Bandits and online optimization in *infinite* spaces. UMD spaces are UMD-learnable.

Isaac Meza Lopez
Department of IEOR
-University of California at Berkeley
isaac.meza@berkeley.edu

Abstract

We give a characterization for a large class of adversarial bandits in infinite dimensional spaces that achieve sublinear regret (i.e. *learnable*).

In this paper we explore the extension of finite arm bandits to the infinite case. Since the pioneer work on this problem [Agr95], a growing literature has emerged with particular emphasis on Lipschitz bandits [MCP14; LWHZ19; WYGR20; BSY11], and Gaussian bandits [SJ18; SKKS09; DKC13]. This paper focuses particularly on the case in which the decision set is a Banach Space [SST11; BKTB16; DT14]. We exploit the geometric structure of such spaces as tools to analyze this bandits.

Another contribution is to give an unification of Mirror Descent and Dual Averaging. In [McM11] an equivalence between regularized dual averaging, and composite objective mirror descent is established as instantiations of a general follow the regularized leader (FTLR) update. However, the unification is different to the one in [JKM20], and so to the one appearing here: the difference between mirror descent and dual averaging appears as a result of having regularizers/mirror maps which vary over time. The unification by [McM11] is then achieved by tweaking the way the time-varying regularizers/mirror maps are defined.

1 Adversarial Bandits in Banach spaces

Our primary object of study are going to be adversarial bandits in infinite dimensional spaces, more precisely in a Banach space $\mathcal{B}$. We denote by $\mathcal{B}^*$ the dual space of $\mathcal{B}$, and the evaluation functional will be denoted by $\langle \cdot, \cdot \rangle$. Consider the sequential decision problem in which the setting is the following:

1.1. *For each round $t \in [n]$:*

1. *Learner picks $x_t \in \mathcal{X} \subseteq \mathcal{B}$*
2. *Adversary picks a convex cost $x_t^* : \mathcal{X} \longmapsto \mathbb{R}$ from a class $\mathcal{F}_{\mathcal{X}}$*
3. *Learner pays cost $\langle x_t, x_t^* \rangle := x_t^*(x_t)$ for the chosen arm*

We will refer to the above setting as the adversarial bandit $(\mathcal{X}, \mathcal{F})$. An online algorithm $\mathcal{A}$ is a sequence of mappings $\mathcal{A}_t : \mathcal{F}_{\mathcal{X}}^{t-1} \longmapsto \mathcal{X}$, and the regret is defined by

$$R_n(\mathcal{A}, x_1^*, \dots, x_n^*) = \sum_{t=1}^{n} \langle x_t^*, \mathcal{A}_t(x_{1:t-1}^*) \rangle - \inf_{x \in \mathcal{X}} \sum_{t=1}^{n} \langle x_t^*, x \rangle$$

The goal is to minimize regret. Associated to the regret is the concpet of the value of the problem, which is the minimax value defined as:

$$\mathcal{V}_n(\mathcal{X}, \mathcal{F}) := \inf_{\mathcal{A}} \sup_{x_{1:n}^* \in \mathcal{F}_{\mathcal{X}}} R_n(\mathcal{A}, x_{1:n}^*)$$

Preprint. Under review.

We will say that an online convex problem is *learnable* whenever the minimax value is sublinear:

$$\mathcal{V}_n(\mathcal{X}, \mathcal{F}) = o(n)$$

Our fundamental question will be to characterize the spaces $\mathcal{X}$, and $\mathcal{F}$ where $(\mathcal{X}, \mathcal{F})$ is learnable.

2 Preliminaries

2.1 Convexity

The key tools in the analysis of UMD will be the geometric structure of Banach spaces, closely related to its convexity. The modulus of convexity of a norm $||\cdot||$ on X is defined for $\epsilon \in [0, 2]$ by

$$\delta_{||\cdot||}(\epsilon) := \inf\{1 - \frac{1}{2}||x + y|| \mid ||x|| = ||y|| = 1 \;, ||x - y|| \geq \epsilon\}$$

We say that the space X (and the norm) is uniformly convex if $\delta_{||\cdot||}(\epsilon) > 0$ for all $\epsilon \in (0, 2]$; additionally we say that $||\cdot||$ has modulus of convexity of power type q or that it is q-uniformly convex if there exists $C > 0$ so that $\delta_{||\cdot||}(\epsilon) \geq C\epsilon^q$ for all $\epsilon \in [0, 2]$.

Definition 1 (q-uniformly convex function)**.** *A function $h : \mathcal{B} \longmapsto \mathbb{R}$ is q-uniformly convex w.r.t. $||\cdot||$ in $\mathcal{X} \subseteq \mathcal{B}$ whenever*

$$\forall\, x, x' \;,\; \forall\, \alpha \in [0, 1] \;, \qquad h(\alpha x + (1 - \alpha)x') \leq \alpha h(x) + (1 - \alpha)h(x') - \frac{\alpha(1 - \alpha)}{q}||x - x'||^q$$

The following result gives the existence of a q-uniformly convex function in a uniformly convex space.

