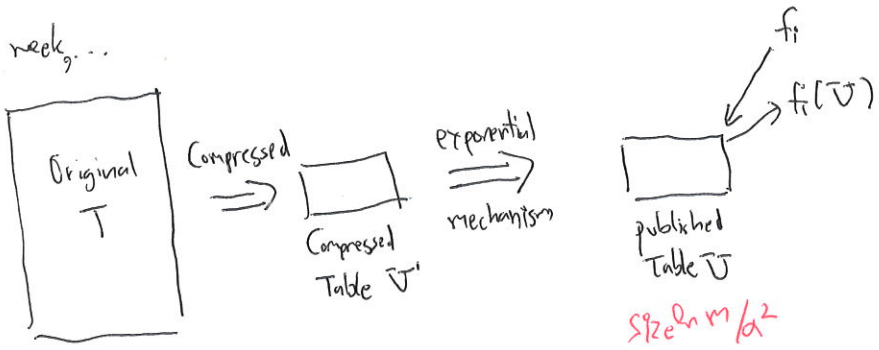
From last week...



Theorem For any T and $f_1, \dots, f_m$, there exists a table $\bar{T}$ with size $\ln m/\alpha^2$ where

$$\max_i |f_i(T) - f_i(\bar{T})| \le \alpha$$

$$\text{Utility}(T, \bar{T}) = -\max_i |f_i(T) - f_i(\bar{T})|$$

There exists $\bar{T}^*$ such that $\text{Utility}(T, \bar{T}^*) \ge -\alpha$

$$OPT \ge -\alpha$$

By the assumption on f_i, we have $\Delta\text{Utility} = \frac{1}{\|T\|}$

Theorem # choices (# possible tables) = (# possible values at each record)$^{\#\text{records}}$ $\ln m/\alpha^2$

Theorem $\Pr[E \le OPT - \frac{2 \cdot \Delta\text{Utility}}{\varepsilon}(\ln(\#\text{choices}) + t)] \le e^{-t}$

when E is utility for publishing table $\bar{T}$.

$$\Pr[E \le -\alpha - \frac{2}{\varepsilon\|T\|}[\ln(\#\text{possible}^{\#\text{records}}) + t]] \le e^{-t} \quad \ln m/\alpha^2 \quad t := -\ln\beta$$

$$\Pr[E \le -\alpha - \frac{2}{\varepsilon\|T\|}(\frac{\ln m}{\alpha^2} \cdot \ln(\#\text{possible}) - \ln\beta)] \le \exp[+\ln\beta] = \beta \quad \sigma := 2\cdot\alpha$$

α : $\sigma/2$; $\frac{\ln m}{\alpha^2}$: $\alpha\sigma/2$; $-\sigma$

$$+\sigma \le -\alpha + \sigma/2 + \frac{2}{\varepsilon\|T\|}[\frac{\ln m}{\alpha^2} \cdot \ln(\#\text{possible}) - \ln\beta] \quad \alpha^2\sigma^2/4$$

$$\sigma/2 \ge \frac{2}{\varepsilon\|T\|}[\frac{\ln m}{\alpha^2\sigma^2/4} \cdot \ln(\#\text{possible}) - \ln\beta]$$

$$\|T\| \ge \frac{4}{\varepsilon\cdot\sigma}[\frac{\ln m}{\sigma^2/4}\ln(\#\text{possible}) - \ln\beta]$$

when $\|T\| \ge \frac{4}{\varepsilon\cdot\sigma}[\frac{\ln m}{\sigma^2/4}\ln(\#\text{possible}) - \ln\beta]$, $\Pr[E \le -\sigma] \le \beta$.

Theorem When T is a table with size $\frac{\ln m}{\varepsilon^3/\alpha^2}$ selected using exponential mechanism, and $|T|$ is large enough.

$$\Pr[E \le \alpha] \le \beta$$

The difference in every result is no more than α.

PAC Learning (Formal Definition)

Definition ~~For any~~ Our algorithm has access to $n = poly(1/\alpha, \log(1/\beta))^{V}$ independent random samples from any distribution $\mathcal{D}$. The algorithm is PAC learning if

$$\Pr[\text{error}(\text{learning output}) > \alpha] \le 1-\beta.$$

for any α and β.

Private PAC Learning Our algorithm has access to $n = poly(1/\varepsilon, 1/\alpha, \log(1/\beta))$ random samples from any distribution $\mathcal{D}$. The algorithm is private PAC learning if

1) The release of learning results is ε-differentially private

2) $\Pr[\text{error}(\text{learning results}) > \alpha] \le 1-\beta$

for any α, β, and ε.

