`dp-stats`: A Python Library for Differentially-private Statistics and Machine Learning Algorithms

Sijie Xiong and Hafiz Imtiaz

Advisor: Anand D. Sarwate

Rutgers University

Motivation

- **Goal**: to prepare a publicly available Python library of commonly used differentially-private (DP) statistics and machine learning algorithms.

What's in the Package

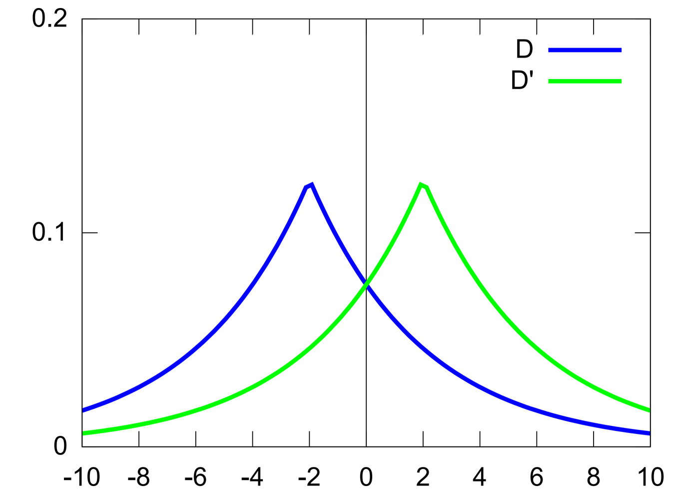
- **Functions:** Mean, Variance, Histogram, Principal Component Analysis (PCA), Support Vector Machines (SVM), Logistic Regression.
- iPython notebook tutorials for each function.
- Package installation and setup guidelines.

Differential Privacy

- Algorithm $\mathcal{A}(\mathbb{D})$ taking values in an output space $\mathbb{T}$ provides (ϵ, δ)-differential privacy [2] if

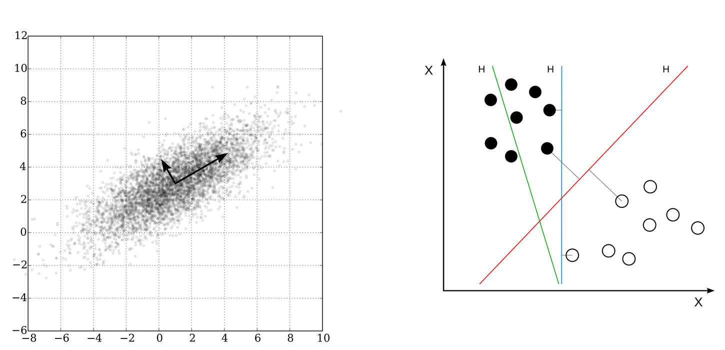
$$\Pr(\mathcal{A}(\mathbb{D}) \in \mathbb{S}) \leq \exp(\epsilon)\Pr(\mathcal{A}(\mathbb{D}') \in \mathbb{S}) + \delta,$$

  for all measurable $\mathbb{S} \subseteq \mathbb{T}$ and all *neighboring* data sets $\mathbb{D}$ and $\mathbb{D}'$ differing in a single entry.
- ϵ and δ - privacy parameters.
- Low ϵ and δ ensure more privacy.



Basics of PCA and SVM

- **PCA:** is a statistical procedure to convert a set of samples of possibly correlated variables into a set of linearly uncorrelated variables using orthogonal transformation.
- **SVM:** given a set of labeled training samples, SVM builds a model (separating hyperplane) that can assign labels to new samples.



Figures are from [3] and [4].

