
Adaptive Algorithms for Online Convex Optimization with Long-term Constraints

Rodolphe Jenatton*, Jim C. Huang†, Cedric Archambeau* {JENATTON,HUANGJIM,CEDRICA}@AMAZON.COM
Amazon, *Berlin, Germany, †Seattle, USA

Abstract

We present an adaptive online gradient descent algorithm to solve online convex optimization problems with long-term constraints, which are constraints that need to be satisfied when accumulated over a finite number of rounds T , but can be violated in intermediate rounds. For some user-defined trade-off parameter $\beta \in (0, 1)$, the proposed algorithm achieves cumulative regret bounds of $\mathcal{O}(T^{\max\{\beta, 1-\beta\}})$ and $\mathcal{O}(T^{1-\beta/2})$, respectively for the loss and the constraint violations. Our results hold for convex losses, can handle arbitrary convex constraints and rely on a single computationally efficient algorithm. Our contributions generalize over the best known cumulative regret bounds of Mahdavi et al. (2012a), which are respectively $\mathcal{O}(T^{1/2})$ and $\mathcal{O}(T^{3/4})$ for general convex domains, and respectively $\mathcal{O}(T^{2/3})$ and $\mathcal{O}(T^{2/3})$ when the domain is further restricted to be a polyhedral set. We supplement the analysis with experiments validating the performance of our algorithm in practice.

1. Introduction

Online convex optimization (OCO) plays a key role in machine learning applications, such as adaptive routing in networks (Awerbuch and Kleinberg, 2008) and online display advertising (Agrawal and Devanur, 2015). An OCO problem can be viewed as a sequential, repeated game between a learner and an adversary. In each round t , the learner first plays a vector $\mathbf{x}_t \in \mathcal{X} \subseteq \mathbb{R}^d$, where \mathcal{X} is a convex set corresponding to the set of possible actions. The learner then incurs a loss $f_t(\mathbf{x}_t)$ for playing vector \mathbf{x}_t . The function $f_t : \mathcal{X} \rightarrow \mathbb{R}_+$ is defined by the adversary and can vary in each round. We say that f_t is *strongly convex* with modulus $\sigma > 0$ if $f_t(\mathbf{x}) \leq f_t(\mathbf{y}) + \nabla f_t(\mathbf{x})^\top (\mathbf{x} - \mathbf{y}) - \frac{\sigma}{2} \|\mathbf{x} - \mathbf{y}\|_2^2$

for any $\mathbf{x}, \mathbf{y} \in \mathcal{X}$, where the notation $\nabla f_t(\mathbf{x})$ refers to any (sub-)gradient of f_t at \mathbf{x} . If $\sigma = 0$, we say that f_t is *convex*. The learner’s objective is to generate a sequence of vectors $\mathbf{x}_t \in \mathcal{X}$ for $t = 1, 2, \dots, T$ that minimizes the cumulative regret over T rounds relative to the optimal vector \mathbf{x}^* :

$$\text{Regret}_T(\mathbf{x}^*) \triangleq \sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x}^*). \quad (1)$$

This regret measures the difference between the cumulative loss of the learner’s sequence of vectors $\{\mathbf{x}_t\}_{t=1}^T$ and the accumulated loss that would be incurred if the sequence of loss functions f_t would be known in advance and the learner could choose the best vector \mathbf{x}^* in hindsight. Several algorithms have been developed over the past decade that achieve sub-linear cumulative regret in the OCO setting. The problem was formalized by Zinkevich (2003), who introduced an online algorithm based on projected subgradients (Bertsekas and Tsitsiklis, 1989). The algorithm guarantees a cumulative regret of $\mathcal{O}(T^{1/2})$ when the set \mathcal{X} is convex and the loss functions are Lipschitz-continuous over \mathcal{X} . Hazan et al. (2007) and Shalev-Shwartz and Kakade (2009) introduced algorithms with logarithmic regret bounds for strongly convex loss functions, e.g., online gradient descent has an $\mathcal{O}(\log(T))$ regret bound for appropriate choices of the step size.

In the aforementioned works, the constraint on vector \mathbf{x}_t is assumed to hold for each round t , such that a projection step is applied every round to enforce the feasibility of each \mathbf{x}_t . For general convex sets \mathcal{X} , the projection step may require solving an auxiliary optimization problem, which can be computationally expensive (e.g., projections onto the semi-definite cone). More importantly, in practical applications, the learner may in fact only be concerned with satisfying long-term constraints, that is, the *cumulative* constraint violations resulting from the sequence of vectors $\{\mathbf{x}_t\}_{t=1}^T$ should not exceed a certain amount by the final round T . An example of such an application is in online display advertising, where \mathbf{x}_t is a vector of ad budget allocations and the learner is primarily concerned in enforcing the long-term constraint that each ad fully consumes its budget over the lifetime of the ad. Another example is from wireless

communications (Mannor and Tsitsiklis, 2006), where \mathbf{x}_t is a vector of power allocations across multiple devices, and the learner must satisfy average power consumption constraints per device over several rounds.

In this work, we consider OCO problems where the learner’s cumulative regret is defined by (1) and it is required to satisfy long-term constraints—the notion of long-term constraint is made formal in Section 2. This class of problems was studied previously in (Mahdavi et al., 2012a;b). In particular, Mahdavi et al. (2012b) restricted their study to online exponentially weighted averages with linear loss and linear constraints, while Mahdavi et al. (2012a) presented online algorithms based on projected subgradients and the mirror-prox method (Nemirovski, 2004). The authors derived cumulative regret bounds for the cumulative loss and cumulative constraint violations, respectively of $\mathcal{O}(T^{1/2})$ and $\mathcal{O}(T^{3/4})$ in the case of online projected subgradients, and respectively of $\mathcal{O}(T^{2/3})$ and $\mathcal{O}(T^{2/3})$ in the case of mirror-prox. To our knowledge, these are the best-known regret bounds for OCO with long-term constraints. The analysis of Mahdavi et al. (2012a) relies on two different algorithms to achieve the two aforementioned guarantees, and the mirror-prox method additionally requires the constraint set \mathcal{X} to be polyhedral. Some convex domains of interest, such as spectral constraints or norm balls, are difficult or cannot be represented as the intersection of a finite number of linear constraints.

The concept of long-term constraints enables us to avoid the computation of potentially expensive projections onto the domain \mathcal{X} in each round. This is related in spirit to recent work in stochastic optimization, where one aims to minimize the number of expensive projection steps (Mahdavi et al., 2012c), or online approaches based on Frank-Wolfe (Hazan and Kale, 2012). Our work only assumes access to the function values and (sub)gradients, and does not make further assumptions about other computational oracles (e.g., maximization of linear functions for Frank-Wolfe, or access to some proximal operator). The guarantees sought in our analysis have also similarities with those obtained for the online alternating direction method (Wang and Banerjee, 2013), where regret bounds are provided for the violation of equality constraints. Finally, long-term constraints bear resemblance with recent work on bandits under global constraints (Badanidiyuru et al., 2013).

