ECE 543 Final Report
Minimax Bounds for Online Learning

Amr Martini

May 2019

1 Introduction

In this report, I will outline a few various proofs of lower bounds for online learning. A natural choice for lower bound analysis in online learning is minimax analysis, since the framework is already there, and it also represents the greatest lower bound that could be achieved. Since the focus of online learning is about regret, we should not be concerned with just any lower bound, because regret is naturally bounded below by 0 (detailed in 2). Rather, finding the greatest lower bound should be of particular interest as it finishes the search for lower bounds.

In section 2, I detail my first attempt in obtaining a lower bound based on duality and also what ended up being a rigorous way of showing that regret is bounded below by 0. Then in section 2.1, I show an instance of how using a duality technique can offer an alternate proof of generalization of online algorithms. In section 3, I show a proof of a lower bound that asymptotically close to the online gradient descent upper bound. In 4, I show a proof that I modified slightly Finally, in 5

The proofs in the subsequent sections all focus on the case of linear loss functions, which can be shown to be the hardest case for the player in the min-max sense [1]. That is, if the adversary had a choice between linear loss functions and non-linear convex functions or strongly convex functions, the adversary would choose the linear ones. Also, while the notation of the proofs is based on [1], note that ultimately it is not the same. I introduced and modified small steps and some larger steps, removed needless complexity or added complexity where it was needed, and tried to not deviate too far from the notation used in the course.

2 First Attempt at a Lower Bound

The goal in online learning is to minimize regret. The proof aimed to derive a lower bound for regret via minimax analysis, and in the context of game theory, this means solving for the minimax value of the game. So let,

- $\mathcal{X} \subset \mathbb{R}^d$, $D := \max_{x,x'} ||x - x'||$, so $\mathcal{X}$ is a compact, convex subset of $\mathbb{R}^d$
- $\mathcal{F}$ be some function space of loss functions (later we restrict this to linear functions)
- T be the number of rounds that the player and adversary play (i.e. $1 \leq t \leq T$)
- the player choose $x_t \in \mathcal{X}$ upon round t
- the adversary then choose $f_t \in \mathcal{F}$ upon round t
- $||\nabla f|| \leq L$ for all $f \in \mathcal{F}$
- the game $\mathcal{G}$ be defined by the tuple of strategies for both players, $(\mathcal{X}, \mathcal{F})$, and the regret function: $R_T = \sum_{t=1}^{T} f_t(x_t) - \inf_{x_t \in \mathcal{X}} \sum_t f_t(x)$

As usual, the regret is defined as the difference in objective value between a dynamic strategy and the best fixed strategy up to time T.

Since we are trying to solve for the minimax value of a repeated game, we can be more explicit about what we mean by minimax. Specifically, we are trying to solve:

$$\inf_{x_1 \in \mathcal{X}} \sup_{f_1 \in \mathcal{F}} \dots \inf_{x_T \in \mathcal{X}} \sup_{f_T \in \mathcal{F}} \Big[\sum_{t=1}^{T} f_t(x_t) - \inf_{x \in \mathcal{X}} \sum_{t=1}^{T} f_t(x) \Big]$$

For the specific case of linear functions, the function space is restricted to:

$$\mathcal{F} = \{ f(x) = w \cdot x : ||w|| < L \}$$

or to simplify subsequent notation, we can define $\mathcal{W}$, the ball of radius L centered at 0:

$$\mathcal{W} = \{ w : ||w|| \leq L \}$$

For the first derivation, we require that the set of player strategies is also restricted to a ball, with radius $\frac{D}{2}$, diameter D, centered at 0:

$$\mathcal{X} = \{ x : ||x|| \leq \frac{D}{2} \}$$

The minimax value of the game becomes,

$$V_1 = \inf_{x_1 \in \mathcal{X}} \sup_{w_1 \in \mathcal{W}} \dots \inf_{x_T \in \mathcal{X}} \sup_{w_T \in \mathcal{W}} \Big[\sum_{t=1}^{T} w_t \cdot x_t - \inf_{x \in \mathcal{X}} x \cdot \sum_{t=1}^{T} w_t \Big]$$

