mathbf{x}_t^\top \ell_t \leq \sum_{t=1}^{T} p^\top \ell_t + 2\sqrt{T \log N}.$$

Since $\ell_t := \text{sign}(q^\top \tilde{p}_t - q^\top p) \in [-1, 1]^{|X|}$ and $\mathbf{x}_t := \tilde{p}_t$ can be considered as elements of Δ_X, we can apply this corollary to the non-private MW algorithm by direct substitution. We get

$$\sum_{t=1}^{T} (x_t - p)^\top \ell_t := \sum_{t=1}^{T} (\tilde{p}_t - p)^\top q_t \times \text{sign}(q_t^\top \tilde{p}_t - q_t^\top p) \tag{4.1}$$

$$= \sum_{t=1}^{T} |q_t^\top \tilde{p}_t - q_t^\top p| \tag{4.2}$$

$$\leq 2\sqrt{T \log |X|}. \tag{4.3}$$

On the other hand, we enter the if statement (i.e., the error is above threshold) iff $|q^\top \tilde{p}_t - q^\top p| \geq \alpha$, which implies

$$T\alpha \leq \sum_{t=1}^{T} |q_t^\top \tilde{p}_t - q_t^\top p|.$$

Combine the bounds, we get $T\alpha \leq \sum_{t=1}^{T} |q_t^\top \tilde{p}_t - q_t^\top p_t| \leq 2\sqrt{T \log |X|}$. Thus $T \leq \frac{log|X|}{\alpha^2}$.

For this algorithm, privacy is left unprotected when the algorithm decides which queries answers are above the threshold error, and when it updates the synthetic dataset using information of $y := q^\top \tilde{p}$. Once these two non-private procedures are identified, we are now ready to upgrade it to a private version by applying Above threshold to the former, and Laplace mechanism to the latter. Both of which are the building blocks of privacy which we have seen in previous lectures.

Algorithm: Private Multiplicative Weights

- True data $p = \frac{x}{n}$, initialize synthetic dataset $\tilde{p}_i = \frac{1}{|X|}$, set $\tilde{\alpha} = \alpha + \text{Lap}(\frac{2}{n\epsilon_0})$.
- Adversary selects an online sequence of queries
- If $|q^\top \tilde{p}_t - q^\top p| + \text{Lap}(\frac{4}{n\epsilon_0}) \geq \tilde{\alpha}$:
  1. Privately release $y = q^\top p + \text{Lap}(\frac{1}{n\epsilon_0})$
  2. Set the loss vector to be $\ell_t := \text{sign}(q^\top \tilde{p}_t - y) \times q$.
  3. Update $\tilde{p}_{t+1} = \text{Normalize}(\tilde{p}_t e^{-\eta \ell_t})$.
  4. Increment t, i.e., $t = t + 1$. Break if $t > N$.
  5. Refresh threshold noise: $\hat{\alpha} = \alpha + \text{Lap}(\frac{2}{n\epsilon_0})$
- Else: release $q^\top \tilde{p}_t$.

The key idea of Private MW is analogous to the sparse vector technique. In SVT, we achieve a better dependency on the number of queries by answering only a small number of "interesting" queries that are above a predefined threshold. In Private MW, we will achieve the same by answering only the "interesting" queries that the synthetic data fails to answer accurately, while answering the rest with no privacy cost (using synthetic data). Since the multiplicative weight algorithm ensures that our synthetic dataset is able to learn the true dataset quickly and privately, we know there is only a small number of "interesting" queries that degrade the privacy.

The choice of unspecified parameters N and ϵ_0 will be derived naturally by going through the privacy and accuracy analysis. For utility analysis, we "de-randomize" the algorithm by hiding the undesirable events under the small δ probability, and perform analysis on events that happen with high probability in a deterministic fashion.

Given queries of size k, the number of Laplace random variables invoked by *Private MW* is upper bounded by $2N + k$, due to the applications of Above threshold and Laplace mechanism. Let $\{Z_i\}$ be the set of all Laplace random variables released throughout the run of the algorithm. Since we only care about upper bounding $|Z_i|$ with large probability, we treat each Z_i as a $\text{Lap}(\frac{4}{n\epsilon_0})$ random variable, which has the largest Laplace parameter among the entire algorithm. By union bound,

$$\mathbb{P}(\forall i, |Z_i| \leq \frac{4}{n\epsilon_0} \log(\frac{2N+k}{\delta})) \geq 1 - \delta.$$

This allows us to treat $|Z_i| \leq \frac{4}{n\epsilon_0} \log(\frac{3k}{\delta})$ as a deterministic event in the rest of the analysis. All statements below simultaneously hold with probability at least $1 - \delta$.