Contributions. We propose an algorithm based on a saddle-point formulation of the OCO problem with long-term constraints. The resulting online algorithm is *adaptive* in that the step sizes are different for the primal/dual variables and depend on the round t , as does the regularisation parameter. As a result, we obtain an algorithm with a faster practical convergence as shown in the experiments.

In the case of convex losses, we show that, for any convex

constraint set \mathcal{X} and some user-defined trade-off parameter $\beta \in (0, 1)$, our algorithm achieves regret bounds of $\mathcal{O}(T^{\max\{\beta, 1-\beta\}})$ and $\mathcal{O}(T^{1-\beta/2})$ for the cumulative loss and the cumulative constraint violations. Hence, we recover the $\mathcal{O}(T^{1/2})$ and $\mathcal{O}(T^{3/4})$ guarantees from Mahdavi et al. (2012a) when equating β to $1/2$, but we enable the user to trade-off cumulative loss for cumulative constraint violations depending on his/her application. Moreover, we extend the regret bounds of $\mathcal{O}(T^{2/3})$ achieved by Mahdavi et al. (2012a) beyond polyhedral sets, proposing an arguably less involved approach than mirror-prox. While Mahdavi et al. (2012a) consider only the convex setting, we also study the case of strongly convex losses for which we prove tighter regret bounds, but without improving their leading terms. All the bounds we obtain are valid for adversarially-generated sequences, in particular, we do not assume the losses are generated in an i.i.d. fashion.

Finally, we supplement our theoretical results with an empirical study and compare our algorithm to the ones proposed by Mahdavi et al. (2012a). We consider i) the online estimation of doubly stochastic matrices and ii) the online learning of sparse logistic regression with the elastic net penalty (Zou and Hastie, 2005). The empirical validation fills a gap in previous work (Mahdavi et al., 2012a;b) and uncovers the unexpected results that the practical benefit of having adaptive regularization and step sizes leads to algorithms with faster vanishing cumulative regrets, even when not fully captured by the theory.

2. Problem statement

Consider m convex functions $g_j : \mathbb{R}^d \mapsto \mathbb{R}$ which induce a convex constraint set

$$\mathcal{X} \triangleq \left\{ \mathbf{x} \in \mathbb{R}^d : \max_{j \in \{1, \dots, m\}} g_j(\mathbf{x}) \leq 0 \right\}.$$

We assume that the set \mathcal{X} is bounded so that it is included in some Euclidean ball \mathcal{B} with radius $R > 0$

$$\mathcal{X} \subseteq \mathcal{B} \triangleq \left\{ \mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 \leq R \right\}.$$

Along with the functions g_j , we consider a sequence of convex functions $f_t : \mathbb{R}^d \mapsto \mathbb{R}_+$ such that

$$F \triangleq \max_{t \in \{1, \dots, T\}} \max_{\mathbf{x}, \mathbf{x}' \in \mathcal{B}} |f_t(\mathbf{x}) - f_t(\mathbf{x}')| > 0.$$

As is typically assumed in online learning (Cesa-Bianchi and Lugosi, 2006), the functions g_j and f_t shall be taken to be Lipschitz continuous. In particular, for some finite $G > 0$ the (sub-)gradients of f and g_j are bounded

$$\max_{j \in \{1, \dots, m\}} \max_{\mathbf{x} \in \mathcal{B}} \|\nabla g_j(\mathbf{x})\|_2 \leq G, \quad \max_{t \in \{1, \dots, T\}} \max_{\mathbf{x} \in \mathcal{B}} \|\nabla f_t(\mathbf{x})\|_2 \leq G.$$

We take the same constant G for both g_j and f_t for simplicity as we can always take the maximum between that

of g_j and that of f_t . We notably have $F \leq 2RG$. Also, we do not generally assume g_j and f_t to be differentiable. Finally, we assume that there exists a finite $D > 0$ such that the constraint functions are bounded over \mathcal{B} :

$$\max_{j \in \{1, \dots, m\}} \max_{\mathbf{x} \in \mathcal{B}} |g_j(\mathbf{x})| \leq D.$$

The set of assumptions enumerated above match the ones in [Mahdavi et al. \(2012a\)](#).

2.1. Online Optimization with Long-term Constraints

Let $\{\mathbf{x}_t\}_{t=1}^T$ be the sequence of vectors played by the learner from the set \mathcal{B} and $\{f_t(\mathbf{x}_t)\}_{t=1}^T$ the corresponding sequence of incurred losses. We aim at minimizing the cumulative regret subject to *long-term constraint*, that is:

$$\sum_{t=1}^T f_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^T f_t(\mathbf{x}) \quad \text{s.t.} \quad \max_{j \in \{1, \dots, m\}} \sum_{t=1}^T g_j(\mathbf{x}_t) \leq 0.$$

3. Adaptive Online Algorithms based on a Saddle-point Formulation

Following [Mahdavi et al. \(2012a\)](#), we consider a saddle-point formulation of the optimization problem. For any $\lambda \in \mathbb{R}^+$, $\mathbf{x} \in \mathcal{B}$ we define the following function:

$$\mathcal{L}_t(\mathbf{x}, \lambda) \triangleq f_t(\mathbf{x}) + \lambda g(\mathbf{x}) - \frac{\theta_t}{2} \lambda^2, \quad (2)$$

where $g(\mathbf{x}) \triangleq \max_{j \in \{1, \dots, m\}} g_j(\mathbf{x})$ and $\{\theta_t\}_{t=1}^T$ is a sequence of positive numbers to be specified later. The role of g is to aggregate the m constraints into a single function. It otherwise preserves the same properties as those of individual g_j 's (sub-differentiability, bounded (sub-)gradients and bounded values; see Proposition 6 in ([Mahdavi et al., 2012a](#)) or Section 2.3 in ([Borwein and Lewis, 2006](#))).

In the saddle-point formulation (2), we will alternate between minimizing with respect to the primal variable \mathbf{x} and maximizing with respect to the dual parameter λ . A closer look at the function $\lambda \mapsto \mathcal{L}_t(\mathbf{x}, \lambda)$ indicates that we penalize the violation of the constraint $g(\mathbf{x}) \leq 0$:

$$\frac{1}{2\theta_t} [g(\mathbf{x})]_+^2 = \sup_{\lambda \in \mathbb{R}^+} \left[\lambda g(\mathbf{x}) - \frac{\theta_t}{2} \lambda^2 \right], \quad (3)$$

where $[u]_+ \triangleq \max\{0, u\}$. This penalty is commonly used when transforming constrained to unconstrained problems (see e.g., [Nocedal and Wright \(2006\)](#), Section 17.1). Also, we can see from (3) that θ_t acts as a regularization parameter. We note that [Mahdavi et al. \(2012a\)](#) make use of a single θ that is constant in all rounds, equal to the product of a constant step size times a constant scaling factor.

In the sequel, we study the following online algorithm where we alternate between primal and dual updates:

- Initialize $\mathbf{x}_1 = \mathbf{0}$ and $\lambda_1 = 0$.

