

An Online Algorithm for Smoothed Online Convex Optimization

Gautam Goel

Adam Wierman *

ABSTRACT

We consider Online Convex Optimization (OCO) in the setting where the costs are m -strongly convex and the online learner pays a switching cost for changing decisions between rounds. We show that the recently proposed Online Balanced Descent (OBD) algorithm is constant competitive in this setting, with competitive ratio $3 + O(1/m)$, irrespective of the ambient dimension. We demonstrate the generality of our approach by showing that the OBD framework can be used to construct competitive a algorithm for LQR control.

1. INTRODUCTION

In this paper we study the problem of *smoothed online convex optimization* (SOCO), a variant of OCO where the online learner incurs a *switching cost* for changing its actions between rounds. More concretely, the online learner plays a series of rounds $t = 1 \dots T$. In each round, the learner receives a convex loss function f_t , picks a point x_t from a convex action space $\chi \subset \mathbb{R}^d$, and pays a *hitting cost* $f_t(x_t)$ as well as a *switching cost* $c(x_t, x_{t-1})$ which penalizes the learner for changing its action between rounds.

This problem was first introduced in the context of the dynamic management of service capacity in data centers [14], where the switching costs represent the performance and wear-and-tear costs associated with changing server configurations. Since then, SOCO has attracted considerable interest, both theoretical and applied, due to its use in dozens of applications across learning, distributed systems, networking, and control, such as speech animation [11], video streaming [9], management of electric vehicle charging [10], geographical load balancing [13], and multi-timescale control [6]. See [5] for an extensive list of applications.

Unfortunately, despite a large and growing literature, all existing results identifying competitive algorithms for SOCO either (i) place strong restrictions on the action space, (ii), place strong restrictions on the class of loss functions, or (iii) require algorithms to make use of predictions of future cost functions. For example, a series of papers [14], [2] developed competitive algorithms for one-dimensional action spaces.

*This is an extended abstract of [7]. Readers can refer to the full version for more details and proofs. Gautam Goel and Adam Wierman are with the Department of Computing and Mathematical Sciences, California Institute of Technology. Email: {ggoel, adamw}@caltech.edu

Until earlier this year there were no known algorithms that were competitive for SOCO beyond one dimension without requiring the use of predictions. Finally, [5] presented the first algorithm that is constant-competitive beyond one dimension, but the algorithm was shown to be constant competitive only in the case of polyhedral cost functions, a restrictive class that does not include most loss functions used in machine learning. Beyond this result, the most general positive results all assume predictions of future cost functions are available, e.g. [13], [3], [4], [12].

The existing work on SOCO highlights a crucial open question: *Does there exist a competitive algorithm for high-dimensional SOCO problems with cost functions that capture standard losses for online learning problems, e.g., logistic loss or least-squares loss?*

In this paper we answer this question by proving that the recently introduced Online Balanced Descent (OBD) algorithm is constant-competitive for SOCO with strongly convex costs. Additionally, highlighting the importance of the class of strongly convex costs, we show that the OBD framework can be used to construct the first competitive algorithm for LQR control, which was not possible with previous approaches.

Related work. There is a vast literature on OCO; for a recent survey see [8]. OCO with switching costs was first studied in the scalar setting in [14], which used SOCO to model dynamic right-sizing in data centers and gave a 3-competitive algorithm. In subsequent work, [2] improved the competitive ratio to 2, also in the scalar setting. The first constant-competitive algorithm beyond one dimension was given in [5], which introduced the OBD framework and showed that it was competitive for SOCO with polyhedral costs. The results in this paper highlight that OBD is also constant-competitive for strongly convex cost functions, a class that is particularly important for learning and control applications, and is wholly disjoint from the class of polyhedral cost functions when the minimizer of the cost function is zero.