Claim 1. All selected answers (i.e., $|\tilde{p}^\top q_t - p^\top q_t| + Z_i > \tilde{\alpha}$) are accurate.

If the answer is above threshold, the algorithm returns the perturbed true answer $p^\top q_t + Z_k$. We get

$$|(p^\top q_t + Z_k) - p^\top q_t| = |Z_k| < \frac{4}{n\epsilon_0} \log(\frac{3k}{\delta}).$$

Claim 2. All answers that are not selected are accurate.

If the answer is below threshold, then $|\tilde{p}^\top q_t - p^\top q_t| + Z_i \leq \alpha + Z_j$ and we get

$$|\tilde{p}^\top q_t - p^\top q_t| \leq \alpha + |Z_i| + |Z_j| < \alpha + \frac{8}{n\epsilon_0}\log(\frac{3k}{\delta}).$$

Claim 3. From the regret bound of MW, the number of iterations is small.

Recall that in the non-private version, we have the bound $T\alpha \leq \sum_{t=1}^{T}(\tilde{p}_t - p)\cdot q \ \text{sign}(\tilde{p}_t^\top q_t - p^\top q_t) \leq 2\sqrt{T\log|X|}$. We want the same bound to hold in the privatized version of this algorithm (to arrive at the conclusion of $T \in O(\frac{\log|X|}{\alpha^2})$), meaning that we want a bound of the form

$$\Theta(T\alpha) \leq \sum_{t=1}^{T}(\tilde{p}_t - p)\top \cdot q_t \ \text{sign}(\tilde{p}_t^\top q_t - (p^\top q_t + Z_i)) \leq 2\sqrt{T\log|X|}.$$

This bound will be readily available if the sign of $\tilde{p}^\top q_t - p^\top q_t$ does not change under the perturbation by the Laplace random variable Z_i. A sufficient condition for invariance of sign is $|\tilde{p}q_t - p^\top q_t| > |Z_k|$, and it remains to choose an appropriate parameter for Z_i for this inequality to hold. For an answer to be selected, we must have $|\tilde{p}q - p^\top q_t| + Z_i > \alpha + Z_j$, meaning that $|\tilde{p}^\top q_t - p^\top q_t| > \alpha + Z_j - Z_i$. To ensure the sign is invariant, we also require $|\tilde{p}^\top q_t - p^\top q_t| > \alpha + Z_j - Z_i \overset{set}{>} |Z_k|$. This inequality holds if we pick ϵ_0 so that $|Z_k| < \frac{4}{n\epsilon_0}\log(\frac{3k}{\delta}) \overset{set}{=} \frac{\alpha}{4}$, so we set $\epsilon_0 := \frac{16\log(\frac{3k}{\delta})}{n\alpha}$.

With this choice of ϵ_0, it is guaranteed that $|\tilde{p}^\top q_t - p^\top q_t| > \frac{\alpha}{2}$, so we adjust the lower bound and obtain

$$T(\frac{\alpha}{2}) \leq \sum_{t=1}^{T}(\tilde{p} - p)\cdot q_t \ \text{sign}(\tilde{p}^\top q_t - (p^\top q_t + Z_i)) \leq 2\sqrt{T\log|X|},$$

meaning that $T \leq \frac{16\log|X|}{\alpha^2}$. This completes the proof of *claim* 3.

Claim 1. and *claim 2.* together shows that the errors of query answers are uniformly small. In particular, one can easily check that under the specified choice of ϵ_0 in *claim 3.*, for all $q \in Q$ we have

$$|\tilde{p}^\top q_t - p^\top q_t| < 1.25\alpha.$$

Since we visit the synthetic dataset by at most $\frac{16\log|X|}{\alpha^2}$ times, we set $N = \frac{16\log|X|}{\alpha^2}$.

In regards to privacy, we apply Above threshold and Laplace mechanism for each round we update our synthetic dataset, which happens for at most N rounds. With the choice of ϵ_0 and N specified above, by basic composition theorem the privacy budget needs to be at least

$$\epsilon_{total} = 2\epsilon_0 N = \frac{512\log(\frac{3k}{\delta})\log(|X|)}{n\alpha^3}.$$

To sum up, we have proven the following theorem.