- For $t \in \{1, \dots, T-1\}$:

$$\mathbf{x}_{t+1} = \Pi_{\mathcal{B}}(\mathbf{x}_t - \eta_t \nabla_{\mathbf{x}} \mathcal{L}_t(\mathbf{x}_t, \lambda_t)),$$

$$\lambda_{t+1} = \Pi_{\mathbb{R}^+}(\lambda_t + \mu_t \nabla_{\lambda} \mathcal{L}_t(\mathbf{x}_t, \lambda_t)),$$

where $\Pi_{\mathcal{C}}$ stands for the Euclidean projection onto the set \mathcal{C} , while $\{\eta_t\}_{t=1}^T$ and $\{\mu_t\}_{t=1}^T$ are sequences of non-negative step sizes that respectively drive the update of \mathbf{x} and λ . The algorithm resembles the ones proposed by [Mahdavi et al. \(2012a\)](#); [Koppel et al. \(2014\)](#), but it is adaptive. The step sizes, which are different for the updates of \mathbf{x} and λ , are listed in Table 1 and result from the analysis we provide in the next section. We also derive sub-linear regret bounds associated to these instantiations of the sequences $\{\theta_t\}_{t=1}^T$, $\{\eta_t\}_{t=1}^T$ and $\{\mu_t\}_{t=1}^T$.

3.1. Main Results

We begin by listing three sufficient conditions for obtaining sub-linear regret bounds for the proposed algorithm:

(C1): For any $t \geq 2$, $\frac{1}{\mu_t} - \frac{1}{\mu_{t-1}} - \theta_t \leq 0$.

(C2): For any $t \geq 2$, $\eta_t G^2 + \mu_t \theta_t^2 - \frac{1}{2} \theta_t \leq 0$.

(C3): For some finite $U_\eta > 0$, $\sum_{t=2}^T \left[\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} - \sigma \right] \leq U_\eta$.

Conditions **C1** and **C3** impose constraints on the rate of decrease of the step sizes. We note that there is an asymmetry between μ_t and η_t : while we will always be able to control the norm of the variables \mathbf{x}_t (by design, they must lie in \mathcal{B}), the sequence $\{\lambda_t\}_{t=1}^T$ is not directly upper-bounded in the absence of further assumptions on the gradient of g , hence the most stringent condition **C1** is to avoid any dependencies on λ_t . Condition **C2** couples the behaviour of the three sequences to guarantee their validity. Finally, **C1**, **C2** and **C3** are expressed for $t \geq 2$ because of our choice for the initial conditions $\mathbf{x}_1 = \mathbf{0}$ and $\lambda_1 = 0$.

Our main result is described in the following theorem, whose proof is detailed in Section 3.2:

Theorem 1. Consider the choices of the sequences μ_t, η_t and θ_t for some $\beta \in (0, 1)$, as summarized in Table 1. Let $\mathbf{x}^* \in \arg \min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^T f_t(\mathbf{x})$. For any $T \geq 1$, it holds that

$$\sum_{t=1}^T \Delta f_t \leq \mathcal{R}_T^f \quad \text{and} \quad \sum_{t=1}^T g(\mathbf{x}_t) \leq \sqrt{\frac{C}{1-\beta}} \left(\mathcal{R}_T^f + FT \right) T^{1-\beta}$$

where $\Delta f_t \triangleq f_t(\mathbf{x}_t) - f_t(\mathbf{x}^*)$. The term \mathcal{R}_T^f is defined as

$$\mathcal{R}_T^f \stackrel{(\sigma=0)}{=} \left[RG + \frac{D^2}{6\beta RG} \right] T^\beta + \frac{2RG}{1-\beta} T^{1-\beta} \quad \text{with } C \triangleq 24RG,$$

$$\mathcal{R}_T^f \stackrel{(\sigma>0)}{=} \frac{G^2}{\sigma} (1 + \log(T)) + \frac{D^2 \sigma}{6G^2 \beta} T^\beta \quad \text{with } C \triangleq \frac{24G^2}{\sigma},$$

for respectively the cases $\sigma = 0$ and $\sigma > 0$.

	Convex ($\sigma = 0$)	Strongly convex ($\sigma > 0$)
θ_t	$\frac{6RG}{t^\beta}$	$\frac{6G^2}{\sigma t^\beta}$
η_t	$\frac{R}{G t^\beta}$	$\frac{\sigma}{t}$
μ_t	$\frac{\theta_t(t+1)}{R}$	$\frac{\theta_t(t+1)}{\sigma}$
S_θ	$\frac{6RG}{1-\beta} T^{1-\beta}$	$\frac{6G^2}{\sigma(1-\beta)} T^{1-\beta}$
S_η	$\frac{R}{G(1-\beta)} T^{1-\beta}$	$\frac{1}{\sigma}(1 + \log(T))$
S_μ	$\frac{1}{6\beta RG} T^\beta$	$\frac{\sigma}{6\beta G^2} T^\beta$
U_η	$\frac{G}{R} T^\beta$	0
$\frac{1}{\mu_1} - \theta_1$	$6RG$	$\frac{6G^2}{\sigma}$
$\frac{1}{\eta_1} - \sigma$	$\frac{G}{R}$	0

Table 1. Parameter choices in different regimes, with $\beta \in (0, 1)$.

Theorem 1 can be further simplified in the convex case (i.e., $\sigma = 0$). Forgetting momentarily about the dependencies on $\{D, G, R, F\}$, it can be stated as follows:

$$\sum_{t=1}^T \Delta f_t \leq \mathcal{O}(\max\{T^\beta, T^{1-\beta}\}),$$

$$\sum_{t=1}^T g(\mathbf{x}_t) \leq \mathcal{O}(T^{1-\beta/2}).$$

Setting $\beta = 2/3$ leads to $\sum_{t=1}^T \Delta f_t \leq \mathcal{O}(T^{2/3})$ and $\sum_{t=1}^T g(\mathbf{x}_t) \leq \mathcal{O}(T^{2/3})$, which matches the mirror-prox guarantees of Mahdavi et al. (2012a). However, it is valid for general convex constraint sets \mathcal{X} as opposed to just polyhedral constraint sets. Taking $\beta = 1/2$, we recover the regret bounds $\sum_{t=1}^T \Delta f_t \leq \mathcal{O}(T^{1/2})$ and $\sum_{t=1}^T g(\mathbf{x}_t) \leq \mathcal{O}(T^{3/4})$ achieved in Section 3.1 of (Mahdavi et al., 2012a). Other trade-offs between loss and constraint violations can be selected, e.g., $\beta = 3/4$ with regret bounds of $\mathcal{O}(T^{3/4})$ and $\mathcal{O}(T^{5/8})$ respectively.

Theorem 1 shows cumulative regret bounds in the strongly convex case (i.e., $\sigma > 0$). The bounds have the same leading terms, but are tighter. Condition C2 couples the loss and constraint regrets, suggesting that a logarithmic regret for the loss would result in a non-vanishing regret for the constraint. This non-trivial property is related to how fast θ_t is allowed to decrease. Intuitively, there is a tension between (a) making progress in \mathbf{x} with large η_t 's, and (b) adaptively controlling the constraint violation via $1/\theta_t$. Having $1/\theta_t$ to be too large risks impeding the progress in \mathbf{x} ; setting them too small leads to large constraint violations. As for μ_t 's, making these larger penalizes the violations more. This intuitive trade-off is formalized via C1-2-3.

