JMLR: Workshop and Conference Proceedings vol 23 (2012) 6.21–6.24 25th Annual Conference on Learning Theory

Commentary on "Online Optimization with Gradual Variations"

Satyen Kale SCKALE@US.IBM.COM
IBM T.J. Watson Research Center

Editor: Shie Mannor, Nathan Srebro, Robert C. Williamson

1. Introduction

This commentary is about (Chiang et al., 2012b). This paper is the result of a merge between two papers, (Yang et al., 2012) and (Chiang et al., 2012a). Both papers address the same question: is it possible to obtain regret bounds in various online learning settings that depend on some notion of variation in the costs, rather than the number of periods? Both papers give remarkably similar algorithms for this problem, although the analysis techniques are quite different, and obtain very similar results. While (Yang et al., 2012) gives two algorithms obtaining such regret bounds for general online convex optimization (OCO), (Chiang et al., 2012a) gives a unified framework to obtain such regret bounds for three specific cases of OCO: online linear optimization, online learning with experts, and online expconcave optimization.

2. Variation-bounded regret

The basic question of whether it is possible to obtain regret bounds that depend on variation in cost functions was posed by Cesa-Bianchi et al. (2007). It was first solved by the author of this commentary in collaboration with Elad Hazan in a series of papers: (Hazan and Kale, 2010) for online linear optimization and online learning with experts, (Hazan and Kale, 2011) for online linear optimization in the bandit setting, and (Hazan and Kale, 2012) for online expconcave optimization. In these papers, the following notion of variation of cost functions was considered. At time t, assume that the cost function can be parameterized by a vector $\mathbf{v}_t \in \mathbb{R}^n$, and assume for the sake of normalization that $\|\mathbf{v}_t\| \leq 1$ (ths is the usual Euclidean norm). Then, define the squared deviation variation of the cost functions as

$$Q_{\text{SD}} = \sum_{t=1}^{T} \|\mathbf{v}_t - \boldsymbol{\mu}\|^2,$$

where T is the number of rounds in the online learning problem, and $\boldsymbol{\mu} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{v}_t$ is the mean of the parameter vectors. It is easy to see that

$$Q_{\text{SD}} = O(T).$$

We were able to show regret bounds that had the same dependence on Q_{SD} as that on T in the best previously known regret bounds for all the scenarios mentioned above (i.e.

$O(\sqrt{Q_{\mathrm{SD}}})$ regret instead of $O(\sqrt{T})$ for online linear optimization, $O(\log Q_{\mathrm{SD}})$ regret of $O(\log T)$ for online expconcave optimization, etc.). Clearly these bounds are tighter than previously known bounds since in simple situations[1] we can have $Q_{\mathrm{SD}} \ll T$.

Another natural notion of variation of cost functions is squared change variation, defined as

$$Q_{\mathrm{SC}} = \sum_{t=2}^{T} \|\mathbf{v}_t - \mathbf{v}_{t-1}\|^2.$$

It is easy to check that

$$Q_{\mathrm{SC}} = O(Q_{\mathrm{SD}}),$$

and so regret bounds in terms of Q_{SC} instead of Q_{SD} would be tighter. In fact, for simple situations[2] it is possible that $Q_{\mathrm{SC}} \ll Q_{\mathrm{SD}}$, thus leading to much tighter regret bounds.

When we gave talks on our work, a fairly common question asked was whether it was possible to derive similar regret bounds in terms of Q_{SC} instead. It seemed hard to obtain such regret bounds (various known algorithms could be shown to not achieve this kind of regret guarantee). Thus it was of considerable interest if such regret bounds were possible.

The papers under consideration solve this important open problem. While the work of Chiang et al. (2012a) improves the regret bounds in our work for online linear optimization and online expconcave optimization by replacing the dependence on Q_{SD} by the same dependence on Q_{SC}, the work of Yang et al. (2012) tackles the general OCO problem for which no variation-bound on regret was known, primarily because of the lack of a suitable parameter vector. Yang et al. (2012) define the squared change variation as follows:

$$Q_{\mathrm{SC}} = \sum_{t=2}^{T} \max_{\mathbf{x} \in \mathcal{K}} \|\nabla c_t(\mathbf{x}) - \nabla c_{t-1}(\mathbf{x})\|^2,$$

where $\mathcal{K}$ is the convex, compact domain in the OCO problem, and $c_t : \mathcal{K} \to \mathbb{R}$ is the cost function at time t. Thus, variation is measured by the maximum change over points in the domain in the gradients at that point between iterations. Yang et al. (2012) then show that the regret can be bounded as $O(\sqrt{Q_{\mathrm{SC}}})$, via two different algorithms.

