The Hedge Algorithm on a Continuum

Walid Krichene WALID@EECS.BERKELEY.EDU
University of California, 652 Sutardja Dai Hall, Berkeley, CA 94720 USA

Maximilian Balandat BALANDAT@EECS.BERKELEY.EDU
University of California, 736 Sutardja Dai Hall, Berkeley, CA 94720 USA

Claire Tomlin TOMLIN@EECS.BERKELEY.EDU
University of California, 721 Sutardja Dai Hall, Berkeley, CA 94720 USA

Alexandre Bayen BAYEN@BERKELEY.EDU
University of California, 642 Sutardja Dai Hall, Berkeley, CA 94720 USA

Abstract

We consider an online optimization problem on a compact subset $S \subset \mathbb{R}^n$ (not necessarily convex), in which a decision maker chooses, at each iteration t, a probability distribution $x^{(t)}$ over S, and seeks to minimize a cumulative expected loss, $\sum_{t=1}^{T} \mathbb{E}_{s\sim x^{(t)}}[\ell^{(t)}(s)]$, where $\ell^{(t)}$ is a Lipschitz loss function revealed at the end of iteration t. Building on previous work, we propose a generalized Hedge algorithm and show a $O(\sqrt{t \log t})$ bound on the regret when the losses are uniformly Lipschitz and S is uniformly fat (a weaker condition than convexity). Finally, we propose a generalization to the dual averaging method on the set of Lebesgue-continuous distributions over S.

1. Introduction

We consider the online optimization setting used by Zinkevich (2003) and Hazan et al. (2007), where a decision maker chooses, at each iteration t, a probability distribution $x^{(t)}$ over some compact feasible set $S \subset \mathbb{R}^n$, and incurs a loss $\mathbb{E}_{s\sim x^{(t)}}[\ell^{(t)}(s)]$. When choosing $x^{(t)}$, the decision maker only has access to $(\ell^{(\tau)}(\cdot))_{1\leq\tau\leq t-1}$, i.e. the loss functions up to iteration $t-1$. The cumulative regret is a natural measure of performance in sequential decision problems; it was introduced by Hannan (1957) in the context of repeated games, then later used in the analysis of general sequential decision problems, see for example Cesa-Bianchi & Lugosi (2006) and Bubeck & Cesa-Bianchi (2012). The cumulative regret $R^{(t)}$ at iteration t is defined as the difference between the loss incurred by the decision maker and the loss of the best fixed decision in hindsight, that is,

$$R^{(t)} = \sum_{\tau=1}^{t} \mathbb{E}_{s\sim x^{(t)}}[\ell^{(\tau)}(s)] - \inf_{s\in S} \sum_{\tau=1}^{t} \ell^{(\tau)}(s).$$

Zinkevich (2003) shows that if the feasible set S is convex, and the loss functions $\ell^{(t)}$ are convex with a bounded gradient, a simple online gradient descent algorithm achieves a $\mathcal{O}(\sqrt{t})$ cumulative regret. In 2007, Hazan et al. show that if the loss functions $\ell^{(t)}$ are exp-concave, uniformly in t, a generalized Hedge algorithm achieves a $\mathcal{O}(\log t)$ regret. The Hedge algorithm, also known as the multiplicative weights update (Arora et al., 2012), has been extensively studied in the discrete case, i.e. when S is a discrete set. The Hedge algorithm was introduced as the exponentially weighted average forecaster by Littlestone & Warmuth (1989). It has also been analyzed in the context of convex optimization, and is known as the exponentiated gradient method (Kivinen & Warmuth, 1997), or the entropic descent method (Beck & Teboulle, 2003). The Hedge algorithm is a simple method to implement and to analyze, and achieves sublinear regret in the discrete case whenever the loss functions $\ell^{(t)}$ are uniformly bounded. More precisely, if the action set has size N and the the learning rates have a $1/\sqrt{t}$ decay rate, then the regret is bounded by $\mathcal{O}(\sqrt{t \log N})$, see for example the analysis in (Bubeck & Cesa-Bianchi, 2012).

We seek to generalize this regret bound to a setting in which the action set is a continuum, while making only mild assumptions on the geometry of the set and the class of loss functions. The logarithmic regret bound achieved by Hazan et al. (2007) requires the feasible set to be convex, and the loss functions to be exp-concave. We extend their analy-

Proceedings of the 32^{nd} International Conference on Machine Learning, Lille, France, 2015. JMLR: W&CP volume 37.

Assumptions on $\ell^{(t)}$	convex	α-exp-concave	uniformly L-Lipschitz
Assumptions on S	convex	convex	v-uniformly fat
Method	Gradient descent (Zinkevich)	Hedge (Hazan et al.)	Hedge (this paper)
Learning rates	$1/\sqrt{t}$	α	$1/\sqrt{t}$
$R^{(t)}/t$	$\mathcal{O}(1/\sqrt{t})$	$\mathcal{O}(t^{-1}\log t)$	$\mathcal{O}(\sqrt{t^{-1}\log t})$

Table 1. Some regret upper bounds for different classes of losses.

sis to a less restrictive class of problems, which only requires uniform fatness of the action set (a weaker condition than convexity) and uniform Lipschitz continuity of the loss functions. We show that under such assumptions the Hedge algorithm achieves a $\mathcal{O}(\sqrt{t\log t})$ regret. Table 1 summarizes the regret bounds for these problem classes.

The online optimization model on a continuum has various applications, including machine learning (Hazan et al., 2007), portfolio optimization (Cover, 1991; Blum & Kalai, 1999), pricing with uncertain demand and transmission power control over noisy channels (Cope, 2009). By relaxing the assumptions of convexity of the feasible set and exp-concavity of the loss functions, we extend the class of problems for which the Hedge algorithm provides bounds on the worst-case regret. For example, in the context of portfolio optimization, this would allow for non-convex diversification constraints and non-convex transaction costs (Xidonas & Mavrotas, 2014).

In Section 2, we derive a general regret bound for Lipschitz losses. In Section 3, we specialize this bound to convex feasible sets, then relax the convexity assumption and show that the Hedge algorithm guarantees sublinear regret on uniformly fat sets. In Section 4, we study the dual averaging method and prove a regret bound on the set of Lebesgue-continuous distributions on S, then discuss how one can recover the Hedge regret bound as a special case. In Section 5, we compare the Hedge algorithm to learning on a finite cover of S. Finally, we illustrate these theoretical results with a numerical example in Section 6.

2. A General Regret Bound on Metric Spaces

Consider a compact metric space (S, d) where d is a distance function. Let ν be a reference probability measure on S, and denote by $\Delta_\nu(S)$ the set of probability measures that are absolutely continuous with respect to ν.

