Lecture Notes on Online Learning
DRAFT

Alexander Rakhlin

These lecture notes contain material presented in
the *Statistical Learning Theory* course at UC Berkeley, Spring'08.

Various parts of these notes have been discovered together with
J. Abernethy, A. Agarwal, P. Bartlett, E. Hazan, and A. Tewari

January 13, 2009

CONTENTS

CHAPTER

ONE

INTRODUCTION

The past two decades witnessed a surge of activity on prediction and learning methods in adversarial environments. Progress on this topic has been made in various fields, with many methods independently discovered and rediscovered. In their recent book, Nicolò Cesa-Bianchi and Gábor Lugosi [5] collected and organized many of these results under a common umbrella. We are indebted to this book for our own interest in the field, which seemed very fragmented before Nicolò and Gábor's effort. That being said, we feel that it might be beneficial to organize the ideas in a manner different from [5]. The purpose of these lecture notes is to stress the role of regularization as a common umbrella for some of the known online learning methods. While many of the results mentioned here are not novel, we hope to give the reader a fresh perspective through a very natural formulation.

We start with the time-varying potential method of Chapter 11.6 of [5], which, we feel, is one of the most general results of the book. The versatility of the method is diminished by the fact that it is hidden in the middle of a chapter on "linear pattern recognition". In contrast, we would like to bring out this result in a generic setting of convex loss functions and show how various other algorithms arise from this formulation.

Another motivation for this note is the realization that the time-varying potential method is nothing more than a sequence of regularized empirical error minimizations. The latter is the basis for most of the batch machine learning methods, such as SVM, Lasso, etc. It is, therefore, very natural to *start* with an algorithm which minimizes the regularized empirical loss at every step of the online interaction with the environment. This provides a connection between online and batch learning which is conceptually important.

We also point the reader to the recent thesis of Shai Shalev-Shwartz [9, 10]. The primal-dual view of online updates is illuminating and leads to new algorithms; however, the focus of these notes is slightly different.

The General Setting

Let $\mathcal{K} \subseteq \mathbb{R}^n$, the set of moves of the player, be a closed convex set. Let $\mathcal{F}$, the set of moves of the adversary, contain *convex* functions from $\mathbb{R}^n$ to $\mathbb{R}$. The following repeated game is the object of study of these notes.

Online Convex Optimization (OCO) Game

At each time step $t = 1$ to T,
- Player chooses $\boldsymbol{w}_t \in \mathcal{K}$
- Adversary chooses $\ell_t \in \mathcal{F}$
- Player suffers loss $\ell_t(\boldsymbol{w}_t)$ and observes feedback $\mathfrak{F}$

The **goal of the Player** (our online learning algorithm) is to minimize the *regret*, a notion studied in decision theory:

$$R_T := \sum_{t=1}^{T} \ell_t(\boldsymbol{w}_t) - \min_{\mathbf{u} \in \mathcal{K}} \sum_{t=1}^{T} \ell_t(\mathbf{u}). \tag{1.1}$$

That is, regret is the cost accumulated throughout the game minus the hypothetical cumulative cost of the best fixed decision. Having settled on the notion of regret, we can ask the following questions.

- What kind of feedback can one use for minimizing regret?
- How does one design algorithms which enjoy $R_T \to 0$ as $T \to \infty$ (called *Hannan consistency*)?
- What are the best rates of convergence, in terms of T, for the given sets $\mathcal{K}, \mathcal{F}$ and a certain type of feedback?
- How does the dimension n enter into the bounds?
- What are the minimax-optimal algorithms and lower bounds on R_T?
- Which algorithms are more efficient? Is there a trade-off between regret and efficiency?

Some of these questions will be addressed in these notes.

We remark that our notation "$\ell_t(\boldsymbol{w}_t)$" for the cost of the player's and adversary's moves is slightly more general than "$\ell(\boldsymbol{w}_t, \boldsymbol{\gamma}_t)$" for a fixed loss function ℓ, where $\boldsymbol{\gamma}_t$ is a move of the adversary in some parameter set $\boldsymbol{\Gamma}$. Our notation, which allows the loss function to change, is motivated by the optimization literature, whereas the fixed-loss-function notation is more natural for game theory, statistics, and machine learning. For the most part of the notes, however, one can think of $\ell_t(\boldsymbol{w}_t) := \ell(\boldsymbol{w}_t, \boldsymbol{\gamma}_t)$.

We distinguish the *full-information* and *bandit* versions of OCO. In the full-information version, considered in Section 2, the Player may observe the entire function ℓ_t as his feedback $\mathfrak{F}$ and can exploit this in making his decisions (i.e. using first-order or second-order optimization methods). In the *bandit* setting, considered in Section 4, the feedback $\mathfrak{F}$ provided to the player on round t is only the scalar value $\ell_t(\boldsymbol{w}_t)$ (i.e. only the zeroth-order method can be used).

Without any assumptions on $\mathcal{F}$, we cannot hope to say anything meaningful about the regret. One thing is certain: we can hope to attain smaller regret as the adversary restricts $\mathcal{F}$ more and more. Just as in optimization, curvature of functions in $\mathcal{F}$ plays a crucial role. As we will observe, the worst convex functions are linear, and much of the focus of these notes is indeed on $\mathcal{F}$ being a set of linear functions with a bounded norm.

We also remark that the requirement $\mathcal{K} \subset \mathbb{R}^n$ can be easily relaxed to, for example, fields. In fact, in Section **??** we consider OCO on matrices.

As a warm-up, let us consider the following standard example.

Example 1. Suppose Player's moves come from the set $\mathcal{K} = [-1, 1]$ and $\mathcal{F} = \{\ell : \ell(x) = \alpha x,\ \alpha = [-1, 1]\}$. In other words, the player chooses points in the interval $[-1, 1]$ and the adversary responds with cost functions with a slope in the same interval. Suppose we use the following naïve algorithm for choosing $\boldsymbol{w}_t$:

Algorithm 1: Follow the Leader (FTL)

On the first round output any $\boldsymbol{w}_1 \in \mathcal{K}$. On a round $t > 1$ output

$$\boldsymbol{w}_{t+1} = \arg\min_{\boldsymbol{w} \in \mathcal{K}} \sum_{s=1}^{t} \ell_s(\boldsymbol{w}).$$

Suppose the adversary chooses the following sequence of cost functions: $0.5x, -x, x, -x, x, \ldots$. On this sequence, for $t > 1$, FTL outputs $\boldsymbol{w}_t = 1$ if t is odd and $\boldsymbol{w}_t = -1$ if even. The cumulative cost grows as T, whereas the cost of a fixed decision $\boldsymbol{w} = 0$ is 0. Hence, FTL fails to achieve non-trivial regret guarantees (Hannan consistency).

We start these lectures with a philosophical point. Recent notes by Olivier Bousquet [3] present a simplified, yet interesting point of view on successful *batch* algorithms in machine learning: they can be roughly collected under the umbrella of "regularized loss minimization" (with the notable exception of "random projection" methods). We would like to say the same about online learning; this is, at-large, the motivation for these notes.

The idea of regularization is age-old. In their seminal work, Tikhonov and Arsenin [11, 12] develop regularization methods for solving ill-posed problems. One can think of "learning" the underlying phenomenon from the scarce observed data is an ill-posed inverse problem: out of many possible hypotheses that explain the data, which one should we choose? To restore uniqueness[1] and reinforce the choice of simple models, regularization is the method that comes to mind. Support Vector Machines and many other successful algorithms arise from these considerations.

On the surface, it is not obvious why regularization methods would have anything to do with *online* learning. Indeed, the game described above does not aim at reconstructing some hidden phenomenon, as in the batch learning case. However, it is becoming apparent that regularization is indeed very natural. Just as regularization presents a cure to overfitting in the batch setting, so does regularization allow the online algorithm to avoid being fooled by an adversary. Indeed, as we just saw in the above example, blindly following the best decision given the past data implies, in some cases, playing into adversary's hands. Regularization is a way to choose "safe" decisions. Randomization is another way to regularize the algorithm, and this technique will be discussed in Section 3.