2. SMOOTHED ONLINE CONVEX OPTIMIZATION

An instance of SOCO consists of a convex action set $\chi \subset \mathbb{R}^d$, an initial point $x_0 \in \chi$, a sequence of non-negative convex costs $f_1 \dots f_T : \mathbb{R}^d \rightarrow \mathbb{R}^+$, and a non-negative function $c : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^+$. In each round t , the online learner observes the cost function f_t , picks a point x_t , and pays the sum of the *hitting cost* $f_t(x_t)$ and the *movement* or *switching cost* $c(x_t, x_{t-1})$. The switching cost acts as a regularizer,

penalizing the online learner for changing its decisions between rounds. The goal of the online learner is to minimize its aggregate cost so as to approximate the offline optimal cost:

$$\min_{x_1 \dots x_T \in \chi} \sum_{t=1}^T f_t(x_t) + c(x_t, x_{t-1})$$

More generally, x_t could be matrix-valued and f_t, c could be functions on matrices. Note that we make no restrictions on the sequence of cost functions $f_1 \dots f_T$ other than strong convexity; they could be adversarial, or even adaptively chosen to hurt the online learner.

We emphasize that SOCO differs from OCO in two important ways. Firstly, unlike in OCO, the costs incurred in each round of SOCO depend on the previous choice, coupling the online learner's decisions across rounds. Secondly, the online learner can observe the cost function f_t before picking x_t . This is a standard assumption in the SOCO literature, e.g. in [2], [14], [5] and isolates the complexity of SOCO onto the coupling across timesteps due to the switching costs instead of the uncertainty in the costs.

In this paper, we measure the performance of OBD in terms of its *dynamic regret* and *competitive ratio*. The dynamic regret is defined as

$$\sum_{t=1}^T f_t(x_t) + c(x_t, x_{t-1}) - \left[\sum_{t=1}^T f_t(x_t^*) + c(x_t^*, x_{t-1}^*) \right]$$

Here $x_1 \dots x_T$ are the points picked by the online learner and $x_1^* \dots x_T^*$ are the offline optimal points. We note that this is a more natural performance metric for SOCO than static regret, since the main motivation for SOCO is to understand the effects of switching costs on online learning. In contrast, in the static regret setting the comparator never moves and hence incurs no switching cost, making it a less ideal performance metric for SOCO.

Instead of using an additive metric the competitive ratio uses a multiplicative metric:

$$\frac{\sum_{t=1}^T f_t(x_t) + c(x_t, x_{t-1})}{\sum_{t=1}^T f_t(x_t^*) + c(x_t^*, x_{t-1}^*)}$$

We note that [1] showed that, in general, no online algorithm can have both sublinear static regret and constant competitive ratio.

Much attention has been focused on the setting where the switching cost is a norm: $c(x_t, x_{t-1}) = \|x_t - x_{t-1}\|$, e.g. [14], [2]. Note that in the one-dimensional setting, all ℓ_p norms are identical, making the choice of norm somewhat vacuous. The first algorithm to work beyond the one dimensional setting was proposed in [5], which considered a setting where the switching cost is given by the Euclidean distance and the loss functions are *polyhedral*, meaning that they grow at least linearly as one moves away from the minimizer.

We instead focus on the setting where the cost functions $f_1 \dots f_T$ are m -strongly convex with respect to the Euclidean norm and the switching cost is quadratic:

$$c(x_t, x_{t-1}) = \frac{1}{2} \|x_t - x_{t-1}\|_2^2$$

In Section ??, we show that OBD can be used with many important loss functions, such as the least-squares loss and

Algorithm 1 Online Balanced Descent (OBD)

- 1: **for** $t = 1, \dots, T$ **do**
 - 2: Receive f_t . Let v_t be the minimizer of f_t .
 - 3: Let $x(l) = \Pi_{K_t^l}(x_{t-1})$. Initialize $l = f_t(v_t)$. Here K_t^l is the l -sublevel set of f_t , i.e., $K_t^l = \{x \mid f_t(x) \leq l\}$.
 - 4: Increase l . Stop either when $x(l) = v_t$ or $\frac{1}{2} \|x(l) - x_{t-1}\|_2^2 = \beta l$.
 - 5: Set $x_t = x(l)$.
 - 6: **end for**
-

the ℓ_2 regularized logistic loss, none of which could be handled by previous work.