Let $\ell^{(t)} \in C^0(S, \mathbb{R}_+)$ denote the loss function at iteration t. We assume that the losses are bounded uniformly in t, i.e. $\exists\, M > 0$ such that $\ell^{(t)}(s) \in [0, M]$ for all $t \in \mathbb{N}$ and all $s \in S$. The decision maker chooses, at iteration t, a distribution over S, i.e. an element of $\Delta_\nu(S)$. Its density w.r.t. ν will be denoted by $x^{(t)}$. The Hedge algorithm with initial density $x^{(0)}$ and learning rates (η_t), is defined by the sequence $(x^{(t)})$ of densities as follows:

$$x^{(t+1)}(s) = \frac{1}{\bar{Z}^{(t)}} x^{(0)}(s) \exp\left(-\eta_{t+1} \textstyle\sum_{\tau=1}^{t} \ell^{(\tau)}(s)\right) \quad (1)$$

where $\bar{Z}^{(t)}$ is the appropriate normalization constant, i.e., $\bar{Z}^{(t)} = \mathbb{E}_{s\sim x^{(0)}}\left[\exp(-\eta_{t+1}\sum_{\tau=1}^{t}\ell^{(\tau)}(s))\right]$. The Hedge algorithm is summarized in Algorithm 1.

Algorithm 1 Hedge algorithm with initial density $x^{(0)}$ and learning rates (η_t).

for $t \in \mathbb{N}$ **do**
  Choose action $s \sim x^{(t)}$
  Observe loss function $\ell^{(t)}$
  Update $x^{(t+1)}(s) \propto x^{(0)}(s) \exp\left(-\eta_{t+1} \sum_{\tau=1}^{t} \ell^{(\tau)}(s)\right)$

Define $r^{(t)}(s) = \mathbb{E}_{u\sim x^{(t)}}[\ell^{(t)}(u)] - \ell^{(t)}(s)$ as the instantaneous regret function at iteration t, and

$$\begin{aligned} R^{(t)} &= \sup_{s\in S} \textstyle\sum_{\tau=1}^{t} r^{(\tau)}(s) \\ &= \textstyle\sum_{\tau=1}^{t} \mathbb{E}_{s\sim x^{(\tau)}}[\ell^{(\tau)}(s)] - \inf_{s\in S} \sum_{\tau=1}^{t} \ell^{(\tau)}(s) \end{aligned} \quad (2)$$

as the cumulative regret. The first term in the above expression is the expected cumulative loss of the decision maker. The second term is the infimum of the cumulative loss function, which will be denoted by

$$L^{(t)}(s) := \textstyle\sum_{\tau=1}^{t} \ell^{(\tau)}(s)$$

Since S is compact and $L^{(t)}$ is continuous, the infimum of $L^{(t)}$ is attained on S. We write $s_t^\star \in \arg\min_{s\in S} L^{(t)}(s)$ for the minimizer.

We start by giving a first regret bound for Lipschitz-continuous losses. This bound can be obtained as a consequence of Theorem 4.6 of (Audibert, 2009), and a similar result is proved by (Dalalyan & Salmon, 2012).

Lemma 1. *The Hedge algorithm with non-increasing learning rates (η_t) guarantees*

$$R^{(t)} \leq \frac{M^2}{8} \sum_{\tau=1}^{t} \eta_\tau + \left(\xi(\eta_t, L^{(t)}) - L^{(t)}(s_t^\star)\right) \quad (3)$$

where $\xi : \mathbb{R}_+ \times L^2(S) \to \mathbb{R}$ is given by

$$\xi(\eta, L) = -\frac{1}{\eta} \log \int x^{(0)}(s) \exp(-\eta L(s)) \nu(ds)$$

$$\sum_{\tau=1}^{t}\mathbb{E}_{x^{(\tau)}}[\ell^{(\tau)}] \leq \frac{M^2}{8}\sum_{\tau=1}^{t}\eta_\tau + \sum_{\tau=1}^{t-1}\Big(\xi(\eta_\tau, L^{(\tau)}) - \xi(\eta_{\tau+1}, L^{(\tau)})\Big) + \xi(\eta_t, L^{(t)}) \tag{4}$$

$$\begin{aligned}
\frac{\partial \xi(\eta, f)}{\partial \eta} &= \frac{1}{\eta^2}\log\int x^{(0)}(s)\exp(-\eta f(s))\,\nu(ds) - \frac{1}{\eta}\frac{\int -f(s)x^{(0)}(s)\exp(-\eta f(s))\,\nu(ds)}{\int x^{(0)}(s)\exp(-\eta f(s))\,\nu(ds)} \\
&= -\frac{1}{\eta^2}\log\frac{1}{Z_f} - \frac{1}{\eta}\int -f(s)x_f(s)\nu(ds) = -\frac{1}{\eta^2}\int\log\frac{1}{Z_f}x_f(s)\nu(ds) - \frac{1}{\eta^2}\int\log\exp(-\eta f(s))x_f(s)\nu(ds) \\
&= -\frac{1}{\eta^2}\int\log\frac{\exp(-\eta f(s))}{Z_f}x_f(s)\nu(ds) = -\frac{1}{\eta^2}\int\log\frac{x_f(s)}{x^{(0)}(s)}x_f(s)\nu(ds) = -D_{KL}(x_f, x^{(0)})
\end{aligned} \tag{5}$$

Proof. We have

$$\begin{aligned}
&\xi(\eta_{t+1}, L^{(t+1)}) - \xi(\eta_{t+1}, L^{(t)}) \\
&\quad = -\frac{1}{\eta_{t+1}}\log\frac{\int x^{(0)}(s)\exp(-\eta_{t+1}\sum_{\tau=1}^{t+1}\ell^{(\tau)}(s))\,\nu(ds)}{\int x^{(0)}(u)\exp(-\eta_{t+1}\sum_{\tau=1}^{t}\ell^{(\tau)}(u))\,\nu(du)} \\
&\quad = -\frac{1}{\eta_{t+1}}\log\int x^{(t+1)}(s)\exp(-\eta_{t+1}\ell^{(t+1)}(s))\,\nu(ds) \\
&\quad = -\frac{1}{\eta_{t+1}}\log\mathbb{E}_{x^{(t+1)}}\left[\exp(-\eta_{t+1}\ell^{(t+1)})\right] \\
&\quad \geq \mathbb{E}_{x^{(t+1)}}\left[\ell^{(t+1)}\right] - \frac{\eta_{t+1}M^2}{8}
\end{aligned}$$

where the last inequality follows from Hoeffding's lemma. Summing the inequalities, we find that

$$\begin{aligned}
&\textstyle\sum_{\tau=1}^{t}\left(\xi(\eta_\tau, L^{(\tau)}) - \xi(\eta_\tau, L^{(\tau-1)})\right) \\
&\qquad\qquad \geq \textstyle\sum_{\tau=1}^{t}\mathbb{E}_{x^{(\tau)}}[\ell^{(\tau)}] - \frac{M^2}{8}\sum_{\tau=1}^{t}\eta_\tau
\end{aligned}$$

Rearranging, and observing that $\xi(\eta, L^{(0)}) = \xi(\eta, 0) = -\frac{1}{\eta}\log\int x^{(0)}(s)\,\nu(ds) = 0$, we obtain equation (4) at the top of the page.

