A Single Recipe for Online Submodular Maximization with Adversarial or Stochastic Constraints

Omid Sadeghi[1], Prasanna Raut[2], Maryam Fazel[1]

1. ECE Department, University of Washington

2. ME Department, University of Washington

Online Protocol

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

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

$$\underset{x_t \in \mathcal{X}}{\text{maximize}} \sum_{t=1}^T f_t(x_t), \qquad \text{subject to } \frac{1}{T}\sum_{t=1}^T h_t(x_t) \le B.$$

DR-Submodular Functions. A differentiable function $f: \mathcal{X} \to \mathbb{R}, \mathcal{X} \subset \mathbb{R}_+^n$, is called DR-submodular if:

$$x \succcurlyeq y \to \nabla f(x) \preccurlyeq \nabla f(y).$$

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

Motivating Application: Online Welfare Maximization with Production Cost

- A supplier has n types of products for allocation to T online customers that may be produced on demand using the budget BT.
- At step $t \in [T]$, an agent (customer) arrives online and the seller has to assign a bundle $x_t \in \mathcal{X} = \{x \in \mathbb{R}_+^n : 0 \preccurlyeq x \preccurlyeq 1\}$ of products to the agent.
- Then, the production cost for this assignment, $h_t(x_t) = x_t^T P_t x_t$, is revealed where P_t is an entry-wise non-negative and positive definite matrix. The agent has an unknown DR-submodular valuation function $f_t(x) = \text{logdet}\,(\text{diag}(x)(L_t - I) + I)$ over the items, where L_t is a positive semidefinite matrix and the DR-submodularity property characterizes the diversity of the assigned bundle.

Algorithm 1

Input: $\mathcal{X}$ is the constraint set, T is the horizon, $V > 0$, $\alpha > 0$ and $K \in \mathbb{N}$.
Output: $\{x_t : 1 \le t \le T\}$.
Initialize $\lambda_1^{(k)} = v_0^{(k)} = x_0^{(k)} = 0\ \forall k \in [K]$.
for $t = 1$ **to** T **do**
  $x_t^{(1)} = 0$.
  for $k = 1$ **to** K **do**
    $v_t^{(k)} = \arg\max_{x \in \mathcal{X}} \left(\langle V \nabla f_{t-1}(x_{t-1}^{(k)}) - \lambda_t^{(k)} \nabla h_{t-1}(v_{t-1}^{(k)}), x\rangle - \alpha \|x - v_{t-1}^{(k)}\|^2\right)$,
    $x_t^{(k+1)} = x_t^{(k)} + \frac{1}{K} v_t^{(k)}$.
  end for
  Set $x_t = x_t^{(K+1)}$ and play x_t.
  Observe the utility function f_t and the constraint function h_t.
  for $k = 1$ **to** K **do**
    $\lambda_{t+1}^{(k)} = [\lambda_t^{(k)} + h_{t-1}(v_{t-1}^{(k)}) + \langle \nabla h_{t-1}(v_{t-1}^{(k)}), v_t^{(k)} - v_{t-1}^{(k)}\rangle - B]_+$.
  end for
end for

Two Online Models

Model 1. For all $t \in [T]$, both f_t and h_t are chosen adversarially.

- Mannor et al. (2009) provided a counterexample showing that obtaining sub-linear regret (against a benchmark action satisfying the budget constraint) and C_T bounds simultaneously is impossible.
- Alternative regret metric: The $(1 - \frac{1}{e})$-regret is defined as:

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

where:

$$x_W^* = \text{argmax}_{x \in \mathcal{X}_W} \sum_{t=1}^T f_t(x),$$
$$\mathcal{X}_W = \{x \in \mathcal{X} : \sum_{\tau=t}^{t+W-1} h_\tau(x) \le WB, 1 \le t \le T - W + 1\}.$$

In other words, the benchmark action is required to satisfy the budget constraint proportionally over any window of length $W \in [1, T]$.

- For the special case that h_t is linear, Sadeghi and Fazel (2020) proposed the OSPHG algorithm which obtains $\mathcal{O}(\sqrt{WT})$ and $\mathcal{O}(W^{1/4}T^{3/4})$ bounds for R_T and C_T respectively.

Algorithm 1 obtains $\mathcal{O}(\sqrt{WT})$ bounds for both R_T and C_T for general convex constraint functions h_t.