Whenever $\mathcal{X}$ and $\mathcal{W}$ are balls, or more generally, whenever the set of unit vectors in $\mathcal{W}$ is contained in $\mathcal{X}$, we can simplify the fixed strategy term in the regret function.

$$\begin{aligned} - \inf_{x \in \mathcal{X}} x \cdot \sum_{t=1}^{T} w_t &= \max_{x \in \mathcal{X}} ||x|| \Big|\Big| \sum_{t} w_t \Big|\Big| \\ &= \frac{D}{2} \Big|\Big| \sum_{t=1}^{T} w_t \Big|\Big| \end{aligned}$$

The minimax regret becomes:

$$V_1 = \inf_{x_1 \in \mathcal{X}} \sup_{w_1 \in \mathcal{W}} \dots \inf_{x_T \in \mathcal{X}} \sup_{w_T \in \mathcal{W}} \Big[\sum_{t=1}^{T} w_t \cdot x_t + \frac{D}{2} \Big|\Big| \sum_{t=1}^{T} w_t \Big|\Big| \Big] \tag{1}$$

Next, we invoke duality at the last stage to obtain a lower bound, and we repeat this process for each stage:

$$
\begin{aligned}
V_1 &\geq \inf_{x_1\in\mathcal{X}} \sup_{w_1\in\mathcal{W}} \dots \sup_{w_{T-1}\in\mathcal{W}} \inf_{x_{T-1}\in\mathcal{X}} \sup_{w_T\in\mathcal{W}} \inf_{x_T\in\mathcal{X}} \Big[\sum_{t=1}^{T} w_t \cdot x_t + \frac{D}{2}\Big|\Big|\sum_{t=1}^{T} w_t\Big|\Big|\Big] \\
&= \inf_{x_1\in\mathcal{X}} \sup_{w_1\in\mathcal{W}} \dots \sup_{w_{T-1}\in\mathcal{W}} \inf_{x_{T-1}\in\mathcal{X}} \sup_{w_T\in\mathcal{W}} \Big[\sum_{t=1}^{T-1} w_t \cdot x_t - ||w_T||\frac{D}{2} + \frac{D}{2}\Big|\Big|\sum_{t=1}^{T} w_t\Big|\Big|\Big] \\
&= \inf_{x_1\in\mathcal{X}} \sup_{w_1\in\mathcal{W}} \dots \sup_{w_{T-1}\in\mathcal{W}} \inf_{x_{T-1}\in\mathcal{X}} \Big[\sum_{t=1}^{T-1} w_t \cdot x_t + \frac{D}{2}\Big|\Big|\sum_{t=1}^{T-1} w_t\Big|\Big|\Big] \\
&\dots \\
&\geq 0
\end{aligned}
$$

In the third step, we use the fact that $\Big|\Big|\sum_{t=1}^{T} w_t\Big|\Big| - ||w_t|| \leq \Big|\Big|\sum_{t=1}^{T} w_t - w_t\Big|\Big| = \Big|\Big|\sum_{t=1}^{T-1} w_t\Big|\Big|$, and this inequality is tight if w_t points in the same direction as $\Big|\Big|\sum_{t=1}^{T} w_t\Big|\Big|$. Due to the fact that $\mathcal{W}$ is a ball, we can always find such a w_t.

The interpretation of this derivation is that, the max-min game is, firstly, when the adversary goes first. The adversary can guarantee a 0 regret by consistently playing the same strategy, forcing the player into a fixed strategy. Because the player is playing optimally, naturally the regret is zero.

2.1 A new result through min-max duality

The conclusion above is not useful, but it turns out that there are additional assumptions on the game that allow for strong duality to hold, and we can get a meaningful interpretation out of the max-min game. Specifically, this analysis is used in a recent paper to redefine min-max regret as worst-case regret over all adversarial distributions [2]. To define their theorem, we redefine some sets and quantities we've used earlier.