Proposition 1 (Theorem 2.3 in [BGHV09])**.** *Let $(\mathcal{B}, ||\cdot||)$ be a Banach space. For $q \in [2, \infty)$, the following are equivalent.*

(i) $(\mathcal{B}, ||\cdot||)$ *is q-uniformly convex*

(ii) *The function $f = ||\cdot||^q$ is q-uniformly convex*

2.2 Regularizers

Definition 2 (Regularizer)**.** *Let $h : \mathcal{X} \longmapsto \mathbb{R} \cup \{\infty\}$ be a function. h is a pre-regularizer if it is convex, lower-semicontinuous, and if $\operatorname{cl} \operatorname{dom} h = \mathcal{X}$. Moreover, if $\operatorname{dom} h^* = \mathcal{X}^*$, with $h^*(x^*) := \sup_{x \in \mathcal{X}}\{\langle x^*, x\rangle - f(x)\}$ as its convex conjugate, then h is said to be a $\mathcal{X}$-regularizer*

Lemma 2. *Let $h : \mathcal{X} \longmapsto \mathbb{R} \cup \{\infty\}$ be a pre-regularizer; if it is q-uniformly convex, then it is a regularizer.*

Lemma 3 (Proposition 7.34 in [FHH+10])**.** *Let f be a proper function from a Banach space $\mathcal{B}$ into $\mathbb{R} \cup \{\infty\}$. Let $x \in \mathcal{B}$, $x^* \in \mathcal{B}^*$ and f^* be the convex conjugate. The following are equivalent*

(i) $f(x) + f^*(x^*) = \langle x^*, x\rangle$

(ii) $x \in \partial f^*(x^*)$

(iii) $x^* \in \partial f(x)$

2.3 Unconditional Martingale Difference spaces

We will see that the precise geometric structure we are looking for is the so-called UMD property, which means that $\mathcal{B}$-valued martingale difference sequences are unconditional in $\mathbb{L}_p(\mathcal{B})$.

Definition 3 (Type and Cotype)**.** *A Banach Space $\mathcal{B}$ has martingale type $p \in [1, 2]$ if there exists a constant $\tau \geq 0$ such that for all finite $\mathbb{L}_p$ martingales $(f_n)_{n=1}^N$*

$$||f_N||_{\mathbb{L}_p(S;\mathcal{B})} \leq \tau \left(||f_0||^p_{\mathbb{L}_p(S;\mathcal{B})} + \sum_{n=1}^{N} ||df_n||^p_{\mathbb{L}_p(S;\mathcal{B})} \right)^{\frac{1}{p}}$$

and we say it has type $p \in [1,2]$ whenever for all finite sequences $(x_n)_{n=1}^N$ and $(\epsilon_n)_{n\geq 1}$ a Rademacher sequence, there is a constant $\tau \geq 0$ such that

$$\left(\mathbb{E}\left|\left|\sum_{n=1}^N \epsilon_n x_n\right|\right|^p\right)^{\frac{1}{p}} \leq \tau \left(\sum_{n=1}^N ||x_n||^p\right)^{\frac{1}{p}}$$

The space $\mathcal{B}$ has martingale cotype $q \in [2,\infty]$ if there exists a constant $c \geq 0$ such that for all finite $\mathbb{L}_q$ martingales $(f_n)_{n=1}^N$

$$\left(||f_0||_{\mathbb{L}_q(S;\mathcal{B})}^q + \sum_{n=1}^N ||df_n||_{\mathbb{L}_q(S;\mathcal{B})}^q\right)^{\frac{1}{q}} \leq c\, ||f_N||_{\mathbb{L}_q(S;\mathcal{B})}$$

and we say it has cotype $q \in [2,\infty]$ whenever for all finite sequences $(x_n)_{n=1}^N$ and $(\epsilon_n)_{n\geq 1}$ a Rademacher sequence, there is a constant $c \geq 0$ such that

$$\left(\sum_{n=1}^N ||x_n||^q\right)^{\frac{1}{q}} \leq c \left(\mathbb{E}\left|\left|\sum_{n=1}^N \epsilon_n x_n\right|\right|^q\right)^{\frac{1}{q}}$$

It is known that whenever $\mathcal{B}$ has (martingale) type p, then it also has (martingale) type $p' \leq p$. We then define the index type as $p(\mathcal{B}) := \sup\{p \in [1,2] \mid \mathcal{B} \text{ is of type } p\}$, and the index cotype as $q(\mathcal{B}) := \inf\{p \in [2,\infty] \mid \mathcal{B} \text{ is of cotype } q\}$ and similarly for the martingale type.