Next, we show that each term of the second sum in (4) is non-positive. Since $\eta_{t+1} \leq \eta_t$ by assumption, it suffices to show that, for any bounded Lipschitz function f, $\eta \mapsto \xi(\eta, f)$ is decreasing. Calculating the partial derivative w.r.t. η (using Dominated Convergence to differentiate under the integral) we obtain (5), where we use x_f to denote the density function $x_f(s) = Z_f^{-1}x^{(0)}(s)\exp(-\eta f(s))$, with Z_f the corresponding normalization constant. Thus $\partial_\eta\xi(\eta, f)$ is proportional to the negative Kullback-Leibler divergence of x_f with respect to $x^{(0)}$, therefore $\eta \mapsto \xi(\eta, f)$ is non-increasing. The bound (4) then reduces to

$$\textstyle\sum_{\tau=1}^{t}\mathbb{E}_{x^{(\tau)}}[\ell^{(\tau)}] \leq \frac{M^2}{8}\sum_{\tau=1}^{t}\eta_\tau + \xi(\eta_t, L^{(t)})$$

and we conclude by subtracting $L^{(t)}(s_t^\star) = \min_{s\in S}L^{(t)}(s)$ from both sides. □

Next, we refine this regret bound by bounding the difference $\xi(\eta_t, L^{(t)}) - L^{(t)}(s_t^\star)$.

For any measurable subset $A \subseteq S$, we define the *diameter* $D(A) := \sup_{s,s'\in A} d(s, s')$ and the *generalized volume* $V_{x^{(0)}}(A) := \int_A x^{(0)}(s)\,\nu(ds)$.

Lemma 2. *Suppose that the loss functions $\ell^{(t)}$ are* L*-Lipschitz uniformly in t. Consider a sequence (S_t) of measurable subsets of S, such that $s_t^\star \in S_t$ for all t. Then the Hedge algorithm with non-increasing learning rates (η_t) guarantees the following bound on the regret:*

$$R^{(t)} \leq \frac{M^2}{8}\sum_{\tau=1}^{t}\eta_\tau + t\,\mathrm{L}D(S_t) - \frac{\log V_{x^{(0)}}(S_t)}{\eta_t} \tag{6}$$

Proof. Since the loss functions $\ell^{(t)}$ are uniformly L-Lipschitz, we have for all $s \in S_t$, $|\ell^{(\tau)}(s) - \ell^{(\tau)}(s_t^\star)| \leq \mathrm{L}d(s, s_t^\star) \leq \mathrm{L}D(S_t)$. Hence, $\ell^{(\tau)}(s) \leq \ell^{(\tau)}(s_t^\star) + \mathrm{L}D(S_t)$, and $L^{(t)}(s) \leq L^{(t)}(s_t^\star) + t\mathrm{L}D(S_t)$. Therefore

$$\begin{aligned}
&\xi(\eta_t, L^{(t)}) \leq -\tfrac{1}{\eta_t}\log\textstyle\int_{S_t}x^{(0)}(s)\exp\big(-\eta_t L^{(t)}(s)\big)\nu(ds) \\
&\quad \leq -\tfrac{1}{\eta_t}\log\textstyle\int_{S_t}x^{(0)}(s)\exp\big(-\eta_t(L^{(t)}(s_t^\star) + t\mathrm{L}D(S_t))\big)\nu(ds) \\
&\quad = L^{(t)}(s_t^\star) + t\mathrm{L}D(S_t)\mathrm{L}D(S_t) - \tfrac{1}{\eta_t}\log\textstyle\int_{S_t}x^{(0)}(s)\,\nu(ds) \\
&\quad = L^{(t)}(s_t^\star) + t\mathrm{L}D(S_t) - \tfrac{1}{\eta_t}\log V_{x^{(0)}}(S_t)
\end{aligned}$$

Combining this with Lemma 1 concludes the proof. □

Lemma 2 provides a regret bound in terms of any sequence (S_t) of subsets of S, with each S_t containing $s_t^\star$, an optimal decision in hindsight. However, this bound is only useful if one can construct such a sequence with appropriate relative decay rates of the diameters and generalized volumes. More precisely, we have the following corollary.

Corollary 1. *Consider the Hedge algorithm with learning rates (η_t) such that $\sum_{\tau=1}^{t}\eta_\tau = o(t)$. Suppose that there exists a sequence (S_t) of subsets with $s_t^\star \in S_t, \forall t$, and such that $D(S_t) = o(1)$ and $\log V_{x^{(0)}}(S_t) = o(t\eta_t)$. Then the regret grows sublinearly, i.e. $\limsup_{t\to\infty} R^{(t)}/t \leq 0$.*

In the next section, we give sufficient conditions on the action set S that guarantee the existence of such a sequence.

3. Sublinear regret in $\mathbb{R}^n$

We now restrict our attention to finite dimensional Euclidean spaces. Let S be a compact subset of $\mathbb{R}^n$. We start with the simple case of convex S, and construct a sequence (S_t) using a homothetic transformation centered at $s_t^\star$, similarly to (Blum & Kalai, 1999). Unless stated otherwise, we make the following assumption for the remainder of this paper:

Assumption 1. *The reference measure is the Lebesgue measure λ, and the initial distribution $x^{(0)}$ is the Lebesgue-uniform distribution over S, i.e. $x^{(0)}(s) = \frac{1}{\lambda(S)}$.*

3.1. Sublinear Regret on Convex Sets

Lemma 3. *If S is a convex compact subset of $\mathbb{R}^n$ and $x^{(0)}$ is the Lebesgue-uniform probability density over S, then the set S_t defined by the homothetic transformation*

$$S_t = \left\{s_t^\star + d_t(s - s_t^\star), s \in S\right\} \tag{7}$$

has diameter $D(S_t) = d_t D(S)$ and generalized volume $V_{x^{(0)}}(S_t) = \frac{\lambda(S_t)}{\lambda(S)} = d_t^n$.

Proof. We have

$$\begin{aligned} D(S_t) &= \sup_{s,s' \in S_t} \|s - s'\| \\ &= \sup_{s,s' \in S} \|s_t^\star + d_t(s - s_t^\star) - s_t^\star - d_t(s' - s^\star)\| \\ &= d_t \sup_{s,s' \in S} \|s - s'\| \end{aligned}$$

Furthermore, using a change of variable $s = s_t^\star + d_t(s' - s_t^\star)$, $s' \in S$, we can write

$$\begin{aligned} V_{x^{(0)}}(S_t) &= \int_{S_t} x^{(0)}(s)\lambda(ds) = \frac{1}{\lambda(S)} \int_{S_t} \lambda(ds) \\ &= \frac{1}{\lambda(S)} \int_S |\det(d_t I_n)| \lambda(ds') = d_t^n \end{aligned}$$

□

Theorem 1 (Hedge on convex compact subsets of $\mathbb{R}^n$)**.** *Let $S \subset \mathbb{R}^n$ be convex and compact, and suppose that the $\ell^{(t)}$ are* L*-Lipschitz uniformly in t. Then under the Hedge algorithm with learning rates $\eta_t = \theta t^{-\alpha}\sqrt{\log t}$, $\alpha \in [0, 1)$, we have*

$$\frac{R^{(t)}}{t} \leq \frac{M^2\theta}{8(1-\alpha)} \frac{\sqrt{\log t}}{t^\alpha} + \frac{\mathrm{L}D(S)}{t} + \frac{n}{\theta} \frac{\sqrt{\log t}}{t^{1-\alpha}} \tag{8}$$

In particular, the per-round regret is $\mathcal{O}\big(t^{-\bar{\alpha}}\sqrt{\log t}\big)$, where $\bar{\alpha} = \min(\alpha, 1 - \alpha)$.