[1] To be precise, we should distinguish *regularization* from *penalization*. The former refers to methods which restore uniqueness of the solution, while the latter aims at reinforcing simpler models.

CHAPTER

TWO

FULL-INFORMATION PROBLEMS: THE REGULARIZATION APPROACH

In this section we study the full-information problem and develop *regularization* algorithms for attaining low regret. To this end, let $\mathcal{R}$ be a regularizer, a strictly-convex differentiable function $\mathcal{R}$. In line with the *Follow the Leader* algorithm introduced above, we give the name *Follow the Regularized Leader* to the following family of algorithms

Algorithm 2: Follow the Regularized Leader (FTRL)

Given $\eta > 0$ and $\mathcal{R}$, compute

$$\boldsymbol{w}_{t+1} = \arg\min_{\boldsymbol{w}\in\mathcal{K}} \left[\eta \sum_{s=1}^{t} \ell_s(\boldsymbol{w}) + \mathcal{R}(\boldsymbol{w}) \right]. \tag{2.1}$$

Here the choices of the regularizer and the learning rate η are under our control. It is understood that $\boldsymbol{w}_1 = \arg\min_{\boldsymbol{w}\in\mathcal{K}} \mathcal{R}(\boldsymbol{w})$.

The key question is: given $\mathcal{K}$ and $\mathcal{F}$, the sets of moves of the player and the adversary, how do we choose $\mathcal{R}$ and η to achieve low regret, if this is at all possible? We will see that FTRL-type algorithms enjoy nontrivial regret guarantees under very general assumptions on $\mathcal{K}$ and $\mathcal{F}$. As a bibliographic remark, we note that in the classification setting, Shalev-Shwartz and Singer [10] analyze the above family of algorithms from the dual perspective.

We also remark that η can depend on t, and a general FTRL algorithm can be written as

$$\boldsymbol{w}_{t+1} = \arg\min_{\boldsymbol{w}\in\mathcal{K}} \left[\sum_{s=1}^{t} \eta_s \ell_s(\boldsymbol{w}) + \mathcal{R}(\boldsymbol{w}) \right]. \tag{2.2}$$

However, we shall first consider the case of a fixed $\eta_t = \eta$, as it greatly simplifies the exposition. Time-varying η_t will be considered in Section 2.8.

FTRL can be written in a form which is more familiar to the statistical learning audience,

$$\boldsymbol{w}_{t+1} = \arg\min_{\boldsymbol{w}\in\mathcal{K}} \left[\frac{1}{t} \sum_{s=1}^{t} \ell_s(\boldsymbol{w}) + (t\eta)^{-1} \mathcal{R}(\boldsymbol{w}) \right].$$

To make the connection to regularization methods in statistical learning theory, the role of $\boldsymbol{w}_t$ is that of a function (or parameter) to be learned or estimated, while the role of ℓ_t is that of the data-dependent loss in the batch case.[1]

[1] For instance, $\ell_t(\boldsymbol{w}) = (y_t - \mathbf{x_t}^\top \boldsymbol{w})^2$ for the input-output pair $(\mathbf{x_t}, y_t)$.

For convenience, we define $\Phi_0(\boldsymbol{w}) = \mathcal{R}(\boldsymbol{w})$ and $\Phi_t = \Phi_{t-1} + \eta\ell_t$. Hence, FTRL can be succinctly written as $\boldsymbol{w}_{t+1} = \arg\min_{\boldsymbol{w}\in\mathcal{K}} \Phi_t(\boldsymbol{w})$. The *batch solution* to the problem is simply $\boldsymbol{w}_{T+1} = \arg\min_{\boldsymbol{w}\in\mathcal{K}} \Phi_T(\boldsymbol{w})$, the minimizer based on the complete sequence of data.

The case $\mathcal{K} = \mathbb{R}^n$ plays a special role in these notes. We denote the unconstrained minimizer of Φ_t by

$$\tilde{\boldsymbol{w}}_{t+1} = \arg\min_{\boldsymbol{w}\in\mathbb{R}^n} \Phi_t(\boldsymbol{w}) \tag{2.3}$$

and call this procedure *unconstrained-FTRL*.

We try to separate results for constrained and unconstrained minimization: this is precisely the reason for using different symbols $\boldsymbol{w}_t$ and $\tilde{\boldsymbol{w}}_t$. The unconstrained problem is typically easier to analyze, and, if predicting in a particular $\mathcal{K}$ is essential, it makes sense to define $\mathcal{R}$ to go to infinity for any sequence approaching the boundary of $\mathcal{K}$ (and infinite outside of $\mathcal{K}$). This way we sidestep the possibility of the minimizer (2.1) falling outside of the set by defining the regularizer appropriate for the geometry of $\mathcal{K}$. Such an assumption is employed in Chapter 11 of [5] to avoid the issue of projections. However, we do not impose such a requirement on $\mathcal{R}$ at the moment, as it allows us to unify some known methods which rely on projections. For instance, the method of Zinkevich [13] employs a quadratic regularizer for any convex set and, therefore, requires projections.

A central notion throughout these notes is that of Bregman Divergences. Recall that, given a strictly convex function F, the Bregman Divergence is defined by

$$D_F(\boldsymbol{w}, \mathbf{y}) := F(\boldsymbol{w}) - F(\mathbf{y}) - \nabla F(\mathbf{y})^\intercal(\boldsymbol{w} - \mathbf{y}).$$

In other words, it is the tail of F beyond the first-order Taylor expansion at $\mathbf{y}$. We invite the reader to think of Bregman Divergences as a "shorthand". Properties of Bregman Divergences are typically easy to verify, and we briefly list some important ones. The proofs can be found in [5], for instance.

- Divergences are non-negative.
- $D_{A+B}(x, y) = D_A(x, y) + D_B(x, y)$ (if both A and B are convex and differentiable)
- The "three-point equality" follows directly from the definition:

$$D_R(u, v) + D_R(v, w) = D_R(u, w) + (u - v)(\nabla R(w) - \nabla R(v))$$

- The Bregman projection onto a convex set $\mathcal{K}$ exists and is unique: $w' = \arg\min_{v\in\mathcal{K}} D_R(v, w)$
- Generalized Pythagorean Theorem: for all $u \in \mathcal{K}$,

$$D_R(u, w) \geq D_R(u, w') + D_R(w', w)$$

  where w' is the Bregman projection above.
- Denoting by R^* the dual of R, it holds that $\nabla R^* = (\nabla R)^{-1}$, i.e. the gradient of the dual is the inverse function of the gradient.
- $D_R(u, v) = D_{R^*}(\nabla R(v), \nabla R(u))$
- $D_{R+f}(x, y) = D_R(x, y)$ if f is linear.
- $\nabla_x D_R(x, y) = \nabla R(x) - \nabla R(y)$

We warn the reader that the next three sections contain very general statements, which, on the surface, do not imply Hannan consistency of the regularization approach. We urge patience and promise that such statements are coming in a few pages.

2.1 Unconstrained Minimization: The (Rare) Case of Equality

The following result, being an equality, is the most general one for unconstrained-FTRL.