Finally, conditions C1-2-3 make it possible to instantiate the sequences $\{\theta_t, \eta_t, \mu_t\}$ without prior knowledge of the time horizon T . In contrast, Mahdavi et al. (2012a) use sequences depending on T , requiring to resort to the doubling trick (e.g., Section 2.3.1 in Shalev-Shwartz (2011)).

3.2. Analysis and Proofs

The analysis is analogous to Mahdavi et al. (2012a), and we provide it for self-containedness. We first introduce a series of lemmas and close the section by proving Theorem 1.

We begin by upper-bounding the variations of \mathcal{L}_t with respect to its two arguments. In particular, the following lemma takes advantage of the fact that the partial function $\lambda \mapsto \mathcal{L}_t(\mathbf{x}_t, \lambda)$ is not only concave as considered in (Mahdavi et al., 2012a), but *strongly* concave with parameter θ_t . This observation, together with separate step sizes for \mathbf{x} and λ forms the basis of our improved regret bounds.

Lemma 1 (Upper bound of $\mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}_t, \lambda_t)$). *Consider $b_t \triangleq (\lambda - \lambda_t)^2$. For $\mathcal{L}_t(\mathbf{x}, \lambda)$ as defined in (2), and non-negative η_t, θ_t, μ_t , the term $\mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}_t, \lambda_t)$ is upper bounded by*

$$\frac{1}{2\mu_t} [b_t - b_{t+1}] - \frac{\theta_t}{2} b_t + \frac{\mu_t}{2} [\nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t)]^2.$$

Proof. The argument can be found in (Hazan et al., 2007). We expand $b_{t+1} = (\lambda - \lambda_{t+1})^2$ into

$$\begin{aligned} &= \left(\lambda - \Pi_{\mathbb{R}^+}(\lambda_t + \mu_t \nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t)) \right)^2 \\ &\leq \left(\lambda - (\lambda_t + \mu_t \nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t)) \right)^2 \\ &= b_t - 2\mu_t(\lambda - \lambda_t) \nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t) + \mu_t^2 (\nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t))^2. \end{aligned}$$

By strong concavity of $\mathcal{L}_t(\mathbf{x}_t, \lambda)$ with respect to λ ,

$$\mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}_t, \lambda_t) \leq (\lambda - \lambda_t) \nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t) - \frac{\theta_t}{2} b_t.$$

Substituting the inequality for $\mu_t(\lambda - \lambda_t) \nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t)$ completes the proof. \square

We omit the proof for $\mathbf{x} \mapsto \mathcal{L}_t(\mathbf{x}, \lambda_t)$ that follows similar arguments, leading to a bound on $\mathcal{L}_t(\mathbf{x}_t, \lambda_t) - \mathcal{L}_t(\mathbf{x}, \lambda_t)$. We now turn to a lower-bound of the variations of \mathcal{L}_t .

Lemma 2. *Let $\mathbf{x}^* \in \arg \min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^T f_t(\mathbf{x})$. The term $\sum_{t=1}^T \mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}^*, \lambda_t)$ is lower bounded by*

$$\sum_{t=1}^T \Delta f_t + \lambda \sum_{t=1}^T g(\mathbf{x}_t) - \frac{\lambda^2}{2} \sum_{t=1}^T \theta_t + \frac{1}{2} \sum_{t=1}^T \theta_t \lambda_t^2.$$

Proof. We have $\mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}^*, \lambda_t)$ equal to

$$f_t(\mathbf{x}_t) - f_t(\mathbf{x}^*) + \lambda g(\mathbf{x}_t) - \lambda_t g(\mathbf{x}^*) - \frac{\theta_t}{2} (\lambda^2 - \lambda_t^2).$$

We simply notice that $g(\mathbf{x}^*) \leq 0$ to obtain a lower bound $-g(\mathbf{x}^*) \sum_{t=1}^T \lambda_t$ to complete the proof after summing both sides over rounds $t = 1, \dots, T$. \square

Lemma 3. Assume **C1** and **C3** hold. Let $a_t \triangleq \|\mathbf{x}_t - \mathbf{x}\|^2$. For any $\sigma, \theta_t \geq 0$, we have

$$\sum_{t=1}^T \frac{1}{2\eta_t} [a_t - a_{t+1}] - \frac{\sigma}{2} a_t \leq \frac{R^2}{2} \delta_\eta + \frac{1}{2} \sum_{t=2}^T a_t \left[\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} - \sigma \right] \leq \frac{R^2}{2} \delta_\eta + R^2 U_\eta,$$

and

$$\sum_{t=1}^T \frac{1}{2\mu_t} [b_t - b_{t+1}] - \frac{\theta_t}{2} b_t \leq \frac{\lambda^2}{2} \delta_\mu + \frac{1}{2} \sum_{t=2}^T b_t \left[\frac{1}{\mu_t} - \frac{1}{\mu_{t-1}} - \theta_t \right] \leq \frac{\lambda^2}{2} \delta_\mu,$$

where we have used $\delta_\eta \triangleq \frac{1}{\eta_1} - \sigma$ and $\delta_\mu \triangleq \frac{1}{\mu_1} - \theta_1$.

Proof. Shifting indices in the sums for terms that depend on $a_{t+1}/\eta_t, b_{t+1}/\mu_t$ and collecting terms that depend on a_t, b_t , we then use $a_1 = \|\mathbf{x}_1 - \mathbf{x}\|_2^2 = \|\mathbf{x}\|_2^2 \leq R^2$, $a_t \leq 2R^2$ for $t > 1$ and $b_1 = (\lambda - \lambda_1)^2 = \lambda^2$ to conclude. \square

We now present the key lemma of the analysis.

Lemma 4. [Cumulative regret bound] Let $\mathbf{x}^* \in \arg \min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^T f_t(\mathbf{x})$ and assume **C1**, **C2** and **C3** hold. Define $\mathcal{R}_T^f \triangleq \frac{R^2}{2} \delta_\eta + G^2 S_\eta + D^2 S_\mu + R^2 U_\eta$, where we have introduced $S_\eta \triangleq \sum_{t=1}^T \eta_t$, $S_\mu \triangleq \sum_{t=1}^T \mu_t$ and $S_\theta \triangleq \sum_{t=1}^T \theta_t$. Then, it holds that

$$\sum_{t=1}^T \Delta f_t \leq \mathcal{R}_T^f, \quad \text{and} \\ \sum_{t=1}^T g(\mathbf{x}_t) \leq \sqrt{2(S_\theta + \delta_\mu)(\mathcal{R}_T^f + FT)}.$$

