Diff. Private Submodular Max. Under Matroid and Knapsack Constraints

Omid Sadeghi†, Maryam Fazel†
†University of Washington

Submodular Functions

Submodular set functions. A set function \( f : 2^V \to \mathbb{R} \) over the ground set \( V \) is submodular if for all \( f \in V \) and for all sets \( A \subseteq B \subseteq V \), the following holds:

\[
F(A U \{j\}) - F(A) \geq F(B U \{j\}) - F(B).
\]

• Examples: cut functions of graphs and hypergraphs, rank functions of matroids and covering functions.

• Applications in viral marketing, feature selection for classification, image segmentation and document summarization.

The multilinear extension \( f : [0,1]^{|V|} \to \mathbb{R} \) of \( f \) is defined as

\[
f(x) = \sum_{S \subseteq V} F(S) \prod_{j \in S} x^j (1 - x^j) = E_{S \subseteq V}[F(S)].
\]

DR-submodular functions. We say that a differentiable function \( f : \mathcal{X} \to \mathbb{R}, \mathcal{X} \subseteq \mathbb{R}^m \), is DR-submodular if its gradient is an order-reversing mapping, i.e.,

\[
x, y \geq 0 \implies \nabla f(x) \preceq \nabla f(y).
\]

A twice differentiable function \( f \) is DR-submodular if and only if its Hessian matrix \( \nabla^2 f \) is entry-wise non-positive.

• Multilinear extension of submodular set functions are DR-submodular.

• Although DR-submodularity and concavity are equivalent for the special case of \( m = 1 \), DR-submodular functions are generally non-concave.

An important consequence of DR-submodularity is concavity along non-negative directions, i.e., for all \( x, y \) such that \( x \leq y \), we have \( f(y) \leq f(x) + \nabla f(x)^T (y - x) \).

Motivating Application: Feature Selection

Let \( D = \{(x_t, c_t)\}_{t=1}^T \) be a sensitive dataset consisting of a feature vector \( x_t \in \mathbb{R}^m \) for each individual \( t \in [T] \) along with a binary class label \( C_t \). The goal is to select a small subset of the \( m \) features that provide a good classifier for the labels. Determining the likelihood of an individual having a certain disease using a representative collection of his or her features (such as height, age and weight) could be cast in this framework.

In order to solve this problem, [2] proposed a non-private algorithm based on maximizing a submodular function capturing the mutual information between a subset of the features and the class label of interest. However, in this setting, along with obtaining the most relevant subset of features, it is crucial to ensure that the privacy of any individual included in the dataset is not compromised.

Diff. Private Offline Submodular Max.

Let \( f : [0,1]^{|V|} \to \mathbb{R} \) be the multilinear extension of the monotone submodular set function \( f : 2^V \to \mathbb{R} \). Set \( \lambda = \epsilon^2 2^{|V|} \geq 0 \) where \( \epsilon = \min_{x \in \mathbb{R}^{|V|}} \nabla f(x) = F(V) - F(V \setminus \{j\}) \) and \( x^* \in 2^V \) is the optimal point corresponding to the optimal solution \( S^* \subseteq V \). Also, let \( g(x) = f(x) - \epsilon^2 x \).

Algorithm \( \kappa \)-Differentially Private Continuous Greedy (\( \kappa \)-DPCG) algorithm

Input: \( \kappa, \epsilon, \delta \geq 0 \), the constraint set \( \mathcal{P} \), and the noise distribution \( D \).

Initialization: \( x^{(0)} = 0 \).

for \( k = 1, 2, . . ., K \) do

Draw \( y^{(k)} \) \~\( D \).

Set \( y^{(k)} = \arg \max_{y \geq 0} g(x^{(k)} + \epsilon \nabla f(y^{(k)}) + y^{(k)}) \).

Set \( x^{(k+1)} = x^{(k)} + \frac{1}{\epsilon} y^{(k)} \).

end for

Output: \( x = x^{(K+1)} \).

Theorem 1. Let \( \kappa_F = 1 - \min_{x \neq y} \frac{\nabla f(x)^T (y - x)}{\epsilon^2} \) be the total curvature of \( f \). Setting \( K = \Theta\left(\frac{\epsilon^2}{\epsilon^2 + \kappa_F} \frac{\ln(1/\delta)}{\ln(1/\delta)}\right) \), the \( \kappa \)-DPCG Algorithm is \( \epsilon, \delta \)-differentially and has the following approximation guarantees for matroid and knapsack constraints respectively:

\[
E[f(x)] \geq \left(1 - \frac{\kappa_F}{\epsilon^2}\right) f(x^*) - \Theta\left(\frac{\epsilon^2}{\epsilon^2 + \kappa_F} \frac{\kappa_F}{\kappa_F + \epsilon} \frac{\epsilon}{\epsilon + \delta}\right).
\]

Diff. Private Online Submodular Max.

The \( (1 - \frac{1}{e}) \)-regret is defined as \( R_T = (1 - \frac{1}{e}) \max_{x \in \mathcal{X}} \sum_{t=1}^T F_t(S_t) - \sum_{t=1}^T F_t(S_t) \).

Algorithm: Differentially Private Meta-Frank-Wolfe (DPMFW) algorithm

Input: \( K \), the constraint set \( \mathcal{P} \).

Output: \( x_T \).

Initialize \( \mathcal{K} \) instances \((\mathcal{C}_t, k_t)_{t=1}^K\) of the \( \frac{\epsilon^2}{\epsilon^2 + \kappa_F} \)-differentially private algorithm of [1] with noise distribution \( D \) and regularizer \( R(x) = \sum_{t=1}^K x_t \ln(x_t) \) for online linear optimization over \( \mathcal{P} \).

for \( t = 1, 2, . . ., T \) do

\( x_t^{(0)} = x_t \).

for \( k = 1, 2, . . ., K \) do

Let \( x_t^{(k)} \) be the output of \( \mathcal{C}_t \) for round \( t \).

Set \( x_t^{(k+1)} = x_t^{(k)} + \frac{1}{\epsilon} y_t^{(k)} \).

end for

Play \( x_t = x_T^{(K+1)} \).

for \( k = 1, 2, . . ., K \) do

Feedback \( \nabla f(x_t^{(k)}) \) to \( \mathcal{C}_t \) as the linear utility vector observed at round \( t \).

end for

Theorem 2. Let \( 0 < \epsilon < 1 \) and \( \delta > 0 \). If \( D = \mathcal{L}^{|V|}(\mu) \), setting \( \mu = \frac{2\epsilon^2 |V| / \kappa_F}{e \sqrt{\epsilon^2 + \kappa_F} \ln(1/\delta)} \) and \( K = \Theta\left(\sqrt{T} \right) \), the DPMFW algorithm is \( \epsilon, \delta \)-differentially private and has the following expected regret bounds for matroid and knapsack constraints respectively:

\[
E[R_T] \leq O\left(\frac{\kappa_F}{\epsilon^2} \sqrt{T} \ln(1/\delta) \right) + \Theta\left(\frac{\kappa_F}{\epsilon^2 + \kappa_F} \frac{\kappa_F}{\kappa_F + \epsilon} \frac{\epsilon}{\epsilon + \delta}\right).
\]

Reference