Lemma 1. *Suppose $\mathcal{K} = \mathbb{R}^n$. Then the regret of unconstrained-FTRL (2.3) against any $\mathbf{u} \in \mathcal{K}$,*

$$\eta \sum_{t=1}^{T} \ell_t(\tilde{\boldsymbol{w}}_t) - \ell_t(\mathbf{u}) = D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_1) - D_{\Phi_T}(\mathbf{u}, \tilde{\boldsymbol{w}}_{T+1}) + \sum_{t=1}^{T} D_{\Phi_t}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}).$$

Proof. For the unconstrained minimum, $\nabla \Phi_t(\tilde{\boldsymbol{w}}_{t+1}) = 0$ and

$$D_{\Phi_t}(\mathbf{u}, \tilde{\boldsymbol{w}}_{t+1}) = \Phi_t(\mathbf{u}) - \Phi_t(\tilde{\boldsymbol{w}}_{t+1}) = \Phi_{t-1}(\mathbf{u}) + \eta \ell_t(\mathbf{u}) - \Phi_t(\tilde{\boldsymbol{w}}_{t+1})$$

for any $\mathbf{u} \in \mathcal{K}$. Hence,

$$\eta \ell_t(\mathbf{u}) = D_{\Phi_t}(\mathbf{u}, \tilde{\boldsymbol{w}}_{t+1}) + \Phi_t(\tilde{\boldsymbol{w}}_{t+1}) - \Phi_{t-1}(\mathbf{u})$$

and

$$\eta \ell_t(\boldsymbol{w}_t) = D_{\Phi_t}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}) + \Phi_t(\tilde{\boldsymbol{w}}_{t+1}) - \Phi_{t-1}(\tilde{\boldsymbol{w}}_t).$$

Combining,

$$\begin{aligned} \eta(\ell_t(\tilde{\boldsymbol{w}}_t) - \ell_t(\mathbf{u})) &= D_{\Phi_t}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}) + \Phi_t(\tilde{\boldsymbol{w}}_{t+1}) - \Phi_{t-1}(\tilde{\boldsymbol{w}}_t) - D_{\Phi_t}(\mathbf{u}, \tilde{\boldsymbol{w}}_{t+1}) - \Phi_t(\tilde{\boldsymbol{w}}_{t+1}) + \Phi_{t-1}(\mathbf{u}) \\ &= D_{\Phi_t}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}) + D_{\Phi_{t-1}}(\mathbf{u}, \tilde{\boldsymbol{w}}_t) - D_{\Phi_t}(\mathbf{u}, \tilde{\boldsymbol{w}}_{t+1}) \end{aligned}$$

Summing over $t = 1 \dots T$,

$$\eta \sum_{t=1}^{T} \ell_t(\tilde{\boldsymbol{w}}_t) - \ell_t(\mathbf{u}) = D_R(\mathbf{u}, \tilde{\boldsymbol{w}}_1) - D_{\Phi_T}(\mathbf{u}, \tilde{\boldsymbol{w}}_{T+1}) + \sum_{t=1}^{T} D_{\Phi_t}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}).$$

□

As with any equality, one can argue that nothing is gained by representing the regret in this form. The following simple result gives a handle on the terms on the right-hand side, justifying the particular decomposition. Note that $\nabla \Phi_t(\tilde{\boldsymbol{w}}_{t+1}) = 0$ and $\nabla \Phi_t(\tilde{\boldsymbol{w}}_t) = \nabla \Phi_{t-1}(\tilde{\boldsymbol{w}}_t) + \eta \nabla \ell_t(\tilde{\boldsymbol{w}}_t) = \eta \nabla \ell_t(\tilde{\boldsymbol{w}}_t)$. From the properties of divergences,

$$D_{\Phi_t}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}) = D_{\Phi_t^*}(\nabla \Phi_t(\tilde{\boldsymbol{w}}_{t+1}), \nabla \Phi_t(\tilde{\boldsymbol{w}}_t)) = D_{\Phi_t^*}(0, \eta \nabla \ell_t(\tilde{\boldsymbol{w}}_t)) \tag{2.4}$$

If we are able to calculate the dual of Φ_t, the above equality tells us that size of the "steps" $\eta \nabla \ell_t(\tilde{\boldsymbol{w}}_t)$ (as measured with respect to Φ_t^*) precisely amounts to the regret.

A reader might encounter an algorithm different from (2.1) while reading most of the original papers on the topic from 90's and more recently, including [5]. This algorithm motivates the updates by the idea of balancing loss minimization and staying close to the previous decision. We now show that, under a mild condition on $\mathcal{R}$, this algorithm is equivalent to unconstrained-FTRL.

Algorithm 3: Equivalent form of unconstrained-FTRL.

Choose $\boldsymbol{w}_1$ s.t. $\nabla \mathcal{R}(\boldsymbol{w}_1) = \mathbf{0}$. Otherwise,

$$\tilde{\boldsymbol{w}}_{t+1} = \arg\min_{\boldsymbol{w} \in \mathbb{R}^n} \left[\eta \ell_t(\boldsymbol{w}) + D_{\Phi_{t-1}}(\boldsymbol{w}, \tilde{\boldsymbol{w}}_t) \right] \tag{2.5}$$

Lemma 2. *The definition in (2.3) is equivalent to (2.5), i.e.*

$$\arg\min_{\boldsymbol{w}\in\mathbb{R}^n}\left[\eta\ell_t(\boldsymbol{w})+D_{\Phi_{t-1}}(\boldsymbol{w},\tilde{\boldsymbol{w}}_t)\right]=\arg\min_{\boldsymbol{w}\in\mathbb{R}^n}\left[\eta\sum_{s=1}^{t}\ell_s(\boldsymbol{w})+\mathcal{R}(\boldsymbol{w})\right]$$

Proof. Note that

$$\eta\ell_t(\boldsymbol{w})=\Phi_t(\boldsymbol{w})-\Phi_{t-1}(\boldsymbol{w})$$

and thus

$$\eta\ell_t(\boldsymbol{w})+D_{\Phi_{t-1}}(\boldsymbol{w},\tilde{\boldsymbol{w}}_t)=\Phi_t(\boldsymbol{w})-\Phi_{t-1}(\boldsymbol{w})+D_{\Phi_{t-1}}(\boldsymbol{w},\tilde{\boldsymbol{w}}_t).$$

Recall that

$$\nabla_{\boldsymbol{w}}D_{\Phi_{t-1}}(\boldsymbol{w},\tilde{\boldsymbol{w}}_t)=\nabla_{\boldsymbol{w}}\Phi_{t-1}(\boldsymbol{w})-\nabla_{\boldsymbol{w}}\Phi_{t-1}(\tilde{\boldsymbol{w}}_t).$$

Then, setting the gradient of the objective to zero, we observe that $\tilde{\boldsymbol{w}}_{t+1}$ defined as (2.5) satisfies the zero-gradient equation

$$\nabla\Phi_t(\tilde{\boldsymbol{w}}_{t+1})=\nabla\Phi_{t-1}(\tilde{\boldsymbol{w}}_t)$$

Thus, $\nabla\Phi_t(\tilde{\boldsymbol{w}}_{t+1})=\nabla\Phi_0(\tilde{\boldsymbol{w}}_1)=\nabla\mathcal{R}(\boldsymbol{w}_1)=\boldsymbol{0}$. We conclude that $\tilde{\boldsymbol{w}}_{t+1}$ minimizes Φ_t and thus it is equivalent to (2.3). □

We mention that the requirement $\nabla\mathcal{R}(\boldsymbol{w}_1)=\boldsymbol{0}$ is not restrictive, as we can always subtract $\boldsymbol{w}\cdot\nabla\mathcal{R}(\boldsymbol{w}_1)$ from the objective, effectively shifting $\mathcal{R}$ by a linear function.

Remark 3. *It is instructive to write regret bounds, such as that of Lemma 1, in the following form:*

$$\sum_{t=1}^{T}\ell_t(\tilde{\boldsymbol{w}}_t)\le\inf_{\mathbf{u}\in\mathcal{K}}\left[\sum_{t=1}^{T}\ell_t(\mathbf{u})+\eta^{-1}D_{\mathcal{R}}(\mathbf{u},\tilde{\boldsymbol{w}}_1)\right]+\eta^{-1}\sum_{t=1}^{T}D_{\Phi_t}(\tilde{\boldsymbol{w}}_t,\tilde{\boldsymbol{w}}_{t+1}).$$

This has the form of Oracle *inequalities, where the performance of the procedure is related to the performance of any comparator penalized by its complexity.*

2.2 Constrained Minimization with Bregman Projections

First, we show that the constrained minimization problem is equivalent to the unconstrained one followed by a projection. This implies that if $\mathcal{K}\neq\mathbb{R}^n$, we can keep track and update the unconstrained solutions $\tilde{\boldsymbol{w}}_t$ while predicting with the projected versions $\boldsymbol{w}_t$. This has also been termed the "Lazy Projection Algorithm". Define the projection function as

$$\Pi_{\Phi_t,\mathcal{K}}(\tilde{\boldsymbol{w}}_{t+1})=\arg\min_{\boldsymbol{w}\in\mathcal{K}}D_{\Phi_t}(\boldsymbol{w},\tilde{\boldsymbol{w}}_{t+1}).$$

Algorithm 4: Equivalent form of FTRL.