Proof. Constructing the sequence S_t as in Lemma 3, it follows from the regret bound (6) that

$$\frac{R^{(t)}}{t} \leq \frac{M^2\theta}{8} \frac{\sum_{\tau=1}^t \tau^{-\alpha}}{t} + \mathrm{L}D(S)d_t - \frac{n \log d_t}{\theta\, t^{1-\alpha}}$$

Bounding $\sum_{\tau=1}^t \eta_\tau \leq \theta\sqrt{\log t}\sum_{\tau=1}^t \tau^{-\alpha} \leq \theta\sqrt{\log t}\int_0^t \tau^{-\alpha} d\tau = \frac{\theta t^{1-\alpha}\sqrt{\log t}}{1-\alpha}$, we have

$$\frac{R^{(t)}}{t} \leq \frac{M^2\theta}{8(1-\alpha)} \frac{\sqrt{\log t}}{t^\alpha} + \mathrm{L}D(S)d_t + \frac{n}{\theta} \frac{\log 1/d_t}{t^{1-\alpha}\sqrt{\log t}}$$

Now (8) follows by taking $d_t = 1/t$. □

Corollary 2. *Under the assumptions of Theorem 1, with $\eta_t = \eta$ constant, we have $\frac{R^{(t)}}{t} \leq \frac{M^2\eta}{8} + \frac{\mathrm{L}D(S)}{t} + \frac{n \log t}{\eta\, t}$. For a given horizon T, we can choose $\eta = M^{-1}\sqrt{8n\log T/T}$ to minimize this bound, for which*

$$\frac{R^{(T)}}{T} \leq \frac{\mathrm{LD(S)}}{T} + M\sqrt{\frac{n\log T}{2T}} \tag{9}$$

3.2. Sublinear Regret on Uniformly Fat Sets

The convexity assumption can, in fact, be relaxed, while keeping the same asymptotic rate of the regret. Intuitively, to be able to use the sequence (S_t) of sets as constructed in Lemma 3, it suffices to find, for each $s_t^\star$, a convex set K_t containing $s_t^\star$ such that its volume $V_{x^{(0)}}(K_t) = \frac{\lambda(K_t)}{\lambda(S)}$ is uniformly bounded below. This motivates the following relaxation of convexity.

Definition 3.1 (Uniform fatness)**.** *A set $S \subset \mathbb{R}^n$ is v-uniformly fat w.r.t. the density x if, for all $s \in S$, there exists a convex set $K_s \subseteq S$ such that $s \in K_s$ and $V_x(K_s) = \int_{K_s} x(s)\lambda(ds) \geq v$.*

Intuitively, the uniform fatness property ensures that there is sufficient volume around any point of the set, so that it is possible to assign sufficient probability mass around the optimal point $s_t^\star$ in particular. Note that uniform fatness excludes isolated points, but does not require the set to be connected.

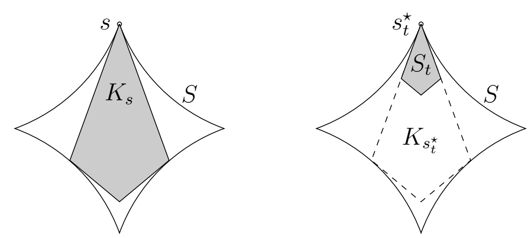

Figure 1. Illustration of the uniform fatness condition (left) and the construction of the set $S_t = s_t^\star + d_t(K_{s_t^\star} - s_t^\star)$ in the proof of Theorem 2 (right).

We are now ready to give a regret bound for the Hedge algorithm on uniformly fat sets.

Theorem 2. *Let $x^{(0)}$ be Lebesgue uniform, and suppose that S is v-uniformly fat w.r.t. $x^{(0)}$ and that the loss functions are* L*-Lipschitz uniformly in time. Then the regret of the Hedge algorithm with learning rates $\eta_t = \theta\, t^{-\alpha}\sqrt{\log t}$, $\alpha \in [0, 1)$, satisfies*

$$\frac{R^{(t)}}{t} \leq \frac{M^2\theta}{8(1-\alpha)} \frac{\sqrt{\log t}}{t^\alpha} + \frac{\mathrm{L}D(S)}{t} + \frac{n\log t + \log\frac{1}{v}}{\theta\, t^{1-\alpha}\sqrt{\log t}} \tag{10}$$

In particular, if S is convex, then it is 1-uniformly fat, and Theorem 1 becomes a special case of Theorem 2.

Proof. Since S is v-uniformly fat, for all t, there exists a convex measurable subset $K_{s_t^\star} \subset S$ with $s_t^\star \in K_{s_t^\star}$ and $V_{x^{(0)}}(K_{s_t^\star}) \geq v$. Similarly to (7), define S_t as the homothetic transformation of $K_{s_t^\star}$ with center $s_t^\star$ and ratio d_t. By Lemma 3, we have $D(S_t) = d_t D(K_t^\star) \leq d_t D(S)$ and $V_{x^{(0)}}(S_t) = d_t^n V_{x^{(0)}}(K_t) \geq d_t^n v$. Applying the regret bound (6) with $\eta_t = \theta t^{-\alpha}\sqrt{\log t}$, we have

$$\begin{aligned}\frac{R^{(t)}}{t} &\leq \frac{M^2\theta}{8(1-\alpha)}\frac{\sqrt{\log t}}{t^\alpha} + \mathrm{L}D(S_t) - \frac{\log V_{x^{(0)}}(S_t)}{\theta t^{1-\alpha}\sqrt{\log t}} \\ &\leq \frac{M^2\theta}{8(1-\alpha)}\frac{\sqrt{\log t}}{t^\alpha} + d_t\mathrm{L}D(S) - \frac{\log(d_t^n v)}{\theta t^{1-\alpha}\sqrt{\log t}}\end{aligned}$$

and we conclude by taking $d_t = 1/t$. □

Corollary 3. *Under the assumptions of Theorem 2, with constant learning rate $\eta_t = \eta$, we have $\frac{R^{(t)}}{t} \leq \frac{M^2\eta}{8} + \frac{\mathrm{L}\,D(S)}{t} + \frac{n\log t - \log v}{\eta\, t}$. For a given horizon T, we can choose $\eta_T = M^{-1}\sqrt{8(n\log T - \log v)/T}$ to minimizes this bound, for which*

$$\frac{R^{(T)}}{T} \leq \frac{\mathrm{L}\,D(S)}{T} + M\sqrt{\frac{n\log T - \log v}{2T}} \tag{11}$$

Remark 1. *If $S \subset \mathbb{R}^n$ is a lower-dimensional manifold, it is not uniformly fat with respect to the Lebesgue measure on $\mathbb{R}^n$. However, if it is homeomorphic to a uniformly fat $S' \subset \mathbb{R}^m$, $m < n$, then one can run the Hedge algorithm on S' instead.*

4. Dual Averaging on $L^2(S)$

In this Section, we study a more general family of algorithms based on the dual averaging method.

