Regret Bounds for Online Gradient Descent with $d+1$ point Bandit Feedback

Dhamma Kimpara

December 8, 2020

1 Introduction

Definition 1. Regret:

$$\mathbb{E}\frac{1}{k}\sum_{t=1}^{T}\sum_{i=1}^{k}\ell_t(y_{t,i}) - \mathbb{E}\sum_{t=1}^{T}\min_{x_t^*\in\mathcal{K}}\ell_t(x_t^*)$$

2 Preliminaries

2.1 Problem Setting

First we formally introduce the problem. Our objective is:

$$\min_{x_t\in\mathcal{K}}\ell_t(x_t)$$

over rounds $t = 1, ...T$. Where $\mathcal{K} \subset \mathbb{R}^d$ is our action set and ℓ_t are adversarially chosen time varying loss functions. The key in our setting is that we do not have first order gradient information $\nabla\ell_t$ but we are able to get zeroth-order (bandit) feedback with $k = d + 1$ points. First we outline our assumptions.

We assume that $\mathcal{K}$ is compact and has a nonempty interior (otherwise project $\mathcal{K}$ to a lower dimensional space). For this work, when we implicitly or explicitly refer to norms, we will be using the euclidean norm. We also have the following assumptions for each loss function ℓ_t:

Assumption 1. Let $\mathcal{B}$ denote the unit ball centered at the origin. There exists $r, D > 0$ such that

$$r\mathcal{B} \subseteq \mathcal{K} \subseteq D\mathcal{B}$$

Assumption 2. The gradient of ℓ_t over $\mathcal{K}$ is bounded:

$$\|\nabla\ell_t(x)\| \le G \quad \forall t, \forall x \in \mathcal{K}$$

In other terms, ℓ_t is G-Lipschitz

Assumption 3. Strong Convexity. For $\mu \ge 0$, ℓ_t is μ-strongly convex over the set $\mathcal{K}$:

$$\ell_t(x) \ge \ell_t(y) + \nabla\ell_t(y)^\intercal(x-y) + \frac{\mu}{2}\|x-y\|^2, \quad \forall x, y \in \mathcal{K}$$

Assumption 4. ℓ_t is L-smooth on $\mathcal{K}$ if it is differentiable on an open set containing $\mathcal{K}$ and its gradient is Lipschitz continuous with constant L:

$$\|\nabla\ell_t(x) - \nabla\ell_t(y)\| \le L\|x-y\|, \quad \forall x, y \in \mathcal{K}$$

.

We also use the notation $\mathbb{E}_t$ to denote the conditional expectation conditioned on all randomness in the first $t-1$ rounds.

3 Projected Gradient Descent with k queries per round

We now present that main result, generic for k randomized query feedback.

Lemma 1. Assume that the functions ℓ_t are strongly convex and L-smooth. Assume further that the gradient norms are bounded. Then, the dynamic regret $Regret_T^d$ for the sequence of actions x_t generated by OGD is bounded by

$$Regret_T^d(OGD) \leq GK_1 \sum_{t=2}^{T} \|x_t^* - x_{t-1}^*\| + GK_2$$

where the constants K_1 and K_2 are explicitly given by

$$K_1 := \frac{\|x_1 - x_1^*\| - \rho\|x_T - x_T^*\|}{(1-\rho)}, \quad K_2 := \frac{1}{1-\rho}.$$

Where $0 \leq \rho := (1-\eta\mu)^{1/2} < 0$. Is our linear convergence constant.