By considering Rademacher differences $df_n = \epsilon_n x_n$ one sees that martingale type p (cotype q) implies type p (cotype q). We will be interested where this two notions agree[1].

As we will see, a crucial condition for learnability is when this two *type* concepts coincide. We then introduce a class of spaces where they both agree.

Definition 4 ((UMD) spaces)**.** *A Banach space $\mathcal{B}$ is said to have the property of unconditional martingale difference (UMD) if for all $p \in (1,\infty)$ there exists a finite constant $\beta(p,\mathcal{B}) \geq 0$ such that the following holds:*

Whenever $(S,\mathcal{A},\mu)$ is a σ-finite measure space, $(\mathcal{F}_n)_{n=1}^N$ a σ-finite filtration and $(f_n)_{n=1}^N$ a finite martingale in $\mathbb{L}_p(S;\mathcal{B})$, then for all scalars $|\epsilon_n| = 1$, $n = 1\ldots,N$ we have

$$\left|\left|\sum_{n=1}^N \epsilon_n df_n\right|\right|_{\mathbb{L}_p(S;\mathcal{B})} \leq \beta \left|\left|\sum_{n=1}^N df_n\right|\right|_{\mathbb{L}_p(S;\mathcal{B})}$$

where $df_n = f_n - f_{n-1}$. If this condition holds, then $\mathcal{B}$ is said to be a UMD space.

Proposition 4 (Proposition 4.3.13 in [HvNVW16])**.** *Every UMD space has a non-trivial type (which in turn implies finite cotype), which then also implies non-trivial martingale type (and finite martingale cotype).*

Remark 1. *Another important class of spaces where type and martingale type definitions agree are the Banach lattices of type $p > 1$.*

The importance of the previous proposition to our purposes shows up when combined with the following Theorem characterizing spaces with non-trivial martingale type.

Theorem 5 (Theorem 19.5 in [SC06])**.** *The following properties of a Banach space $\mathcal{B}$ are equivalent*

(i) $\mathcal{B}$ is superreflexive

(ii) $\mathcal{B}$ is uniformly convexifiable, i.e. has an equivalent norm with $\delta_{||\cdot||}(\epsilon) \geq C\epsilon^q$

(iii) $\mathcal{B}$ is of non-trivial martingale cotype q for some q.

(iv) $\mathcal{B}$ is of non-trivial martingale type p for some p.

[1] In general, it is possible to find a uniformly convex space $\mathcal{B}$ for which the index of type $p(\mathcal{B})$ differs from the corresponding index for the martingale type. Similarly for the cotype.

3 Unified Mirror Descent

We extend the general class of algorithms called Unified Mirror Descent, which will also be denoted[2 3] as (UMD), introduced in [JKM20], and which generalize Mirror Descent and Dual Averaging, to an infinite dimensional space $\mathcal{X}$.

We will give geometric conditions on the set $\mathcal{X}$ such that (UMD) is well defined and prove its theoretical guarantees.

Definition 5 ((UMD) algorithm)**.** *Let h be a $\mathcal{X}$-*regularizer*, and $\xi := (\xi_t)_{t\geq 1}$ be a sequence in $\mathcal{F}$. We say that $(x_t, \theta_t)_{t\geq 1}$ is a $\mathrm{UMD}(h, \xi)$ sequence if for $t \geq 1$:*

(i) $x_t = \nabla h^*(\theta_t)$

(ii) For all $x \in \mathcal{X}$, $\langle \theta_{t+1} - \theta_t + \xi_t, x - x_{t+1} \rangle \geq 0$

We call $(x_t)_{t\geq 1}$ and $(\theta_t)_{t\geq 1}$ a sequence of primal and dual iterates with dual increments $(\xi_t)_{t\geq 1}$.

Given a $\mathcal{X}$-regularizer h and a sequence of dual increments $(\xi_t)_{t\geq 1}$, observe that the $\mathrm{UMD}(h, \xi)$ iterates always exist. From the definition of regularizer, there is a point $x_1 \in \mathcal{X}$ such that $\partial h(x_1) \neq \emptyset$; in other words, there exists (x_1, θ_1) such that $X_1 = \nabla h^*(\theta_1)$. Then, for $t \geq 1$ we define

(i) $\theta_{t+1} := \theta_t - \xi_t$

(ii) $x_{t+1} := \nabla h^*(\theta_{t+1})$

It can be proven that this corresponds to the iterates of the Dual Averaging algorithm [Xia09]. Before we state our main results, we need one more definition where we extend the concept of *Bregman divergence* to include subdifferentials.

