Optimal Strategies and Minimax Lower Bounds for Online Convex Games

Jacob Duncan Abernethy
Peter Bartlett
Alexander Rakhlin
Ambuj Tewari

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2008-19

http://www.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-19.html

February 22, 2008

Optimal Strategies and Minimax Lower Bounds for Online Convex Games

Jacob Abernethy
Peter L. Bartlett
Alexander Rakhlin
Division of Computer Science
UC Berkeley
{jake,bartlett,rakhlin}@cs.berkeley.edu

Ambuj Tewari
TTI Chicago
tewari@tti-c.org

February 22, 2008

Abstract

A number of learning problems can be cast as an Online Convex Game: on each round, a learner makes a prediction x from a convex set, the environment plays a loss function f, and the learner's long-term goal is to minimize regret. Algorithms have been proposed by Zinkevich, when f is assumed to be convex, and Hazan et al., when f is assumed to be strongly convex, that have provably low regret. We consider these two settings and analyze such games from a minimax perspective, proving minimax strategies and lower bounds in each case. These results prove that the existing algorithms are essentially optimal.

1 Introduction

It is rather unfortunate that the benefit of hindsight is only available post factum. Let us consider any scenario in which we are repeatedly asked to make a choice from some fixed set of options. Had we the foresight to know what choice would reap the largest long-term benefit, we would select this choice always without regret. Realistically, such prescience is not available to us and we must make decisions, on the fly, as information is given to us.

We may pose this problem more precisely as follows. We are given a set X and some set of functions $\mathcal{F}$. On each round $t = 1, \ldots, T$, we must choose some $\mathbf{x}_t$ from a set X. After we have made this choice, the *environment* chooses a function $f_t \in \mathcal{F}$. We incur a cost (loss) $f_t(\mathbf{x}_t)$, and the game proceeds to the next round. Of course, had we the fortune of perfect foresight and had access to the sum $f_1 + \ldots + f_T$, we would know the optimal choice $\mathbf{x}^* = \arg\min_{\mathbf{x}} \sum_{t=1}^T f_t(\mathbf{x})$. Instead, at time t, we will have only seen $f_1, \ldots, f_{t-1}$, and we must make the decision $\mathbf{x}_t$ with only historical knowledge. Thus, a natural long-term goal is to minimize the *regret*, defined as

$$\sum_{t=1}^T f_t(\mathbf{x}_t) - \inf_{\mathbf{x} \in X} \sum_{t=1}^T f_t(\mathbf{x}).$$

A special case of this setting is when the decision space X is a convex set and $\mathcal{F}$ is some set of convex functions on X. In the literature, this framework has been referred to as Online Convex Optimization (OCO), since our goal is to minimize a global function, i.e. $f_1 + f_2 + \cdots + f_T$, while this objective is revealed to us but one function at a time. Online Convex Optimization has attracted much interest in recent years [4, 9, 6, 1], as it provides a general analysis for a number of standard online learning problems including, among others, online classification and regression, prediction with expert advice, the portfolio selection problem, and online density estimation.

While instances of OCO have been studied over the past two decades, the general problem was first analyzed by Zinkevich [9], who showed that a very simple and natural algorithm, online gradient descent, elicits a bound on the regret that is on the order of $\sqrt{T}$. Online gradient descent can be described simply by the update $\mathbf{x}_{t+1} =$

$\mathbf{x}_t - \eta \nabla f_t(\mathbf{x}_t)$, where η is some parameter of the algorithm. This regret bound only required that f_t be smooth, convex, and with bounded derivative.

A regret bound of order $O(\sqrt{T})$ is not surprising: a number of online learning problems give rise to similar bounds. More recently, however, Hazan et al. [4] showed that when $\mathcal{F}$ consists of *curved* functions, i.e. f_t is strongly convex, then we get a bound of the form $O(\log T)$. It is quite surprising that curvature gives such a great advantage to the player. Curved loss functions, such as square loss or logarithmic loss, are very natural in a number of settings.

Finding algorithms that can guarantee low regret is, however, only half of the story; indeed, it is natural to ask "do better algorithms exist?" Without knowing whether such bounds are tight it remains to be seen if we can obtain even lower regret. The goal of the present paper is to address these questions, in some detail, for several classes of such online optimization problems.

This is achieved by a game-theoretic analysis: if we pose the above online optimization problem as a game between a Player who chooses $\mathbf{x}_t$ and an Adversary who chooses f_t, we may consider the regret achieved when each player is playing optimally. This is typically referred to as the *value* V_T of the game. In general, computing the value of zero-sum games is difficult, as we may have to consider exponentially many, or even uncountably many, strategies of the Player and the Adversary. Ultimately we will show that this value, as well as the optimal strategies of both the player and the adversary, can be computed *exactly and efficiently* for certain classes of online optimization games.

The central results of this paper are as follows:

- When the adversary plays *linear* loss functions, we use a known randomized argument to lower bound the value V_T. We include this mainly for completeness.
- We show that indeed this same linear game can be solved *exactly* for the case when the input space X is a ball, and we provide the optimal strategies for the player and adversary.
- We perform a similar analysis for the *quadratic game*, that is where the adversary must play quadratic functions. We describe the adversary's strategy, and we prove that the well-known Follow the Leader strategy is optimal for the player.
- We show that the above results apply to a much wider class of games, where the adversary can play either convex or strongly convex functions, suggesting that indeed the linear and quadratic games are the "hard cases".

2 Online Convex Games

The general optimization game we consider is as follows. We have two agents, a player and an adversary, and the game proceeds for T rounds with T known in advance to both agents. The player's choices will come from some convex set $X \subset \mathbb{R}^d$, and the adversary will choose functions from the class $\mathcal{F}$. To consider the game in full generality, we assume that the adversary's "allowed" functions may change on each round, and thus we imagine there is a sequence of allowed sets $L_1, L_2, \ldots, L_T \subset \mathcal{F}$.

Online Convex Game

$\mathcal{G}(X, \{L_t\})$:

for $t = 1$ to T **do**

  Player chooses (predicts) $\mathbf{x}_t \in X$.

  Adversary chooses a function $f_t \in L_t$.

end for

Player suffers regret

$$R_T = \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \inf_{\mathbf{x} \in X} \sum_{t=1}^{T} f_t(\mathbf{x}).$$

From this general game, we obtain each of the examples above with appropriate choice of $X, \mathcal{F}$ and the sets $\{L_t\}$. We define a number of particular games in the definitions below.

It is useful to prove regret bounds within this model as they apply to any problem that can be cast as an Online Convex Game. The known general upper bounds are as follows:

- **Zinkevich [9]:** If $L_1 = \ldots = L_T = \mathcal{F}$ consist of continuous twice differentiable functions f, where $\|\nabla f\| \le G$ and $\nabla^2 f \succeq \mathbf{0}$, then[1]
$$R_T \le \frac{1}{2} DG\sqrt{T}.$$
where $D := \max_{\mathbf{x},\mathbf{y}\in X} \|\mathbf{x}-\mathbf{y}\|$ and G is some positive constant.

- **Hazan et al. [4]:** If $L_1 = \ldots = L_T = \mathcal{F}$ consist of continuous twice differentiable functions f, where $\|\nabla f\| \le G$ and $\nabla^2 f \succeq \sigma I$, then
$$R_T \le \frac{1}{2}\frac{G^2}{\sigma} \log T,$$
where G and σ are positive constants.

- **Bartlett et al. [1]:** If L_t consists of continuous twice differentiable functions f, where $\|\nabla f\| \le G_t$ and $\nabla^2 f \succeq \sigma_t I$, then
$$R_T \le \frac{1}{2} \sum_{t=1}^{T} \frac{G_t^2}{\sum_{s=1}^{t} \sigma_s},$$
where G_t and σ_t are positive constants. Moreover, the algorithm does not need to know G_t, σ_t on round t.

All three of these games posit an upper bound on $\|\nabla f\|$ which is required to make the game nontrivial (and is natural in most circumstances). However, the first requires only that the second derivative be nonnegative, while the second and third game has a strict positive lower bound on the eigenvalues of the Hessian $\nabla^2 f$. Note that the bound of Bartlett et al recovers the logarithmic regret of Hazan et al whenever G_t and σ_t do not vary with time.