Dual averaging (Nesterov, 2009) is a general method for solving constrained optimization problems. It was applied to online learning on a convex subset of $\mathbb{R}^n$, for example in (Xiao, 2010) and (Bubeck, 2014). Building on these ideas, we propose to apply it to our problem of learning on uniformly fat sets.

Consider a Hilbert space E, and a feasible set $\mathcal{X} \subset E$, assumed closed and convex. Given a sequence $(\ell^{(t)})$ of linear functionals in the dual space E^*, the method projects, at each step, the cumulative dual vector $L^{(t)} = \sum_{\tau=1}^t \ell^{(\tau)}$ onto the feasible set, using a Bregman projection. This is summarized in Algorithm 2. In constrained convex optimization, one seeks to minimize a convex function f over $\mathcal{X}$, and the dual vectors $\ell^{(t)}$ are taken to be subgradients of f at the current iterate, but dual averaging provides regret guarantees without requiring $\ell^{(t)}$ to be subgradient vectors. The function ψ in Algorithm 2 is assumed to be ℓ_ψ-strongly convex with respect to a norm[1] $\|\cdot\|$ on E, that is, $\psi(x) \geq \psi(y) + \langle\nabla\psi(y), x-y\rangle + \frac{\ell_\psi}{2}\|x-y\|^2$ for all $x, y \in \mathcal{X}$. To simplify the discussion, we also assume, without loss of generality, that $\inf_{x\in\mathcal{X}}\psi(x) = 0$.

[1] The reference norm $\|\cdot\|$ need not necessarily be the norm induced by the inner product on E.

Algorithm 2 Dual averaging method with input sequence $(\ell^{(t)})$ and learning rates (η_t)

1: **for** $t \in \mathbb{N}$ **do**
2: $\quad L^{(t)} = \sum_{\tau=1}^t \ell^{(\tau)}$
3: $\quad$ Update

$$x^{(t+1)} = \arg\min_{x\in\mathcal{X}} \left\langle L^{(t)}, x\right\rangle + \frac{1}{\eta_{t+1}}\psi(x) \tag{12}$$

The dual averaging update (12) can be written in terms of the Legendre-Fenchel transform of ψ: Let $\psi^*(L) = -\inf_{x\in\mathcal{X}}\psi(x) - \langle L, x\rangle$. Note that the minimum is attained and the minimizer is unique since ψ is strongly convex and $\mathcal{X}$ is closed and convex (Theorem 11.9 in (Bauschke & Combettes, 2011)). Then $\nabla\psi^*(L) = \arg\min_{x\in\mathcal{X}}\psi(x) - \langle L, x\rangle$, and the update (12) can be written

$$x^{(t+1)} = \nabla\psi^*(-\eta_{t+1}L^{(t)}).$$

To apply the dual averaging method to our problem of learning on S, let $E = L^2(S)$, the Lebesgue space of square integrable functions[2] on S, endowed with the inner product $\langle f, g\rangle = \int_S f(s)g(s)\lambda(ds)$, and the feasible set

$$\mathcal{X} := \left\{f \in L^2(S) : f \geq 0 \text{ a.e. and } \textstyle\int_S f(s)\lambda(ds) = 1\right\}.$$

Note that while $\mathcal{X}$ is closed and convex, it may be unbounded[3]. An element $f \in \mathcal{X}$ will be identified with the probability distribution with density f. The dual space is $E^* = L^2(S)$, and since S is compact, E^* contains, in particular, the set $C^0(S)$ of continuous functions on S.

We next show that the regret of the dual averaging method grows sublinearly under appropriate assumptions on the feasible set S and the regularizer ψ. The result extends the regret bound of (Nesterov, 2009) to $L^2(S)$. We will use the following Lemma, which can be proved following Lemma 1 in (Nesterov, 2009), mutatis mutandis.

Lemma 4. *If ψ is ℓ_ψ-strongly convex w.r.t. $\|\cdot\|$. Then ψ^* is $\frac{1}{\ell_\psi}$-smooth w.r.t. $\|\cdot\|_*$, that is, for all x, y*

$$\psi^*(x) - \psi^*(y) - \langle\nabla\psi^*(y), x-y\rangle \leq \tfrac{1}{2\ell_\psi}\|x-y\|_*^2.$$

Lemma 5 (Dual averaging regret bound)**.** *Suppose that there exists $M > 0$ such that for all t, $\|\ell^{(t)}\|_* \leq M$. Then under the dual averaging method with non-increasing learning rates (η_t), for all $x \in \mathcal{X}$,*

[2] An element of E is, in fact, an equivalence class of functions equal almost everywhere.

[3] Consider for example the simple case $S = [0,1]$, and the sequence $f_n = n1_{[0,\frac{1}{n}]}$, for which $\|f_n\|_1 = 1$ but $\|f_n\|_2 = \sqrt{n}$.

$$\sum_{\tau=1}^{t} \left\langle \ell^{(\tau)}, x^{(\tau)} - x \right\rangle \leq \frac{1}{\eta_t}\psi(x) + \frac{M^2}{2\ell_\psi}\sum_{\tau=1}^{t}\eta_\tau \qquad (13)$$

Proof. We use a similar argument to the proof of Lemma 1. Define the potential function $\xi : \mathbb{R}_+ \times L^2(S) \to \mathbb{R}$

$$\xi(\eta, L) = -\tfrac{1}{\eta}\psi^*(-\eta L) = \tfrac{1}{\eta}\inf_{x\in\mathcal{X}} \langle \eta L, x\rangle + \psi(x)$$

We first show the following inequality:

$$\left\langle x^{(t)}, \ell^{(t)} \right\rangle \leq \xi(\eta_t, L^{(t)}) - \xi(\eta_{t-1}, L^{(t-1)}) + \frac{\eta_t}{2\ell_\psi}\|\ell^{(t)}\|_*^2 \qquad (14)$$

Since ψ is ℓ_ψ-strongly convex, by Lemma 4, ψ^* is $\frac{1}{\ell_\psi}$-smooth, therefore

$$\begin{aligned}\psi^*(-\eta_t L^{(t)}) &- \psi^*(-\eta_t L^{(t-1)}) \\ &\leq \left\langle \nabla\psi^*(-\eta_t L^{(t-1)}), -\eta_t \ell^{(t)} \right\rangle + \frac{1}{2\ell_\psi}\|\eta_t \ell^{(t)}\|_*^2 \\ &= -\eta_t \left\langle x^{(t)}, \ell^{(t)} \right\rangle + \frac{\eta_t^2}{2\ell_\psi}\|\ell^{(t)}\|_*^2\end{aligned}$$

Dividing by η_t and rearranging, we have

$$\left\langle x^{(t)}, \ell^{(t)} \right\rangle \leq \xi(\eta_t, L^{(t)}) - \xi(\eta_t, L^{(t-1)}) + \frac{\eta_t}{2\ell_\psi}\|\ell^{(t)}\|_*^2.$$

