Differentially Private Matrix Completion Revisited
Prateek Jain, Om Thakkar, Abhradeep Thakurta

1Microsoft Research, 2Boston University, 3University of California-Santa Cruz

INTRODUCTION
Low-Rank Matrix Completion: Given an incomplete matrix $X \in R$, s.t. $R$ is low-rank, output $P$, such that $Y = R$.

Main contributions:

A joint DP matrix completion algorithm that provides the first non-trivial generalization error guarantee, and better empirical performance than state-of-the-art private algorithms.

MAIN CONTRIBUTION

Joint DP Frank-Wolfe (FW) is an iterative algorithm in which each iteration $t \in [T]$ can be broadly divided into two parts:

1) Local (user-side) computation:

- Let $Y_{ij}$ denote $Y$ without user $i$'s output.
- For randomized algorithm $A$, we have $Pr(A(Y_{-ij}) = Y_{ij}) = Pr_{ij}$ for all $i$, $j$.
- If $X$'s rank is $d$, then $A$ will output a rank-$d$ $X$.
- Distinction from standard DP: (a) joint DP guarantee for each user $i$ (b) $orall X$, s.t. $d(X,X') = 1$ and user $i$'s data changes, $Y_{ij}$ is in Range($A_{ij}$).

2) Global (server-side) computation:

- Privacy parameters: $(\epsilon,\delta)$.
- Global private update $\hat{Y}$ is $\hat{Y} = \hat{Y} + \delta(X,Y)$.
- global update $\hat{Y}_{ti}$ is $\hat{Y}_{ti} = \frac{1}{1+\delta} Y_{ti} + \frac{1}{1+\delta} \hat{Y}_{ti}$.
- Global update $\hat{Y}_{t+1}$ is $\hat{Y}_{t+1} = \hat{Y}_{t+1} + \delta(X,Y)$.

PRIVATE OJA'S ALGORITHM

Let $Z = \sum_{i=0}^{n} x_i$. Server update in Joint DP Frank-Wolfe:

- Update: $v_t = v_{t-1} + \epsilon (v_{t-1} \cdot Z + \text{Gaussian}(\epsilon,\delta,\Delta))$
- Normalize: $v_t = \frac{v_t}{\|v_t\|}$
- Return $v_t, \hat{d} = \|v_t \cdot Z\|^2 + \text{Gaussian}(\epsilon,\delta,\Delta)$

UTILITY OF JOINT DP FRANK-WOLFE

If $\Omega$ is set of non-zero indices in $X$, $\max_{\Omega \subseteq \{1\}} \|v_t\| \leq L$, for all $t$, and we run $(\epsilon,\delta)$ joint DP Frank-Wolfe for $T$ iterations, then w.h.p:

- $\mathbb{E}[\|v_t\|^2] = O(\epsilon^{-2} \cdot n \cdot T)$
- $\mathbb{E}[\|v_t\|] = O(\epsilon \cdot n \cdot T)$

EXPERIMENTAL RESULTS

- $n =$ number of users, $d =$ number of items, $\delta = 10^{-8}$
- Unless specified, each rating in $[0.5]$, sample 80 ratings per user (Top 400) - Selecting the 400 most rated items.
- Algorithms: DP Frank-Wolfe, DP SVD after cleansing (Mironov, DP Projected Gradient Descent (PGD) [CCS’16, ITCS’14, ACG’16].

REFERENCES

[CCS’16] Catriona Sherr, SoFPT’16.
[DMM’09] Dwork McSherry Mironov Smith, ITCS’06.