Given the unconstrained-FTRL solutions $\tilde{\boldsymbol{w}}_t$, define

$$\boldsymbol{w}_t=\Pi_{\Phi_t,\mathcal{K}}(\tilde{\boldsymbol{w}}_t).$$

Lemma 4. *Algorithm 4 is equivalent to Algorithm 2, i.e.*

$$\Pi_{\Phi_t,\mathcal{K}}\left(\arg\min_{\boldsymbol{w}\in\mathbb{R}^n}\left[\eta\sum_{s=1}^{t}\ell_s(\boldsymbol{w})+\mathcal{R}(\boldsymbol{w})\right]\right)=\arg\min_{\boldsymbol{w}\in\mathcal{K}}\left[\eta\sum_{s=1}^{t}\ell_s(\boldsymbol{w})+\mathcal{R}(\boldsymbol{w})\right].$$

Hence, also

$$\Pi_{\Phi_t,\mathcal{K}}\left(\arg\min_{\boldsymbol{w}\in\mathbb{R}^n}\left[\eta\ell_t(\boldsymbol{w})+D_{\Phi_{t-1}}(\boldsymbol{w},\tilde{\boldsymbol{w}}_t)\right]\right)=\arg\min_{\boldsymbol{w}\in\mathcal{K}}\left[\eta\sum_{s=1}^{t}\ell_s(\boldsymbol{w})+\mathcal{R}(\boldsymbol{w})\right].$$

Proof. Let $\boldsymbol{w}'_{t+1}=\Pi_{\Phi_t,\mathcal{K}}(\tilde{\boldsymbol{w}}_{t+1})$. By definition,

$$\Phi_t(\boldsymbol{w}_{t+1})\le\Phi_t(\boldsymbol{w}'_{t+1}).$$

On the other hand, $\tilde{\boldsymbol{w}}_{t+1}$ is the unconstrained minimum of a convex function and thus the gradient of Φ_t is zero at $\tilde{\boldsymbol{w}}_{t+1}$. Hence, $D_{\Phi_t}(\boldsymbol{w},\tilde{\boldsymbol{w}}_{t+1})=\Phi_t(\boldsymbol{w})-\Phi_t(\tilde{\boldsymbol{w}}_{t+1})$. By definition, $D_{\Phi_t}(\boldsymbol{w}'_{t+1},\tilde{\boldsymbol{w}}_{t+1})\le D_{\Phi_t}(\boldsymbol{w}_{t+1},\tilde{\boldsymbol{w}}_{t+1})$. Thus,

$$\Phi_t(\boldsymbol{w}'_{t+1})\le\Phi_t(\boldsymbol{w}_{t+1}).$$

Since $\mathcal{R}$ is strictly convex and ℓ_t's are convex, $\boldsymbol{w}_{t+1}=\boldsymbol{w}'_{t+1}$. □

It is only a small modification of the proof of Lemma 1 for the unprojected case, but we provide the proof of a guarantee for Algorithm 2 for completeness. Note that the bound is no longer an equality (although the equality is kept until the end when the negative terms are dropped). We admit that this guarantee is very loose and not practical for actually obtaining meaningful bounds. In later sections, we will prove better bounds.

Lemma 5. *Suppose $\boldsymbol{w}_t$ is defined as in (2.1) and $\tilde{\boldsymbol{w}}_t$ as in (2.3). For any $\mathbf{u}\in\mathcal{K}$,*

$$\eta\sum_{t=1}^{T}\ell_t(\boldsymbol{w}_t)-\ell_t(\mathbf{u})\le D_{\Phi_0}(\mathbf{u},\boldsymbol{w}_1)-D_{\Phi_T}(\mathbf{u},\tilde{\boldsymbol{w}}_{T+1})+\sum_{t=1}^{T}D_{\Phi_t}(\boldsymbol{w}_t,\tilde{\boldsymbol{w}}_{t+1})$$

Proof. For the unconstrained minimum, $\nabla\Phi_t(\tilde{\boldsymbol{w}}_{t+1})=0$ and

$$D_{\Phi_t}(\mathbf{u},\tilde{\boldsymbol{w}}_{t+1})=\Phi_t(\mathbf{u})-\Phi_t(\tilde{\boldsymbol{w}}_{t+1}).$$

Moreover,

$$\Phi_t(\mathbf{u})=\Phi_{t-1}(\mathbf{u})+\eta\ell_t(\mathbf{u}).$$

Combining the above,

$$\eta\ell_t(\mathbf{u})=D_{\Phi_t}(\mathbf{u},\tilde{\boldsymbol{w}}_{t+1})+\Phi_t(\tilde{\boldsymbol{w}}_{t+1})-\Phi_{t-1}(\mathbf{u})$$

and

$$\eta\ell_t(\boldsymbol{w}_t)=D_{\Phi_t}(\boldsymbol{w}_t,\tilde{\boldsymbol{w}}_{t+1})+\Phi_t(\tilde{\boldsymbol{w}}_{t+1})-\Phi_{t-1}(\boldsymbol{w}_t).$$

Thus,

$$\begin{aligned}\eta(\ell_t(\boldsymbol{w}_t)-\ell_t(\mathbf{u}))&=D_{\Phi_t}(\boldsymbol{w}_t,\tilde{\boldsymbol{w}}_{t+1})+\Phi_t(\tilde{\boldsymbol{w}}_{t+1})-\Phi_{t-1}(\boldsymbol{w}_t)-D_{\Phi_t}(\mathbf{u},\tilde{\boldsymbol{w}}_{t+1})-\Phi_t(\tilde{\boldsymbol{w}}_{t+1})+\Phi_{t-1}(\mathbf{u})\\&=D_{\Phi_t}(\boldsymbol{w}_t,\tilde{\boldsymbol{w}}_{t+1})-\Phi_{t-1}(\boldsymbol{w}_t)-D_{\Phi_t}(\mathbf{u},\tilde{\boldsymbol{w}}_{t+1})+\Phi_{t-1}(\mathbf{u})\\&=D_{\Phi_t}(\boldsymbol{w}_t,\tilde{\boldsymbol{w}}_{t+1})+D_{\Phi_{t-1}}(\mathbf{u},\tilde{\boldsymbol{w}}_t)-D_{\Phi_t}(\mathbf{u},\tilde{\boldsymbol{w}}_{t+1})+(\Phi_{t-1}(\tilde{\boldsymbol{w}}_t)-\Phi_{t-1}(\boldsymbol{w}_t))\end{aligned}$$

Summing over $t = 1 \dots T$ and using $\boldsymbol{w}_1 = \tilde{\boldsymbol{w}}_1$,

$$\begin{aligned}\eta \sum_{t=1}^{T} \ell_t(\boldsymbol{w}_t) - \ell_t(\mathbf{u}) &= D_{\Phi_0}(\mathbf{u}, \boldsymbol{w}_1) - D_{\Phi_T}(\mathbf{u}, \tilde{\boldsymbol{w}}_{T+1}) + \sum_{t=1}^{T} D_{\Phi_t}(\boldsymbol{w}_t, \tilde{\boldsymbol{w}}_{t+1}) + \sum_{t=1}^{T} (\Phi_{t-1}(\tilde{\boldsymbol{w}}_t) - \Phi_{t-1}(\boldsymbol{w}_t)) \\ &\leq D_{\Phi_0}(\mathbf{u}, \boldsymbol{w}_1) - D_{\Phi_T}(\mathbf{u}, \tilde{\boldsymbol{w}}_{T+1}) + \sum_{t=1}^{T} D_{\Phi_t}(\boldsymbol{w}_t, \tilde{\boldsymbol{w}}_{t+1})\end{aligned}$$

□

2.3 The Joy of Knowing the Future

Suppose we actually knew the function ℓ_t that is to be played by the adversary on round t. Can we devise a strategy to achieve a low regret? The answer is quite simple: one can just use the FTL algorithm with the "extra" cost function. Moreover, FTRL with the extra information also achieves a low regret, as we show now. We start with a simple observation that, by definition of FTRL,

$$\sum_{s=1}^{t} \ell_s(\boldsymbol{w}_{t+1}) + \eta^{-1}\mathcal{R}(\boldsymbol{w}_{t+1}) \leq \sum_{s=1}^{t} \ell_s(\boldsymbol{w}_t) + \eta^{-1}\mathcal{R}(\boldsymbol{w}_t)$$

and

$$\sum_{s=1}^{t-1} \ell_s(\boldsymbol{w}_t) + \eta^{-1}\mathcal{R}(\boldsymbol{w}_t) \leq \sum_{s=1}^{t-1} \ell_s(\boldsymbol{w}_{t+1}) + \eta^{-1}\mathcal{R}(\boldsymbol{w}_{t+1}).$$

Adding the two equations and canceling the terms, we obtain

$$\ell_t(\boldsymbol{w}_{t+1}) \leq \ell_t(\boldsymbol{w}_t). \tag{2.6}$$

In other words, including ℓ_t in the minimization objective immediately implies a lower loss on this function. This, in fact, leads to a dramatic decrease in regret. We formally state the following "hypothetical" algorithm, although the only difference from FTRL is the (illegal) inclusion of the function to be played by the adversary.