Definition 6 (Bregman Divergence)**.** *Let $h : \mathcal{B} \longmapsto \mathbb{R}$ be a $\mathcal{X}$-regularizer. For $x_0 = \in \mathcal{X}$ such that $\partial h(x_0) \neq \emptyset$, $x \in \mathcal{X}$, and $\theta \in \partial h(x_0)$, we define the Bregman divergence from x_0 to x with subdifferentials θ as*

$$D_h(x, x_0; \theta) := h(x) - h(x_0) - \langle \theta, x - x_0 \rangle$$

4 Main results

We now state the main results of this paper.

Theorem 6. *Let $\mathcal{X} \subseteq \mathcal{B}$ be a closed convex symmetric[4], bounded subset of a UMD space $\mathcal{B}$. Then there exists an $\mathcal{X}$-*regularizer*, which we call h. Moreover, if (x_t, θ_t) is a sequence of $\mathrm{UMD}(h, \zeta)$ iterates with $\zeta := (\eta\xi_t)_{t\geq 1}$, then for all $n \geq 1$,*

$$R_n(\mathrm{UMD}, \zeta) \leq \frac{\sup_{x\in\mathcal{X}} D_h(x, x_1; \theta_1)}{\eta} + \frac{\eta^{p-1}}{p} \sum_{t=1}^{n} ||\xi_t||_*^p$$

where p is the type of the Banach space.

The above theorem shows when a $\mathcal{X}$-*regularizer* exists, and then, when UMD learns $(\mathcal{X}, cvx)$. Note that it includes Theorem 6 in [BKTB16] , Theorem 9 in [SST11] and Corollary 4.11 in [JKM20]

Observe that if (ζ) is uniformly bounded, i.e. $||\xi_t||_* \leq M$ for all $t \geq 1$, then the value of the problem is sublinear.

[2]We do not distinguish between the two abbreviations and will be clear from context which we are referring to.

[3]Our main result shows this is a fortunate coincidence.

[4]Without loss of generality we can consider the unit ball of $\mathcal{B}$, since then the Minkowski functional

$$\rho_{\mathcal{X}}(x) := \inf\{r > 0 \mid x \in r\mathcal{X}\}$$

is a norm.

Corollary 7. *Let $\mathcal{X} \subseteq \mathcal{B}$ be a closed convex symmetric, bounded subset of a UMD space $\mathcal{B}$. If $||\xi_t||_*^p \leq M$ for all $t \geq 1$, putting $\eta = (pD/Mn)^{\frac{1}{p}}$ where $D := \sup_{x \in \mathcal{X}} D_h(x, x_1; \theta_1)$, then the adversarial bandit problem $(\mathcal{X}, cvx)$ is learnable by UMD. In particular*

$$R_n(\mathrm{UMD}, \zeta) \leq (pMnD)^{\frac{1}{p}}$$

We have a partial converse result. Denote by cvx the class of functions $f : \mathcal{X} \longmapsto \mathbb{R}$ such that f is convex and non-expansive, and lin the class of functions that are linear and non-expansive.

Theorem 8. *Let $\mathcal{X} \subseteq \mathcal{B}$ be a closed convex symmetric, bounded subset of a UMD space $\mathcal{B}$. If $\mathcal{V}_n(\mathcal{X}, cvx) = o(n)$, i.e. the adversarial bandit problem $(\mathcal{X}, cvx)$ is learnable, then $\mathcal{B}$ has non-trivial martingale type (and thus non-trivial martingale cotype).*

Moreover, we have a stronger result

Theorem 9. *Let $\mathcal{X} \subseteq \mathcal{B}$ be a closed convex symmetric, bounded subset of a Banach space $\mathcal{B}$. If $\mathcal{V}_n(\mathcal{X}, cvx) = o(n)$, i.e. the adversarial bandit problem $(\mathcal{X}, cvx)$ is learnable, then $\mathcal{B}$ has non-trivial martingale type (and thus non-trivial martingale cotype).*

4.1 Proofs of results

Proposition 10. *Let $(x_t, \theta_t)_{t \geq 1}$ be an $\mathrm{UMD}(h, \xi)$ sequence. Then for all $t \geq 1$,*

(i) $\theta_t \in \partial h(x_t)$

(ii) $\theta_t - \xi_t \in \partial h(x_{t+1})$ *and* $x_{t+1} = \nabla h^*(\theta_t - \xi_t)$