We assume that the domain χ is all of $\mathcal{R}^d$. Note that this presents no real restriction, since we can always define $f_t(x) = \infty$ for all $x \notin \chi$. The objective becomes

$$\min_{x_1 \dots x_T \in \mathcal{R}^d} \sum_{t=1}^T f_t(x_t) + \frac{1}{2} \|x_t - x_{t-1}\|_2^2 \quad (1)$$

Notation. We use $\|\cdot\|$ to denote the ℓ_2 norm. We often use H_t and M_t to denote the hitting cost $f_t(x_t)$ and the movement cost $\frac{1}{2} \|x_t - x_{t-1}\|_2^2$, respectively. The offline costs H_t^* and M_t^* are defined analogously. We let ALG denote the total cost incurred by OBD across all rounds and define OPT to be the analogous offline cost. We let v_t denote the minimizer of the cost function f_t .

3. A COMPETITIVE ALGORITHM

Our main technical result shows that a recently proposed algorithm, Online Balanced Descent (OBD), is constant competitive for SOCO problems with strongly convex cost functions.

OBD was introduced in [5], where it was analyzed for the class of polyhedral costs. The detailed workings of OBD are summarized in Algorithm 1. The key insight of OBD is to exploit the full geometry of the level sets of the current cost function f_t when choosing the point x_t in such a way as to take switching costs into account.

OBD works by iteratively projecting the previously chosen point onto a level set of the current cost function. The level set K_t picked by OBD is the level set such that the switching cost incurred while traveling from x_{t-1} to K_t is equal to $\beta f_t(x_t)$, where x_t is the projection onto K_t and β is the *balance parameter* which can be tuned to get different performance guarantees. We note that OBD can be efficiently implemented via a binary search over the level sets [5].

We can now state our main result, a bound on the competitive ratio of OBD for strongly convex costs.

Theorem 1. *OBD is competitive for the problem (1) for all $\beta > \frac{4}{m}$. Furthermore, if β is set to be $2 + \frac{10}{m}$, the competitive ratio of OBD is at most $3 + O(1/m)$, irrespective of the ambient dimension.*

We note that [5] proved a bound on the competitive ratio of OBD of the form $3 + O(1/\alpha)$ where α measures the “steepness” of the costs. While this superficially resembles the bound in Theorem 1, we emphasize that the settings are quite different; their work applied to the class of polyhedral cost functions while we focus on strongly convex cost functions. In the case where the cost functions have minimum

value zero these classes are wholly disjoint. We are led to consider strongly convex costs due to the fact many common learning and control problems have loss functions that are strongly convex (e.g., see Section 5). Until this paper, there existed no competitive algorithms for SOCO problems with strongly convex costs.

4. LINEAR QUADRATIC REGULATOR (LQR) CONTROL

We give an application from the controls community. Consider the classical problem of LQR control:

$$\min_{u_1 \dots u_T} \sum_{t=1}^T \frac{1}{2} x_t' Q_t x_t + \frac{1}{2} u_t' R u_t$$

with dynamics given by

$$x_{t+1} = Ax_t + Bu_t + w_t$$

Here u is a control action, x a state variable, and Q_t, R are assumed to be positive definite. Usually, the noise increments w_t are assumed to be i.i.d. Gaussian, and the goal is to design a control policy to minimize the expected cost. Instead of an in-expectation result, we can use OBD to design a controller with a strong pathwise guarantee, with no distributional or boundedness assumptions on the noise. The key observation is that the LQR problem can be rewritten as a SOCO problem after a change of variables. We focus on the setting where $A = I$, i.e. the system is stationary in the absence of noise or control actions.