Algorithm 5: Hypothetical Be-The-Regularized-Leader or Be-The-Leader

$$\boldsymbol{w}_t = \arg\min_{\boldsymbol{w} \in \mathcal{K}} \left[\eta \sum_{s=1}^{t} \ell_s(\boldsymbol{w}) + \mathcal{R}(\boldsymbol{w}) \right] \tag{2.7}$$

The following lemma says that the regret is constant (does not depend on T) if we know the future and play $\boldsymbol{w}_{t+1}$ instead of $\boldsymbol{w}_t$ on round t. Again, keep in mind that such algorithm is only hypothetical, as it does not adhere to the protocol of the OCO game. We also remark that in the following lemma, $\mathcal{R}$ is not necessarily *strictly* convex, and, therefore, can be set to zero to recover the so-called Be-The-Leader algorithm.

Lemma 6. *For any* $\mathbf{u} \in \mathcal{K}$,

$$\sum_{t=1}^{T} \ell_t(\boldsymbol{w}_{t+1}) - \sum_{t=1}^{T} \ell_t(\mathbf{u}) \leq \eta^{-1} \left(\mathcal{R}(\mathbf{u}) - \mathcal{R}(\boldsymbol{w}_1) \right). \tag{2.8}$$

Here $\boldsymbol{w}_1 = \arg\min_{\boldsymbol{w} \in \mathcal{K}} \mathcal{R}(\boldsymbol{w})$.

Proof. The proof proceeds by induction. The base case of $T = 0$ holds by the definition of $\boldsymbol{w}_1$. Now, suppose the statement holds for $T-1$, i.e.

$$\sum_{t=1}^{T-1} \ell_t(\boldsymbol{w}_{t+1}) + \eta^{-1}\mathcal{R}(\boldsymbol{w}_1) \leq \sum_{t=1}^{T-1} \ell_t(\mathbf{u}) + \eta^{-1}\mathcal{R}(\mathbf{u}).$$

Since it holds for any $\mathbf{u} \in \mathcal{K}$, it holds for $\mathbf{u} = \boldsymbol{w}_{T+1}$:

$$\sum_{t=1}^{T-1} \ell_t(\boldsymbol{w}_{t+1}) + \eta^{-1}\mathcal{R}(\boldsymbol{w}_1) \leq \sum_{t=1}^{T-1} \ell_t(\boldsymbol{w}_{T+1}) + \eta^{-1}\mathcal{R}(\boldsymbol{w}_{T+1}).$$

Adding $\ell_T(\boldsymbol{w}_{T+1})$ to both sides,

$$\sum_{t=1}^{T} \ell_t(\boldsymbol{w}_{t+1}) + \eta^{-1}\mathcal{R}(\boldsymbol{w}_1) \leq \sum_{t=1}^{T} \ell_t(\boldsymbol{w}_{T+1}) + \eta^{-1}\mathcal{R}(\boldsymbol{w}_{T+1}).$$

The induction step follows because $\boldsymbol{w}_{T+1}$ is defined as the minimizer of the right-hand side. □

An immediate consequence of this simple lemma is the following corollary:

Corollary 7. *FTRL enjoys, for any* $\mathbf{u} \in \mathcal{K}$*,*

$$\sum_{t=1}^{T} \ell_t(\boldsymbol{w}_t) - \ell_t(\mathbf{u}) \leq \sum_{t=1}^{T} \left(\ell_t(\boldsymbol{w}_t) - \ell_t(\boldsymbol{w}_{t+1})\right) + \eta^{-1}\left(\mathcal{R}(\mathbf{u}) - \mathcal{R}(\boldsymbol{w}_1)\right). \tag{2.9}$$

By this point, the reader is probably wondering if any of the statements in the past three sections are useful. After all, we have not proved a single statement, which one could eye-ball and confidently say "Yes, for so-and-so sets $\mathcal{K}, \mathcal{F}$, the FTRL algorithm is Hannan consistent." We now specialize the very general results of previous sections to the case of linear cost functions $\mathcal{F}$ and arrive at some satisfying results.

2.4 Linear Losses: FTRL Algorithm

There are good reasons for studying linear losses. First, the regret is easy to analyze. Second, if non-linear (convex) functions are played by the adversary, we can construct linear approximations to these functions at the point we play and pretend as if the functions were indeed linear. As we show in the next section, the regret of playing against convex functions can be upper-bounded by the regret of the linearized versions. Hence, in some sense, the linear losses are hardest to play against. If no curvature restriction is provided for the choice of moves $\mathcal{F}$, the adversary will, in general, play linear functions. If $\mathcal{F}$ contains only "curved" functions, the fact can be exploited by the player to achieve smaller regret, as will be shown in later sections.

Now, let us write the linear costs as $\ell_t(\mathbf{u}) = \mathbf{f}_t^{\mathsf{T}}\mathbf{u}$. Going back to statements in the previous sections, we observe that the divergence terms simplify. Indeed, divergence with respect to a function $\mathcal{R}$ is not changed by adding linear functions, i.e. $D_{\Phi_t} = D_{\Phi_0} = D_{\mathcal{R}}$. We, therefore, have the following statements corresponding to Lemma 1 and Lemma 5.

Corollary 8. *Suppose* $\mathcal{K} = \mathbb{R}^n$*, i.e. the problem is unconstrained. Then unconstrained-FTRL satisfies, for any* $\mathbf{u} \in \mathcal{K}$*,*

$$\eta \sum_{t=1}^{T} \mathbf{f}_t^{\mathsf{T}}(\tilde{\boldsymbol{w}}_t - \mathbf{u}) = D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_1) - D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_{T+1}) + \sum_{t=1}^{T} D_{\mathcal{R}}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}).$$

Corollary 9. *Suppose $\boldsymbol{w}_t$ is defined as in (2.1) and $\tilde{\boldsymbol{w}}_t$ as in (2.3). Then FTRL satisfies, for any $\mathbf{u} \in \mathcal{K}$,*

$$\eta \sum_{t=1}^{T} \mathbf{f}_t^\mathsf{T}(\boldsymbol{w}_t - \mathbf{u}) \leq D_\mathcal{R}(\mathbf{u}, \boldsymbol{w}_1) - D_\mathcal{R}(\mathbf{u}, \tilde{\boldsymbol{w}}_{T+1}) + \sum_{t=1}^{T} D_\mathcal{R}(\boldsymbol{w}_t, \tilde{\boldsymbol{w}}_{t+1})$$

The statement of the following Lemma is almost identical to that of Corollary 7 (since $\nabla\mathcal{R}(\tilde{\boldsymbol{w}}_1) = 0$, we have $D_\mathcal{R}(\mathbf{u}, \tilde{\boldsymbol{w}}_1) = \mathcal{R}(\mathbf{u}) - \mathcal{R}(\tilde{\boldsymbol{w}}_t)$) in the unconstrained case. The proof, however, reveals the looseness of Corollary 7 as compared to the best possible guarantee of Corollary 8.