In the present paper, we analyze each of these games with the goal of obtaining the exact minimax value of the game, defined as:

$$V_T(\mathcal{G}(X, \{L_t\})) = \inf_{\mathbf{x}_1 \in X} \sup_{f_1 \in L_1} \ldots \inf_{\mathbf{x}_T \in X} \sup_{f_T \in L_T} \left(\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \inf_{\mathbf{x}\in X} \sum_{t=1}^{T} f_t(\mathbf{x}) \right).$$

The quantity $V_T(\mathcal{G})$ tells us the worst case regret of an *optimal* strategy in this game.

First, in the spirit of [1], we consider V_T for the games where constants G and σ, which respectively bound the first and second derivatives of f_t, can change throughout the game. That is, the Adversary is given two sequences before the game begins, $\langle G_1, \ldots, G_T\rangle$ and $\langle \sigma_1, \ldots, \sigma_T\rangle$. We also require only that the gradient of f_t is bounded *at the point* $\mathbf{x}_t$, i.e. $\|\nabla f_t(\mathbf{x}_t)\| \le G_t$, as opposed to the global constraint $\|\nabla f_t(\mathbf{x})\| \le G_t$ for all $\mathbf{x} \in X$. We may impose both of the above constraints by carefully choosing the sets $L_t \subseteq \mathcal{F}$, and we note that these sets will depend on the choices $\mathbf{x}_t$ made by the Player.

We first define the Linear and Quadratic Games, which are the central objects of this paper.

Definition 2.1. *The Linear Game $\mathcal{G}_{lin}(X, \langle G_t\rangle)$ is the game $\mathcal{G}(X, \{L_t\})$ where*
$$L_t = \{f : f(\mathbf{x}) = v^\top(\mathbf{x} - \mathbf{x}_t) + c, v \in \mathbb{R}^n, c \in \mathbb{R}; \|v\| \le G_t\}.$$

Definition 2.2. *The Quadratic Game $\mathcal{G}_{quad}(X, \langle G_t\rangle, \langle \sigma_t\rangle)$ is the game $\mathcal{G}(X, \{L_t\})$ where*
$$L_t = \{f : f(\mathbf{x}) = v^\top(\mathbf{x} - \mathbf{x}_t) + \frac{\sigma_t}{2}\|\mathbf{x} - \mathbf{x}_t\|^2 + c,$$
$$v \in \mathbb{R}^n, c \in \mathbb{R}; \|v\| \le G_t\}.$$

[1] This bound can be obtained by a slight modification of the analysis in [9].

The functions in these definitions are parametrized through $\mathbf{x}_t$ to simplify proofs of the last section. In Section 4, however, we will just consider the standard parametrization $f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x}$.

We also introduce more general games: the Convex Game and the Strongly Convex Game. While being defined with respect to a much richer class of loss functions, we show that these games are indeed no harder than the Linear and the Quadratic Games defined above.

Definition 2.3. *The Convex Game $\mathcal{G}_{conv}(X, \langle G_t \rangle)$ is the game $\mathcal{G}(X, \{L_t\})$ where*

$$L_t = \{f : \|\nabla f(\mathbf{x}_t)\| \leq G_t, \nabla^2 f \succeq 0\}.$$

Definition 2.4. *The Strongly Convex Game $\mathcal{G}_{st\text{-}conv}(X, \langle G_t \rangle, \langle \sigma_t \rangle)$ is the game $\mathcal{G}(X, \{L_t\})$ where*

$$L_t = \{f : \|\nabla f(\mathbf{x}_t)\| \leq G_t, \nabla^2 f - \sigma_t I \succeq 0\}.$$

We write $\mathcal{G}(G)$ instead of $\mathcal{G}(\langle G_t \rangle)$ when all values $G_t = G$ for some fixed G. This holds similarly for $\mathcal{G}(\sigma)$ instead of $\mathcal{G}(\langle \sigma_t \rangle)$. Furthermore, we suppose that $\sigma_1 > 0$ throughout the paper.

3 Previous Work

Several lower bounds for various online settings are available in the literature. Here we review a number of such results relevant to the present paper and highlight our primary contributions.

The first result that we mention is the lower bound of Vovk in the online linear regression setting [8]. It is shown that there exists a randomized strategy of the Adversary such that the expected regret is at least $\left[(n-\epsilon)G^2 \ln T - C_\epsilon\right]$ for any $\epsilon > 0$ and C_ϵ a constant. One crucial difference between this particular setting and ours is that the loss functions of the form $(y_t - \mathbf{x}_t \cdot \mathbf{w}_t)^2$ used in linear regression are curved in only one direction and linear in all other, thus this setting does not quite fit into any of the games we analyze. The lower bound of Vovk scales roughly as $n \log T$, which is quite interesting given that n does not enter into the lower bound of the Strongly Convex Game we analyze.

The lower bound for the log-loss functions of Ordentlich and Cover [5] in the setting of Universal Portfolios is also logarithmic in T and linear in n. Log-loss functions are parameterized as $f_t(\mathbf{x}) = -\log(\mathbf{w} \cdot \mathbf{x})$ for $\mathbf{x}$ in the simplex, and these fit more generally within the class of "exp-concave" functions. Upper bounds on the class of log-loss functions were originally presented by Cover [3] whereas Hazan et al. [4] present an efficient method for competing against the more general exp-concave functions. The log-loss lower bound of [5] is quite elegant yet, contrary to the minimax results we present, the optimal play is not efficiently computable.

The work of Takimoto and Warmuth [7] is most closely related to our results for the Quadratic Game. The authors consider functions $f(\mathbf{x}) = \frac{1}{2}||\mathbf{x} - \mathbf{y}||^2$ corresponding to the log-likelihood of the datapoint $\mathbf{y}$ for a unit-variance Gaussian with mean $\mathbf{x}$. The lower bound of $\frac{1}{2}D^2(\ln T - \ln\ln T + O(\ln\ln T/\ln T))$ is obtained, where D is the bound on the norm of adversary's choices $\mathbf{y}$. Furthermore, they exhibit the minimax strategy which, in the end, corresponds to a biased maximum-likelihood solution. We emphasize that these results differ from ours in several ways. First, we enforce a constraint on the size of the gradient of f_t whereas [7] constrain the location of the point $\mathbf{y}$ when $f_t(\mathbf{x}) = \frac{1}{2}||\mathbf{x} - \mathbf{y}||^2$. With our slightly weaker constraint, we can achieve a regret bound of the order $\log T$ instead of the $\log T - \log\log T$ of Takimoto and Warmuth. Interestingly, the authors describe the "$-\log\log T$" term of their lower bound as "surprising" because many known games "were shown to have $O(\log T)$ upper bounds". They conjecture that the apparent slack is due to the learner being unaware of the time horizon T. In the present paper, we resolve this issue by noting that our slightly weaker assumption erases the additional term; it is thus the limit on the adversary, and not knowledge of the horizon, that gives rise to the slack. Furthermore, the minimax strategy of Takimoto and Warmuth, a biased maximum likelihood estimate on each round, is also an artifact of their assumption on the boundedness of adversary's choices. With our weaker assumption, the minimax strategy is *exactly* maximum likelihood (generally called "Follow The Leader").

All previous work mentioned above deals with "curved" functions. We now discuss known lower bounds for the Linear Game. It is well-known that in the expert setting, it is impossible to do better than $O(\sqrt{T})$. The lower bound in Cesa-Bianchi and Lugosi [2], Theorem 3.7, proves an asymptotic bound: in the limit of $T \to \infty$, the value of the game behaves as $\sqrt{(\ln N)T/2}$, where N is the number of experts. We provide a similar randomized argument, which has been sketched in the literature (e.g. Hazan et al [4]), but our additional minimax analysis indeed gives the tightest bound possible for any T.

Finally, we provide reductions between Quadratic and Strongly Convex as well as Linear and Convex Games. While apparent that Adversary does better by playing linear approximations instead of convex functions, it requires a careful analysis to show that this holds for the minimax setting.

4 The Linear Game

In this section we begin by providing a relatively standard proof of the $O(\sqrt{T})$ lower bound on regret when competing against linear loss functions. The more interesting result is our *minimax* analysis which is given in Section 4.2.

