ISYE/CS 8803: Foundations of Data Privacy **Fall 2017**

Lecture 11: Private Multiplicative Weights

Lecturer: Rachel Cummings *Scribe: Austin Brooks, Oct 4, 2017*

1 Combining MW Update Rule w/ Sparse Vector

The queries to Sparse Vector will be about the error $E_i = \left|f_i(x) - f_i(x^t)\right|$. Sparse Vector can only notify if something is above threshold. Since Sparse Vector cannot capture the absolute value, two copies of Sparse Vector must be run to ask about $f_i(x) - f_i(x^t)$ and $f_i(x^t) - f_i(x)$.

Algorithm 1 Private MW(x, $\{f_i\}$,ϵ,δ,α,β,Q)

Let $c = \dfrac{4\ln|X|}{\alpha^2}$,
if $\delta = 0$ **then**
  Let $T = \dfrac{18c(\ln(2|Q|) + \ln(\frac{4c}{\beta}))}{\varepsilon\|x\|}$
else
  Let $T = \dfrac{(2+32\sqrt{2})\sqrt{c\ln(2/\delta)}(ln(k) + ln(\frac{4c}{\beta}))}{\varepsilon\|x\|}$
end if
Initialize Sparse(x,$\{f_i\}$,T,c,ε,δ), Outputting $\{E_i\}$
Let $t = 0$, $x^o \in \Delta([X])$ s.t. $x_i^o = \dfrac{1}{|X|}$ $\forall i \in [\![X|]$
for all Queries f_i **do**
  Let $f'_{2i-1}(x) = f_i - f_i(x^t)$
  Let $f'_{2i}(x) = f_i(x^t) - f_i(x)$
  if $E_{2i-1} = \perp$ and $E_{2i} = \perp$ **then**
    Let $a_i = f_i(x^t)$
  else
    if $E_{2i-1} \in \mathbb{R}$ **then**
      Let $a_i = f_i(x^t) + E_{2i-1}$
    else
      Let $a_i = f_i(x^t) - E_{2i}$
    end if
    Let $x^(t+1) = MW(x^t, f_i, a_i)$
    Let $t = t + 1$
  end if
end for

Theorem 1. *Private MW is (ε,δ)-DP*

Proof. Follows immediately from privacy of Sparse Vector because Private MW only accesses the data through Sparse Vector. Everything else is post-processing. □

Theorem 2. *With probability* $(1-\beta)$, $\forall f_i$, *Private MW returns an answer* a_i *s.t.* $\left|f_i(x) - a_i \leq 3\alpha\right|$ *for:*
($\delta = 0$),

$$\alpha = \left(\frac{36\ln|X|\left[\ln|Q| + \ln\left(\frac{32\ln|X|^{\frac{1}{3}}\|x\|^{\frac{2}{3}}}{\beta}\right)\right]}{\|x\|^2\,\varepsilon}\right)^{\frac{1}{3}}$$

($\delta > 0$),

$$\alpha = \left(\frac{(2+32\sqrt{2})\sqrt{\ln|X|\ln\frac{2}{\delta}(\ln|Q| + \ln\frac{32\|x\|}{\beta})}}{\|x\|\,\varepsilon}\right)^{\frac{1}{2}}$$

Remark 1. *When $\delta > 0$, accuracy is better in terms of $\|x\|$ because of the better composition theorems for (ε, δ)-DP.*

Proof. Use Sparse Vector accuracy theorem to show that w.h.p,

1. MW update rule is only called when $\left|f_i(x) - f_i(x^t)\right|$ is large,
2. and the released noisy approximation to $f_i(x)$ is accurate.

Recall from Lecture 10, these were the two conditions needed to prove that MW converged quickly. Then use MW convergence theorem to show that after $c = \dfrac{4\ln|X|}{\alpha^2}$ updates, Private MW answers all queries in Q approximately correctly. □

Sparse Vector gives error that scales like $O\left(\dfrac{\ln|X|\left[\ln|Q| + \ln\frac{c}{\beta}\right]}{\varepsilon\|x\|\,\alpha^2}\right)$ for k-sensitivity $\dfrac{1}{\|x\|}$ queries, where at most c are above threshold. The only "payment" in the privacy budget is for those c queries, so noise $O\left(\dfrac{c}{\varepsilon\|x\|}\right)$ can be added to each query.

In this case where $k = |Q|$ and $c = \dfrac{4\ln|X|}{\alpha^2}$. These bounds are achieved by plugging this into the Sparse Vector guarantee

$$\alpha = O\left(\frac{\ln|X|\left[\ln|Q| + \ln\frac{c}{\beta}\right]}{\varepsilon\|x\|\,\alpha^2}\right)$$

and solving.