Corollary 1. *Suppose that $A = I$ and B is invertible, and the matrices Q_t each have their lowest eigenvalue bounded below by $\lambda > 0$. The LQR problem can be rewritten as a SOCO problem, and the competitive ratio of OBD is*

$$3 + O\left(\frac{\lambda_{\max}(R)}{\lambda_{\min}(B)^2 \lambda}\right).$$

Note that $\lambda_{\min}(B)$ can be interpreted as a lower bound on the gain of the control action u ; intuitively, systems with high control gain are easier to regulate, since each control action gets amplified. Similarly, it is intuitive that as $\lambda_{\max}(R)$ decreases the competitive ratio improves, since R controls the cost incurred by using the controller.

5. REFERENCES

- [1] L. Andrew, S. Barman, K. Ligett, M. Lin, A. Meyerson, A. Roytman, and A. Wierman. A tale of two metrics: Simultaneous bounds on competitiveness and regret. In *Conference on Learning Theory*, pages 741–763, 2013.
- [2] N. Bansal, A. Gupta, R. Krishnaswamy, K. Pruhs, K. Schewior, and C. Stein. A 2-competitive algorithm for online convex optimization with switching costs. In *LIPICs-Leibniz International Proceedings in Informatics*, volume 40, 2015.
- [3] N. Chen, A. Agarwal, A. Wierman, S. Barman, and L. L. Andrew. Online convex optimization using predictions. In *ACM SIGMETRICS Performance Evaluation Review*, volume 43, pages 191–204. ACM, 2015.
- [4] N. Chen, J. Comden, Z. Liu, A. Gandhi, and A. Wierman. Using predictions in online optimization: Looking forward with an eye on the past. *ACM SIGMETRICS Performance Evaluation Review*, 44(1):193–206, 2016.
- [5] N. Chen, G. Goel, and A. Wierman. Smoothed online convex optimization in high dimensions via online balanced descent. In *Proceedings of the 31st Conference On Learning Theory*, volume 75 of *Proceedings of Machine Learning Research*, pages 1574–1594. PMLR, 2018.
- [6] G. Goel, N. Chen, and A. Wierman. Thinking fast and slow: Optimization decomposition across timescales. In *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on*, pages 1291–1298. IEEE, 2017.
- [7] G. Goel and A. Wierman. An online algorithm for smoothed regression and lqr control. In K. Chaudhuri and M. Sugiyama, editors, *Proceedings of Machine Learning Research*, volume 89 of *Proceedings of Machine Learning Research*, pages 2504–2513. PMLR, 16–18 Apr 2019.
- [8] E. Hazan et al. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.
- [9] V. Joseph and G. de Veciana. Jointly optimizing multi-user rate adaptation for video transport over wireless systems: Mean-fairness-variability tradeoffs. In *2012 Proceedings IEEE INFOCOM*, pages 567–575, March 2012.
- [10] S.-J. Kim and G. B. Giannakis. Real-time electricity pricing for demand response using online convex optimization. In *Innovative Smart Grid Technologies Conference (ISGT), 2014 IEEE PES*, pages 1–5. IEEE, 2014.
- [11] T. Kim, Y. Yue, S. Taylor, and I. Matthews. A decision tree framework for spatiotemporal sequence prediction. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 577–586. ACM, 2015.
- [12] Y. Li, G. Qu, and N. Li. Using predictions in online optimization with switching costs: A fast algorithm and a fundamental limit. In *2018 Annual American Control Conference (ACC)*, pages 3008–3013. IEEE, 2018.
- [13] M. Lin, Z. Liu, A. Wierman, and L. L. H. Andrew. Online algorithms for geographical load balancing. In *2012 International Green Computing Conference (IGCC)*, pages 1–10, June 2012.
- [14] M. Lin, A. Wierman, L. L. H. Andrew, and E. Thereska. Dynamic right-sizing for power-proportional data centers. In *2011 Proceedings IEEE INFOCOM*, pages 1098–1106, April 2011.