To prove (14), it suffices to show that $\eta \mapsto \xi(\eta, L)$ is decreasing. Taking the derivative with respect to η,

$$\begin{aligned}\partial_\eta \xi(\eta, L) &= \frac{1}{\eta^2}\psi^*(-\eta L) - \frac{1}{\eta}\langle -L, \nabla\psi^*(-\eta L)\rangle \\ &= \frac{1}{\eta^2}\left(\psi^*(-\eta L) + \langle \eta L, \nabla\psi^*(-\eta L)\rangle\right) \\ &\leq \frac{1}{\eta^2}\psi^*(0) \qquad \text{by convexity of } \psi^* \\ &= -\frac{1}{\eta^2}\inf_{x\in\mathcal{X}}\psi(x) = 0\end{aligned}$$

which proves inequality (14). Summing over $\tau \in \{1, \dots, t\}$, and using the bound on $\|\ell^{(t)}\|_*$, we have

$$\sum_{\tau=1}^{t} \left\langle \ell^{(\tau)}, x^{(\tau)} \right\rangle \leq \xi(\eta_t, L^{(t)}) - \xi(\eta_0, L^{(0)}) + \frac{M^2}{2\ell_\psi}\sum_{\tau=1}^{t}\eta_\tau$$

Finally, by definition of ξ,

$$\xi(\eta_0, L^{(0)}) = \frac{1}{\eta_0}\inf_{x\in\mathcal{X}}\psi(x) = 0$$

$$\xi(\eta_t, L^{(t)}) = \frac{1}{\eta_t}\inf_{x\in\mathcal{X}}\left\langle \eta_t L^{(t)}, x\right\rangle + \psi(x) \leq \left\langle L^{(t)}, x\right\rangle + \frac{1}{\eta_t}\psi(x)$$

Therefore

$$\sum_{\tau=1}^{t} \left\langle \ell^{(\tau)}, x^{(\tau)} \right\rangle \leq \left\langle L^{(t)}, x \right\rangle + \frac{1}{\eta_t}\psi(x) + \frac{M^2}{2\ell_\psi}\sum_{\tau=1}^{t}\eta_\tau$$

which proves the claim. □

Lemma 5 gives an upper bound on the regret with respect to elements of $\mathcal{X}$, the set of Lebesgue-continuous densities. This can also provide a bound on the regret (2), with respect to elements of S, as defined in Section 2, by observing that when S is uniformly fat[4],

$$\begin{aligned}R^{(t)} &= \textstyle\sum_{\tau=1}^t \left\langle \ell^{(\tau)}, x^{(\tau)}\right\rangle - \min_{s\in S}\sum_{\tau=1}^t \ell^{(\tau)}(s) \\ &= \textstyle\sum_{\tau=1}^t \left\langle \ell^{(\tau)}, x^{(\tau)}\right\rangle - \inf_{x\in\mathcal{X}}\left\langle \sum_{\tau=1}^t \ell^{(\tau)}, x\right\rangle.\end{aligned}$$

In particular, if ψ is bounded on $\mathcal{X}$, then it follows from Lemma 5 that

$$R^{(t)} \leq \frac{1}{\eta_t}\sup_{x\in\mathcal{X}}\psi(x) + \frac{M^2}{2\ell_\psi}\sum_{\tau=1}^{t}\eta_\tau$$

which implies that the regret is sublinear for $\eta_t = \Theta(t^{-\alpha})$, $\alpha \in (0, 1)$, for *any* sequence of continuous losses. However, when $\mathcal{X}$ is unbounded, so is ψ by strong convexity. But one can still obtain a sublinear bound on the regret for *Lipschitz* losses, as stated in the following Theorem.

Theorem 3 (Dual averaging regret for Lipschitz losses)**.** *Suppose that $\ell^{(t)}$ is* L*-Lipschitz, and $\|\ell^{(t)}\|_* \leq M$, uniformly in t. Then the dual averaging method with learning rates (η_t) guarantees the following bound on the regret: For any positive sequence (d_t),*

$$\frac{R^{(t)}}{t} \leq \frac{M^2}{2\ell_\psi}\frac{\sum_{\tau=1}^t \eta_\tau}{t} + \mathrm{L}D(S)d_t + \frac{1}{t\eta_t}\inf_{x\in\mathcal{B}_t}\psi(x) \qquad (15)$$

where $\mathcal{B}_t \subset \mathcal{X}$ denotes the set of Lebesgue-continuous densities supported on $B(s_t^\star, D(S)d_t)$.

Proof. Since the losses are L-Lipschitz, we have, $\forall\, x \in \mathcal{B}_t$,

$$\begin{aligned}\left\langle \textstyle\sum_{\tau=1}^t \ell^{(\tau)}, x\right\rangle &= \int_{B_t}\textstyle\sum_{\tau=1}^t \ell^{(\tau)}(s)x(s)\lambda(ds) \\ &\leq \int_{B_t}\textstyle\sum_{\tau=1}^t (\ell^{(\tau)}(s_t^\star) + \mathrm{L}D(S)d_t)x(s)\lambda(ds) \\ &= L^{(t)}(s_t^\star) + t\,\mathrm{L}D(S)d_t\end{aligned}$$

Thus, for all $x \in \mathcal{B}_t$,

$$\begin{aligned}R^{(t)} &= \textstyle\sum_{\tau=1}^t \left\langle \ell^{(\tau)}, x^{(\tau)}\right\rangle - L^{(t)}(s_t^\star) \\ &\leq \textstyle\sum_{\tau=1}^t \left\langle \ell^{(\tau)}, x^{(\tau)} - x\right\rangle + t\,\mathrm{L}D(S)d_t \\ &\leq \tfrac{1}{\eta_t}\psi(x) + \tfrac{M^2}{2\ell_\psi}\textstyle\sum_{\tau=1}^t \eta_\tau + t\,\mathrm{L}D(S)d_t\end{aligned}$$

where the last inequality uses Lemma 5. We conclude by dividing by t and taking the infimum over $x \in \mathcal{B}_t$. □

In the finite dimensional case, the Hedge algorithm is known to be an instance of the dual averaging method,

[4] Proof in the supplementary material.

when the feasible set $\mathcal{X}$ is the simplex, and the distance generating function ψ is the negative entropy function, as observed by Nesterov (2009), Beck & Teboulle (2003), and many others. This is also true in the infinite dimensional case, as discussed for example in (Audibert, 2009) and in Chapter 5 in (Catoni, 2004). Next, we apply the regret bound of Theorem 3 to the Hedge algorithm.

Example 1 (Hedge algorithm or the entropic dual averaging)**.** *Let the reference measure be the Lebesgue measure on S, denoted by λ, and let ψ be the generalized negative entropy, defined by*

$$\psi(f) = \int_S f(s) \log f(s) \lambda(ds) + \log \lambda(S) \tag{16}$$

Note that $\psi(f)$ is the Kullback-Leibler divergence of f with respect to the Lebesgue-uniform distribution. By Pinsker's inequality, ψ is 1-strongly convex with respect to the total variation norm $\|f\| = \int |f(s)|\lambda(ds)$. Furthermore, ψ is nonnegative since it is minimal on the Lebesgue uniform distribution, for which it takes value 0. One can show that $\psi^*(f) = \log \frac{1}{\lambda(S)} \int_S \exp(f(s))\lambda(ds)$, thus the definitions of $\xi(\eta, L)$ in Lemmas 1 and 5 coincide.