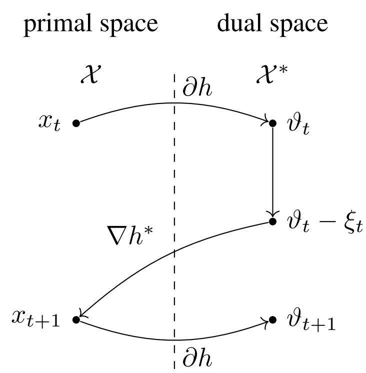

Figure 1: Iterates of UMD

Proof. Let $t \geq 1$, by definition of UMD iterates, $x_t = \nabla h^*(\theta_t)$, then by Lemma (3) $\theta_t \in \partial h(x_t)$, which establishes (i). For all $x \in \mathcal{X}$, we deduce from $\theta_{t+1} \in \partial h(x_{t+1})$ that

$$h(x) - h(x_{t+1}) \geq \langle \theta_{t+1}, x - x_{t+1} \rangle \geq \langle \theta_t - \xi_t, x - x_{t+1} \rangle$$

where in the last inequality we used the variational condition (ii) of UMD iterates. Then we have that $\theta_t + \xi_t \in \partial h(x_{t+1})$, i.e. property (ii) in this proposition. ☐

Remark 2. *We can consider the following alternative definition of UMD iterates. Let $\Pi_h : \mathcal{X} \rightrightarrows \mathcal{X} \times \mathcal{F}$ be a multi-valued prox-mapping defined as follows.*

$$\Pi_h(\zeta) := \{(x, \theta) \mid x = \nabla h^*(\zeta) \ , \theta \in \partial h(x) \ , \langle \theta - \zeta, x' - x \rangle \geq 0 \ , \forall x' \in \mathcal{X}\}$$

Then $(x_t, \theta_t)_{t \geq 1}$ is a sequence of $\mathrm{UMD}(h, \xi)$ *iterates*[5] *if and only if*

$$\begin{aligned} \theta_1 \in& \partial h(x_1) \\ (x_{t+1}, \theta_{t+1}) \in& \Pi_h(\theta_t - \xi_t) \ , \qquad t \geq 1 \end{aligned}$$

[5]We know that UMD spaces are superreflexive. A problem in Fixed Point Theory asks whether superrefelxive spaces have the fixed point property (fpp) or the weak fixed point property (w-fpp). If the answer is positive, we can then hope for an extension of UMD iterates by generalizations of this mapping. It is interesting to ask if such mapping is non-expansive, and since its domain is supposed to be bounded, closed, and convex, then it has a fixed point. We might also ask what the meaning of such point is in this context.

Proof of Theorem 6. By Theorem (5) a UMD space of co-type q is isomorphic to a q-uniformly convex space. Let $||\cdot||$ be the norm (of power type q) of such space. By Theorem 1, the function $h = ||\cdot||^p$ is q-uniformly convex, then by (2) it is an $\mathcal{X}$-regularizer, so $\mathrm{UMD}(h, \xi)$ algorithm is well defined. Now, for any $x \in \operatorname{dom} h$

$$\begin{aligned}\langle \xi_t, x_{t+1} - x\rangle \leq & \langle \theta_{t+1} - \theta_t, x - x_{t+1}\rangle \\ =& \langle \theta_{t+1}, x - x_{t+1}\rangle - \langle \theta_t, x - x_{t+1}\rangle + \langle \theta_t, x_{t+1} - x_t\rangle \\ =& \left(h(x) - h(x_t) - \langle \theta_t, x - x_t\rangle\right) \\ & \quad - \left(h(x) - h(x_{t+1}) - \langle \theta_{t+1}, x - x_{t+1}\rangle\right) \\ & \quad - \left(h(x_{t+1}) - h(x_t) - \langle \theta_t, x_{t+1} - x_t\rangle\right) \\ =& D_h(x, x_t; \theta_t) - D_h(x, x_{t+1}; \theta_{t+1}) - D_h(x_{t+1}, x_t; \theta_t) \end{aligned} \tag{1}$$

where the last equality follows from the definition of Bregman divergence since, by Proposition (10) $\theta_t \in \partial h(x_t)$, and $\theta_{t+1} \in \partial h(x_{t+1})$.

Now, from the definition of Bregman Divergence

$$\langle \xi_t, x_t - x_{t+1}\rangle = D_h(x_{t+1}, x_t; \theta_t) + D_h(x_t, x_{t+1}; \theta_t - \xi_t)$$

and the latter divergence is well defined since $\theta_t - \xi_t \in \partial h(x_{t+1})$, again by Proposition (10). Moreover

$$D_h(x_t, x_{t+1}; \theta_t - \xi_t) = D_{h^*}(\theta_t - \xi_t, \theta_t; x_t) = D_h^*(\theta_t - \xi_t, \theta_t)$$

where the first equality is a consequence of Lemma (3), and the second comes from the differentiability of h^*. Then, combining the two previous expressions and adding to (1) gives

$$\begin{aligned}\langle \xi_t, x_t - x\rangle &\leq D_h(x, x_t; \theta_t) - D_h(x, x_{t+1}; \theta_t) + D_{h^*}(\theta_t - \xi_t, \theta_t) \\ &\leq D_h(x, x_t; \theta_t) - D_h(x, x_{t+1}; \theta_t) + \frac{1}{p}||\xi_t||_*^p\end{aligned}$$