Theorem 4.3. *With probability at least $1-\delta$, the Private MW algorithm calibrated to achieve $\epsilon-DP$ is able to answer any online sequence of $|Q|$ linear queries and a max error bounded by*

$$1.25\alpha \leq 1.25\Big(512\Big(\frac{\log(\frac{3|Q|}{\delta})\log(|X|)}{n\epsilon}\Big)^{\frac{1}{3}}\Big)$$

In comparison to two error bounds obtained through Laplace mechanism, Private MW has weaker dependence on n but stronger dependence on either $|X|$ or $|Q|$.

4.2 Exponential mechanism

Laplace mechanism serves little purpose when 1) the output space are discrete set or objects 2) instead of accuracy, we want to maximize the value of a predefined utility function. Consider the following examle. Suppose there n consumers each interested in purchasing one copy of an item, and the seller has an unlimited supply of items. Each consumer is willing to purchase the item at his/her maximum price $p_i \in \mathbb{N}$. Furthermore, all of the consumers are only to willing to pay an integer price for the item.

As the seller, we want a price $p \in \mathbb{N}$ that maximizes the total revenue $\sum_{p_i \leq p} p$. The optimal price p cannot be released directly, so one can think of applying Laplace mechanism to preserve privacy. However, two challenges arises: 1) the price provided by Laplace mechanism are numerical values, not integers 2) even if we ignore the integer price constraint, the perturbed price may plummet the total revenue, making the mechanism of little value.

The exponential algorithm serves to deal with these two challenges. First we formally define utility function and its sensitivity. A utility function is a mapping $u : \mathbb{N}^{|\mathcal{X}|} \times \mathcal{R} \to \mathbb{R}$. We define its sensitivity to be

$$\Delta u \equiv \max_{r \in \mathcal{R}} \max_{x,y:\|x-y\|_1 \leq 1} |u(x,r) - u(y,r)|.$$

The exponential algorithm is defined to output each $r \in \mathcal{R}$ with probability proportional to $\exp(\epsilon u(r,x))/\Delta u)$.

Theorem 4.4. *The exponential mechanism preserves $(\varepsilon, 0)$ differential privacy.*

Proof.

$$\begin{aligned}
\frac{\Pr\left[\mathcal{M}_E(x,u,\mathcal{R})=r\right]}{\Pr\left[\mathcal{M}_E(y,u,\mathcal{R})=r\right]} &= \frac{\left(\frac{\exp\left(\frac{\varepsilon u(x,r)}{2\Delta u}\right)}{\sum_{r'\in\mathcal{R}}\exp\left(\frac{\varepsilon u(x,r')}{2\Delta u}\right)}\right)}{\left(\frac{\exp\left(\frac{\varepsilon u(y,r)}{2\Delta u}\right)}{\sum_{r'\in\mathcal{R}}\exp\left(\frac{\varepsilon u(y,r')}{2\Delta u}\right)}\right)} \\
&= \left(\frac{\exp\left(\frac{\varepsilon u(x,r)}{2\Delta u}\right)}{\exp\left(\frac{\varepsilon u(y,r)}{2\Delta u}\right)}\right)\cdot\left(\frac{\sum_{r'\in\mathcal{R}}\exp\left(\frac{\varepsilon u(y,r')}{2\Delta u}\right)}{\sum_{r'\in\mathcal{R}}\exp\left(\frac{\varepsilon u(x,r')}{2\Delta u}\right)}\right) \\
&= \exp\left(\frac{\varepsilon\left(u\left(x,r'\right)-u\left(y,r'\right)\right)}{2\Delta u}\right) \\
&\quad\cdot\left(\frac{\sum_{r'\in\mathcal{R}}\exp\left(\frac{\varepsilon u(y,r')}{2\Delta u}\right)}{\sum_{r'\in\mathcal{R}}\exp\left(\frac{\varepsilon u(x,r')}{2\Delta u}\right)}\right) \\
&\leq \exp\left(\frac{\varepsilon}{2}\right)\cdot\exp\left(\frac{\varepsilon}{2}\right)\cdot\left(\frac{\sum_{r'\in\mathcal{R}}\exp\left(\frac{\varepsilon u(x,r')}{2\Delta u}\right)}{\sum_{r'\in\mathcal{R}}\exp\left(\frac{\varepsilon u(x,r')}{2\Delta u}\right)}\right) \\
&= \exp(\varepsilon).
\end{aligned}$$

The utility analysis will be discussed in the next lecture.