Lemma 2. For any point $x \in \mathcal{K}$,

$$\frac{1}{k}\sum_{i=1}^{k} \ell_t(y_{t,i}) - \ell_t(x) \leq \ell_t(x_t) - \ell_t((1-\xi)x) + G\delta + GD\xi.$$

Proof. By assumption of Lipschitz continuity,

$$\ell_t(y_{t,i}) \leq \ell_t(x_t) + G\delta.$$

We also have that by the Lipschitz property and $\|x\| \leq D$, for all $x \in \mathcal{K}$,

$$\ell_t((1-\xi)x) \leq \ell_t(x) + GD\xi.$$

Combining the above two inequalities we get

$$\frac{1}{k}\sum_{i=1}^{k} \ell_t(y_{t,i}) + \ell_t((1-\xi)x) \leq +\ell_t(x_t) + \ell_t(x) + G\delta + GD\xi.$$

Rearranging terms gives us the Lemma.

□

Theorem 1. Assume that the assumptions hold. Suppose on round t the algorithm plays k random queries $y_{t,1}, ..., y_{t,k}$, constructs a gradient estimator $\tilde{g}_t$ and uses the the algorithmic step $x_{t+1} = proj_{(1-\xi)\mathcal{K}}(x_t - \eta\tilde{g}_t)$ with $1/\eta \geq L$, $\delta = \frac{\log(T)}{T}$, and $\xi = \frac{\delta}{r}$. If the gradient estimator satisfies the following conditions for all $t \geq 1$:

1. $\|x_t - y_{t,i}\| \leq \delta$ for $i = 1, ...k$.
2. $\|\tilde{g}_t\| \leq G_1$ for some constant G_1.
3. $\|\mathbb{E}_t\tilde{g}_t - \nabla\ell_t(x_t)\| \leq c\delta$ for some constant c.

Then for any sequence $\{x_t^*\}_{t=1}^T \in \mathcal{K}^T$ we have

$$\mathbb{E}\frac{1}{k}\sum_{t=1}^{T}\sum_{i=1}^{k} \ell_t(y_{t,i}) - \mathbb{E}\sum_{t=1}^{T}\min_{x_t^* \in \mathcal{K}} \ell_t(x_t^*) \leq G_1K_1\sum_{t=2}^{T}\|x_t^* - x_{t-1}^*\| + G_1K_2 + G_1\log(T)\big(1 + 2c + \frac{D}{r}\big).$$

where the constants K_1 and K_2 are explicitly given by

$$K_1 := \frac{\|x_1 - x_1^*\| - \rho\|x_T - x_T^*\|}{(1-\rho)}, \quad K_2 := \frac{1}{1-\rho}.$$

Where $0 \leq \rho := (1-\eta\mu)^{1/2} < 0$. Is our linear convergence constant.

Proof. Start by defining $h_t(x) = \ell_t(x) + (\tilde{g}_t - \nabla \ell_t(x))^\intercal x$. Then h_t has the same convexity properties as ℓ_t. Note also that $\nabla h_t(x_t) = \tilde{g}_t$. So the algorithm is actually performing gradient descent, as if with full information) on the functions h_t restricted to $(1-\xi)\mathcal{K}$. Using the regret bound from Lemma 1 we have that

$$\mathbb{E}\sum_{t=1}^{T}\frac{1}{k}\sum_{i=1}^{k}\ell_t(y_{t,i}) - \mathbb{E}\sum_{t=1}^{T}\ell_t(x_t^*) \leq G_1K_1\sum_{t=2}^{T}\left\|x_t^* - x_{t-1}^*\right\| + G_1K_2 := \mathit{Regret}_T^d(\mathit{OGD}).$$

Then taking expectations,

$$\begin{aligned}
\mathbb{E}\sum_{t=1}^{T}[\ell_t(x_t) - \ell_t(x_t^*)] &= \mathbb{E}\sum_{t=1}^{T}[h_t(x_t) - h_t(x_t^*)] + \mathbb{E}\sum_{t=1}^{T}[\ell_t(x_t) - h_t(x_t) - \ell_t(x_t^*) + h_t(x_t^*)] \\
&\leq \mathit{Regret}_T^d(\mathit{OGD}) + \mathbb{E}\sum_{t=1}^{T}(\mathbb{E}_t\tilde{g}_t - \nabla\ell_t(x_t))^\intercal(x_t - x_t^*) \\
&\leq \mathit{Regret}_T^d(\mathit{OGD}) + 2c\delta DT.
\end{aligned}$$