4.1 The Randomized Lower Bound

Lower bounds for games with linear loss functions have appeared in the literature though often not in detail. The rough idea is to imagine a randomized Adversary and to compute the Player's expected regret. This generally produces an $O(\sqrt{T})$ lower bound yet it is not fully satisfying since the analysis is not tight. In the following section we provide a much improved analysis with minimax strategies for both the Player and Adversary.

Theorem 4.1. *Suppose $X = [-D/(2\sqrt{n}), D/(2\sqrt{n})]^n$, so that the diameter of X is D. Then*

$$V_T(\mathcal{G}_{lin}(X, \langle G_t \rangle)) \geq \frac{D}{2\sqrt{2}} \sqrt{\sum_{t=1}^{T} G_t^2}$$

Proof. Define the scaled cube

$$\mathcal{C}_t = \{-G_t/\sqrt{n}, G_t/\sqrt{n}\}^n.$$

Suppose the Adversary chooses functions from

$$\hat{L}_t = \{f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} : \mathbf{w} \in \mathcal{C}_t\}.$$

Note that $\|\nabla f\| = \|\mathbf{w}_t\| = G_t$ for any $f \in \hat{L}_t$.

Since we are restricting the Adversary to play linear functions with restricted $\mathbf{w}$,

$$\begin{aligned} V_T(\mathcal{G}_{\text{lin}}(X, \langle G_t \rangle)) \geq V_T(\mathcal{G}(X, \hat{L}_1, \ldots, \hat{L}_T)) &= \inf_{\mathbf{x}_1 \in X} \sup_{f_1 \in \hat{L}_1} \ldots \inf_{\mathbf{x}_T \in X} \sup_{f_T \in \hat{L}_1} \left[\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \inf_{\mathbf{x} \in X} \sum_{t=1}^{T} f_t(\mathbf{x}) \right] \\ &= \inf_{\mathbf{x}_1 \in X} \sup_{\mathbf{w}_1 \in \mathcal{C}_1} \ldots \inf_{\mathbf{x}_T \in X} \sup_{\mathbf{w}_T \in \mathcal{C}_T} \left[\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{x}_t - \inf_{\mathbf{x} \in X} \mathbf{x} \cdot \sum_{t=1}^{T} \mathbf{w}_t \right] \\ &\geq \inf_{\mathbf{x}_1 \in X} \mathbb{E}_{\mathbf{w}_1} \ldots \inf_{\mathbf{x}_T \in X} \mathbb{E}_{\mathbf{w}_T} \left[\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{x}_t - \inf_{\mathbf{x} \in X} \mathbf{x} \cdot \sum_{t=1}^{T} \mathbf{w}_t \right], \end{aligned}$$

where $\mathbb{E}_{\mathbf{w}_t}$ denotes expectation with respect to any distribution over the set $\mathcal{C}_t$. In particular, it holds for the uniform distribution, i.e. when the coordinates of $\mathbf{w}_t$ are $\pm G_t/\sqrt{n}$ with probability $1/2$. Since in this case $\mathbb{E}_{\mathbf{w}_T} \mathbf{w}_T \cdot \mathbf{x}_T = 0$

for any $\mathbf{x}_T$, we obtain

$$
\begin{aligned}
V_T(\mathcal{G}_{\text{lin}}(X, \langle G_t\rangle)) &\geq \inf_{\mathbf{x}_1\in X} \mathbb{E}_{\mathbf{w}_1} \cdots \inf_{\mathbf{x}_{T-1}\in X} \mathbb{E}_{\mathbf{w}_{T-1}} \inf_{\mathbf{x}_T\in X} \mathbb{E}_{\mathbf{w}_T} \left[\sum_{t=1}^{T} \mathbf{w}_t\cdot\mathbf{x}_t - \inf_{\mathbf{x}\in X} \mathbf{x}\cdot\sum_{t=1}^{T}\mathbf{w}_t\right] \\
&= \inf_{\mathbf{x}_1\in X} \mathbb{E}_{\mathbf{w}_1} \cdots \inf_{\mathbf{x}_{T-1}\in X} \mathbb{E}_{\mathbf{w}_{T-1}} \inf_{\mathbf{x}_T\in X} \left[\sum_{t=1}^{T-1} \mathbf{w}_t\cdot\mathbf{x}_t - \mathbb{E}_{\mathbf{w}_T} \inf_{\mathbf{x}\in X} \mathbf{x}\cdot\sum_{t=1}^{T}\mathbf{w}_t\right] \\
&= \inf_{\mathbf{x}_1\in X} \mathbb{E}_{\mathbf{w}_1} \cdots \inf_{\mathbf{x}_{T-1}\in X} \mathbb{E}_{\mathbf{w}_{T-1}} \left[\sum_{t=1}^{T-1} \mathbf{w}_t\cdot\mathbf{x}_t - \mathbb{E}_{\mathbf{w}_T} \inf_{\mathbf{x}\in X} \mathbf{x}\cdot\sum_{t=1}^{T}\mathbf{w}_t\right],
\end{aligned}
$$

where the last equality holds because the expression no longer depends on $\mathbf{x}_T$. Repeating the process, we obtain

$$
V_T(\mathcal{G}_{\text{lin}}(X, \langle G_t\rangle)) \geq -\mathbb{E}_{\mathbf{w}_1,\ldots,\mathbf{w}_T} \inf_{\mathbf{x}\in X} \mathbf{x}\cdot\sum_{t=1}^{T}\mathbf{w}_t = -\mathbb{E}_{\{\epsilon_{i,t}\}} \min_{\mathbf{x}\in\left\{-\frac{D}{2\sqrt{n}},\frac{D}{2\sqrt{n}}\right\}^n} \left(\mathbf{x}\cdot\sum_{t=1}^{T}\mathbf{w}_t\right),
$$

where $\mathbf{w}_t(i) = \epsilon_{i,t}G_t/\sqrt{n}$, with i.i.d. Rademacher variables $\epsilon_{i,t} = \pm 1$ with probability $1/2$. The last equality is due to the fact that a linear function is minimized at the vertices of the cube. In fact, the dot product is minimized by matching the sign of $\mathbf{x}(i)$ with that of the ith coordinate of $\sum_{t=1}^{T}\mathbf{w}_t$. Hence,

$$
V_T(\mathcal{G}_{\text{lin}}(X, \langle G_t\rangle)) \geq -\mathbb{E}_{\{\epsilon_{i,t}\}} \sum_{i=1}^{n} -\frac{D}{2\sqrt{n}}\left|\sum_{t=1}^{T}\epsilon_{i,t}\frac{G_t}{\sqrt{n}}\right| = \frac{D}{2}\mathbb{E}_{\{\epsilon_{i,t}\}}\left|\sum_{t=1}^{T}\epsilon_{i,t}G_t\right| \geq \frac{D}{2\sqrt{2}}\sqrt{\sum_{t=1}^{T}G_t^2},
$$

where the last inequality follows from the Khinchine's inequality [2]. □

4.2 The Minimax Analysis

While in the previous section we found a lower bound on $V_T(\mathcal{G}_{\text{lin}})$, here we present a complete minimax analysis for the particular case when X is a ball of dimension at least 3. We are indeed able to compute exactly the value

$$
V_T(\mathcal{G}_{\text{lin}}(X, \langle G_t\rangle))
$$

and we provide the simple minimax strategies for both the Player and the Adversary. The unit ball, while a special case, is a very natural choice for X as it is the *largest* convex set of diameter 2.

For the remainder of this section, let $f_t(\mathbf{x}) := \mathbf{w}_t\cdot\mathbf{x}$ where $\mathbf{w}_t \in \mathbb{R}^n$ with $\|\mathbf{w}_t\| \leq G_t$. Also, we define $\mathbf{W}_t = \sum_{s=1}^{t}\mathbf{w}_s$, the cumulative functions chosen by the Adversary.

Theorem 4.2. *Let $X = \{\mathbf{x} : \|\mathbf{x}\|_2 \leq D/2\}$ and suppose the Adversary chooses functions from*

$$
L_t = \{f(\mathbf{x}) = \mathbf{w}\cdot\mathbf{x} : \|\mathbf{w}\|_2 \leq G_t\}.
$$

Then the value of the game

$$
V_T(\mathcal{G}_{lin}(X, \langle G_t\rangle)) = \frac{D}{2}\sqrt{\sum_{t=1}^{T}G_t^2}.
$$

Furthermore, the optimal strategy for the player is to choose

$$
\mathbf{x}_{t+1} = \left(-\frac{D}{2\sqrt{\|\mathbf{W}_t\|^2 + G_{t+1}}}\right)\mathbf{W}_t.
$$

To prove the theorem, we will need a series of short lemmas.