3. Technical contributions

The main technical contribution, in my opinion, is the introduction of a new style of regret minimizing algorithm based on mirror prox method (Nemirovski, 2004) and the related mirror descent method (Beck and Teboulle, 2003). Interestingly, both papers rely on using *two* Bregman projections (and two points in the domain) per round (rather than the usual one). It seems crucial for the sake of analysis to have the intermediate step.

For the one setting that is common to the two papers, viz. online linear optimization, the second algorithm of (Yang et al., 2012) is identical to the algorithm in (Chiang et al., 2012a), although the analyses in the two papers seems superficially dissimilar.

1. E.g. if $\mathbf{v}_t = \mathbf{u}$ for all t, where $\mathbf{u}$ is a unit vector, then $Q_{\mathrm{SD}} = 0$.
2. For example, consider the case when $\mathbf{v}_t = \mathbf{u}$ for $t \leq T/2$ and $\mathbf{v}_t = -\mathbf{u}$ for $t > T/2$, where $\mathbf{u}$ is a unit vector. Then $Q_{\mathrm{SC}} = O(1)$ whereas $Q_{\mathrm{SD}} = \Theta(T)$.

4. Future work

The main remaining questions are the following:

- Is it possible to obtain $\tilde{O}(\sqrt{Q_{\mathrm{SC}}})$ [3] regret for bandit online linear optimization, improving on the $\tilde{O}(\sqrt{Q_{\mathrm{SC}}})$ bound of Hazan and Kale (2012)? Yang et al. (2012) do get such bounds for bandit settings where multi-point queries are allowed, but the pure bandit case is still open.

- The bound of Chiang et al. (2012a) for online learning with experts is not directly comparable to the bound of Hazan and Kale (2010). Is it possible to get bounds similar to those obtained by Hazan and Kale (2010), but using squared change rather than squared deviation variation?

- Is this the end of the story? Another natural notion of variation for OCO is simply considering the maximum, over all points in the domain, of the variation (either squared deviation or squared change) of the cost functions at the point. This is analogous to the notion of variation that is appropriate in the online learning with experts problem. Is it possible to obtain *efficient* regret bounds that depend on this notion of variation instead? Inefficiently it is certainly possible by simply discretizing the domain and applying the experts algorithm.

References

Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Oper. Res. Lett.*, 31(3):167–175, 2003.

Nicolò Cesa-Bianchi, Yishay Mansour, and Gilles Stoltz. Improved second-order bounds for prediction with expert advice. *Machine Learning*, 66(2-3):321–352, 2007.

Chao-Kai Chiang, Chia-Jung Lee, and Chi-Jen Lu. Online learning in a gradually evolving world. 2012a.

Chao-Kai Chiang, Tianbao Yang, Chia-Jung Lee, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, and Shenghuo Zhu. Online optimization with gradual variations. In *COLT (to appear)*, 2012b.

Elad Hazan and Satyen Kale. Extracting certainty from uncertainty: regret bounded by variation in costs. *Machine Learning*, 80(2-3):165–188, 2010.

Elad Hazan and Satyen Kale. Better algorithms for benign bandits. *Journal of Machine Learning Research*, 12:1287–1311, 2011.

Elad Hazan and Satyen Kale. An online portfolio selection algorithm with regret logarithmic in price variation. *Mathematical Finance (to appear)*, 2012.

Arkadi Nemirovski. Prox-Method with Rate of Convergence O(1/t) for Variational Inequalities with Lipschitz Continuous Monotone Operators and Smooth Convex-Concave Saddle Point Problems. *SIAM Journal on Optimization*, 15(1):229–251, 2004.

3. The $\tilde{O}(\cdot)$ notation hides dependence on polylogarithmic factors depending on T.

Tianbao Yang, Mehrdad Mahdavi, Rong Jin, and Shenghuo Zhu. Regret bound by variation for online convex optimization. 2012.