Where the first inequality is by the convexity of ℓ_t and h_t. Now we use Lemma 2 and obtain

$$\mathbb{E}\frac{1}{k}\sum_{t=1}^{T}\sum_{i=1}^{k}\ell_t(y_{t,i}) - \mathbb{E}\sum_{t=1}^{T}\min_{x_t^*\in\mathcal{K}}\ell_t(x_t^*) \leq \mathit{Regret}_T^d(\mathit{OGD}) + 2c\delta DT + TG\delta + GDT\xi.$$

To finish the proof, plug in the values for δ and ξ.

□

Algorithm 1 Gradient descent with deterministic estimator based on $d+1$ points

Input: Step size η, exploration parameter δ and shrinkage coefficient ξ.
Set $x_1 = 0$
for $t = 1, ..., T$ **do**
  Observe $\ell_t(x_t)$, $\ell_t(x_t + \delta e_i)$ for $i = 1, ..., d$.
  Set $\tilde{g}_t = \frac{1}{\delta}\sum_{i=1}^{d}(\ell_t(x_t + \delta e_i) - \ell_t(x_t))e_i$.
  Update $x_{t+1} = \mathit{proj}_{(1-\xi)\mathcal{K}}(x_t - \eta\tilde{g}_t)$.
end for

4 $d+1$ point feedback

In this section, we show that we can construct a deterministic gradient estimator using $d+1$ point feedback. Thus we obtain a deterministic version of Theorem 1. Hence, the algorithm is no-regret even against completely adaptive adversaries meaning that the adversary can choose the loss ℓ_t after the algorithm plays x_t. Hence we match the full-information bound.

The algorithm constructs the deterministic gradient estimator

$$\tilde{g}_t = \frac{1}{\delta}\sum_{i=1}^{d}(\ell_t(x_t + \delta e_i) - \ell_t(x_t))e_i.$$

Where e_i's are the standard unit basis vectors. We further need only the assumptions on strong convexity and L-smoothness, since they imply a bound on the gradient which we denote G. We can thus derive a bound on the norm of the estimator

$$\begin{aligned}
\|\tilde{g}_t\| &= \left\| \frac{1}{\delta} \sum_{i=1}^{d} (\ell_t(x_t + \delta e_i) - \ell_t(x_t)) e_i \right\| \\
&\leq \frac{d}{\delta} \max_i \|\ell_t(x_t + \delta e_i) - \ell_t(x_t)\| \\
&\leq \frac{d}{\delta} \delta G \\
&= dG.
\end{aligned}$$

And also the divergence of the estimator and the true gradient

$$\begin{aligned}
\|\tilde{g}_t - \nabla \ell_t(x_t)\| &= \sqrt{\frac{1}{\delta} \sum_{i=1}^{d} |(\ell_t(x_t + \delta e_i) - \ell_t(x_t)) e_i - \langle \nabla \ell_t(x_t), e_i \rangle|^2} \\
&\leq \frac{d}{\delta} \max_i \|\ell_t(x_t + \delta e_i) - \ell_t(x_t)\| \\
&\leq \frac{d}{\delta} \delta G \\
&= \frac{\sqrt{d} L \delta}{2}
\end{aligned}$$

Hence we have that the properties of $\tilde{g}_t$ are the deterministic version of the properties of the estimator outline in theorem 1. Hence we have an algorithm that can guarantee no-regret against a completely adaptive adversary.