Simple Algorithm (Occam's Razor)

Return learning result with smallest error, for the n samples

* We know from the first class that Occam's Razor is PAC learning.

Algorithm with Differential Privacy

$$Utility(T, \ell) := -[\text{error from learning result } \ell]$$

(error → # misclassification.)

$$h_T(\ell) = \exp\left(\frac{\varepsilon \cdot Utility(T, \ell)}{2 \cdot \Delta Utility}\right), \quad \Delta Utility = 1$$

The error will increase or decrease by 1 if one person change his/her data.

return ℓ with prob. $\frac{h_T(\ell)}{\sum_{\ell'} h_T(\ell')}$

(ℓ → all possible learning results.)

It is clear that this algorithm is ε-differentially private.

[by exponential mechanism]

Theorem By $n = O\left((\ln|\#\text{possible}| + \ln 1/\beta) \cdot \max\left\{\frac{1}{\varepsilon\alpha}, \frac{1}{\alpha^2}\right\}\right)$, we have

$$\Pr[\text{error}(\text{learning output}) > \alpha] \le 1-\beta$$

from the differencial privacy's version of Occam's Razor.

Proof

error (learning output) = Prob. that from $\mathcal{D}$ we will have a misclassification.

difference from definition in mechanism.

$\text{error}_{\text{mech.}}(\text{learning output}) = \#\text{samples} \cdot \text{error}_T(\ell)$

$$\text{error}_T(\text{learning output}) = \frac{\#\text{ samples in } T \text{ that is misclassified}}{\#\text{samples}} = \frac{\sum_{i=1}^{n} X_i}{\#\text{samples}} = S$$

X_i: Is sample i misclassified

expected value of $\text{error}_T(\text{learning output}) = \text{error}(\text{learning output})$

By Chernoff's bound,

$$\Pr[|\underbrace{\text{error}_T(\text{learning output})}_{S} - \underbrace{\text{error}(\text{learning output})}_{\mu}| \ge \rho] \le 2\cdot\exp(-2n\rho^2).$$

for all learning output.

$$\Pr[|\text{error}_T(\ell) - \text{error}(\ell)| \ge \rho \text{ for some learning output } \ell]$$
$$\le \sum_{\ell} \underbrace{\Pr[|\text{error}_T(\ell) - \text{error}(\ell)| \ge \rho]}_{\le 2\cdot\exp(-2n\rho^2)} \le |\#\text{possible output}| \cdot 2\cdot\exp(-2n\rho^2).$$

from now, assume that $|\text{error}_T(\ell) - \text{error}(\ell)| < \rho$ for all ℓ. The probability that we output a particular learning result ℓ is

$$\frac{\exp\left(-\frac{\varepsilon}{2}\cdot n\cdot \text{error}_T(\ell)\right)}{\sum_{\ell'}\exp\left(\frac{-\varepsilon}{2}\cdot n\cdot \text{error}_T(\ell')\right)} \le \frac{\exp\left(-\frac{\varepsilon}{2}\cdot n\cdot \text{error}_T(\ell)\right)}{\max_{\ell'}\exp\left(-\frac{\varepsilon}{2}\cdot n\cdot \text{error}_T(\ell')\right)} = \frac{\exp\left(-\frac{\varepsilon}{2}\cdot n\cdot \text{error}_T(\ell)\right)}{\exp\left(\max_{\ell'} -\frac{\varepsilon}{2}\cdot n\cdot \text{error}_T(\ell')\right)}$$

$\sum_{\ell'}\exp\left(\frac{-\varepsilon}{2}\cdot n\cdot \text{error}_T(\ell')\right) \ge \max_{\ell'}\exp\left(\frac{-\varepsilon}{2}\cdot n\cdot \text{error}_T(\ell')\right)$

$$= \exp\left(-\frac{\varepsilon}{2}\cdot n\,\text{error}_T(\ell) - \max_{\ell'}\left[-\frac{\varepsilon}{2}\cdot n\cdot \text{error}_T(\ell')\right]\right)$$

$$= \exp\left(-\frac{\varepsilon}{2}n\left(\text{error}_T(\ell) - \min_{\ell'}\text{error}_T(\ell')\right)\right)$$

$\min_{\ell'}\text{error}_T(\ell') \le \text{error}_T(\ell^*)$

$\le \underbrace{\text{error}(\ell^*)}_{OPT} + \delta$

$$\le \exp\left(-\frac{\varepsilon}{2}n\left(\text{error}_T(\ell) - OPT - \rho\right)\right)$$