Theorem 1. *If $\mathcal{F}$ is a compact, convex set from which the player chooses actions f_t on round t, if $\mathcal{Z}$ is a closed, compact set from which the adversary chooses a probability distribution P that is over this set, and if $l : \mathcal{F}, \mathcal{Z} \to \mathbb{R}$ is the loss function for the game and is convex in the first argument, and with all other constants defined as before, then*

$$
R_T = \sup_P \mathbb{E}\left[\sum_{t=1}^{T} \inf_{f_t\in\mathcal{F}} \mathbb{E}\left[l(f_t, Z_t)|Z^{t-1}\right] - \inf_{f\in\mathcal{F}} \sum_{t=1}^{T} l(f, Z_t)\right] \tag{2}
$$

where Z^t is the set of all adversary choices up to and including time t and Z^0 is taken to be as the empty set.

Using more familiar notation, we can write this equation as,

$$
R_T = \sup_P \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} L(f_t) - L_n^*\right] \tag{3}
$$

where f_t specifically denotes the use of a min-max optimal strategy. Using a mixture strategy $\bar{f}_T = \frac{1}{T}\sum_{t=1}^{T} f$ and the fact that l is convex in the first argument, it is clear that,

$$
\sup_P \mathbb{E}\left[L(\bar{f}_T)\right] - L_n^* \leq R_T
$$

Finally, we can show that if R_T is bounded, then the min-max optimal strategy is consistent.

$$\sup_P \underbrace{\mathbb{E}\left[L(\bar{f}_T) - \frac{1}{T}\sum_{t=1}^{T} l(f_t, Z_t)\right]}_{g_T((f_t))} + \underbrace{\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} l(f_t, Z_t) - L_n*\right]}_{e_T((f_t))} \leq R_T$$

This concludes an alternate proof of consistency and generalization of the min-max optimal online learning strategy. See [2] for more details on Theorem 1.

3 A Lower Bound on Regret

(Based on notation in [1], proof is based on [3]).

- Let $\mathcal{W}_{\mathcal{C}} = \{-L/\sqrt{d}, L/\sqrt{d}\}^d$, a cube of dimension d
- T is the number of rounds that the player and adversary play (i.e. $1 \leq t \leq T$)
- Player chooses $x_t \in \mathcal{X}$ upon round t, and $\mathcal{X}$ has diameter D
- Adversary chooses $w_t \in \mathcal{W}_{\mathcal{C}}$ upon round t
- The condition $||\nabla f|| \leq L$ is still satisfied
- The game $\mathcal{G}$ is defined by the tuple of strategies for both players, $(\mathcal{X}, \mathcal{F})$, and the regret function: $R_T = \sum_{t=1}^{T} f_t(x_t) - \inf_{x_t \in \mathcal{X}} \sum_t f_t(x)$

Following the same notation as in the previous section, we begin by recalling the regret expression, then obtaining lower bound:

$$\begin{aligned}
V_1 &= \inf_{x_1 \in \mathcal{X}} \sup_{w_1 \in \mathcal{W}} \dots \inf_{x_T \in \mathcal{X}} \sup_{w_T \in \mathcal{W}} \left[\sum_{t=1}^{T} w_t \cdot x_t - \inf_{x \in \mathcal{X}} x \cdot \sum_{t=1}^{T} w_t\right] \\
&\geq \inf_{x_1 \in \mathcal{X}} \sup_{w_1 \in \mathcal{W}_{\mathcal{C}}} \dots \inf_{x_T \in \mathcal{X}} \sup_{w_T \in \mathcal{W}_{\mathcal{C}}} \left[\sum_{t=1}^{T} w_t \cdot x_t - \inf_{x \in \mathcal{X}} x \cdot \sum_{t=1}^{T} w_t\right] \\
&\geq \inf_{x_1 \in \mathcal{X}} \mathbb{E}_{w_1} \dots \inf_{x_T \in \mathcal{X}} \mathbb{E}_{w_T} \left[\sum_{t=1}^{T} w_t \cdot x_t - \inf_{x \in \mathcal{X}} x \cdot \sum_{t=1}^{T} w_t\right]
\end{aligned}$$

