HW 4B: Stochastic Gradient Descent and Lipschitz Extensions

CS 208 Applied Privacy for Data Science, Spring 2019

Version 1.2: Due Tuesday, April 30, 11:59pm.

Instructions: Submit a single PDF file containing your solutions, plots, and analyses. Make sure to thoroughly explain your process and results for each problem. Also include your documented code and a link to a public repository with your code (such as GitHub/GitLab). Make sure to list all collaborators and references.

1. For each of the following sets $\mathcal{G}$ of datasets and neighbor relations $\sim$, hypotheses $\mathcal{H} \subseteq \mathcal{G}$, and functions $f : \mathcal{G} \to \mathbb{R}$, calculate (i) the global sensitivity of f (denoted GS_f or ∂f), (ii) the minimum local sensitivity of f, i.e. $\min_{x \in \mathcal{G}} \mathrm{LS}_f(x)$, and (iii) the restricted sensitivity of f (denoted $\partial_{\mathcal{H}} f$ or $\mathrm{RS}_f^{\mathcal{H}}$). For Part 1a, also describe an explicit Lipschitz extension of f from $\mathcal{H}$ to all of $\mathcal{G}$.

   (a) $\mathcal{G} = \mathbb{R}^n$ where $x \sim x'$ if x and x' differ on one row, $\mathcal{H} = [a, b]^n$ for real numbers $a \leq b$, and $f(x) = (1/n) \sum_{i=1}^n x_i$.

   (b) $\mathcal{G} = \mathbb{R}^n$ where $x \sim x'$ if x and x' differ on one row, $\mathcal{H} = [a, b]^n$ for real numbers $a \leq b$, and $f(x) = \mathrm{median}(x_1, \ldots, x_n)$.

   (c) $\mathcal{G}$ = the set of undirected graphs (without self-loops) on vertex set $\{1, \ldots, n\}$ where $x \sim x'$ if x and x' are identical except for the neighborhood of a single vertex (i.e. node privacy), $\mathcal{H}$ = the set of graphs in $\mathcal{G}$ in which every vertex has degree at most d for a parameter $2 \leq d \leq n-1$, and $f(x)$ = the number of isolated (i.e. degree 0) vertices in x.

2. In our code example,[1] we saw how to release an estimated Logistic regression using differentially private stochastic gradient descent (DP-SGD) to optimize the log-likelihood loss function under the centralized model. Convert this code to once again release the probability of marriage given education level, but using DP-SGD under the *local* model.[2] Recall that local DP does not satisfy privacy amplification by subsampling, but you can achieve a similar effect by rotating through disjoint batches, so that each individual partipates in at most $\lceil T \cdot B/n \rceil$ batches, where T is the number of iterations and B is the batch size.[3] Evaluate the performance of your method as a function of ϵ (fixing $\delta = 1 \times 10^{-6}$), by showing the classification error over ϵ, compared to the RMSE of the coefficients compared to the non-privacy preserving estimates.

[1] See `https://github.com/privacytoolsproject/cs208/blob/master/examples/wk7_localmodel/privateSGD.r` and `privateSGD.ipynb`.

[2] For some useful guidance, look at the class notes for April 8th at `http://people.seas.harvard.edu/~salil/cs208/spring19/MLwithDP-lecture.pdf`.

[3] Note, in the code example, $T = \sqrt{n}$, $B = \sqrt{n}$.