Lemma 4.1. *When X is the unit ball $B = \{\mathbf{x} : \|\mathbf{x}\| = 1\}$, the value V_T can be written as*

$$\inf_{\mathbf{x}_1 \in B} \sup_{\mathbf{w}_1 \in L_1} \ldots \inf_{\mathbf{x}_T \in B} \sup_{\mathbf{w}_T \in L_T} \left[\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{x}_t + \|\mathbf{W}_T\| \right] \tag{1}$$

In addition, if we choose a larger radius D, the value of the game will scale linearly with this radius and thus it is enough to assume $X = B$.

Proof. The last term in the regret

$$\inf_{\mathbf{x} \in B} \sum_t f_t(\mathbf{x}) = \inf_{\mathbf{x} \in B} \mathbf{W}_T \cdot \mathbf{x} = -\|\mathbf{W}_T\|$$

since the infimum is obtained when $\mathbf{x} = \frac{\mathbf{W}_T}{\|\mathbf{W}_T\|}$. This implies equation (1). The fact that the bound scales linearly with $D/2$ follows from the fact that both the norm $\|\mathbf{W}_T\|$ will scale with $D/2$ as well as the terms $\mathbf{w}_t \cdot \mathbf{x}_t$. □

For the remainder of this section, we simply assume that $X = B$, the unit ball with diameter $D = 2$.

Lemma 4.2. *Regardless of the Player's choices, the Adversary can always obtain regret at least*

$$\sqrt{\sum_{t=1}^{T} G_t^2} \tag{2}$$

whenever the dimension n is at least 3.

Proof. Consider the following adversarial strategy and assume $X = B$. On round t, after the Player has chosen $\mathbf{x}_t$, the adversary chooses $\mathbf{w}_t$ such that $\|\mathbf{w}_t\| = G_t$, $\mathbf{w}_t \cdot \mathbf{x}_t = 0$ and $\mathbf{w}_t \cdot \mathbf{W}_{t-1} = 0$. Finding a vector of length G_t that is perpendicular to two arbitrary vectors can always be done when the dimension is at least 3. With this strategy, it is guaranteed that $\sum_t \mathbf{w}_t \cdot \mathbf{x}_t = 0$ and we claim also that

$$\|\mathbf{W}_T\| = \sqrt{\sum_{t=1}^{T} G_t^2}.$$

This follows from a simple induction. Assuming $\|\mathbf{W}_{t-1}\| = \sqrt{\sum_{s=1}^{t-1} G_s^2}$, then

$$\|\mathbf{W}_t\| = \|\mathbf{W}_{t-1} + \mathbf{w}_t\| = \sqrt{\|\mathbf{W}_{t-1}\|^2 + \|\mathbf{w}_t\|^2},$$

implying the desired conclusion. □

The result of the last lemma is quite surprising: the adversary need only play some vector with length G_t which is perpendicular to both $\mathbf{x}_t$ and $\mathbf{W}_{t-1}$. Indeed, this lower bound has a very different flavor from the randomized argument of the previous section. All that remains is to show that the Adversary can *do no better!*

Lemma 4.3. *Let $\mathbf{w}_0 = \mathbf{0}$. If the player always plays the point*

$$\mathbf{x}_t = \frac{-\mathbf{W}_{t-1}}{\sqrt{\|\mathbf{W}_{t-1}\|^2 + \sum_{s=t}^{T} G_s^2}} \tag{3}$$

then

$$\sup_{\mathbf{w}_1} \sup_{\mathbf{w}_2} \ldots \sup_{\mathbf{w}_T} \left[\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{x}_t + \|\mathbf{W}_T\| \right] \leq \sqrt{\sum_{t=1}^{T} G_t^2}$$

i.e., the regret can be no greater than the value in (2).

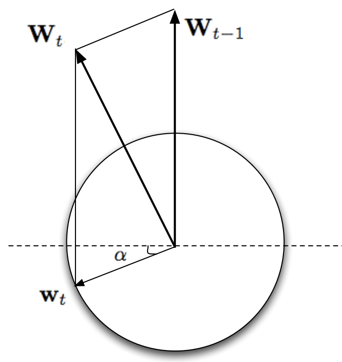

Figure 1: Illustration for the proof of the minimax strategy for the ball. We suppose that $\mathbf{x}_t$ is aligned with $\mathbf{W}_{t-1}$ and depict the plane spanned by $\mathbf{W}_{t-1}$ and $\mathbf{w}_t$. We assume that $\mathbf{w}_t$ has angle α with the line perpendicular to $\mathbf{W}_{t-1}$ and show that $\alpha = 0$ is optimal.

Proof. As before, $\mathbf{W}_t = \sum_{s=1}^t \mathbf{w}_s$. Define $\Gamma_t^2 = \sum_{s=t}^T G_s^2$, the forward sum, with $\Gamma_{T+1} = 0$. Define

$$\Phi_t(\mathbf{w}_1, \ldots, \mathbf{w}_{t-1}) = \sum_{s=1}^{t-1} \mathbf{x}_s \cdot \mathbf{w}_s + \sqrt{\|\mathbf{W}_{t-1}\|^2 + \Gamma_t^2}$$

where $\mathbf{x}_t$ is as defined in (3) and Φ_1 is $\sqrt{\sum_{t=1}^T G_t^2}$. Let

$$V_t(\mathbf{w}_1, \ldots, \mathbf{w}_{t-1}) = \sup_{\mathbf{w}_t} \ldots \sup_{\mathbf{w}_T} \left[\sum_{t=1}^T \mathbf{w}_t \cdot \mathbf{x}_t + \|\mathbf{W}_T\| \right]$$

be the optimum payoff to the adversary given that he plays $\mathbf{w}_1, \ldots, \mathbf{w}_{t-1}$ in the beginning and then plays optimally. The player plays according to (3) throughout. Note that the value of the game is V_1.

We prove by backward induction that, for all $t \in \{1, \ldots, T\}$,

$$V_t(\mathbf{w}_1, \ldots, \mathbf{w}_{t-1}) \leq \Phi_t(\mathbf{w}_1, \ldots, \mathbf{w}_{t-1})$$

The base case, $t = T + 1$ is obvious. Now assume it holds for $t + 1$ and we will prove it for t. We have

$$\begin{aligned} V_t(\mathbf{w}_1, \ldots, \mathbf{w}_{t-1}) &= \sup_{\mathbf{w}_t} V_{t+1}(\mathbf{w}_1, \ldots, \mathbf{w}_t) \\ \text{(induc.)} \quad &\leq \sup_{\mathbf{w}_t} \Phi_{t+1}(\mathbf{w}_1, \ldots, \mathbf{w}_t) \\ &= \sum_{s=1}^{t-1} \mathbf{x}_s \cdot \mathbf{w}_s + \underbrace{\sup_{\mathbf{w}_t} \left[\frac{-\mathbf{w}_t \cdot \mathbf{W}_{t-1}}{\sqrt{\|\mathbf{W}_{t-1}\|^2 + \Gamma_t^2}} + \sqrt{\|\mathbf{W}_{t-1} + \mathbf{w}_t\|^2 + \Gamma_{t+1}^2} \right]}_{(*)} \end{aligned}$$

Let us consider the final supremum term above. If we can show that it is no more than

$$\sqrt{\|\mathbf{W}_{t-1}\|^2 + \Gamma_t^2} \tag{4}$$

then we will have proved $V_t \leq \Phi_t$ thus completing the induction. This is the objective of the remainder of this proof.

