Online Continuous DR-Submodular Maximization with Long-Term Budget Constraints [1]

Omid Sadeghi, Maryam Fazel

University of Washington, Seattle

Diminishing Returns (DR) Property

Definition

A differentiable function $F : K \to \mathbb{R}$, $K \subset \mathbb{R}^n_+$, satisfies the Diminishing Returns (DR) property if:

$$x \succeq y \Rightarrow \nabla F(x) \preceq \nabla F(y)$$

- If F is twice differentiable, the DR property is equivalent to $[\nabla^2 F(x)]_{i,j} \leq 0 \ \forall i, j \in [n], x \in K$.
- For $n = 1$, the DR property is equivalent to concavity, however, for $n > 1$, they are not equivalent.

Functions which satisfy the DR property are called "smooth submodular" and "DR-submodular" in the literature.

Introduction

Motivating Application: Online Ad Placement

$$\begin{aligned} &\text{maximize}_{x_t \in \mathcal{X}} && \textstyle\sum_{t=1}^{T} f_t(x_t) \\ &\text{subject to} && \textstyle\sum_{t=1}^{T} \langle p_t, x_t \rangle \leq B_T \end{aligned}$$

- At round $t \in [T]$, an advertiser should choose an investment vector $x_t \in \mathcal{X}$ over n different websites where $[x_t]_i$ denotes the amount that the advertiser is willing to pay per each click on the ad on the i-th website.
- The cost of an investment is $\langle p_t, x_t \rangle$ where $[p_t]_i$ is the number of clicks the ad on the i-th website receives.
- $p_t \ \forall t \in [T]$ is not known in advance and could be adversarial.
- The advertiser needs to balance her total investment against an allotted long-term budget B_T.
- At round $t \in [T]$, the advertiser's utility function $f_t(x_t)$, quantifying overall amount of impressions of the ads, is monotone DR-submodular, i.e., making an ad more visible will attract proportionally fewer extra viewers because each website shares a portion of its visitors with other websites.

Performance Metric

Definition (Regret Metric)

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

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

where:

$$x_W^* = \arg\max_{x \in \mathcal{X}_W} \sum_{t=1}^{T} f_t(x)$$

$$\mathcal{X}_W = \{x \in \mathcal{X} : \sum_{\tau=t}^{t+W-1} \langle p_\tau, x \rangle \leq \frac{W}{T} B_T, \ 1 \leq t \leq T-W+1\}$$

Definition (Total Budget Violation Metric)

The total budget violation is defined as follows:

$$C_T = \sum_{t=1}^{T} \langle p_t, x_t \rangle - B_T$$

Goal: Design an online algorithm which achieves sub-linear bounds for both the $(1-\frac{1}{e})$-regret R_T and the budget violation C_T.

Algorithm

Algorithm 1 OSPHG Algorithm

Input: Domain set $\mathcal{X}$, horizon T, μ, δ and K

Output: $\{x_t : 1 \leq t \leq T\}$

Initialize K instances $\mathcal{E}_k \ \forall k \in [K]$ of Online Gradient Ascent with step size μ for online maximization of linear functions over $\mathcal{X}$

$\lambda_1 = 0$

$\mathbf{v}_0^{(k)} = 0 \ \forall k \in [K]$

for $t = 1$ **to** T **do**

  $x_t^{(1)} = 0$

  for $k = 1$ **to** K **do**

    $v_t^{(k)} = \mathcal{P}_{\mathcal{X}}\left(v_{t-1}^{(k)} + \mu \nabla_x \mathcal{L}_{t-1}(x_{t-1}^{(k)}, \lambda_{t-1})\right)$

    $x_t^{(k+1)} = x_t^{(k)} + \frac{1}{K} v_t^{(k)}$

  end for

  Play $x_t = x_t^{(K+1)}$ and observe the function $\mathcal{L}_t(x_t, \lambda_t) = f_t(x_t) - \lambda_t g_t(x_t) + \frac{\delta\mu}{2}\lambda_t^2$

  for $k = 1$ **to** K **do**

    Feedback $\langle v_t^{(k)}, \nabla_x \mathcal{L}_t(x_t^{(k)}, \lambda_t) \rangle$ as the payoff to be received by $\mathcal{E}_k$

  end for

  $\lambda_{t+1} = [\lambda_t - \mu \nabla_\lambda \mathcal{L}_t(x_t, \lambda_t)]_+$

end for

Main Lemma

For $\mu = \frac{R}{\beta\sqrt{WT}}$, $\delta = 4\beta^2$ and any $\lambda \geq 0$, if T is large enough, we have:

$$\begin{aligned} R_T + C_T\lambda - \frac{\delta\mu}{2}T\lambda^2 - \frac{\lambda^2}{\mu} \leq &(F + \beta R)(W-1) + \frac{G}{2}(G+\beta R)\mu(W-1)(T-1) \\ &+ \frac{R^2}{\mu} + (G^2+\beta^2)\mu T + \frac{G^2}{2}\mu(W-1)(T-W+1) + \frac{LR^2}{2K}(T-W+1) \end{aligned}$$