Proof. Using the triangle inequality, we have $\|\nabla_{\mathbf{x}} \mathcal{L}_t(\mathbf{x}, \lambda_t)\|_2^2 \leq 2G^2(1 + \lambda_t^2)$ and $(\nabla_{\lambda} \mathcal{L}_t(\mathbf{x}, \lambda_t))^2 \leq 2(D^2 + \theta_t^2 \lambda_t^2)$. Starting from $\mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}^*, \lambda_t) = \mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}_t, \lambda_t) + \mathcal{L}_t(\mathbf{x}_t, \lambda_t) - \mathcal{L}_t(\mathbf{x}^*, \lambda_t)$, we sum over $t = 1 \dots T$, and combine Lemmas 1-2 so as to upper bound $\sum_{t=1}^T [\mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}^*, \lambda_t)]$ by

$$\mathcal{R}_T^f + \sum_{t=1}^T \lambda_t^2 \left[\eta_t G^2 + \mu_t \theta_t^2 \right] + \frac{\lambda^2}{2} \delta_\mu.$$

Lemma 3 along with the previous equation leads to

$$\sum_{t=1}^T \Delta f_t + \lambda \sum_{t=1}^T g(\mathbf{x}_t) - \frac{\lambda^2}{2} \left[S_\theta + \delta_\mu \right] \leq \mathcal{R}_T^f + \sum_{t=1}^T \lambda_t^2 \left[\eta_t G^2 + \mu_t \theta_t^2 - \frac{1}{2} \theta_t \right] \triangleq \mathcal{R}_T^f + \mathcal{M}_T.$$

Maximizing the left-hand side with respect to $\lambda \geq 0$, we get

$$\sum_{t=1}^T \Delta f_t + \left[\sum_{t=1}^T g(\mathbf{x}_t) \right]_+^2 / (2[S_\theta + \delta_\mu]) \leq \mathcal{R}_T^f + \mathcal{M}_T,$$

where we have used $[s]_+^2 / 2\alpha = \max_{v \geq 0} \{sv - \alpha v^2 / 2\}$ for $\alpha > 0$. The regret bound on the loss is obtained by using

C2 and $[\sum_{t=1}^T g(\mathbf{x}_t)]_+^2 / (2[S_\theta + \delta_\mu]) \geq 0$. The bound on constraint violations is obtained as above, but by substituting the lower bound $\sum_t \Delta f_t \geq -FT$. \square

In order to discuss the scaling of our regret bounds, we state the next simple lemma without proof:

Lemma 5. Let $\beta \in (0, 1)$. Then $\sum_{t=1}^T \frac{1}{t^\beta} \leq \frac{T^{1-\beta}}{1-\beta}$.

With the above lemmas, we now prove Theorem 1:

Proof of Theorem 1. For the proposed choices of θ_t, μ_t and η_t , we can verify that **C1**, **C2** and **C3** hold. Here we focus on the convex case (the strongly convex one follows along the same lines). First, we can easily see that **C1** is true as long as θ_t is non-increasing. Then, we can notice that, given the choice of μ_t , condition **C2** is implied by the stronger condition $\eta_t \leq \frac{\theta_t}{6G^2}$ (satisfied by the choice of η_t and θ_t in Table 1). This results in

$$S_\mu = \sum_{t=1}^T \frac{1}{\theta_t(t+1)} \leq \sum_{t=1}^T \frac{t^\beta}{6RGt} \leq \frac{T^\beta}{6\beta RG}. \\ S_\eta = \sum_{t=1}^T \eta_t \leq \frac{R}{G} \frac{T^{1-\beta}}{1-\beta}, \quad S_\theta = \sum_{t=1}^T \frac{6RG}{t^\beta} \leq \frac{6RG}{1-\beta} T^{1-\beta},$$

along with $1/\mu_1 - \theta_1 = 6RG$ and $1/\eta_1 - \sigma = G/R$. The term U_η can be obtained by summing the series $1/\eta_t - 1/\eta_{t-1} = (G/R)(t^\beta - (t-1)^\beta)$ over t , which directly simplifies by telescoping for the $\sigma = 0$ case, and is identically equal to zero for $\sigma > 0$. As a result, we obtain from Lemma 4 for both the cost and constraint

$$\sum_{t=1}^T \Delta f_t \leq \mathcal{R}_T^f \triangleq \left[RG + \frac{D^2}{6\beta RG} \right] T^\beta + \frac{RG}{1-\beta} T^{1-\beta} + \frac{RG}{2} \\ \sum_{t=1}^T g(\mathbf{x}_t) \leq \sqrt{2(\mathcal{R}_T^f + FT) \left[\frac{6RG}{1-\beta} T^{1-\beta} + 6RG \right]}.$$

We obtain the desired conclusion by noticing that for any $T \geq 1$ and $\beta \in (0, 1)$, we have $\frac{T^{1-\beta}}{1-\beta} \geq 1$. \square

3.3. Towards No Violation of Constraints

Next, we show that our results apply to the specific case considered in Section 3.2 of (Mahdavi et al., 2012a), where additional assumptions on the gradient of g can translate into the absence of constraint violations. Assume that there exist $\gamma \geq 0$ and $r > 0$ such that the variations of g are lower bounded as

$$\min_{\mathbf{x} \in \mathbb{R}^d: g(\mathbf{x}) + \gamma = 0} \|\nabla g(\mathbf{x})\|_2 \geq r. \quad (4)$$

Let us denote $\mathcal{X}_\gamma \triangleq \{\mathbf{x} \in \mathbb{R}^d : g(\mathbf{x}) + \gamma \leq 0\} \subseteq \mathcal{X}$. Assumption (4) is better understood in light of the optimality conditions of the offline problem: λ can be shown to be

inversely proportional to the norm of the gradient defined in (4), hence lower bounding this norm makes it possible to upper bound λ . In turn, the gap between the optimal value of the original optimization problem and that of the problem over \mathcal{X}_γ can be well-controlled as a function of (γ, r) , as proved by Theorem 7 in Mahdavi et al. (2012a)

$$\left| \sum_{t=1}^T f_t(\mathbf{x}^*) - \sum_{t=1}^T f_t(\mathbf{x}_\gamma) \right| \leq \frac{G}{r} T \gamma, \quad (5)$$

where \mathbf{x}^* and \mathbf{x}_γ are solutions of $\min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^T f_t(\mathbf{x})$ and $\min_{\mathbf{x} \in \mathcal{X}_\gamma} \sum_{t=1}^T f_t(\mathbf{x})$ respectively. Examples where (4) holds include the positive semi-definite cone, as described in Section 4 of Mahdavi et al. (2012c). For lack of space, we state our result in a simplified form and only provide a sketch of proof along the same lines as the proof in Section 3.2.

Corollary 1. *Assume (4) holds. Consider the convex case ($\sigma = 0$) and some instantiations of the sequences μ_t, η_t and θ_t for some $\beta \in (0, 1)$, differing from Table 1 up to constants. There exist c_0 and c_1 depending on $\{D, G, R, F, r\}$ such that setting $\gamma \triangleq c_1 T^{-\beta/2}$, we have for any $T \geq c_0$*

$$\sum_{t=1}^T \Delta f_t \leq \mathcal{O}(\max\{T^\beta, T^{1-\beta}, T^{1-\beta/2}\}),$$

and no constraint violations $\sum_{t=1}^T g(\mathbf{x}_t) \leq 0$.