With this choice of $\mathcal{X}$ and ψ, the dual averaging update can be shown to be equal to the Hedge density (1):

Proposition 1. *Let $L^{(t)} \in E^*$, and consider the dual averaging iteration (12) with ψ the negative entropy (16). Then the solution $x^{(t+1)}$ is given by the Hedge update rule*

$$x^{(t+1)}(s) = \frac{1}{\bar{Z}^{(t)}} e^{-\eta_{t+1} L^{(t)}(s)}$$

with normalization constant $\bar{Z}^{(t)} = \int_S e^{-\eta_{t+1} L^{(t)}(s)} \lambda(ds)$.

A proof is provided in the supplementary material for completeness.

By Theorem 3, the regret of the Hedge algorithm with Lipschitz losses is then bounded as follows

$$\frac{R^{(t)}}{t} \leq \frac{M^2}{2} \frac{\sum_{\tau=1}^t \eta_\tau}{t} + \mathrm{L}D(S)d_t + \frac{1}{t\eta_t} \inf_{x \in \mathcal{B}_t} \psi(x)$$

In order to bound the last term, consider a subset $S_t \subset B(s_t^\star, D(S)d_t)$, and take x to be the uniform distribution over S_t (thus $x \in \mathcal{B}_t$). Then

$$\psi(x) = \log \lambda(S) + \int_{S_t} \frac{1}{\lambda(S_t)} \log \frac{1}{\lambda(S_t)} \lambda(ds) = \log \frac{\lambda(S)}{\lambda(S_t)}$$

Now, if S is v-uniformly fat, then there exists a convex set K_t containing $s_t^\star$, such that $\lambda(K_t)/\lambda(S) \geq v$, and constructing S_t as in the proof of Theorem 1, as the homothetic transform of K_t with center $s_t^\star$ and ratio d_t we have $D(S_t) \leq D(S)d_t$ and $\lambda(S_t)/\lambda(S) \geq v d_t^n$, therefore $\inf_{x \in \mathcal{B}_t} \psi(x) \leq -\log(v d_t^n)$, and taking $d_t = 1/t$,

$$\frac{R^{(t)}}{t} \leq \frac{M^2}{2} \frac{\sum_{\tau=1}^t \eta_\tau}{t} + \frac{\mathrm{L}D(S)}{t} + \frac{n \log t - \log v}{t\, \eta_t}$$

which results in a regret bound similar to Theorem 2.

5. Learning on a finite cover

In this section, we briefly compare our method to a related idea: for a given horizon, compute a finite cover of the set, such that the maximum difference of losses on each element of the cover is small enough, and then perform a discrete learning algorithm on the finite cover.

More precisely, fixing a horizon T and a constant $\epsilon_T > 0$, suppose that we can compute a finite cover $\mathcal{A}_T$ of S such that, for all $S_T \in \mathcal{A}_T$,

$$\sup_{s, s' \in S_T} |\ell^{(t)}(s) - \ell^{(t)}(s')| \leq \epsilon_T. \tag{17}$$

If we call $\tilde{R}^{(t)}$ the regret with respect to the discrete set, then running the discrete Hedge algorithm on this finite cover with learning rate η guarantees (Bubeck & Cesa-Bianchi, 2012) that

$$\frac{\tilde{R}^{(T)}}{T} \leq \frac{M^2}{8}\eta + \frac{\log |\mathcal{A}|}{T\eta}$$

and the optimal η given the horizon T yields the regret bound $\tilde{R}^{(T)}/T = \mathcal{O}\big(\sqrt{\log |\mathcal{A}|\,/T}\big)$. Since we incur at most ϵ_T additional per-round regret due to the variation of losses within each element of the cover, we have that $R^{(T)}/T = \mathcal{O}\big(\sqrt{\log |\mathcal{A}|\,/\,T} + \epsilon_T\big)$.

Since the loss functions are L-Lipschitz, a sufficient condition for (17) to hold is to have each element S_T of the cover have diameter $D(S_T) \leq \epsilon_T/\mathrm{L}$. Thus, the size of the cover is typically $|\mathcal{A}_T| = \mathcal{O}(1/\epsilon_T^n)$. Under this estimate of the size of $\mathcal{A}_T$, we have $R^{(T)}/T = \mathcal{O}\big(\sqrt{-n \log \epsilon_T\,/\,T} + \epsilon_T\big)$, and choosing $\epsilon_T = 1/\sqrt{T}$, we obtain the bound

$$R^{(T)}/T = \mathcal{O}\big(\sqrt{n \log T / T}\,\big)$$

which matches the regret rate of Corollary 3.

The Hedge algorithm on uniformly fat sets is conceptually similar to the idea of working with a finite cover. This is most visible in the proof of Theorem 2, where we rely on the existence of a set around $s_t^\star$ with an appropriate relationship between diameter and volume. To apply the Hedge algorithm, one needs to sample from the distributions $x^{(t)}$, without having to explicitly construct a cover. Sampling from the distribution can be more tractable, as is the case in the example of Section 6.

6. Numerical results

We test our algorithm on a numerical example in $\mathbb{R}^2$ with convex quadratic loss functions of the form

$$\ell^{(t)}(s) = \frac{1}{2}(s - \mu_t)^\top Q_t (s - \mu_t) + c_t$$

restricted to the domain $S \subset \mathbb{R}^2$ shown in Figure 2.

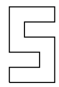

Figure 2. The set S for the numerical example.

If $x^{(0)}$ is the uniform distribution over S one can show that

$$x^{(t+1)}(s) \propto |\tilde{Q}_t|^{-1/2} \exp\Big(-\frac{1}{2}(s-\tilde{\mu}_t)^\top \tilde{Q}_t (s-\tilde{\mu}_t)\Big)$$

on S, so $x^{(t+1)}$ is a multivariate Gaussian distribution restricted to S with covariance matrix $\tilde{Q}_t = \eta_t \sum_{\tau \le t} Q_\tau$ and mean $\tilde{\mu}_t = \tilde{Q}_t^{-1} \eta_t \sum_{\tau \le t} Q_\tau \mu_\tau$. Hence the size of the parameter space required to represent the cumulative loss (and thus the Hedge densities) is independent of the horizon.

Remark 2. *Since the Hedge distributions are, in this case, multivariate Gaussian, sampling from these distributions can be done efficiently. This example is one instance of a problem in which one can directly sample from the Hedge distributions without having to maintain a discrete cover. More generally, the complexity of the Hedge algorithm depends on the complexity of sampling from the Hedge distributions.*

For a simulation horizon of $T = 10^4$, we randomly generated the parameters μ_t, Q_t and c_t of the loss functions subject to the uniform bounds $M = 10$ and L $= 5$. The set S has diameter $D = 5.83$ and is v-uniformly fat with $v = 0.273$. We simulated the algorithm 2500 times for each of the different choices of the learning rates. Figure 3 shows means (solid lines), regret bounds (dashed lines) and regions between the 10% and 90% quantiles (shaded) of the per-round cumulative regret over these simulations. To determine the regret, the computation of the best choice in hindsight for each period t was performed by solving multiple quadratic programs on a convex decomposition of S.