where the last inequality follows from the q-uniformly convexity of h, and p is the type. Finally, summing over $t = 1 : n$ yields,

$$\begin{aligned}\sum_{t=1}^{n} \langle \xi_t, x_t - x\rangle \leq & D_h(x, x_1; \theta_1) - D_h(x, x_{n+1}; \theta_n) + \frac{1}{p}\sum_{t=1}^{n} ||\xi_t||_*^p \\ \leq & \sup_{x \in \mathcal{X}} D_h(x, x_1; \theta) + \frac{1}{p}\sum_{t=1}^{n} ||\xi_t||_*^p\end{aligned}$$

and the result follows once we take the infimum over $\mathcal{X}$ in the LHS. $\square$

Lemma 11. *Let $\mathcal{X}$ be the unit ball of $\mathcal{B}$, then*

$$\sup_{x^*_{1:n} \in \mathcal{X}^*} \mathbb{E}\left[\left|\left|\sum_{t=1}^{n} \epsilon_t x_t^*\right|\right|\right] = \sup_{x^*_{1:n} \in \mathcal{X}^*} \mathbb{E}\left[\sup_{x \in \mathcal{X}} \left(\sum_{t=1}^{n} \epsilon_t \langle x_t^*, x\rangle\right)\right] \leq \mathcal{V}_n(\mathcal{X}, \mathcal{X}^*)$$

with $(\epsilon_t)_{t=1}^n$ a Rademacher sequence.

Proof. Let Q be a distribution over $\mathcal{X}^{*n}$. Note that

$$\mathcal{V}_n(\mathcal{X}, \mathcal{X}^*) \geq \sup_{Q} \inf_{\mathcal{A}} \mathbb{E}_{x^*_{1:n} \sim Q}[R_n(\mathcal{A}, x^*_{1:n})]$$

Define $V_Q = \inf_{\mathcal{A}} \mathbb{E}_{x^*_{1:n} \sim Q}[R_n(\mathcal{A}, x^*_{1:n})]$. Consider the following distribution Q: fix $x^{*\prime}_{1:n} \in \mathcal{X}^*$ and choose iid Rademacher random variables $\epsilon_{1:n}$. Let $x^*_t = \epsilon_t x^{*\prime}_t$. Under this distribution, for any

$x_t \in \mathcal{X}$, $\mathbb{E}_{t-1}[-\langle x_t^*, x_t\rangle] = -\langle x_t^{*\prime}, x_t\rangle \mathbb{E}_{t-1}[\epsilon_t] = 0$. Using the definition of regret, and Q as above:

$$\begin{aligned} V_Q &\geq \inf_{\mathcal{A}} \mathbb{E}\left[\sum_{t=1}^n \langle x_t^*, \mathcal{A}_t(x_{1:t-1}^*)\rangle\right] - \mathbb{E}\inf_{x\in\mathcal{X}}\sum_{t=1}^n \langle x_t^*, x\rangle \\ &= \inf_{\mathcal{A}}\left[\sum_{t=1}^n \mathbb{E}\mathbb{E}_{t-1}\langle x_t^*, \mathcal{A}_t(x_{1:t-1}^*)\rangle\right] - \mathbb{E}\inf_{x\in\mathcal{X}}\sum_{t=1}^n \langle x_t^*, x\rangle \\ &= \sum_{t=1}^n \inf_{\mathcal{A}} \mathbb{E}\mathbb{E}_{t-1}\langle x_t^*, \mathcal{A}_t(x_{1:t-1}^*)\rangle - \mathbb{E}\inf_{x\in\mathcal{X}}\sum_{t=1}^n \langle x_t^*, x\rangle \\ &= \sum_{t=1}^n \mathbb{E}\left[\inf_{x_t\in\mathcal{X}} \mathbb{E}_{t-1}\langle x_t^*, x_t\rangle\right] - \mathbb{E}\inf_{x\in\mathcal{X}}\sum_{t=1}^n \langle x_t^*, x\rangle \\ &= -\sum_{t=1}^n \mathbb{E}\left[\sup_{x_t\in\mathcal{X}} \mathbb{E}_{t-1} - \langle x_t^*, x_t\rangle\right] + \mathbb{E}\sup_{x\in\mathcal{X}}\sum_{t=1}^n -\langle x_t^*, x\rangle \\ &= \mathbb{E}\sup_{x\in\mathcal{X}}\sum_{t=1}^n -\langle \epsilon_t x_t^{*\prime}, x\rangle \end{aligned}$$

Then we have

$$\mathcal{V}_n(\mathcal{X}, \mathcal{X}^*) \geq \sup_Q V_Q \geq \sup_{x_{1:n}^{*\prime}} \mathbb{E}\sup_{x\in\mathcal{X}} \sum_{t=1}^n -\langle \epsilon_t x_t^{*\prime}, x\rangle \geq \sup_{x_{1:n}^{*\prime}} \mathbb{E}\sup_{x\in\mathcal{X}} \sum_{t=1}^n \langle \epsilon_t x_t^{*\prime}, x\rangle$$

and the first identity in the Lemma is common and follows from Hahn-Banach.