Sketch of proof. We can apply the same analysis as before to the function $g_\gamma(\mathbf{x}) \triangleq g(\mathbf{x}) + \gamma$, replacing D by $D + \gamma$ and adapting the constants in both **C2** (i.e., $\eta_t G^2 + \frac{3}{2} \mu_t \theta_t^2 - \frac{1}{2} \theta_t$) as well as for the instantiations of μ_t, η_t and θ_t . The regret bound on $\sum_{t=1}^T \Delta f_t$ is identical as earlier, with additional additive terms $3\gamma^2 S_\mu/2$ and $GT\gamma/r$ introduced as a result of (5). As for $\sum_{t=1}^T g(\mathbf{x}_t)$, the term $[\sum_{t=1}^T g(\mathbf{x}_t)]_+^2$ becomes here $[\sum_{t=1}^T g(\mathbf{x}_t) + \gamma T]_+^2$, which in turn leads to the same regret bound as previously stated, minus the contribution $-\gamma T$. We cancel out the constraint violations—scaling in $\mathcal{O}(T^{1-\beta/2})$ according to Theorem 1—by choosing $\gamma = c_1 T^{-\beta/2}$. Note that c_0 is set by studying when the extra term $3\gamma^2 S_\mu/2$ is upper bounded by those in \mathcal{R}_T^f . \square

The regret bound presented in Corollary 1 is minimized for $\beta = 2/3$, leading to a regret of $\mathcal{O}(T^{2/3})$ with no constraint violations. This result extends Theorem 8 and Corollary 13 from (Mahdavi et al., 2012a) in that it holds for general convex domains \mathcal{X} (as opposed to only polyhedral ones).

4. Experiments

We ran two sets of experiments to assess the performance for our adaptive algorithms for OCO with long-term constraints and compare to the algorithms proposed by Mahdavi et al. (2012a). First, we examine the problem of online estimation of doubly-stochastic matrices where the convex

domain of interest \mathcal{X} is polyhedral but whose projection operator is difficult to compute (Helmbold and Warmuth, 2009; Fogel et al., 2013). Second, we consider the problem of sparse online binary classification based on the elastic net penalty (Zou and Hastie, 2005). This is a standard benchmark problem for large-scale online learning, where the constraint set is defined over a non-polyhedral domain.

We shall refer to our adaptive online gradient descent for convex f_t (i.e., $\sigma = 0$) as `Convex A-OGD` and for strongly convex f_t (i.e., $\sigma > 0$) as `Strongly convex A-OGD`, which enjoy the same regret guarantees of $\mathcal{O}(T^{2/3})$ for the loss and constraint for $\beta = 2/3$. The method of Mahdavi et al. (2012a, see Algorithm 1), which can handle general convex domains \mathcal{X} , will be referred to as `Convex OGD`. The mirror prox method of Mahdavi et al. (2012a, see Algorithm 2), which is only applicable to polyhedral domains, will be referred to as `Convex mirror prox`. The parameters of `Convex OGD` and `Convex mirror prox` are instantiated according to (Mahdavi et al., 2012a). The sequence of losses $\{f_t\}_{t=1}^T$ in the experiments are generated stochastically.

4.1. Doubly-Stochastic Matrices

Doubly-stochastic (DS) matrices appear in many machine learning and optimization problems, such as clustering applications (Zass and Shashua, 2006) or learning permutations (Helmbold and Warmuth, 2009; Fogel et al., 2013). Consider a sequence of random permutation matrices $\{\mathbf{Y}_t\}_{t=1}^T$ in $\mathbb{R}^{p \times p}$. Since permutation matrices are known to constitute the extreme points of the set of DS matrices (Birkhoff, 1946), we try to find the closest DS matrix \mathbf{X} by solving the following optimization problem online:

$$\min_{\mathbf{X} \in \mathbb{R}^{p \times p}} \sum_{t=1}^T \frac{1}{2} \|\mathbf{Y}_t - \mathbf{X}\|_F^2 \quad (6)$$

subject to the (doubly-stochastic) linear constraints $\mathbf{X}\mathbf{1} = \mathbf{1}$, $\mathbf{X}^\top \mathbf{1} = \mathbf{1}$ and the element-wise inequality $\mathbf{X} \geq \mathbf{0}$. We have $d = p^2$, $f_t(\mathbf{X}) = \frac{1}{2} \|\mathbf{Y}_t - \mathbf{X}\|_F^2$ and $m = p^2 + 4p$ to describe all the linear constraints. More specifically there are p^2 non-negativity constraints, along with $4p$ inequalities to model the $2p$ equality constraints. This leads to the following instantiations of the parameters controlling f_t and g : $R = \sqrt{p}$, $G = 2R$ and $D = R$. Note that we can apply a) `Strongly convex A-OGD` (since f_t is by construction strongly convex with parameter $\sigma = 1$), and b) `Convex mirror prox` since \mathcal{X} is polyhedral. The cumulative regret for the loss and the long-term constraint are shown in Figures 1 and 2. They are computed over $T = 1000$ iterations with $d = 64$, and are averaged over 10 random sequences $\{\mathbf{Y}_t\}_{t=1}^T$. The standard deviations are not shown as we found experimentally that they were negligible. The offline solutions of (6) required for various

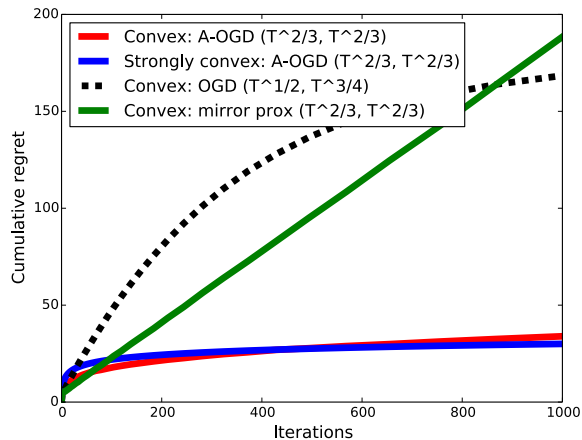


Figure 1. Cumulative regret of the loss function for the estimation of doubly-stochastic matrices. We display the mean computed over 10 random sequences $\{\mathbf{Y}_t\}_{t=1}^T$. (Best viewed in colour.)

$t \in \{1, \dots, T\}$ to compute the regret are obtained using CVXPY (Diamond et al., 2014).

The results shown in Figure 1 and 2 indicate that although the cumulative regret bounds for Strongly convex A-OGD were not demonstrated to be tighter in our analysis than those for Convex A-OGD, they achieve a better cumulative regret for this problem, especially with respect to the long-term constraint. Also, while Convex mirror prox and Convex A-OGD should theoretically exhibit the same behavior, the results suggest that mirror prox is not able to decrease cumulative regret at the same rate as our proposed method. We surmise that this is due to the fact that the guarantees for Convex mirror prox only hold for very large T : Theorem 12 in (Mahdavi et al., 2012a) requires $T \geq 164(m+1)^2$, which translates here into $T > 10^7$.