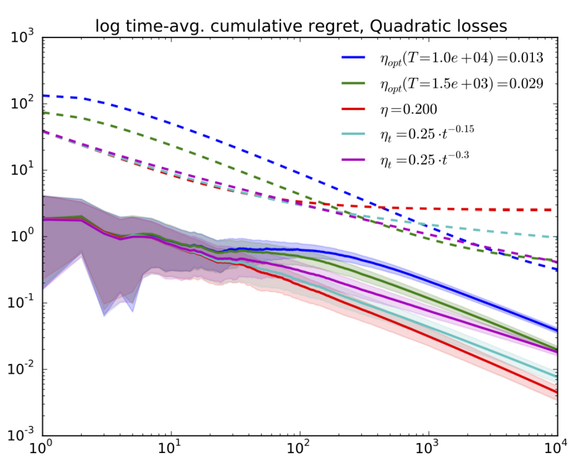

Figure 3. Mean time-average cumulative regret (solid), 10% and 90% quantiles (shaded regions) and worst-case bounds (dashed).

We first verify from Figure 3 that the regret bound (10) is satisfied. Since the loss functions are generated randomly (and not adversarially), the regrets observed in simulation are much smaller than the theoretical regret bounds. We further note that the optimal constant learning rates η_{opt} from Corollary 3 are outperformed by higher constant learning rates[5], an observation familiar from the discrete case (Even-Dar et al., 2008; Koolen et al., 2014).

Finally, Table 2 compares the decay rates of the per-round cumulative regret (solid lines in Figure 3), estimated using a linear regression on $\log \frac{R^{(t)}}{t}$ as a function of $\log t$, with those of the corresponding theoretical bound (10) (dashed lines). The observed rates of decay are higher than those of the theoretical bounds.

η_t	simulation	bound	bound ($t \to \infty$)
$t^{-0.15}$	-0.778	-0.173	-0.150
$t^{-0.3}$	-0.644	-0.397	-0.300

Table 2. Decay rates of the per-round regret.

7. Concluding remarks

We studied an online optimization problem over a compact subset S of $\mathbb{R}^n$. Previous work shows that when S is convex, one can achieve $\mathcal{O}(\sqrt{t})$ cumulative regret using a gradient descent method when the losses are convex, and $\mathcal{O}(\log t)$ cumulative regret using a generalized Hedge algorithm when the losses are uniformly exp-concave. We consider Lipschitz losses, and relax the convexity assumption of S. In particular, we show that as long as the set is uniformly fat, i.e. there exists a convex set of minimal volume v around all points of the set, then the generalized Hedge algorithm achieves $\mathcal{O}(\sqrt{t \log t})$ regret. We further proved a regret bound for dual averaging method, a generalization of the Hedge algorithm. A question which remains open is whether the uniform fatness condition is necessary for a given class of algorithms to achieve sublinear regret.

A related problem, which is not studied here, is bandit learning on a continuum, in which only the current loss value $\ell^{(t)}(s^{(t)})$ is revealed. This problem is studied for example by Bubeck et al. (2011) for Lipschitz losses and general topological spaces. A hierarchical algorithm is proposed, which achieves a $\mathcal{O}(t^{d_1}(\log t)^{d_2})$ where $d_1, d_2 \le 1$ depend on the geometry of the problem. The algorithm requires explicitly computing a hierarchical cover of the set. One question is whether one can generalize the Hedge algorithm to such a bandit setting, so that sublinear regret can be achieved without the need to explicitly maintain a cover.

[5] Due to the rather weak assumptions on the action set and the loss functions, the optimal learning rate $\eta_{opt}(T)$ for the known horizon T derived from Corollary (3) is very small.

References

Arora, Sanjeev, Hazan, Elad, and Kale, Satyen. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.

Audibert, Jean-Yves. Fast learning rates in statistical inference through aggregation. *The Annals of Statistics*, 37 (4):pp. 1591–1646, 2009. ISSN 00905364.

Bauschke, Heinz H. and Combettes, Patrick L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer, 2011.

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

Blum, Avrim and Kalai, Adam. Universal portfolios with and without transaction costs. *Machine Learning*, 35(3): 193–205, 1999.

Bubeck, Sébastien. Theory of convex optimization for machine learning. *ArXiv*, 2014.

Bubeck, Sébastien and Cesa-Bianchi, Nicolò. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends in Machine Learning*, 5(1):1–122, 2012.

Bubeck, Sébastien, Munos, Rémi, Stoltz, Gilles, and Szepesvari, Csaba. X-armed bandits. *Journal of Machine Learning Research (JMLR)*, 12(12):1587–1627, 2011.

Catoni, Olivier. *Statistical learning theory and stochastic optimization, Ecole d'Eté de Probabilités de Saint-Flour XXXI- 2001*, volume 1851. Springer, 2004.

Cesa-Bianchi, Nicolò and Lugosi, Gábor. *Prediction, learning, and games*. Cambridge University Press, 2006.

Cope, Eric W. Regret and convergence bounds for a class of continuum-armed bandit problems. *Automatic Control, IEEE Transactions on*, 54(6):1243–1253, June 2009.

Cover, Thomas M. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.

Dalalyan, Arnak S. and Salmon, Joseph. Sharp oracle inequalities for aggregation of affine estimators. *Ann. Statist.*, 40(4):2327–2355, 08 2012.

Even-Dar, Eyal, Kearns, Michael, Mansour, Yishay, and Wortman, Jennifer. Regret to the best vs. regret to the average. *Machine Learning*, 72(1-2):21–37, 2008.

Hannan, James. Approximation to Bayes risk in repeated plays. *Contributions to the Theory of Games*, 3:97–139, 1957.

Hazan, Elad, Agarwal, Amit, and Kale, Satyen. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.

Kivinen, Jyrki and Warmuth, Manfred K. Exponentiated gradient versus gradient descent for linear predictors. *Information and Computation*, 132(1):1 – 63, 1997.

Koolen, Wouter M., Erven, Tim Van, and Grünwald, Peter D. Learning the learning rate for prediction with expert advice. In *Advances in Neural Information Processing Systems (NIPS) 27*, pp. 2294–2302, Dec 2014.

Littlestone, Nick and Warmuth, Manfred K. The weighted majority algorithm. In *Foundations of Computer Science, 1989., 30th Annual Symposium on*, pp. 256–261. IEEE, 1989.

Nesterov, Yurii. Primal-dual subgradient methods for convex problems. *Mathematical Programming*, 120(1): 221–259, 2009.

Xiao, Lin. Dual averaging methods for regularized stochastic learning and online optimization. *J. Mach. Learn. Res.*, 11:2543–2596, December 2010. ISSN 1532-4435.

Xidonas, Panos and Mavrotas, George. Multiobjective portfolio optimization with non-convex policy constraints: Evidence from the eurostoxx 50. *The European Journal of Finance*, 20(11):957–977, 2014.

Zinkevich, Martin. Online convex programming and generalized infinitesimal gradient ascent. In *ICML*, pp. 928–936, 2003.