where in the second line, we get an inequality due to the fact that we restricted the adversary to a smaller set of plays, and in the third line, $\mathbb{E}_{w_t}$ is any valid expectation over the cube $\mathcal{W}_{\mathcal{C}}$, so we can restrict attention to the uniform distribution over the cube. Now we analyze a particular stage so that the expression can recurse backwards into a simple expression:

$$\begin{aligned}
\inf_{x_t \in \mathcal{X}} \mathbb{E}_{w_t}\left[\sum_{k=1}^{t} w_k \cdot x_k - \inf_{x \in \mathcal{X}} x \cdot \sum_{k=1}^{t} w_k\right] &= \inf_{x_t \in \mathcal{X}} \left(\mathbb{E}_{w_t}\left[\sum_{k=1}^{t} w_k \cdot x_k\right] - \mathbb{E}_{w_t}\left[\inf_{x \in \mathcal{X}} x \cdot \sum_{k=1}^{t} w_k\right]\right) \\
&= \inf_{x_t \in \mathcal{X}} \left(\sum_{k=1}^{t-1} w_k \cdot x_k + \mathbb{E}_{w_t}\left[w_t \cdot x_t\right] - \mathbb{E}_{w_t}\left[\inf_{x \in \mathcal{X}} x \cdot \sum_{k=1}^{t} w_k\right]\right) \\
&= \sum_{k=1}^{t-1} w_k \cdot x_k - \mathbb{E}_{w_t}\left[\inf_{x \in \mathcal{X}} x \cdot \sum_{k=1}^{t} w_k\right]
\end{aligned}$$

which has the same form as we started off with but up to time T. Recursing backwards, we get,

$$V_1 \geq -\mathbb{E}_{w_1 \ldots w_T}\Big[\inf_{x \in \mathcal{X}} x \cdot \sum_{t=1}^{T} w_t\Big]$$

Note that we have not made any assumptions about $\mathcal{X}$ so far, besides that it has diameter D. Actually, $\mathcal{X}$ just has to include a cube of at least the dimensionality of $\mathcal{W}_{\mathcal{C}}$. In general and as mentioned earlier, these proof remains identical for linear functions as long as the set $\mathcal{X}$ includes the same directions as $\mathcal{W}$, and the length of each of those directions is $\frac{D}{2}$. However, as will immediately be seen, the restriction on $\mathcal{W}$ should be one such that, if $w \in \mathcal{W}$, then $-w \in \mathcal{W}$.

Finally, we replace the inner product with a sum over coordinates. Since we are guaranteed that every direction in $\mathcal{W}_{\mathcal{C}}$ is also in $\mathcal{X}$, the regret bound becomes:

$$V_1 \geq -\mathbb{E}_{w_1 \ldots w_T}\Big[\sum_{i=1}^{d} -\frac{D}{2\sqrt{d}}\Big|\sum_{t=1}^{T} w_{t,i}\Big|\Big]$$

Instead of marginalizing over the coordinates of $w_1, \ldots, w_T$, we marginalize over Rademacher random variables $\epsilon_{i,t}$ in the first step, and then over new variables ϵ_t in the second step, recalling that for every $w \in \mathcal{W}_{\mathcal{C}}$, $-w \in \mathcal{W}_{\mathcal{C}}$ as well:

$$\begin{aligned} V_1 &\geq \frac{D}{2\sqrt{d}} \sum_{i=1}^{d} \mathbb{E}_{\epsilon_{i,t}}\Big[\Big|\sum_{t=1}^{T} \epsilon_{i,t} \frac{L}{\sqrt{d}}\Big|\Big] \\ &= \frac{DL}{2} \mathbb{E}_{\epsilon_t}\Big[\Big|\sum_{t=1}^{T} \epsilon_t\Big|\Big] \\ &\geq \frac{DL\sqrt{T}}{2\sqrt{2}}, \quad \textbf{(Khintchine inequality)} \end{aligned}$$

4 Tight Lower Bound on Regret

(Based on [1], with the simplification of 4 being my own introduction).

