Online DR-Submodular Maximization: Minimizing Regret and Constraint Violation

Online Protocol

- A convex domain set $X \subseteq \mathbb{R}^n$ and total available budget $BT$.
- At step $t \in [T]$, the player chooses $x_t \in X$.
- Then, a monotone DR-submodular utility function $f_t: X \rightarrow \mathbb{R}$ and a constraint vector $p_t$ are revealed, the player obtains the reward $f_t(x_t)$ and uses $p_t$, $x_t$ of her budget.

Goal: Maximize the overall obtained reward while ensuring the budget constraint is satisfied on average, i.e., the total budget violation $C_T = \frac{1}{T} \sum_{t=1}^{T} p_t x_t > B$ should grow sub-linearly in $T$.

$$\max \sum_{t=1}^{T} f_t(x_t), \quad \text{subject to } \frac{1}{T} \sum_{t=1}^{T} p_t x_t > B.$$ 

DR-Submodular Functions. A differentiable function $f: X \rightarrow \mathbb{R}, X \subseteq \mathbb{R}_n$, is called DR-submodular if:

$$x \rightarrow f(y(x))$$

If $f$ is twice differentiable, DR property is equivalent to the Hessian being element-wise non-positive.

Regret metric. The $(1 - \frac{1}{e})$-regret is defined as:

$$R_T = \frac{1}{1 - \frac{1}{e}} \sum_{t=1}^{T} f_t(x^*) - \sum_{t=1}^{T} f_t(x_t),$$

where:

$$x^* = \arg\max_{x \in \mathbb{R}^n} f_t(x).$$

Motivating Application: Online Ad Placement

- Your company produces some product and has a choice of $n$ different websites where it can advertise under a fixed limited budget $BT$.
- At step $t \in [T]$, you choose a selection $x_t \in X = \{x \in \mathbb{R}^n; 0 \leq x \leq 1\}$ and place your ads on the websites.
- Then, the cost for this assignment, $< p_t, x_t >$, is revealed where the components of $p_t$ capture the traffic received by the ads on the $n$ websites. Also, you get utility $f_t(x_t)$ which corresponds to the number of products sold on that website.

Our Algorithm

- The penalized formulation of the overall optimization problem can be written as:

$$\max_{x \in \mathbb{R}^n} \sum_{t=1}^{T} f_t(x_t) - \frac{1}{2\delta \mu} \left( \sum_{t=1}^{T} p_t x_t > BT \right)^2.$$ 

- Or equivalently as:

$$\max_{x \in \mathbb{R}^n} \sum_{t=1}^{T} f_t(x_t) - \frac{1}{2\delta \mu} \left( \sum_{t=1}^{T} p_t x_t > BT \right)^2 + \delta \mu x^2.$$ 

- Thus, the per round Lagrangian can be written as:

$$L_t(x, \lambda) = f_t(x) - \lambda \left( < p_t, x > - B \right) + \delta \mu x^2.$$ 

Algorithm 1 Online Lagrangian Frank-Wolfe

Input: $X, T$ is the horizon, $\mu > 0$, $\delta > 0$, $\gamma \in [0, 1]$ and $K$. Initialize $K$ instances $\xi_k \forall k \in [K]$ of Online Gradient Ascent with step size $\mu$ for online maximization of linear functions over $X$.

for $t = 1$ to $T$ do

Set $x_t = \frac{1}{|K|} \sum_k \xi_k^{(t)}$.

Let $\bar{p}_t = \frac{1}{|K|} \sum_k p_t \xi_k$, for $t > 1$.

Let (I) $\bar{p}_t = \left\{ \bar{p}_t, \right\} - B$ or (II) $\bar{p}_t = \left\{ \bar{p}_t, \right\} - \gamma$.

Set $\lambda_t = \frac{\delta \mu}{\gamma - \mu}$ for $t > 1$ and 0 otherwise.

Play $x_t$ and observe $\mathcal{L}_t(x_t, \lambda_t) = f_t(x_t) - \lambda_t \bar{p}_t (x_t) + \frac{\delta \mu}{2} x_t^2$.

for $k = 1$ to $K$ do

Feedback $(\xi_k^{(t)}, \nabla \mathcal{L}_t(x_t^{(t)}, \lambda_t))$ as the payoff to $\xi_k$.

end for

end for

Theoretical Results

Theorem (Expectation Bounds): With parameters $\mu = \frac{\delta}{\sqrt{T}}$, $K = \sqrt{T}, \delta = \beta^2$, Algorithm 1 with update rule (I) achieves $O\left( \frac{\beta}{T} \right)$ regret and $O\left( \frac{\beta}{T} \right)$ constraint violation bound in expectation.

Theorem (High Probability Bounds): For every $\gamma > 0$, with parameters $\mu = \frac{\gamma \delta}{\sqrt{T}}, K = \sqrt{T}, \delta = \beta^2, \gamma T = 2\delta + \ln \left( \frac{1}{\eta} \right) \cdot T$, Algorithm 1 with update rule (II) achieves $O\left( \frac{\beta}{T} \right)$ regret and constraint violation bound simultaneously with probability at least $1 - \epsilon$.

Numerical Experiments

Experiment 1: Online joke recommendation using the Jester dataset.

$$f_t(x) = t^2 x + \sum_{i \in [T]} q_{ti} x_i \quad \forall t \in [T],$$

where $0 \leq |q_t| \leq 10$ is the rating of user $t$ for joke $i$ in the dataset.

Experiment 2: Effect of parameter $\delta$ on the performance of the algorithm $f_t(x) = x^T Q_t x$.

where $Q_t$ is a random matrix whose elements are non-positive. $\delta$ is varied from $10^{-1}$ to $10^3$.

Acknowledgements

This work was supported in part by the following grants: NSF TRIPODS grant 1740551, DARPA Lagrange grant FA8650-18-2-7836, ONR MURI grant N00014-16-1-2710.

References

Jester dataset: http://eigentaste.berkeley.edu/dataset.