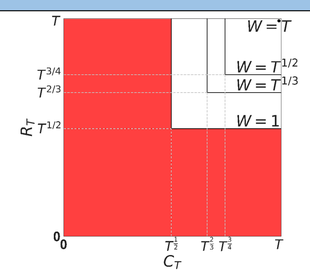

Figure 1. Achievable R_T and C_T bounds using Algorithm 1, for different choices of the window length W. The black curves are the Pareto frontiers for different values of W and all the points north-east of the frontier are achievable.

Model 2. For all $t \in [T]$, f_t is chosen adversarially and h_t is a random i.i.d. sample from an unknown distribution.

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

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

where:

$$x^* = \text{argmax}_{x \in \{x \in \mathcal{X} : \mathbb{E}[h_t(x)] \le B\}} \sum_{t=1}^T f_t(x).$$

- For the special case that h_t is linear, Raut et al. (2020) proposed the OLFW algorithm which obtains $\mathcal{O}(\sqrt{T})$ bounds for C_T, both in expectation and with high probability, and $\mathcal{O}(T^{3/4})$ and $\mathcal{O}(\sqrt{T})$ regret bounds in expectation and with high probability, respectively.

Algorithm 1 obtains $\mathcal{O}(\sqrt{T})$ bounds for both R_T and C_T, in expectation and with high probability, for general convex constraint functions h_t.

Numerical Experiments

Experiment 1: Online joke recommendation using the *Jester* dataset.

$$f_t(x) = r_t^T x + \sum_{i,j:i<j} \theta_{ij}^{(t)} x_i x_j \ \ \forall t \in [T],$$
$$h_t(x) = \langle p_t, x\rangle \ \ \forall t \in [T],$$

where $0 \le [r_t]_i \le 10$ is the rating of user t for joke i in the dataset.

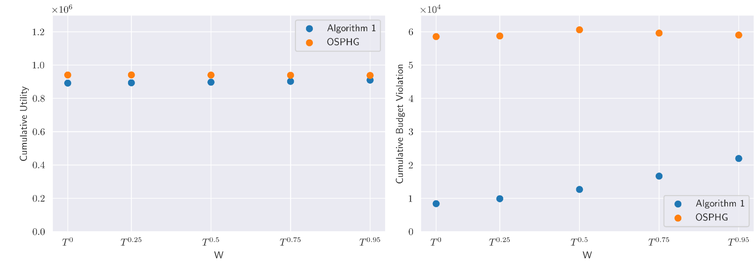

Figure 2. Cumulative utility $\sum_{t=1}^T f_t(x_t)$ and total budget violation C_T for experiment 1.

Experiment 2: Online task assignment in crowdsourcing markets using a synthetic dataset.

$$f_t(x) = \sum_{i=1}^n [u_t]_i \log(1 + x_i) + \sum_{i,j:i<j} \theta_{ij}^{(t)} x_i x_j \ \ \forall t \in [T],$$
$$h_t(x) = \langle p_t, x\rangle \ \ \forall t \in [T].$$

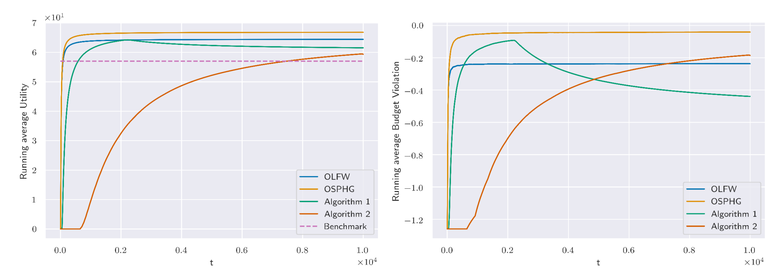

Figure 3. Running average of cumulative utility $\frac{1}{t}\sum_{\tau=1}^t f_t(x_\tau)$ and budget violation $\frac{1}{t}\sum_{\tau=1}^t h_\tau(x_\tau) - B$ for experiment 2.

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 N0014-16-1-2710.

References

Mannor et al. (2009). "Online Learning with Sample Path Constraints". In: Journal of Machine Learning Research, pp. 569-590.

Sadeghi and Fazel (2020). "Online Continuous DR-Submodular Maximization with Long-Term Budget Constraints". In: International Conference on Artificial Intelligence and Statistics, pp. 4410-4419.

Raut el al. (2020). "Online DR-Submodular Maximization with Stochastic Cumulative Constraints". In: arXiv preprint arXiv:2005.14708.

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