We begin by noting two important facts about the expression (*). First, the supremum is taken over a convex function of $\mathbf{w}_t$ and thus the maximum occurs at the boundary, i.e. where $\|\mathbf{w}_t\| = G_t$ exactly. This is easily checked by computing the Hessian with respect to $\mathbf{w}_t$. Second, since $\mathbf{x}_t$ is chosen parallel to $\mathbf{W}_{t-1}$, the only two vectors of interest are $\mathbf{w}_t$ and $\mathbf{W}_{t-1}$. Without loss of generality, we can assume that $\mathbf{W}_{t-1}$ is the 2-dim vector $\langle F, 0\rangle$, where $F = \|\mathbf{W}_{t-1}\|$, and that $\mathbf{w}_t = \langle -G_t \sin\alpha, G_t \cos\alpha\rangle$ for any α. We now rewrite (*) as

$$\sup_\alpha \underbrace{\frac{FG_t \sin\alpha}{\sqrt{F^2 + G_t^2 + \Gamma_{t+1}^2}} + \sqrt{F^2 + G_t^2 + \Gamma_{t+1}^2 - 2FG_t \sin\alpha}}_{\phi(\alpha)}$$

We illustrate this problem in Figure 1. Bounding the above expression requires some care, and thus we prove it in Lemma A.1 found in the appendix. The result of Lemma A.1 gives us that, indeed,

$$\phi(\alpha) \le \sqrt{F^2 + G_t^2 + \Gamma_{t+1}^2} = \sqrt{\|\mathbf{W}_{t-1}\|^2 + \Gamma_t^2}.$$

Since (*) is exactly $\sup_\alpha \phi(\alpha)$, which is no greater than

$$\sqrt{F^2 + \Gamma_t^2},$$

we are done.

□

We observe that the minimax strategy for the ball is exactly the Online Gradient Descent strategy of Zinkevich [9]. The value of the game for the ball is exactly the upper bound for the proof of Online Gradient Descent if the initial point is the center of the ball. The lower bound of the randomized argument in the previous section differs from the upper bound for Online Gradient Descent by $\sqrt{2}$.

5 The Quadratic Game

As in the last section, we now give a minimax analysis of the game $\mathcal{G}_{\text{quad}}$. Ultimately we will be able to compute the exact value of $V_T(\mathcal{G}_{\text{quad}}(X, \langle G_t\rangle, \langle\sigma_t\rangle))$ and provide the optimal strategy of both the Player and the Adversary. What is perhaps most interesting is that the optimal Player strategy is the well-known Follow The Leader approach. This general strategy can be defined simply as

$$\mathbf{x}_{t+1} = \arg\min_{\mathbf{x}\in X} \sum_{s=1}^{t} f_s(\mathbf{x});$$

that is, we choose the best $\mathbf{x}$ "in hindsight". As has been pointed out by several authors, this strategy can incur $\Omega(T)$ regret when the loss functions are linear. It is thus quite surprising that this strategy is optimal when instead we are competing against quadratic loss functions.

For this section, define $F_t(\mathbf{x}) := \sum_{s=1}^t f_s(\mathbf{x})$ and $\mathbf{x}_t^* := \arg\min_{\mathbf{x}} F_t(\mathbf{x})$. Define $\boldsymbol{\sigma}_{1:t} = \sum_{s=1}^t \sigma_s$. We assume from the outset that $\sigma_1 > 0$. We also set $\boldsymbol{\sigma}_{1:0} = 0$.

5.1 A Necessary Restriction

Recall that the upper bound in Hazan et al. [4] is

$$R_T \le \frac{1}{2}\frac{G^2}{\sigma}\log T$$

and note that this expression has no dependence on the size of X. We would thus ideally like to consider the case when $X = \mathbb{R}^n$, for this would seem to be the "hardest" case for the Player. The unbounded assumption is problematic, however, not because the game is too difficult for the Player, but the game is *too difficult for the Adversary!*. This ought to come as quite a surprise, but arises from the particular restrictions we place on the Adversary.

Proposition 5.1. *For $G, \sigma > 0$, if $\max_{\mathbf{x},\mathbf{y}\in X} \|\mathbf{x}-\mathbf{y}\| = D > 4G/\sigma$, there is an $\alpha > 0$ such that $V_T(\mathcal{G}_{quad}(X, G, \sigma)) \leq -\alpha T$.*

Proof. Fix $\mathbf{x}_o, \mathbf{x}_e \in X$ with $\|\mathbf{x}_o - \mathbf{x}_e\| > 4G/\sigma$. Consider a player that plays $\mathbf{x}_{2k-1} = \mathbf{x}_o$, $\mathbf{x}_{2k} = \mathbf{x}_e$. Then for any $\mathbf{x} \in X$,

$$f_{2k-1}(\mathbf{x}) \geq f_{2k-1}(\mathbf{x}_o) - G\|\mathbf{x} - \mathbf{x}_o\| + \frac{\sigma}{2}\|\mathbf{x} - \mathbf{x}_o\|^2,$$

And similarly for f_{2k} and $\mathbf{x}_e$. Summing over t (assuming that T is even) shows that $V_t(\mathcal{G}_{\text{quad}}(X, G, \sigma))$ is no more than

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{x}) \leq \frac{T}{2}\left(G\|\mathbf{x} - \mathbf{x}_o\| - \frac{\sigma}{2}\|\mathbf{x} - \mathbf{x}_o\|^2 + G\|\mathbf{x} - \mathbf{x}_e\| - \frac{\sigma}{2}\|\mathbf{x} - \mathbf{x}_e\|^2\right).$$

But by the triangle inequality, any $\mathbf{x} \in X$ has $\|\mathbf{x}-\mathbf{x}_o\|+\|\mathbf{x}-\mathbf{x}_e\| \geq D$. Subject to this constraint, plus the constraints $0 \leq \|\mathbf{x} - \mathbf{x}_o\| \leq D, 0 \leq \|\mathbf{x} - \mathbf{x}_e\| \leq D$ shows that $V_t(\mathcal{G}_{\text{quad}}(X, G, \sigma)) \leq T(GD - \sigma D^2/4)/2 \leq -\alpha T$ for some $\alpha > 0$, since $D > 4G/\sigma$. ◻

As we don't generally expect regret to be negative, this example suggests that the Quadratic Game is uninteresting without further constraints on the Player. While an explicit bound on the size of X is a possibility, it is easier for the analysis to place a slightly weaker restriction on the Player.

Assumption 5.1. *Let $\mathbf{x}^*_{t-1}$ be the minimizer of $F_{t-1}(\mathbf{x})$. We assume that the Player must choose $\mathbf{x}_t$ such that*

$$\sigma_t\|\mathbf{x}_t - \mathbf{x}^*_{t-1}\| < 2G_t.$$

This restriction is necessary for non-negative regret. Indeed, it can be shown that if we increase the size of the above ball by only ϵ, the method of Proposition 5.1 above shows that the regret will be negative for large enough T.

5.2 Minimax Analysis

With the above restriction in place, we now simply write the game as $\mathcal{G}'_{\text{quad}}(\langle G_t\rangle, \langle\sigma_t\rangle)$, omitting the input X. We now proceed to compute the value of this game exactly.

Theorem 5.1. *Under Assumption 5.1, the value of the game*

$$V_T(\mathcal{G}'_{quad}(\langle G_t\rangle, \langle\sigma_t\rangle)) = \sum_{t=1}^{T} \frac{G_t^2}{2\boldsymbol{\sigma}_{1:t}}.$$

With uniform G_t and σ_t, we obtain the harmonic series, giving us our logarithmic regret bound. We note that this is *exactly* the upper bound proven in [1, 4], even with the constant.

Corollary 5.1. *For the uniform parameters of the game,*

$$\frac{G}{2\sigma}\log(T+1) \leq V_T(\mathcal{G}'_{quad}(G, \sigma)) \leq \frac{G}{2\sigma}(1 + \log T).$$

The main argument in the proof of Theorem 5.1 boils down to reducing the multiple round game to a single round game. The following lemma gives the value of this single round game. Since the proof is somewhat technical, we postpone it to the Appendix.

Lemma 5.1. *Let $G_t, \sigma_t, \boldsymbol{\sigma}_{1:t-1}$ be arbitrary positive constants. Then:*

$$\inf_{\Delta:||\Delta||\leq\frac{2G_t}{\sigma_t}} \sup_{\delta}\left(G_t\|\Delta - \delta\| - \tfrac{1}{2}\sigma_t\|\Delta - \delta\|^2 - \tfrac{1}{2}\boldsymbol{\sigma}_{1:t-1}\|\delta\|^2\right) = \frac{G_t^2}{2\boldsymbol{\sigma}_{1:t}},$$

and indeed the optimal strategy pair is $\Delta = \mathbf{0}$ and δ any vector for which $\|\delta\| = \frac{G_t}{\boldsymbol{\sigma}_{1:t}}$.