`dp-stats` in Action

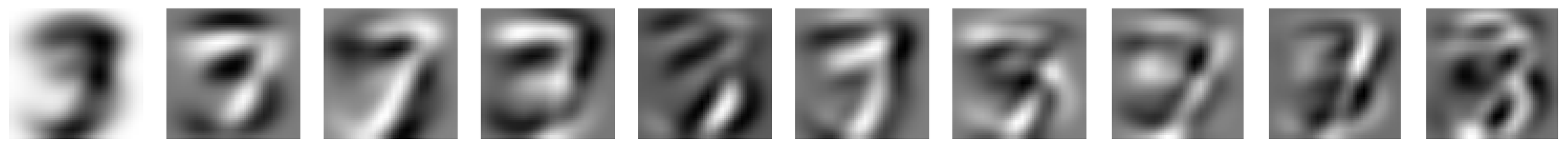
How to use the `dp-stats` package for PCA and SVM?

- $d \times n$ data matrix $X = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$
- $d \times d$ positive semi-definite second-moment matrix $A = XX^\top$
- Data vectors $\mathbf{x}_i \in \mathbb{R}^d$ are bounded $\|\mathbf{x}_i\|_2 \leq 1$

Samples from MNIST [1] Dataset

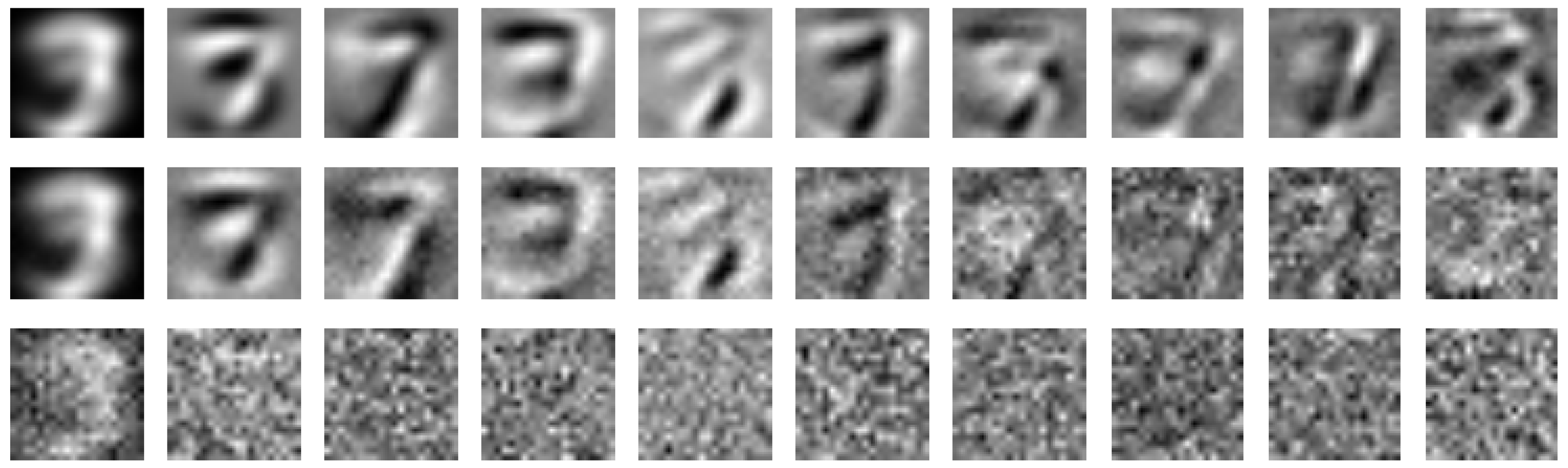
Non-private Principal Components

```
d,n=tr_samples.shape; A=np.dot(tr_samples,tr_samples.T)/n; U,S,V=np.linalg.svd(A)
```



DP Principal Components: Using AG Algorithm [2]

```
Ahat=dps.dp_pca_ag(tr_samples, epsilon=10, delta=0.01); U,S,V=np.linalg.svd(Ahat)
```



In each row, we're showing the top-10 DP principal components for $\epsilon = 10, 2, 0.1$, respectively (with $\delta = 0.01$).