□

Proof of Theorem 8. We claim that

$$n^{\frac{1}{p(\mathcal{B})}} \leq \mathcal{V}_n(\mathcal{X}, cvx)$$

Since $\mathcal{V}_n(\mathcal{X}, cvx)$ is sublinear, then $0 = \lim_n n^{-1}\mathcal{V}_n(\mathcal{X}, cvx) \geq \lim_n n^{\frac{1}{p(\mathcal{B})}-1}$ so $p(\mathcal{B}) > 1$, i.e. $\mathcal{B}$ has non-trivial type. By Theorem (5) it has non-trivial martingale cotype.

We now prove the claim. Without loss of generality we can consider $\mathcal{X}$ to be the unit ball of $\mathcal{B}$. Theorem 14 in [AR08] yields that $\mathcal{V}_n(\mathcal{X}, cvx) = \mathcal{V}_n(\mathcal{X}, lin)$, so it is enough to consider the latter class[6].

By Theorem 3.3 and Remark 3.4 of [Pis86], there exists $x_{1:n}^* \in \mathcal{X}^*$ such that for any $\epsilon_{1:n}$

$$\left|\left|\sum_{t=1}^n \epsilon_t x_t^*\right|\right| \geq \left(\sum_{t=1}^n |\epsilon_t|^{p(\mathcal{B})}\right)^{\frac{1}{p(\mathcal{B})}} = n^{\frac{1}{p(\mathcal{B})}}$$

and we conclude using Lemma (11). □

Proof of Theorem 9. Following Lemma (11), and by hypothesis, we have that for any $n \geq 1$, any Rademacher sequence $\epsilon_{1:n}$, and $\mathcal{X}^*$-valued martingales $(x_t^*)_{t=1}^n$

$$\mathbb{E}\left[\left|\left|\sum_{t=1}^n \epsilon_t x_t^*\right|\right|\right] \leq \mathcal{V}_n(\mathcal{X}, cvx) \leq Cn^{\frac{1}{r}} \leq C'(n+1)^{\frac{1}{r}} \sup_{1\leq t\leq n} ||dx_t^*||_\infty$$

with constants $C, C' \geq 0$ and $r > 1$. We conclude using Lemma 3.1 of [Pis75].

□

[6] In fact, $lin = \mathcal{X}^*$.

5 Conclusion

We have almost characterized the spaces in which convex adversarial bandits achieve sublinear regret. Note that all results hold verbatim if instead of UMD spaces, we consider Banach lattices of type $p > 1$. Moreover, we can consider the class of spaces with non-trivial type such that the type and martingale type index agree. What would be interesting is to characterize, in terms of geometric properties, which spaces are those.

5.1 Future Work

There are several lines for future work. We leave out the obvious extensions like adaptive learning rate as in [FHPF20], or acceleration like in [JKM20]. The first extension we outline is to consider *continuous time*. This in spirit of the results in [BKTB16] about regret minimization in continuous time on reflexive Banach spaces using Dual Averaging. This appears to be straightforward by adjusting the family $\mathcal{F}$ to be (i) locally integrable, meaning for $x^* \in \mathcal{F}$ and for any $x \in \mathcal{X}$, the map $t \longmapsto \langle x_t^*, x\rangle$ be Lebesgue integrable on any compact set $K \subseteq [0, \infty)$; and (ii) to be uniformly bounded, i.e. $\sup_{x\in\mathcal{X}} |\langle x_t^*, x\rangle| \leq M$ for all t.

The second line of future work is to give *high probability bounds* for the regret of UMD. Using concentration bounds from [Pin94; Pis75].

Finally, we want to give a precise formulation of the following Conjecture.

Conjecture 1. *$\mathcal{B}$ is UMD if and only if for a closed convex symmetric bounded subset $\mathcal{X} \subseteq \mathcal{B}$, $(\mathcal{X} \subseteq \mathcal{B}, cvx)$ is UMD-learnable.*

Observe that the sufficiency is given in Theorem (6). Moreover, if $(\mathcal{X} \subseteq \mathcal{B}, cvx)$ is UMD-learnable, then it has both a non-trivial type and martingale type. We end by asking: Do the types coincide without extra assumptions? If not, what extra conditions do we need to impose?

References

[Agr95] Rajeev Agrawal, *The continuum-armed bandit problem*, SIAM journal on control and optimization **33** (1995), no. 6, 1926–1951.