Theorem 2. Suppose a completely adaptive adversary chooses the sequence of loss functions $\{\ell_t\}_{t=1}^{T}$ subject to the same assumptions as above. If Algorithm 1 is run with the $1/\eta \geq L$, $\delta = \frac{\log(T)}{T}$, and $\xi = \frac{\delta}{r}$, then

$$\sum_{t=1}^{T} \frac{1}{d+1} (\ell_t(x_t) + \sum_{i=1}^{d} \ell_t(x_t + \delta e_i)) - \sum_{t=1}^{T} \ell_t(x_t^*) \leq Regret_T^d(OGD) + G \log(T) \big(1 + \frac{\sqrt{d} L \delta}{2} + \frac{D}{r}\big).$$

Proof. This is a modification of the proof of Theorem 1. Define $h_t(x) = \ell_t(x) + (\tilde{g}_t - \nabla \ell_t(x))^\intercal x$. Then $\|\nabla h_t(x_t)\| \leq dG$. Hence from Lemma 1 for any sequence $\{x_t^*\}_{t=1}^{T}$,

$$\sum_{t=1}^{T} h_t(x_t) - h_t(x_t^*) \leq Regret_T^d.$$

Then we use $\|\tilde{g}_t - \nabla \ell_t(x_t)\| \leq \frac{\sqrt{d} L \delta}{2}$. Apply Lemma 2 and plug in our parameters and this gives us the result.

□

References

Scott Aaronson, Greg Kuperberg, and Christopher Granade. The complexity zoo, 2005.

J. Bartholdi, C. A. Tovey, and M. A. Trick. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6(2):157–165, 1989. ISSN 01761714, 1432217X. URL `http://www.jstor.org/stable/41105913`.

Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D Procaccia. *Handbook of computational social choice*. Cambridge University Press, 2016.

Marquis de Condorcet and MJAN de Caritat. An essay on the application of analysis to the probability of decisions rendered by a plurality of votes. *Classics of social choice*, pages 91–112, 1785.

L. A. Hemachandra. The strong exponential hierarchy collapses. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, page 110–122, New York, NY, USA, 1987. Association for Computing Machinery. ISBN 0897912217. doi: 10.1145/28395.28408. URL `https://doi.org/10.1145/28395.28408`.

Edith Hemaspaandra, Lane A. Hemaspaandra, and Jörg Rothe. Exact analysis of dodgson elections: Lewis carroll's 1876 voting system is complete for parallel access to NP. *CoRR*, cs.CC/9907036, 1999. URL `https://arxiv.org/abs/cs/9907036`.

Lane A. Hemaspaandra and Ryan Williams. An atypical survey of typical-case heuristic algorithms. *CoRR*, abs/1210.8099, 2012. URL `http://arxiv.org/abs/1210.8099`.

Christopher M. Homan and Lane A. Hemaspaandra. Guarantees for the success frequency of an algorithm for finding dodgson-election winners. *CoRR*, abs/cs/0509061, 2005. URL `http://arxiv.org/abs/cs/0509061`.

Köbler, Johannes, Schöning, Uwe, and Wagner, Klaus W. The difference and truth-table hierarchies for np. *RAIRO-Theor. Inf. Appl.*, 21(4):419–435, 1987. doi: 10.1051/ita/1987210404191. URL `https://doi.org/10.1051/ita/1987210404191`.

Thomas Lukasiewicz and Enrico Malizia. On the complexity of mcp-nets. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, AAAI'16, page 558–564. AAAI Press, 2016.

Thomas Lukasiewicz and Enrico Malizia. A novel characterization of the complexity class thetakp based on counting and comparison. *Theoretical Computer Science*, 694: 21 – 33, 2017. ISSN 0304-3975. doi: https://doi.org/10.1016/j.tcs.2017.06.023. URL `http://www.sciencedirect.com/science/article/pii/S0304397517305352`.

Christos H. Papadimitriou and Stathis K. Zachos. Two remarks on the power of counting. In Armin B. Cremers and Hans-Peter Kriegel, editors, *Theoretical Computer Science*, pages 269–275, Berlin, Heidelberg, 1982. Springer Berlin Heidelberg. ISBN 978-3-540-39421-1.

Klaus W Wagner. Bounded query classes. *SIAM Journal on Computing*, 19(5):833–846, 1990.

A Code

A.1 Algorithm

A.2 Data Collection