Analyze Gauss (AG) [2]

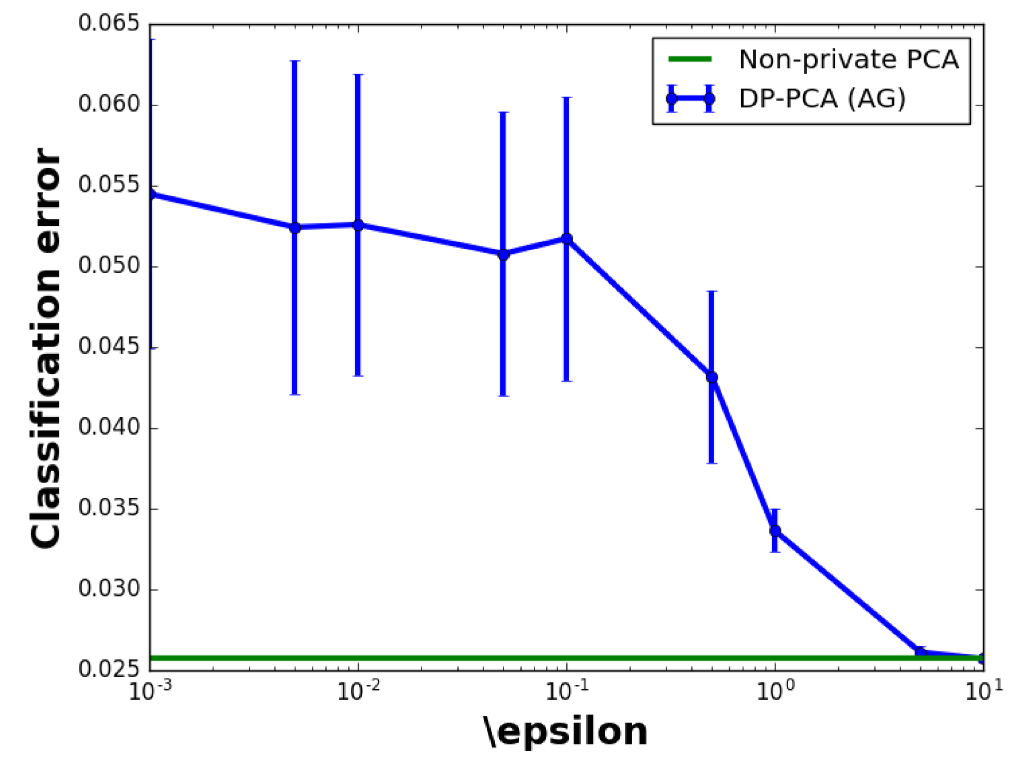
1. Set $\Delta_{\epsilon,\delta} = \frac{1}{\epsilon}\sqrt{2\log(\frac{1.25}{\delta})}$
2. Generate symmetric E of i.i.d. samples from $\mathcal{N}(0, \Delta^2_{\epsilon,\delta})$
3. Compute $\hat{A} = A + E$

Output: Private second-moment matrix $\hat{A}$. Set $\hat{V}_k$ using PCA on $\hat{A}$.

SVM Classification

- Reduce dimension: $X^{(k)} = V_k^\top X$
- Train SVM classifier with $X^{(k)}$: find $f \in \mathbb{R}^k$ for a linear classifier $\mathrm{sgn}(f^\top \mathbf{x}_i^{(k)})$, where $\mathbf{x}_i^{(k)}$ is the i-th k-dimensional sample
- Vary ϵ for DP-PCA and perform SVM classification

Classification Error vs. Privacy



Comments

- DP principal components are noisy.
- When ϵ decreases, the noise added to the principal components increases.
- For $\epsilon = 0.01$, the features are not visually meaningful but are still useful for classification.
- **Open question**: how to optimally allocate a total privacy budget to different stages of an algorithm.

Future Directions

- Extend the package to include more differentially-private algorithms.
- Scan here for more information!
- Contributions are more than welcome!

References

1. LeCun et al., "The MNIST database of handwritten digits." (1998).
2. Dwork et al., *Analyze Gauss: Optimal Bounds for Privacy-preserving Principal Component Analysis*, Annual ACM Symp. Theory of Comp., 2014.
3. `https://commons.wikimedia.org/w/index.php?curid=46871195`.
4. `https://commons.wikimedia.org/w/index.php?curid=22877598`.