4.2. Sparse Online Binary Classification

Next, we examine the application of sparse online binary classification. Our goal is to minimize the log-loss subject to a constrained elastic-net penalty:

$$\min_{\mathbf{x} \in \mathbb{R}^d: \|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{x}\|_2^2 \leq \rho} \sum_{t=1}^T \log(1 + e^{-y_t \mathbf{x}^\top \mathbf{u}_t}), \quad (7)$$

where $\{y_t, \mathbf{u}_t\}_{t=1}^T$ denotes a sequence of label/feature-vector pairs and $\rho > 0$ is a parameter that measures the degree of the sparsity of the solutions of (7). While a penalized formulation would be applicable, constrained formulations with sparsity-inducing terms are sometimes preferred in practice as they express a concrete physical budget (e.g., Xu et al. (2012) follow this route in the context of learning predictors with low-latency).

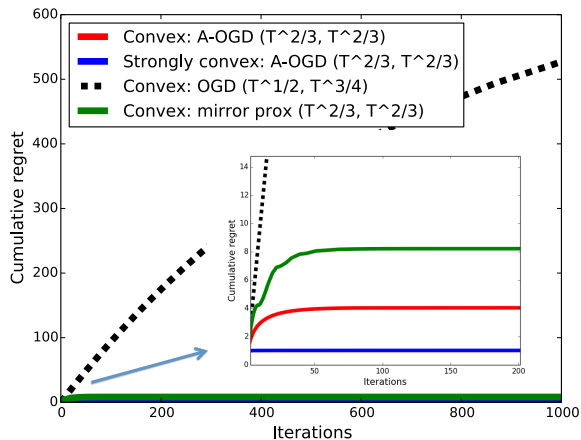


Figure 2. Cumulative regret of the long-term constraint for the estimation of doubly-stochastic matrices. We display the mean computed over 10 random sequences $\{\mathbf{Y}_t\}_{t=1}^T$. The embedded graph is a zoom of the area of interest in the original figure.

We cast the problem in our framework by setting $f_t(\mathbf{x}) = \log(1 + e^{-y_t \mathbf{x}^\top \mathbf{u}_t})$ with $m = 1$, $g(\mathbf{x}) = \|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{x}\|_2^2 - \rho$, $R = \sqrt{1 + 2\rho} - 1$, $G = \max\{\sqrt{d} + R, \max_t \|\mathbf{u}_t\|_2\}$ and $D = \sqrt{d}R + R^2/2$.¹ The sequences $\{y_t, \mathbf{u}_t\}_{t=1}^T$ are generated by drawing pairs at random with replacement.

We solve the above problem using the datasets `ijcnn1` and `covtype`, consisting respectively of 49,990 and 581,012 samples of dimension $d = 22$ and $d = 54$ each.² The parameter ρ is set to obtain approximately 30% of non-zero variables. Moreover, and in order to best display cumulative regret, we compute offline solutions of (7) based on an implementation of Defazio et al. (2014).

The results are reported in Figures 3 and 4 and represent an average over 10 random sequences $\{y_t, \mathbf{u}_t\}_{t=1}^T$. Again, we do not show standard deviations because they were found to be negligible. The number of iterations T is equal to the number of samples in each dataset.

Interestingly, we observe that the constraint is not violated on average (i.e., via a negative cumulative regret) and the iterates \mathbf{x}_t remain feasible within the domain $\|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{x}\|_2^2 \leq \rho$. This tendency is more pronounced for Convex OGD since a closer inspection of the sequence $\{\eta_t\}_{t=1}^T$ shows numerical values smaller than those of our approach Convex A-OGD (by 2 to 3 orders of magnitude). As a result, starting from $\mathbf{x}_1 = \mathbf{0}$, we found that the iterates generated by Convex OGD do not approach the boundary of the domain, hence increasing regret on

¹The value for R is found by noticing that $\|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{x}\|_2^2 \geq \|\mathbf{x}\|_2 + \frac{1}{2}\|\mathbf{x}\|_2^2$ and solving the resulting second-order polynomial inequality.

²www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/binary.html.

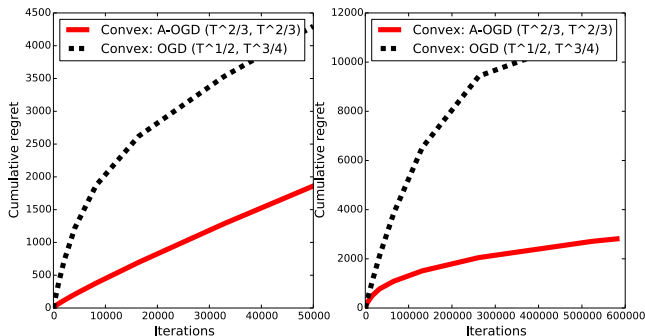


Figure 3. Cumulative regret of the loss function for the sparse online binary classification (left: `ijcnn1`, right: `covtype`). We show the mean computed over 10 random sequences $\{y_t, \mathbf{u}_t\}_{t=1}^T$.

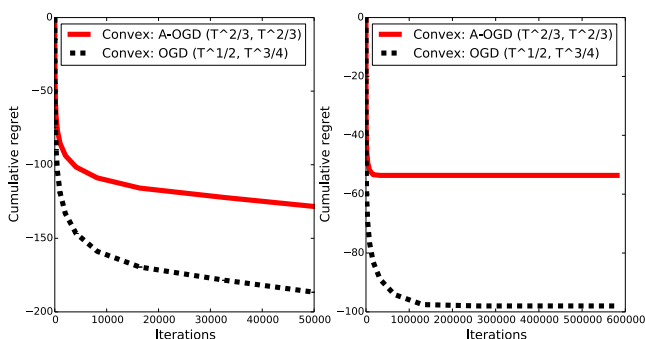


Figure 4. Cumulative regret of the constraint for the sparse online binary classification (left: `ijcnn1`, right: `covtype`). We display the mean computed over 10 random sequences $\{y_t, \mathbf{u}_t\}_{t=1}^T$.

cumulative loss. We also note that the offline solutions of (7) always saturate the constraint. Although our analysis predicts that the cumulative regret of `Convex OGD` associated to the loss (i.e., $\mathcal{O}(T^{1/2})$) should be smaller than that associated to `Convex A-OGD` (i.e., $\mathcal{O}(T^{2/3})$), `Convex A-OGD` achieves a lower cumulative regret. This observation may be explained by the same argument as the one invoked previously, namely that the larger step sizes $\{\eta_t\}_{t=1}^T$ of `Convex A-OGD` enables us to make faster progress.

5. Discussion

We conclude by discussing several generalizations.

Broader families of step sizes: We have assumed that the updates of the primal variable \mathbf{x} are driven by a projected subgradient method step controlled through a single step size η_t . Following (McMahan and Streeter, 2010; Duchi et al., 2011), we could analyze the regret guarantees of our algorithm when there is a diagonal matrix of step sizes, such that each coordinate of \mathbf{x} is updated adaptively. This has for example been proven useful in practice when some

features are less frequently activated than others.

Strong convexity: Condition **C2** couples the loss and constraint regrets, suggesting that a logarithmic regret for the loss would result in a non-vanishing regret for the constraint. We conjecture that logarithmic regrets are unlikely to be achieved, although this conjecture depends on both the saddle-point formulation of the problem and the use of a projected subgradient method. This does not preclude the possibility of achieving logarithmic regret under different and stronger assumptions.