Lemma 10. *Suppose the losses are linear, $\ell_t(\mathbf{u}) = \mathbf{f}_t^\mathsf{T}\mathbf{u}$ and $\mathcal{K} = \mathbb{R}^n$. Then for any $\mathbf{u} \in \mathcal{K}$,*

$$\eta \sum_{t=1}^{T} \mathbf{f}_t^\mathsf{T}(\tilde{\boldsymbol{w}}_t - \mathbf{u}) \leq D_\mathcal{R}(\mathbf{u}, \tilde{\boldsymbol{w}}_1) - D_\mathcal{R}(\mathbf{u}, \tilde{\boldsymbol{w}}_{T+1}) + \eta \sum_{t=1}^{T} \mathbf{f}_t^\mathsf{T}(\tilde{\boldsymbol{w}}_t - \tilde{\boldsymbol{w}}_{t+1}).$$

Proof. By definition, $\tilde{\boldsymbol{w}}_t$ satisfies $\eta \sum_{s=1}^{t-1} \mathbf{f}_s + \nabla\mathcal{R}(\tilde{\boldsymbol{w}}_t) = 0$ and $\tilde{\boldsymbol{w}}_{t+1}$ satisfies $\eta \sum_{s=1}^{t} \mathbf{f}_s + \nabla\mathcal{R}(\tilde{\boldsymbol{w}}_t) = 0$. Subtracting,

$$\nabla\mathcal{R}(\tilde{\boldsymbol{w}}_t) - \nabla\mathcal{R}(\tilde{\boldsymbol{w}}_{t+1}) = \eta\mathbf{f}_t. \tag{2.10}$$

We observe that equation (2.10) implies

$$\begin{aligned} D_{\Phi_t}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}) &= D_\mathcal{R}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}) \\ &\leq D_\mathcal{R}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}) + D_\mathcal{R}(\tilde{\boldsymbol{w}}_{t+1}, \tilde{\boldsymbol{w}}_t) \qquad \text{(introducing looseness)} \\ &= -\nabla\mathcal{R}(\tilde{\boldsymbol{w}}_{t+1})(\tilde{\boldsymbol{w}}_t - \tilde{\boldsymbol{w}}_{t+1}) - \nabla\mathcal{R}(\tilde{\boldsymbol{w}}_t)(\tilde{\boldsymbol{w}}_{t+1} - \tilde{\boldsymbol{w}}_t) \\ &= \eta\mathbf{f}_t^\mathsf{T}(\tilde{\boldsymbol{w}}_t - \tilde{\boldsymbol{w}}_{t+1}). \end{aligned}$$

□

We can expect that the bound of Lemma 10 is off by a multiplicative constant from the actual regret. Nevertheless, we will take it (as well as Corollary 7 for the constrained case) as a starting point for specific bounds on the regret under various assumptions on $\mathcal{F}, \mathcal{K}$.

From Eq. (2.10), we observe that solutions $\tilde{\boldsymbol{w}}_{t+1}$ (defined in (2.3)) have a closed form

$$\tilde{\boldsymbol{w}}_{t+1} = \nabla\mathcal{R}^*(\nabla\mathcal{R}(\tilde{\boldsymbol{w}}_t) - \eta\mathbf{f}_t), \tag{2.11}$$

and, by Lemma 4, $\boldsymbol{w}_t$ (defined in (2.1)) has the form

$$\boldsymbol{w}_{t+1} = \Pi_{\mathcal{R},\mathcal{K}}(\nabla\mathcal{R}^*(\nabla\mathcal{R}(\tilde{\boldsymbol{w}}_t) - \eta\mathbf{f}_t)). \tag{2.12}$$

Here $\mathcal{R}^*$ is the dual function. This procedure is called *Mirror Descent* and goes back to the work of Yudin an Nemirovskii.

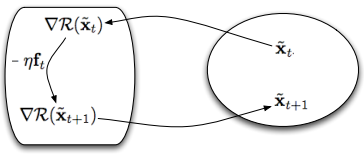

Figure 2.1: Mirror Descent as a gradient descent in the dual.

We are now ready to prove some specific guarantees for FTRL under natural assumptions on $\mathcal{F}$ (see [9]).

Proposition 11. *Fix a norm $\|\cdot\|$ and choose an $\mathcal{R}$ which is strongly convex with respect to this norm. Then FTRL satisfies*

$$\sum_{t=1}^{T} \mathbf{f}_t^{\mathsf{T}}(\boldsymbol{w}_t - \mathbf{u}) \leq \eta \sum_{t=1}^{T} (\|\mathbf{f}_t\|^*)^2 + \eta^{-1}(\mathcal{R}(\mathbf{u}) - \mathcal{R}(\boldsymbol{w}_1)).$$

In particular, if $\mathcal{F} \subseteq B_{\|\cdot\|^*}$, the ball under the dual norm, and $\eta = \sqrt{\mathcal{R}(\mathbf{u})/T}$, then

$$\sum_{t=1}^{T} \mathbf{f}_t^{\mathsf{T}}(\boldsymbol{w}_t - \mathbf{u}) \leq \sqrt{T\mathcal{R}(\mathbf{u})}.$$

Proof. The definition of strong convexity with respect to $\|\cdot\|$ implies that

$$\mathcal{R}(\boldsymbol{w}_t) \geq \mathcal{R}(\boldsymbol{w}_{t+1}) + \langle \nabla\mathcal{R}(\boldsymbol{w}_{t+1}), \boldsymbol{w}_t - \boldsymbol{w}_{t+1}\rangle + \frac{1}{2}\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|^2.$$

Repeating the statement for an expansion around $\boldsymbol{w}_t$ and adding the two, we obtain

$$\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|^2 \leq \langle \nabla\mathcal{R}(\boldsymbol{w}_t) - \nabla\mathcal{R}(\boldsymbol{w}_{t+1}), \boldsymbol{w}_t - \boldsymbol{w}_{t+1}\rangle.$$

Let $\tilde{\boldsymbol{w}}_t$ and $\tilde{\boldsymbol{w}}_{t+1}$ be the unconstrained minimizers, as defined in (2.3). By Lemma 4, $\boldsymbol{w}_t$ and $\boldsymbol{w}_{t+1}$ are projections of these two points, respectively. Hence, by the Kolmogorov's criterion for generalized projections (see [4]),

$$\langle \nabla\mathcal{R}(\boldsymbol{w}_t) - \nabla\mathcal{R}(\boldsymbol{w}_{t+1}), \boldsymbol{w}_t - \boldsymbol{w}_{t+1}\rangle \leq \langle \nabla\mathcal{R}(\tilde{\boldsymbol{w}}_t) - \nabla\mathcal{R}(\tilde{\boldsymbol{w}}_{t+1}), \boldsymbol{w}_t - \boldsymbol{w}_{t+1}\rangle.$$

Combining, applying Hölder's Inequality, and canceling $\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|$, we obtain

$$\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\| \leq \|\nabla\mathcal{R}(\tilde{\boldsymbol{w}}_t) - \nabla\mathcal{R}(\tilde{\boldsymbol{w}}_{t+1})\|^*,$$

where $\|\cdot\|^*$ is the norm dual to $\|\cdot\|$. Hence, strong convexity of $\mathcal{R}$ implies, by (2.10), that $\|\tilde{\boldsymbol{w}}_t - \tilde{\boldsymbol{w}}_{t+1}\| \leq \eta\|\mathbf{f}_t\|$, and with the help of Corollary 7 we arrive at the bound

$$\sum_{t=1}^{T} \mathbf{f}_t^{\mathsf{T}}(\tilde{\boldsymbol{w}}_t - \mathbf{u}) \leq \eta \sum_{t=1}^{T} (\|\mathbf{f}_t\|^*)^2 + \eta^{-1}(\mathcal{R}(\mathbf{u}) - \mathcal{R}(\boldsymbol{w}_1)).$$

□

2.5 Linear Losses: A Tweaked FTRL

Now, let us switch gears a bit. Recall that FTRL can be equivalently written as a projection of the unconstrained minimizer of $\eta\ell_t(\boldsymbol{w}) + D_{\Phi_{t-1}}(\boldsymbol{w}, \tilde{\boldsymbol{w}}_t)$ (see Algorithm 3). What if we put $\boldsymbol{w}_t$ instead of $\tilde{\boldsymbol{w}}_t$ in the above divergence? Specializing this to the linear case, we get the following algorithm, which is not FTRL, but is derived from it (see Zinkevich [13]).