- $\mathcal{X} \subset \mathbb{R}^d$ is a ball of radius $\frac{D}{2}$
- $\mathcal{W} \subset \mathbb{R}^d$ is a ball of radius L
- $d \geq 3$
- All other assumptions hold as before

Starting from equation 1 in the first proof,

$$\begin{aligned} V_1 &= \inf_{x_1 \in \mathcal{X}} \sup_{w_1 \in \mathcal{W}} \ldots \inf_{x_T \in \mathcal{X}} \sup_{w_T \in \mathcal{W}} \Big[\sum_{t=1}^{T} w_t \cdot x_t - \inf_{x \in \mathcal{X}} x \cdot \sum_{t=1}^{T} w_t\Big] \\ &= \inf_{x_1 \in \mathcal{X}} \sup_{w_1 \in \mathcal{W}} \ldots \inf_{x_T \in \mathcal{X}} \sup_{w_T \in \mathcal{W}} \Big[\sum_{t=1}^{T} w_t \cdot x_t + \frac{D}{2}\Big|\Big|\sum_{t=1}^{T} w_t\Big|\Big|\Big] \end{aligned}$$

The theorem is proved by squeezing an upper and lower bound. Namely, if we show that,

1. $\frac{DL\sqrt{T}}{2} \leq V_1$
2. $V_1 \leq \frac{DL\sqrt{T}}{2}$

then we would have proved the result.

The upper bound can be proved using the classic result, but to draw parallels with the derivation of the lower bound, we will highlight a new method to derive the upper bound. The upper and lower bounds can be split up as two different cases:

1. We propose an adversarial strategy and try to find the best player strategy
2. We propose a player strategy and try to find the best adversarial strategy

4.1 Case 1

Let the adversarial strategy be one that satisfies the following constraints:

- $||w_t|| = L$
- $w_t \cdot x_t = 0$
- $w_t \cdot \sum_{k=1}^{t-1} w_k = 0$

If $d \geq 3$, then we are guaranteed to find such a strategy. This strategy implies:

1. $\sum_{t=1}^{T} w_t \cdot x_t = 0$
2. $\left|\left| \sum_{t=1}^{T} w_t \right|\right| = L\sqrt{T}$

(1) implies that player's choice does not affect regret. As we will see, the inf chain collapses. (2) is proven below:

Proof by induction:

Assume $\left|\left| \sum_{t=1}^{T-1} w_t \right|\right| = L\sqrt{T-1}$. Base case is proven by assumptions.

$$\begin{aligned}
\left|\left| \sum_{t=1}^{T} w_t \right|\right| &= \left|\left| w_T + \sum_{t=1}^{T-1} w_t \right|\right| \\
&= \sqrt{||w_T||^2 + \left|\left| \sum_{t=1}^{T-1} w_t \right|\right|^2 + w_T \cdot \sum_{t=1}^{T-1} w_t} \\
&= \sqrt{||w_T||^2 + \left|\left| \sum_{t=1}^{T-1} w_t \right|\right|^2} \\
&= \sqrt{L^2 + L^2(T-1)} \\
&= L\sqrt{T}
\end{aligned}$$

Now we reduce the regret expression:

$$
\begin{aligned}
V_1 &= \inf_{x_1 \in \mathcal{X}} \sup_{w_1 \in \mathcal{W}} \dots \inf_{x_T \in \mathcal{X}} \sup_{w_T \in \mathcal{W}} \Big[\sum_{t=1}^{T} w_t \cdot x_t + \frac{D}{2} \Big|\Big| \sum_{t=1}^{T} w_t \Big|\Big| \Big] \\
&\geq \inf_{x_1 \in \mathcal{X}} \dots \inf_{x_T \in \mathcal{X}} \left[\frac{D}{2} \Big|\Big| \sum_{t=1}^{T} w_t \Big|\Big| \right] \\
&= \frac{D}{2} \Big|\Big| \sum_{t=1}^{T} w_t \Big|\Big| \\
&= \frac{D}{2} L \sqrt{T}
\end{aligned}
$$