We now show how to "unwind" the recursive $\inf\sup$ definition of $V_T(\mathcal{G}'_{\text{quad}}(\langle G_t \rangle, \langle \sigma_t \rangle))$, where the final term we chop off is the object we described in the above lemma.

Proof of Theorem 5.1. Let $\mathbf{x}^*_{t-1}$ be the minimizer of $F_{t-1}(\mathbf{x})$ and $\mathbf{z} \in X$ be arbitrary. Note that F_t is $\boldsymbol{\sigma}_{1:t}$-quadratic, so

$$\begin{aligned} F_t(\mathbf{z}) = F_{t-1}(\mathbf{z}) + f_t(\mathbf{z}) &= F_{t-1}(\mathbf{x}^*_{t-1} + (\mathbf{z} - \mathbf{x}^*_{t-1})) + f_t(\mathbf{z}) \\ &= F_{t-1}(\mathbf{x}^*_{t-1}) + \nabla F_{t-1}(\mathbf{x}^*_{t-1})(\mathbf{z} - \mathbf{x}^*_{t-1}) + \tfrac{1}{2}\boldsymbol{\sigma}_{1:t-1}\|\mathbf{z} - \mathbf{x}^*_{t-1}\|^2 + f_t(\mathbf{z}) \\ &= F_{t-1}(\mathbf{x}^*_{t-1}) + \tfrac{1}{2}\boldsymbol{\sigma}_{1:t-1}\|\mathbf{z} - \mathbf{x}^*_{t-1}\|^2 + f_t(\mathbf{z}), \end{aligned}$$

where the last equality holds by the definition of $\mathbf{x}^*_{t-1}$. Hence,

$$\sum_{s=1}^{t} f_s(\mathbf{x}_s) - F_t(\mathbf{z}) = \left(\sum_{s=1}^{t-1} f_s(\mathbf{x}_s) - F_{t-1}(\mathbf{x}^*_{t-1})\right) + \left(f_t(\mathbf{x}_t) - f_t(\mathbf{z}) - \tfrac{1}{2}\boldsymbol{\sigma}_{1:t-1}\|\mathbf{z} - \mathbf{x}^*_{t-1}\|^2\right).$$

Expanding f_t around $\mathbf{x}_t$,

$$f_t(\mathbf{x}_t) - f_t(\mathbf{z}) = -\nabla f_t(\mathbf{x}_t)(\mathbf{z} - \mathbf{x}_t) - \tfrac{1}{2}\sigma_t\|\mathbf{z} - \mathbf{x}_t\|^2.$$

Substituting,

$$\sum_{s=1}^{t} f_s(\mathbf{x}_s) - F_t(\mathbf{z}) = \left(\sum_{s=1}^{t-1} f_s(\mathbf{x}_s) - F_{t-1}(\mathbf{x}^*_{t-1})\right) + \left(\nabla f_t(\mathbf{x}_t)(\mathbf{x}_t - \mathbf{z}) - \tfrac{1}{2}\sigma_t\|\mathbf{z} - \mathbf{x}_t\|^2 - \tfrac{1}{2}\boldsymbol{\sigma}_{1:t-1}\|\mathbf{z} - \mathbf{x}^*_{t-1}\|^2\right).$$

Then

$$\begin{aligned} V_t &:= \inf_{\mathbf{x}_1}\sup_{f_1}\ldots\inf_{\mathbf{x}_t}\sup_{f_t}\left(\sum_{s=1}^{t} f_s(\mathbf{x}_s) - \inf_{\mathbf{z}} F_t(\mathbf{z})\right) \\ &= \inf_{\mathbf{x}_1}\sup_{f_1}\ldots\inf_{\mathbf{x}_t}\sup_{f_t,\mathbf{z}}\left(\sum_{s=1}^{t} f_s(\mathbf{x}_s) - F_t(\mathbf{z})\right) \\ &= \inf_{\mathbf{x}_1}\sup_{f_1}\ldots\inf_{\mathbf{x}_{t-1}}\sup_{f_{t-1}}\left[\left(\sum_{s=1}^{t-1} f_s(\mathbf{x}_s) - F_{t-1}(\mathbf{x}^*_{t-1})\right) + \right. \\ &\qquad \left. \inf_{\mathbf{x}_t}\sup_{f_t,\mathbf{z}}\left(\nabla f_t(\mathbf{x}_t)(\mathbf{x}_t - \mathbf{z}) - \tfrac{1}{2}\sigma_t\|\mathbf{z} - \mathbf{x}_t\|^2 - \tfrac{1}{2}\boldsymbol{\sigma}_{1:t-1}\|\mathbf{z} - \mathbf{x}^*_{t-1}\|^2\right)\right]. \end{aligned}$$

However, we can simplify the final $\inf\sup$ as follows. We note that the quantity $\nabla f_t(\mathbf{x}_t)(\mathbf{x}_t - \mathbf{z})$ is maximized when $\nabla f_t(\mathbf{x}_t) = G_t \frac{\mathbf{x}_t - \mathbf{z}}{\|\mathbf{x}_t - \mathbf{z}\|}$. Second, we can instead use the variables $\Delta = \mathbf{x}_t - \mathbf{x}^*_{t-1}$ and $\delta = \mathbf{z} - \mathbf{x}^*_{t-1}$ in the optimization. Recall from Assumption 5.1 that $\|\mathbf{x}_t - \mathbf{x}^*_{t-1}\| = \|\Delta\| \le \frac{2G_t}{\sigma_t}$. Then,

$$\begin{aligned} V_t &= \inf_{\mathbf{x}_1}\sup_{f_1}\ldots\inf_{\mathbf{x}_{t-1}}\sup_{f_{t-1}}\left[\left(\sum_{s=1}^{t-1} f_s(\mathbf{x}_s) - F_{t-1}(\mathbf{x}^*_{t-1})\right)\right. \\ &\qquad \left. + \inf_{\Delta:\|\Delta\|\le\frac{2G_t}{\sigma_t}}\sup_{\delta}\left(G_t\|\Delta - \delta\| - \tfrac{1}{2}\sigma_t\|\Delta - \delta\|^2 - \tfrac{1}{2}\boldsymbol{\sigma}_{1:t-1}\|\delta\|^2\right)\right] \\ &= \inf_{\mathbf{x}_1}\sup_{f_1}\ldots\inf_{\mathbf{x}_{t-1}}\sup_{f_{t-1}}\left[\left(\sum_{s=1}^{t-1} f_s(\mathbf{x}_s) - F_{t-1}(\mathbf{x}^*_{t-1})\right) + \frac{G_t^2}{2\boldsymbol{\sigma}_{1:t}}\right] = V_{t-1} + \frac{G_t^2}{2\boldsymbol{\sigma}_{1:t}}, \end{aligned}$$

where the last equality is obtained by applying Lemma 5.1. Unwinding the recursion proves the theorem. $\square$

Corollary 5.2. *The optimal Player strategy is to set $\mathbf{x}_t = \mathbf{x}^*_{t-1}$ on each round.*

Proof. In analyzing the game, we found that the optimal choice of $\Delta = \mathbf{x}_t - \mathbf{x}^*_{t-1}$ was shown to be $\mathbf{0}$ in Lemma 5.1. $\square$

6 General Games

While the minimax results shown above are certainly interesting, we have only shown them to hold for the rather restricted games $\mathcal{G}_{\text{lin}}$ and $\mathcal{G}_{\text{quad}}$. For these particular cases, the class of functions that the Adversary may choose from is quite small: both the set of linear functions and the set quadratic functions can be parameterized by $O(n)$ variables. It would of course be more satisfying if our minimax analyses held for more richer loss function spaces.

Indeed, we prove in this section that both of our minimax results hold much more generally. In particular, we prove that even if the Adversary were able to choose *any* convex function on round t, with derivative bounded by G_t, then he can do no better than if he only had access to linear functions. On a similar note, if the Adversary is given the weak restriction that his functions be σ_t-strongly convex on round t, then he can do no better than if he could only choose σ_t-quadratic functions.