A word of caution: as the following algorithm departs from the FTRL framework, $\boldsymbol{w}_t$ and $\tilde{\boldsymbol{w}}_t$ have a different meaning. We might even want to introduce a different notation for this purpose...

Again, one can show that the above definition is equivalent to

$$\boldsymbol{w}_{t+1} = \arg\min_{\boldsymbol{w}\in\mathcal{K}} \left[\eta\mathbf{f}_t^{\mathsf{T}}\boldsymbol{w} + D_{\mathcal{R}}(\boldsymbol{w}, \boldsymbol{w}_t)\right],$$

but the intermediate solutions $\tilde{\boldsymbol{w}}_{t+1}$ are needed for the analysis.

Algorithm 6: Mirror Descent-style Algorithm

Choose $\tilde{\boldsymbol{w}}_1$ s.t. $\nabla\mathcal{R}(\tilde{\boldsymbol{w}}_1) = \mathbf{0}$. Otherwise,

$$\tilde{\boldsymbol{w}}_{t+1} = \arg\min_{\boldsymbol{w}\in\mathbb{R}} \left[\eta \mathbf{f}_t^\intercal \boldsymbol{w} + D_{\mathcal{R}}(\boldsymbol{w}, \boldsymbol{w}_t)\right],$$

followed by projection

$$\boldsymbol{w}_{t+1} = \Pi_{\mathcal{R},\mathcal{K}} \tilde{\boldsymbol{w}}_{t+1}.$$

Unlike the unprojected version $\tilde{\boldsymbol{w}}_t$ of FTRL, which can be quite far from the set $\mathcal{K}$ (indeed, $\tilde{\boldsymbol{w}}_t = \nabla\mathcal{R}^*(-\eta\sum_{s=1}^{t-1}\mathbf{f}_s)$), the unprojected version $\tilde{\boldsymbol{w}}_t$ of Algorithm 6 is close by. Setting the derivative to zero,

$$\nabla\mathcal{R}(\tilde{\boldsymbol{w}}_{t+1}) - \nabla\mathcal{R}(\boldsymbol{w}_t) = -\eta\mathbf{f}_t,$$

which suggests that the points $\tilde{\boldsymbol{w}}_{t+1}$ and $\boldsymbol{w}_t$ are $O(\eta)$ away from each other.

We can equivalently write

$$\tilde{\boldsymbol{w}}_{t+1} = \nabla\mathcal{R}^*(\nabla\mathcal{R}(\boldsymbol{w}_t) - \eta\mathbf{f}_t)$$

and

$$\boldsymbol{w}_{t+1} = \Pi_{\mathcal{R},\mathcal{K}}\nabla\mathcal{R}^*(\nabla\mathcal{R}(\boldsymbol{w}_t) - \eta\mathbf{f}_t).$$

We observe that this is a Mirror Descent -style algorithm, with a form very similar to FTRL (see Eq. (2.12)).

We start by observing that

$$\eta\mathbf{f}_t^\intercal \boldsymbol{w}_{t+1} + D_{\mathcal{R}}(\boldsymbol{w}_{t+1}, \boldsymbol{w}_t) \le \eta\mathbf{f}_t^\intercal \boldsymbol{w}_t + D_{\mathcal{R}}(\boldsymbol{w}_t, \boldsymbol{w}_t) = \eta\mathbf{f}_t^\intercal \boldsymbol{w}_t.$$

Just as FTRL (see Eq. (2.6)), Algorithm 6 enjoys $\mathbf{f}_t^\intercal \boldsymbol{w}_{t+1} \le \mathbf{f}_t^\intercal \boldsymbol{w}_t$. The question is, again, how much is gained by knowing the future?

Lemma 12 (e.g. [2])**.** *The analogues of Lemma 6 and Corollary 7 hold for Algorithm 6:*

$$\sum_{t=1}^{T} \mathbf{f}_t^\intercal(\boldsymbol{w}_{t+1} - \mathbf{u}) \le \eta^{-1}(\mathcal{R}(\mathbf{u}) - \mathcal{R}(\boldsymbol{w}_1))$$

for any $\mathbf{u} \in \mathcal{K}$*. Thus,*

$$\sum_{t=1}^{T} \mathbf{f}_t^\intercal(\boldsymbol{w}_t - \mathbf{u}) \le \sum_{t=1}^{T} \mathbf{f}_t^\intercal(\boldsymbol{w}_t - \boldsymbol{w}_{t+1}) + \eta^{-1}(\mathcal{R}(\mathbf{u}) - \mathcal{R}(\boldsymbol{w}_1)).$$

Proof. Observe that $\boldsymbol{w}_{t+1}$ is the constrained minimizer of the objective $\eta\mathbf{f}_t^\intercal \boldsymbol{w} + D_{\mathcal{R}}(\boldsymbol{w}, \boldsymbol{w}_t)$. Thus, any direction pointing away from $\boldsymbol{w}_{t+1}$ and into the set should have a positive product with the gradient of the objective at $\boldsymbol{w}_{t+1}$ (for otherwise one could decrease the objective further):

$$\langle \mathbf{u} - \boldsymbol{w}_{t+1}, \eta\mathbf{f}_t + \nabla\mathcal{R}(\boldsymbol{w}_{t+1}) - \nabla\mathcal{R}(\boldsymbol{w}_t)\rangle \ge 0.$$

Rearranging,

$$\eta\mathbf{f}_t^\intercal(\boldsymbol{w}_{t+1} - \mathbf{u}) \le \langle \mathbf{u} - \boldsymbol{w}_{t+1}, \nabla\mathcal{R}(\boldsymbol{w}_{t+1}) - \nabla\mathcal{R}(\boldsymbol{w}_t)\rangle.$$

Using the three-point inequality,

$$\eta\mathbf{f}_t^\intercal(\boldsymbol{w}_{t+1} - \mathbf{u}) \le D_{\mathcal{R}}(\mathbf{u}, \boldsymbol{w}_t) - D_{\mathcal{R}}(\mathbf{u}, \boldsymbol{w}_t) - D_{\mathcal{R}}(\boldsymbol{w}_{t+1}, \boldsymbol{w}_t).$$

Adding over time,

$$\eta\sum_{t=1}^{T} \mathbf{f}_t^\intercal(\boldsymbol{w}_{t+1} - \mathbf{u}) \le D_{\mathcal{R}}(\mathbf{u}, \boldsymbol{w}_1) - D_{\mathcal{R}}(\mathbf{u}, \boldsymbol{w}_{T+1}) - \sum_{t=1}^{T} D_{\mathcal{R}}(\boldsymbol{w}_{t+1}, \boldsymbol{w}_t). \tag{2.13}$$

□

We mention another bound, which will become useful in Chapter 4.