[AR08] Jacob Abernethy and Alexander Rakhlin, *Optimal strategies and minimax lower bounds for online convex games*, In Proceedings of the Nineteenth Annual Conference on Computational Learning Theory, 2008.

[BGHV09] J. Borwein, A. J. Guirao, P. Hájek, and J. Vanderwerff, *Uniformly convex functions on banach spaces*, Proceedings of the American Mathematical Society **137** (2009), no. 3, 1081–1091.

[BKTB16] Maximilian Balandat, Walid Krichene, Claire Tomlin, and Alexandre Bayen, *Minimizing regret on reflexive banach spaces and learning nash equilibria in continuous zero-sum games*, 2016.

[BSY11] Sébastien Bubeck, Gilles Stoltz, and Jia Yuan Yu, *Lipschitz bandits without the lipschitz constant*, Algorithmic Learning Theory (Berlin, Heidelberg) (Jyrki Kivinen, Csaba Szepesvári, Esko Ukkonen, and Thomas Zeugmann, eds.), Springer Berlin Heidelberg, 2011, pp. 144–158.

[DKC13] Josip Djolonga, Andreas Krause, and Volkan Cevher, *High-dimensional gaussian process bandits*, Advances in Neural Information Processing Systems (C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, eds.), vol. 26, Curran Associates, Inc., 2013.

[DT14] R. A. DeVore and V. N. Temlyakov, *Convex optimization on banach spaces*, 2014.

[FHH+10] M. Fabian, P. Habala, P. Hájek, V. Montesinos, and V. Zizler, *Banach space theory: The basis for linear and nonlinear analysis*, CMS Books in Mathematics, Springer New York, 2010.

[FHPF20] Huang Fang, Nicholas J. A. Harvey, Victor S. Portella, and Michael P. Friedlander, *Online mirror descent and dual averaging: keeping pace in the dynamic case*, 2020.

[HvNVW16] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis, *UMD spaces*, Analysis in Banach Spaces, Springer International Publishing, 2016, pp. 267–372.

[JKM20] Anatoli Juditsky, Joon Kwon, and Eric Moulines, *Unifying mirror descent and dual averaging*, 2020.

[LWHZ19] Shiyin Lu, Guanghui Wang, Yao Hu, and Lijun Zhang, *Optimal algorithms for Lipschitz bandits with heavy-tailed rewards*, Proceedings of the 36th International Conference on Machine Learning (Kamalika Chaudhuri and Ruslan Salakhutdinov, eds.), Proceedings of Machine Learning Research, vol. 97, PMLR, 09 2019, pp. 4154–4163.

[McM11] H. Brendan McMahan, *A unified view of regularized dual averaging and mirror descent with implicit updates*, 2011.

[MCP14] Stefan Magureanu, Richard Combes, and Alexandre Proutière, *Lipschitz bandits: Regret lower bounds and optimal algorithms*, CoRR **abs/1405.4758** (2014).

[Pin94] Iosif Pinelis, *Optimum Bounds for the Distributions of Martingales in Banach Spaces*, The Annals of Probability **22** (1994), no. 4, 1679 – 1706.

[Pis75] Gilles Pisier, *Martingales with values in uniformly convex spaces*, Israel Journal of Mathematics **20** (1975), no. 3-4, 326–350.

[Pis86] ______, *Probabilistic methods in the geometry of banach spaces*, Lecture Notes in Mathematics, Springer Berlin Heidelberg, 1986, pp. 167–241.

[SC06] L. Schwartz and P.R. Chernoff, *Geometry and probability in banach spaces*, Lecture Notes in Mathematics, Springer Berlin Heidelberg, 2006.

[SJ18] Shubhanshu Shekhar and Tara Javidi, *Gaussian process bandits with adaptive discretization*, Electronic Journal of Statistics **12** (2018), no. 2, 3829 – 3874.

[SKKS09] Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias W. Seeger, *Gaussian process bandits without regret: An experimental design approach*, CoRR **abs/0912.3995** (2009).

[SST11] Nathan Srebro, Karthik Sridharan, and Ambuj Tewari, *On the universality of online mirror descent*, CoRR **abs/1107.4080** (2011).

[WYGR20] Tianyu Wang, Weicheng Ye, Dawei Geng, and Cynthia Rudin, *Towards practical lipschitz bandits*, Proceedings of the 2020 ACM-IMS on Foundations of Data Science Conference (New York, NY, USA), FODS ’20, Association for Computing Machinery, 2020, p. 129–138.

[Xia09] Lin Xiao, *Dual averaging method for regularized stochastic learning and online optimization*, Advances in Neural Information Processing Systems (Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, and A. Culotta, eds.), vol. 22, Curran Associates, Inc., 2009.