Theorem 6.1. *For fixed $X, \langle G_t\rangle$, and $\langle \sigma_t\rangle$, the values of the Quadratic Game and the Strongly Convex Game are equal*[2]*:*

$$V_T(\mathcal{G}_{\textit{st-conv}}(X, \langle G_t\rangle, \langle \sigma_t\rangle)) = V_T(\mathcal{G}_{\textit{quad}}(X, \langle G_t\rangle, \langle \sigma_t\rangle)).$$

For a fixed X and $\langle G_t\rangle$, the values of the Convex Game and the Linear Game are equal:

$$V_T(\mathcal{G}_{\textit{conv}}(X, \langle G_t\rangle)) = V_T(\mathcal{G}_{\textit{lin}}(X, \langle G_t\rangle)).$$

We need the following lemma whose proof is postponed to the appendix. Define the regret function

$$R(\mathbf{x}_1, f_1, \ldots, \mathbf{x}_T, f_T) = \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \min_{x\in X} \sum_{t=1}^{T} f_t(\mathbf{x}).$$

Lemma 6.1. *Consider a sequence of sets $\{N_s\}_{s=1}^T$ and $M \subseteq N_t$ for some t. Suppose that for all $f_t \in N_t$ and $\mathbf{x}_t \in X$ there exists $f^*_t \in M$ such that for all $(\mathbf{x}_1, f_1, \ldots, \mathbf{x}_{t-1}, f_{t-1}, \mathbf{x}_{t+1}, f_{t+1}, \ldots, \mathbf{x}_T, f_T)$,*

$$R(\mathbf{x}_1, f_1, \ldots, \mathbf{x}_t, f_t, \ldots, \mathbf{x}_T, f_T) \le R(\mathbf{x}_1, f_1, \ldots, \mathbf{x}_t, f^*_t, \ldots, \ldots, \mathbf{x}_T, f_T).$$

Then

$$\begin{aligned}&\inf_{\mathbf{x}_1} \sup_{f_1\in N_1} \ldots \inf_{\mathbf{x}_t} \sup_{f_t\in N_t} \ldots \inf_{\mathbf{x}_T} \sup_{f_T\in N_T} R(\mathbf{x}_1, f_1, \ldots, \mathbf{x}_T, f_T)\\ &= \inf_{\mathbf{x}_1} \sup_{f_1\in N_1} \ldots \inf_{\mathbf{x}_t} \sup_{f_t\in M} \ldots \inf_{\mathbf{x}_T} \sup_{f_T\in N_T} R(\mathbf{x}_1, f_1, \ldots, \mathbf{x}_T, f_T).\end{aligned}$$

Proof of Theorem 6.1. Given the sequences $\langle G_t\rangle, \langle\sigma_t\rangle$, let $L_t(\mathbf{x}_t)$ be defined as for the Strongly Convex Game (Definition 2.3) and $L^*_t(\mathbf{x}_t)$ be defined as for the Quadratic Game (Definition 2.2). Observe that $L^*_t \subseteq L_t$ for any t. Moreover, for any $f_t \in L_t$ and $\mathbf{x}_t \in X$, define $f^*_t(\mathbf{x}) = f_t(\mathbf{x}_t) + \nabla f_t(\mathbf{x}_t)^\top(\mathbf{x} - \mathbf{x}_t) + \frac{1}{2}\sigma_t\|\mathbf{x} - \mathbf{x}_t\|^2$. By definition, $f_t(\mathbf{x}_t) = f^*_t(\mathbf{x}_t)$ and $\nabla f_t(\mathbf{x}_t) = \nabla f^*_t(\mathbf{x}_t)$. Hence, $f^*_t \in L^*_t$. Furthermore, $f_t(\mathbf{x}) \ge f^*_t(\mathbf{x})$ for any $\mathbf{x} \in X$, and $\mathbf{x}^*$ in particular. Hence, for all $(\mathbf{x}_1, f_1, \ldots, \mathbf{x}_{t-1}, f_{t-1}, \mathbf{x}_{t+1}, f_{t+1}, \ldots, \mathbf{x}_T, f_T)$,

$$R(\mathbf{x}_1, f_1, \ldots, \mathbf{x}_t, f_t, \ldots, \mathbf{x}_T, f_T) \le R(\mathbf{x}_1, f_1, \ldots, \mathbf{x}_t, f^*_t, \ldots, \ldots, \mathbf{x}_T, f_T).$$

The statement of the first part of the theorem follows by Lemma 6.1, applied for every $t \in \{1, \ldots, T\}$. The second part is proved by analogous reasoning. $\square$

[2]We note that the computation of V_T for the Quadratic Game required a particular restriction on the player, Assumption 5.1, where here we only consider a fixed domain X.

A Appendix

Proof of Lemma 5.1. We write

$$P_t(\Delta,\delta) := G_t\|\Delta-\delta\| - \tfrac{1}{2}\sigma_t\|\Delta-\delta\|^2 - \tfrac{1}{2}\boldsymbol{\sigma}_{1:t-1}\|\delta\|^2$$

and

$$Q_t(\Delta) := \sup_\delta P_t(\Delta,\delta),$$

then our goal is to obtain $\inf_{\Delta:\|\Delta\|\le\frac{2G_t}{\sigma_t}} Q_t(\Delta)$. We now proceed to show that the choice $\Delta=\mathbf{0}$ is optimal. For this choice,

$$Q_t(\mathbf{0}) = \sup_\delta\ G_t\|\delta\| - \tfrac{1}{2}\boldsymbol{\sigma}_{1:t}\|\delta\|^2 = \frac{G_t^2}{2\boldsymbol{\sigma}_{1:t}}.$$

Here the optimal choice of δ is any vector such that $\|\delta\| = \frac{G_t}{\boldsymbol{\sigma}_{1:t}}$.

Now let us consider the case that $\Delta\neq\mathbf{0}$. First, suppose $\Delta\neq\delta$. Note that the optimum $\sup_\delta P_t(\Delta,\delta)$ will be obtained when the gradient with respect to δ is zero, i.e.

$$-G_t\frac{\Delta-\delta}{\|\Delta-\delta\|} - \sigma_t(\delta-\Delta) - \boldsymbol{\sigma}_{1:t-1}\delta = \mathbf{0}$$

implying that δ is a linear scaling of Δ, i.e. $\delta = c\Delta$. The second case, $\Delta=\delta$, also implies that δ is a linear scaling of Δ. Substituting this optimal form of δ,

$$Q_t(\Delta) = \sup_{c\in\mathbb{R}}\left[G_t|1-c|\cdot\|\Delta\| - \tfrac{1}{2}\sigma_t(1-c)^2\|\Delta\|^2 - \tfrac{1}{2}\boldsymbol{\sigma}_{1:t-1}c^2\|\Delta\|^2\right].$$

We now claim that the supremum over $c\in\mathbb{R}$ occurs at some $c^*\le 1$ for any choice of Δ. Assume by contradiction that $c^*>1$ for some Δ. Then $\tilde{c} = -c^*+2$ achieves at least the same value as c^* since $|1-c^*| = |1-\tilde{c}|$ while $(c^*)^2 > (\tilde{c})^2$, making the last term larger, which is a contradiction. Hence, $c\le 1$ and, collecting the terms,

$$Q_t(\Delta) = \sup_{c\le 1}\left[\left(G_t\|\Delta\| - \tfrac{1}{2}\sigma_t\|\Delta\|^2\right) + c\cdot\left(\sigma_t\|\Delta\|^2 - G_t\|\Delta\|\right) - c^2\cdot\left(\tfrac{1}{2}\boldsymbol{\sigma}_{1:t}\|\Delta\|^2\right)\right].$$