Can we identify a better penalty? In the light of (3), it is tempting to ask if we can find a penalty that would lead to lower cumulative regret guarantees. To this end, we could for example introduce a smooth, 1-strongly-convex function ϕ with domain Ω . The saddle-point formulation of the new problem is then given by

$$\mathcal{L}_t(\mathbf{x}, \lambda) \triangleq f_t(\mathbf{x}) + \lambda g(\mathbf{x}) - \theta_t \phi(\lambda),$$

where $\{\theta_t\}_{t=1}^T$ is, as earlier, a sequence of non-negative numbers to be specified subsequently for any $\lambda \in \Omega, \mathbf{x} \in \mathcal{B}$. Interestingly, it can be shown that condition **C2** becomes a first-order nonlinear ordinary differential inequality in this setting, leading to

$$\eta_t G^2 \lambda^2 + \mu_t \theta_t^2 \left[\frac{d\phi}{d\lambda} \right]^2 - \theta_t \phi(\lambda) \leq 0, \text{ for all } \lambda \in \Omega.$$

Hence, the above differential inequality suggests a family of penalty functions we could use. In particular, we see that ϕ must grow at least quadratically and stay greater than its squared first derivative, which rules out a softmax penalty like $\lambda \mapsto \log(1 + e^\lambda)$. Moreover, the maximization with respect to λ in the last step of Lemma 4 introduces the Moreau envelope (Lemaréchal and Sagastizábal, 1997) of the Fenchel conjugate of ϕ , namely

$$\phi_{S_\theta}^*(u) \triangleq \sup_{\lambda \in \Omega} \left[\lambda u - S_\theta \phi(\lambda) - \delta_\mu \frac{\lambda^2}{2} \right].$$

We can then find a feasible penalty ϕ of which the inverse mapping $u \mapsto [\phi_{S_\theta}^*]^{-1}(u)$ would minimize the regret bound. For instance, the inverse mapping scales as \sqrt{u} when using the squared ℓ_2 norm over $\Omega = \mathbb{R}^+$. We defer to future work the study of this admissible family of penalties.

Which enclosing set for \mathcal{X} ? Our current analysis relies on the idea that instead of having to perform a projection on \mathcal{X} in each update (which could be computationally costly and perhaps intractable in some cases), we restrict the iterates \mathbf{x}_t to remain within a simpler convex set $\mathcal{B} \supseteq \mathcal{X}$. While we assumed *de facto* an Euclidean ball for \mathcal{B} , we could consider sets enclosing \mathcal{X} more tightly, while preserving the appealing computational properties. Having a principled methodology to choose \mathcal{B} and assessing its impact on the regret bounds is an interesting avenue for future research.

References

- Agrawal, S. and Devanur, N. R. (2015). Fast algorithms for online stochastic convex programming. In *SODA 2015 (ACM-SIAM Symposium on Discrete Algorithms)*. SIAM-Society for Industrial and Applied Mathematics.
- Awerbuch, B. and Kleinberg, R. (2008). Online linear optimization and adaptive routing. *Journal of Computer and System Sciences*, 74(1):97–114.
- Badanidiyuru, A., Kleinberg, R., and Slivkins, A. (2013). Bandits with knapsacks. In *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*, pages 207–216.
- Bertsekas, D. P. and Tsitsiklis, J. N. (1989). *Parallel and Distributed Computation: Numerical Methods*. Prentice-Hall, Inc.
- Birkhoff, G. (1946). Tres observaciones sobre el algebra lineal. *Univ. Nac. Tucumán Rev. Ser. A*, 5:147–151.
- Borwein, J. M. and Lewis, A. S. (2006). *Convex Analysis and Nonlinear Optimization: Theory and Examples*. Springer.
- Cesa-Bianchi, N. and Lugosi, G. (2006). *Prediction, learning, and games*. Cambridge University Press.
- Defazio, A., Bach, F., and Lacoste-Julien, S. (2014). Saga: A fast incremental gradient method with support for non-strongly convex composite objectives. Technical report, preprint arXiv:1407.0202.
- Diamond, S., Chu, E., and Boyd, S. (2014). CVXPY: A Python-embedded modeling language for convex optimization, version 0.2. <http://cvxpy.org/>.
- Duchi, J., Hazan, E., and Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12:2121–2159.
- Fogel, F., Jenatton, R., Bach, F., and d’Aspremont, A. (2013). Convex relaxations for permutation problems. In *Advances in Neural Information Processing Systems*, pages 1016–1024.
- Hazan, E., Agarwal, A., and Kale, S. (2007). Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192.
- Hazan, E. and Kale, S. (2012). Projection-free online learning. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 521–528.
- Helmbold, D. P. and Warmuth, M. K. (2009). Learning permutations with exponential weights. *Journal of Machine Learning Research*, 10:1705–1736.
- Koppel, A., Jakubiec, F. Y., and Ribeiro, A. (2014). A saddle point algorithm for networked online convex optimization. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pages 8292–8296. IEEE.
- Lemaréchal, C. and Sagastizábal, C. (1997). Practical aspects of the moreau–yosida regularization: Theoretical preliminaries. *SIAM Journal on Optimization*, 7(2):367–385.
- Mahdavi, M., Jin, R., and Yang, T. (2012a). Trading regret for efficiency: online convex optimization with long term constraints. *Journal of Machine Learning Research*, 13(1):2503–2528.
- Mahdavi, M., Yang, T., and Jin, R. (2012b). Efficient constrained regret minimization. Technical report, preprint arXiv:1205.2265.
- Mahdavi, M., Yang, T., Jin, R., Zhu, S., and Yi, J. (2012c). Stochastic gradient descent with only one projection. In *Advances in Neural Information Processing Systems*, pages 494–502.
- Mannor, S. and Tsitsiklis, J. N. (2006). Online learning with constraints. In *Learning Theory*, pages 529–543. Springer.
- McMahan, H. B. and Streeter, M. (2010). Adaptive bound optimization for online convex optimization. In *Proceedings of the annual conference on Computational Learning Theory (COLT)*, page 244.
- Nemirovski, A. (2004). 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.
- Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. Springer.
- Shalev-Shwartz, S. (2011). Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194.
- Shalev-Shwartz, S. and Kakade, S. M. (2009). Mind the duality gap: Logarithmic regret algorithms for online optimization. In *Advances in Neural Information Processing Systems*, pages 1457–1464.
- Wang, H. and Banerjee, A. (2013). Online alternating direction method. Technical report, preprint arXiv:1306.3721.
- Xu, Z., Weinberger, K., and Chapelle, O. (2012). The greedy miser: Learning under test-time budgets. Technical report, preprint arXiv:1206.6451.

- Zass, R. and Shashua, A. (2006). Doubly stochastic normalization for spectral clustering. In *Advances in Neural Information Processing Systems*, pages 1569–1576.
- Zinkevich, M. (2003). Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the International Conference on Machine Learning (ICML)*.
- Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B*, 67(2):301–320.