4.2 Case 2

Let the player strategy be exactly the following:

$$
x_t = -\frac{\sum_{k=1}^{T-1} w_k}{\sqrt{\phi_t}} D^2/4
$$

where,

$$
\sqrt{\phi_t} = \frac{D}{2} \sqrt{\Big|\Big| \sum_{k=1}^{t-1} w_k \Big|\Big|^2 + L^2(T - t + 1)}
$$

Define $\Phi_t(w_1, \dots, w_{t-1}) = \sum_{k=1}^{t-1} w_k \cdot x_k + \sqrt{\phi_t}$. Thus, $\Phi_1 = \frac{D}{2} L \sqrt{T}$.

Define $V_t(w_1, \dots, w_{t-1}) = \sup_{w_t} \dots \sup_{w_T} \left[\sum_{t=1}^{T} w_t \cdot x_t + \frac{D}{2} \Big|\Big| \sum_{t=1}^{T} w_t \Big|\Big| \right]$. $V_t(w_1, \dots, w_{t-1})$ would be the regret if we fix adversary strategies for $w_1, \dots, w_{t-1}$, and optimally choose $w_t, \dots, w_T$.

Then V_1 is the value of the game that we're interested in, and if $V_1 \leq \Phi_1$, then we have proved (2), and to prove this, it is sufficient to show that $V_t(w_1, \dots, w_{t-1}) \leq \Phi_t(w_1, \dots, w_{t-1})$ for all $t \in [T]$. In the next section, we will prove the claim.

Proof by induction:

Base for $t = T + 1$ should be obvious from definitions. Let's assume that the claim holds for $t + 1$. If we should that it holds for t, then we have proved the claim.

$$
\begin{aligned}
V_t(w_1,\ldots,w_{t-1}) &= \sup_{w_t} V_{t+1}(w_1,\ldots,w_t) \\
&\leq \sup_{w_t} \Phi_{t+1}(w_1,\ldots,w_t) \\
&= \sup_{w_t} \Big[\sum_{k=1}^{t} w_k \cdot x_k + \sqrt{\phi_{t+1}}\Big] \\
&= \sum_{k=1}^{t-1} w_k \cdot x_k + \sup_{w_t}\left[w_t \cdot x_t + \frac{D}{2}\sqrt{\Big|\Big|\sum_{k=1}^{t} w_k\Big|\Big|^2 + L^2(T-t)}\right] \\
&= \sum_{k=1}^{t-1} w_k \cdot x_k + \underbrace{\sup_{w_t}\left[w_t \cdot x_t + \frac{D}{2}\sqrt{\Big|\Big|\sum_{k=1}^{t-1} w_k + w_t\Big|\Big|^2 + L^2(T-t)}\right]}_{*}
\end{aligned}
$$

$$
* = \sup_{w_t}\left[-\frac{w_t \cdot \sum_{k=1}^{t-1} w_t}{\sqrt{\phi_t}}\frac{D^2}{4} + \frac{D}{2}\sqrt{\Big|\Big|\sum_{k=1}^{t-1} w_k\Big|\Big|^2 + L^2(T-t+1) + \underbrace{||w_t||^2 - L^2}_{\substack{\text{cannot be positive}\\ \text{for any choice of } w_t}} + w_t \cdot \sum_{k=1}^{t-1} w_k}\right] \tag{4}
$$