Since we now assume $\|\Delta\|\neq\mathbf{0}$, we see that the supremum is achieved for $c^* = \frac{\sigma_t\|\Delta\|^2 - G_t\|\Delta\|}{\boldsymbol{\sigma}_{1:t}\|\Delta\|^2} = \frac{\sigma_t\|\Delta\|-G_t}{\boldsymbol{\sigma}_{1:t}\|\Delta\|}\le 1$ and

$$\begin{aligned}
Q_t(\Delta) &= \frac{\left(\sigma_t\|\Delta\|^2 - G_t\|\Delta\|\right)^2}{2\boldsymbol{\sigma}_{1:t}\|\Delta\|^2} + (G_t\|\Delta\| - \tfrac{1}{2}\sigma_t\|\Delta\|^2)\\
&= \frac{\sigma_t^2\|\Delta\|^2 - \sigma_t\|\Delta\|G_t + G_t^2}{2\boldsymbol{\sigma}_{1:t}}\|\Delta\| - \tfrac{1}{2}\sigma_t\|\Delta\|^2)\\
&= \frac{\sigma_t}{\boldsymbol{\sigma}_{1:t}}\left(\tfrac{1}{2}\sigma_t\|\Delta\|^2 - \|\Delta\|G_t\right) + (G_t\|\Delta\| - \tfrac{1}{2}\sigma_t\|\Delta\|^2) + \frac{G_t^2}{2\boldsymbol{\sigma}_{1:t}}\\
&= \frac{\boldsymbol{\sigma}_{1:t-1}}{\boldsymbol{\sigma}_{1:t}}\left(G_t - \tfrac{1}{2}\sigma_t\|\Delta\|\right)\|\Delta\| + \frac{G_t^2}{2\boldsymbol{\sigma}_{1:t}} > \frac{G_t^2}{2\boldsymbol{\sigma}_{1:t}},
\end{aligned}$$

where the last inequality holds by because $\|\Delta\|\le\frac{2G_t}{\sigma_t}$. Hence, the value $Q_t(\Delta)$ is strictly larger than $G_t^2/(2\boldsymbol{\sigma}_{1:t})$ whenever $\|\Delta\|>0$ and is equal to this value if $\Delta=\mathbf{0}$. Hence, the optimal choice for the Player is to choose $\Delta=\mathbf{0}$. ☐

Proof of Lemma 6.1. Fix $f_t\in L_t$ and $\mathbf{x}_t\in X$. Let $f_t^*\in M$ be as in the statement of the lemma. Define

$$h_1(\mathbf{x}_1, f_1,\ldots,\mathbf{x}_{t-1}, f_{t-1},\mathbf{x}_{t+1}, f_{t+1},\ldots,\mathbf{x}_T, f_T) := R(\mathbf{x}_1, f_1,\ldots,\mathbf{x}_t, f_t,\ldots,\mathbf{x}_T, f_T)$$

$$h_2(\mathbf{x}_1, f_1, \dots, \mathbf{x}_{t-1}, f_{t-1}, \mathbf{x}_{t+1}, f_{t+1}, \dots, \mathbf{x}_T, f_T) := R(\mathbf{x}_1, f_1, \dots, \mathbf{x}_t, f_t^*, \dots, \mathbf{x}_T, f_T).$$

By assumption, $h_1 \le h_2$. Hence, we can inf/sup over the variables $\mathbf{x}_{t+1}, f_{t+1}, \dots, \mathbf{x}_T, f_T$, obtaining

$$\begin{aligned} &\inf_{\mathbf{x}_{t+1}} \sup_{f_{t+1}\in N_{t+1}} \dots \inf_{\mathbf{x}_T} \sup_{f_T \in N_T} R(\mathbf{x}_1, f_1, \dots, \mathbf{x}_t, f_t, \dots, \mathbf{x}_T, f_T) \\ &\le \inf_{\mathbf{x}_{t+1}} \sup_{f_{t+1}\in N_{t+1}} \dots \inf_{\mathbf{x}_T} \sup_{f_T \in N_T} R(\mathbf{x}_1, f_1, \dots, \mathbf{x}_t, f_t^*, \dots, \mathbf{x}_T, f_T) \end{aligned}$$

for any $(\mathbf{x}_1, f_1, \dots, \mathbf{x}_{t-1}, f_{t-1})$. Hence, since $f_t^* \in M$

$$\begin{aligned} &\sup_{f_t \in N_t} \inf_{\mathbf{x}_{t+1}} \sup_{f_{t+1}\in N_{t+1}} \dots \inf_{\mathbf{x}_T} \sup_{f_T \in N_T} R(\mathbf{x}_1, f_1, \dots, \mathbf{x}_t, f_t, \dots, \mathbf{x}_T, f_T) \\ &\le \sup_{f_t \in M} \inf_{\mathbf{x}_{t+1}} \sup_{f_{t+1}\in N_{t+1}} \dots \inf_{\mathbf{x}_T} \sup_{f_T \in N_T} R(\mathbf{x}_1, f_1, \dots, \mathbf{x}_t, f_t, \dots, \mathbf{x}_T, f_T) \end{aligned}$$

for all $(\mathbf{x}_1, f_1, \dots, \mathbf{x}_{t-1}, f_{t-1}, \mathbf{x}_t)$. Since $M \subseteq N_t$, the above is in fact an equality. Since the two functions of the variables $(\mathbf{x}_1, f_1, \dots, \mathbf{x}_{t-1}, f_{t-1}, \mathbf{x}_t)$ are equal, taking inf's and sup's over these variables we obtain the statement of the lemma. □

Lemma A.1. *The expression*

$$\frac{FG \sin\alpha}{\sqrt{F^2+G^2+K^2}} + \sqrt{F^2+G^2+K^2-2FG\sin\alpha}$$

is no more than $\sqrt{F^2+G^2+K^2}$ *for constants* $F, G, K > 0$ *and any* α.

Proof. We are interested in proving that the supremum of

$$\phi(\alpha) = \frac{FG \sin\alpha}{\sqrt{F^2+G^2+K^2}} + \sqrt{F^2+G^2+K^2-2FG\sin\alpha}$$

over $[-\pi/2, \pi/2]$ is attained at $\alpha = 0$. Setting the derivative of $\Phi(\alpha)$ to zero,

$$\frac{FG \cos\alpha}{\sqrt{F^2+G^2+K^2}} - \frac{FG \cos\alpha}{\sqrt{F^2+G^2+K^2-2FG\sin\alpha}} = 0$$

which implies that either $\cos\alpha = 0$ or $\sin\alpha = 0$, i.e. $\alpha \in \{-\pi/2, 0, \pi/2\}$. Taking the second derivative, we get

$$\phi''(\alpha) = -\frac{FG \sin\alpha}{\sqrt{F^2+G^2+K^2}} - \left(-\frac{FG \sin\alpha}{\sqrt{F^2+G^2+K^2-2FG\sin\alpha}} + \frac{(FG\cos\alpha)(FG\cos\alpha)}{(F^2+G^2+K^2-2FG\sin\alpha)^{3/2}}\right).$$

Thus, $\phi''(0) < 0$. We conclude that the optimum is attained at $\alpha = 0$ and therefore

$$\phi(\alpha) \le \sqrt{F^2+G^2+K^2}$$

□

References

[1] Peter Bartlett, Elad Hazan, and Alexander Rakhlin. Adaptive online gradient descent. In J.C. Platt, D. Koller, Y. Singer, and S. Roweis, editors, *Advances in Neural Information Processing Systems 20*. MIT Press, Cambridge, MA, 2008.

[2] Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

[3] T.M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, January 1991.

[4] Elad Hazan, Adam Kalai, Satyen Kale, and Amit Agarwal. Logarithmic regret algorithms for online convex optimization. In *COLT*, pages 499–513, 2006.

[5] Erik Ordentlich and Thomas M. Cover. The cost of achieving the best portfolio in hindsight. *Math. Oper. Res.*, 23(4):960–982, 1998.

[6] Shai Shalev-Shwartz and Yoram Singer. Convex repeated games and Fenchel duality. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *Advances in Neural Information Processing Systems 19*. MIT Press, Cambridge, MA, 2007.

[7] Eiji Takimoto and Manfred K. Warmuth. The minimax strategy for gaussian density estimation. pp. In *COLT '00: Proceedings of the Thirteenth Annual Conference on Computational Learning Theory*, pages 100–106, San Francisco, CA, USA, 2000. Morgan Kaufmann Publishers Inc.

[8] V. Vovk. Competitive on-line linear regression. In *NIPS '97: Proceedings of the 1997 conference on Advances in neural information processing systems 10*, pages 364–370, Cambridge, MA, USA, 1998. MIT Press.

[9] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *ICML*, pages 928–936, 2003.