Regret and Budget Violation Bound

For $W = o(T)$, if we choose $\mu = \frac{R}{\beta\sqrt{WT}} = \mathcal{O}(\frac{1}{\sqrt{WT}})$ and $K = \mathcal{O}(\sqrt{\frac{T}{W}})$, the $(1-\frac{1}{e})$-regret R_T and budget violation C_T satisfy the following:

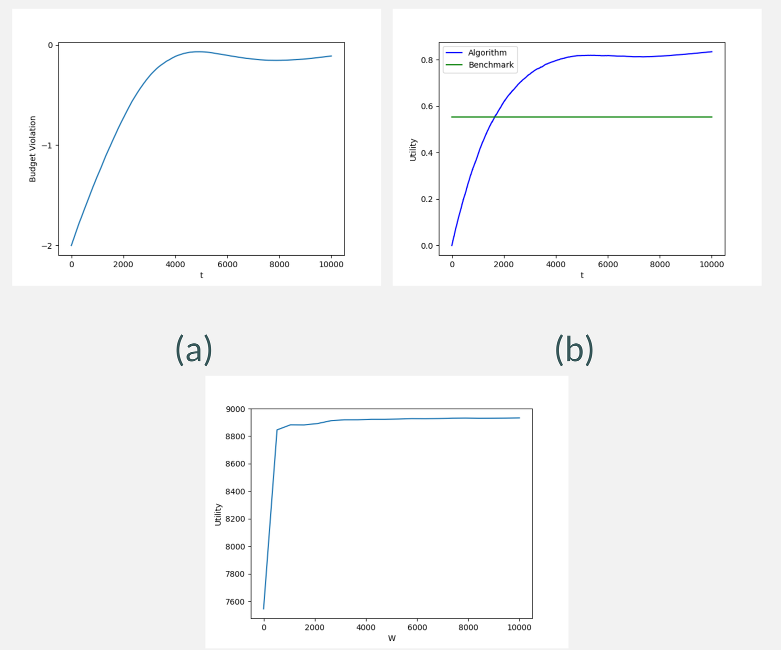
$$R_T \leq \mathcal{O}(\sqrt{WT})$$

$$C_T \leq \mathcal{O}(W^{\frac{1}{4}}T^{\frac{3}{4}})$$

Thus, for $W = T^{1-\epsilon} \ \forall \epsilon > 0$, the $(1-\frac{1}{e})$-regret and budget violation of the OSPHG algorithm is $\mathcal{O}(T^{1-\frac{\epsilon}{2}})$ and $\mathcal{O}(T^{1-\frac{\epsilon}{4}})$ respectively and hence sub-linear.

Experiments

We defined $\mathcal{X} = \{x \in \mathbb{R}^n : 0 \preceq x \preceq \mathbf{1}\}$ and for all $t \in [T]$, we randomly generated monotone non-convex/non-concave quadratic utility functions of the form $f_t(x) = \frac{1}{2}x^T H_t x + h_t^T x$ where $H_t \in \mathbb{R}^{n \times n}$ is a random matrix with uniformly distributed non-positive entries in $[-1, 0]$ and $h_t = -H_t^T \mathbf{1}$ to make the gradient non-negative. Therefore, the utility functions are of the form $f_t(x) = (\frac{1}{2}x - \mathbf{1})^T H_t x$. For all $t \in [T]$, we generated random linear budget functions such that p_t has uniformly distributed entries in $[2, 4]$. We set $T = 10000$, $n = 2$, $B_T = 2T$ and $K = 100$.



(a) (b) (c)

Figure: (a) Budget violation running average $\frac{\sum_{\tau=1}^{t} g_\tau(x_\tau)}{t}$ of OSPHG algorithm for $W = \sqrt{T}$ (b) Utility performance running average $\frac{\sum_{\tau=1}^{t} f_\tau(x_\tau)}{t}$ of OSPHG algorithm for $W = \sqrt{T}$ vs. utility of the benchmark (c) Utility of the benchmark for different window lengths $1 \leq W \leq T$

References

[1] Omid Sadeghi and Maryam Fazel. Online continuous dr-submodular maximization with long-term budget constraints. *arXiv preprint arXiv:1907.00316*, 2019.

ELECTRICAL & COMPUTER ENGINEERING

UNIVERSITY of WASHINGTON