Note that the best choice of w_t must have $||w_t|| = L$, so we have,

$$
\begin{aligned}
* = &\sup_{w_t:||w_t||=L}\left[-\frac{w_t \cdot \sum_{k=1}^{t-1} w_t}{\sqrt{\phi_t}}\frac{D^2}{4} + \frac{D}{2}\sqrt{\frac{\phi_t}{D^2/4} + w_t \cdot \sum_{k=1}^{t-1} w_k}\right] \\
&= \sup_{\alpha}\left[-\frac{L||\sum_{k=1}^{t-1} w_k||}{\phi_t}\frac{D^2}{4}\sqrt{\phi_t}\cos\alpha + \sqrt{\phi_t}\sqrt{1 + \frac{L||\sum_{k=1}^{t-1} w_k||}{\phi_t}\frac{D^2}{4}\cos\alpha}\right] \\
&= \sup_{\alpha}\Big[-\lambda\sqrt{\phi_t}\cos\alpha + \sqrt{\phi_t}\sqrt{1+\lambda\cos\alpha}\Big] \\
&\leq \sup_{\alpha}\Big[-\lambda\sqrt{\phi_t}\cos\alpha + \sqrt{\phi_t}(1+\frac{\lambda}{2}\cos\alpha)\Big], \quad \textbf{(Bernoulli's inequality)} \\
&= \sup_{\alpha}\Big[\cos\alpha(\frac{\lambda}{2}\sqrt{\phi_t} - \lambda\sqrt{\phi_t}) + \sqrt{\phi_t}\Big] \\
&\leq \sqrt{\phi_t}
\end{aligned}
$$

Putting * back into the expression,

$$
V_t(w_1,\ldots,w_{t-1}) \leq \sum_{k=1}^{t-1} w_k \cdot x_k + \sqrt{\phi_t}
$$

which proves the claim.

5 Conclusion

Perhaps the most interesting result from all of this is that the min-max strategy for the assumptions given here "is exactly the Online Gradient Descent strategy of Zinkevich" [1] (actually, we also have this result for non-linear convex loss functions). This means that $\mathcal{X}$ is a ball of radius $\frac{D}{2}$ with dimension $d \geq 3$ and $\mathcal{W}$ is also a ball of radius L. While the authors in [1] did not fully outline the proof of equivalence between the min-max optimal strategy and OGD (online gradient descent) in their original paper, interestingly enough, the min-max value of the game is exactly OGD's upper bound, which means that in the min-max sense and under the assumptions layed out here, OGD is "squeezed" into being optimal.

In the strongly convex case, while the authors do propose a min-max optimal player strategy, they don't claim that it is exactly OGD. On the other hand, their result for the value of the min-max game when loss functions are strongly convex indicate that, under the same assumptions as usual, OGD is again "squeezed" into being optimal. Again, the min-max value of the game is exactly OGD's upper bound for strongly convex functions (i.e. $\frac{1}{2}\frac{L^2}{m}\log T$).

Finally, it would be interesting to see to how much of these assumptions can be relaxed while still obtaining min-max optimality of OGD. In analogy to this, while convexity and strong convexity play their parts, it was shown in [4] that it was not convexity that induces the familiar form of the regret bounds for FTPL. So it is not surprising that existing algorithms are already optimal in various senses. This goes to show that there are many interesting properties of these algorithms that we have yet to show.

References

[1] J. D. Abernethy, P. Bartlett, A. Rakhlin, and A. Tewari, "Optimal strategies and minimax lower bounds for online convex games," EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2008-19, Feb 2008. [Online]. Available: http://www2.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-19.html

[2] J. Abernethy, A. Agarwal, P. Bartlett, and A. Rakhlin, "A stochastic view of optimal regret through minimax duality," *COLT 2009 - The 22nd Conference on Learning Theory*, 04 2009.

[3] N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games.* New York, NY, USA: Cambridge University Press, 2006.

[4] A. S. Suggala and P. Netrapalli, "Online non-convex learning: Following the perturbed leader is optimal," *CoRR*, vol. abs/1903.08110, 2019. [Online]. Available: http://arxiv.org/abs/1903.08110