Lemma 13. *Algorithm 6 enjoys, for any* $\mathbf{u} \in \mathcal{K}$,

$$\eta \sum_{t=1}^{T} \mathbf{f}_t^\intercal (\boldsymbol{w}_t - \mathbf{u}) \le D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_1) + \sum_{t=1}^{T} D_{\mathcal{R}}(\boldsymbol{w}_t, \tilde{\boldsymbol{w}}_{t+1}) \le (\mathcal{R}(\mathbf{u}) - \mathcal{R}(\tilde{\boldsymbol{w}}_1)) + \eta \sum_{t=1}^{T} \mathbf{f}_t^\intercal (\boldsymbol{w}_t - \tilde{\boldsymbol{w}}_{t+1})$$

Proof. Noting that $\eta \mathbf{f}_t = \nabla \mathcal{R}(\boldsymbol{w}_t) - \nabla \mathcal{R}(\tilde{\boldsymbol{w}}_{t+1})$, we can rewrite

$$\begin{aligned} \eta \mathbf{f}_t^\intercal (\boldsymbol{w}_t - \mathbf{u}) &= \langle \nabla \mathcal{R}(\boldsymbol{w}_t) - \nabla \mathcal{R}(\tilde{\boldsymbol{w}}_{t+1}), \boldsymbol{w}_t - \mathbf{u} \rangle \\ &= D_{\mathcal{R}}(\mathbf{u}, \boldsymbol{w}_t) - D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_{t+1}) + D_{\mathcal{R}}(\boldsymbol{w}_t, \tilde{\boldsymbol{w}}_{t+1}) \\ &\le D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_t) - D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_{t+1}) + D_{\mathcal{R}}(\boldsymbol{w}_t, \tilde{\boldsymbol{w}}_{t+1}). \end{aligned}$$

Furthermore,

$$D_{\mathcal{R}}(\boldsymbol{w}_t, \tilde{\boldsymbol{w}}_{t+1}) \le D_{\mathcal{R}}(\boldsymbol{w}_t, \tilde{\boldsymbol{w}}_{t+1}) + D_{\mathcal{R}}(\tilde{\boldsymbol{w}}_{t+1}, \boldsymbol{w}_t) = \langle \nabla \mathcal{R}(\boldsymbol{w}_t) - \nabla \mathcal{R}(\tilde{\boldsymbol{w}}_{t+1}), \boldsymbol{w}_t - \tilde{\boldsymbol{w}}_{t+1} \rangle = \eta \mathbf{f}_t^\intercal (\boldsymbol{w}_t - \tilde{\boldsymbol{w}}_{t+1}).$$

Summing over t results in the bound. □

Just like for FTRL, we can prove specific bounds on the regret for Algorithm 6.

Proposition 14. *Fix a norm* $\|\cdot\|$ *and choose an* $\mathcal{R}$ *which is strongly convex with respect to this norm. Then Algorithm 6 satisfies*

$$\sum_{t=1}^{T} \mathbf{f}_t^\intercal (\boldsymbol{w}_t - \mathbf{u}) \le \eta \sum_{t=1}^{T} (\|\mathbf{f}_t\|^*)^2 + \eta^{-1} (\mathcal{R}(\mathbf{u}) - \mathcal{R}(\tilde{\boldsymbol{w}}_1)).$$

In particular, if $\mathcal{F} \subseteq B_{\|\cdot\|^*}$, *the ball under the dual norm, and* $\eta = \sqrt{\mathcal{R}(\mathbf{u})/T}$, *then*

$$\sum_{t=1}^{T} \mathbf{f}_t^\intercal (\boldsymbol{w}_t - \mathbf{u}) \le \sqrt{T \mathcal{R}(\mathbf{u})}.$$

Proof. As in the proof of Lemma 11, strong convexity implies

$$\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|^2 \le \langle \nabla \mathcal{R}(\boldsymbol{w}_t) - \nabla \mathcal{R}(\boldsymbol{w}_{t+1}), \boldsymbol{w}_t - \boldsymbol{w}_{t+1} \rangle.$$

By the Kolmogorov's criterion,

$$\langle \nabla \mathcal{R}(\boldsymbol{w}_t) - \nabla \mathcal{R}(\boldsymbol{w}_{t+1}), \boldsymbol{w}_t - \boldsymbol{w}_{t+1} \rangle \le \langle \nabla \mathcal{R}(\boldsymbol{w}_t) - \nabla \mathcal{R}(\tilde{\boldsymbol{w}}_{t+1}), \boldsymbol{w}_t - \boldsymbol{w}_{t+1} \rangle.$$

Hence,

$$\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\| \le \|\mathcal{R}(\boldsymbol{w}_t) - \nabla \mathcal{R}(\tilde{\boldsymbol{w}}_{t+1})\|^* = \eta \|\mathbf{f}_t\|^*.$$

Substituting into Lemma 12, we obtain

$$\sum_{t=1}^{T} \mathbf{f}_t^\intercal (\boldsymbol{w}_t - \mathbf{u}) \le \eta \sum_{t=1}^{T} (\|\mathbf{f}_t\|^*)^2 + \eta^{-1} (\mathcal{R}(\mathbf{u}) - \mathcal{R}(\tilde{\boldsymbol{w}}_1)).$$

We note that a more careful analysis can yield a factor $\frac{1}{2}$ in front of $\eta \sum_{t=1}^{T} (\|\mathbf{f}_t\|^*)^2$. This is achieved by taking into account the negative terms in Eq. (2.13). □

Note that the upper bounds of Propositions 11 and 14 are identical. It is an interesting question of whether one algorithm can be shown to outperform the other.

2.6 Approximate solutions for convex losses via linearization

If $\mathcal{F}$ consists of convex functions, we can linearize the costs we observe and pretend the world is flat. This allows us to lift any results of the previous section to the case of general convex functions; however, the resulting bounds might be loose. Observe that by convexity

$$\ell_t(\boldsymbol{w}_t) - \ell_t(\mathbf{u}) \leq \nabla \ell_t(\boldsymbol{w}_t)^\intercal (\boldsymbol{w}_t - \mathbf{u}).$$

Feeding the linear functions $\mathbf{f}_t := \nabla \ell_t(\boldsymbol{w}_t)$ into any algorithm of the previous section completes the reduction.

Below we state algorithms that result from the linearization, along with the guarantees.

Lemma 15. *Suppose that instead of (2.5), we solve an approximate problem*

$$\tilde{\boldsymbol{w}}_{t+1} = \arg\min_{\boldsymbol{w} \in \mathbb{R}^n} \left[\eta \nabla \ell_t(\boldsymbol{w}_t)^\intercal \boldsymbol{w} + D_{\mathcal{R}}(\boldsymbol{w}, \tilde{\boldsymbol{w}}_t)\right] \tag{2.14}$$

followed by

$$\boldsymbol{w}_{t+1} = \Pi_{\mathcal{R},\mathcal{K}} \tilde{\boldsymbol{w}}_{t+1}.$$

Then

$$\eta \sum_{t=1}^T \ell_t(\boldsymbol{w}_t) - \ell_t(\mathbf{u}) \leq D_{\mathcal{R}}(\mathbf{u}, \boldsymbol{w}_1) + \sum_{t=1}^T D_{\mathcal{R}}(\boldsymbol{w}_t, \tilde{\boldsymbol{w}}_{t+1}).$$

The approximate problem enjoys a closed-form solution

$$\boldsymbol{w}_{t+1} = \Pi_{\mathcal{R},\mathcal{K}} \nabla \mathcal{R}^*(\nabla \mathcal{R}(\tilde{\boldsymbol{w}}_t) - \eta \nabla \ell_t(\boldsymbol{w}_t)).$$

Lemma 16. *Define the following update*

$$\tilde{\boldsymbol{w}}_{t+1} = \arg\min_{\boldsymbol{w} \in \mathbb{R}^n} \left[\eta \nabla \ell_t(\boldsymbol{w}_t)^\intercal \boldsymbol{w} + D_{\mathcal{R}}(\boldsymbol{w}, \boldsymbol{w}_t)\right] \tag{2.15}$$

followed by

$$\boldsymbol{w}_{t+1} = \Pi_{\mathcal{R},\mathcal{K}} \tilde{\boldsymbol{w}}_{t+1}.$$

Then

$$\eta \sum_{t=1}^T \ell_t(\boldsymbol{w}_t) - \ell_t(\mathbf{u}) \leq D_{\mathcal{R}}(\mathbf{u}, \boldsymbol{w}_1) + \sum_{t=1}^T D_{\mathcal{R}}(\boldsymbol{w}_t, \boldsymbol{w}_{t+1}).$$

The approximate problem enjoys a closed-form solution

$$\boldsymbol{w}_{t+1} = \Pi_{\mathcal{R},\mathcal{K}} \nabla \mathcal{R}^*(\nabla \mathcal{R}(\boldsymbol{w}_t) - \eta \nabla \ell_t(\boldsymbol{w}_t)).$$

2.7 Examples

2.7.1 Online Gradient Descent

Suppose $\mathcal{R}(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|^2$, half the Euclidean norm, and WLOG suppose that $\boldsymbol{w}_1 = \mathbf{0} \in \mathcal{K}$. Since $\mathcal{R}$ is strongly convex with respect to the Euclidean norm, we can use either FTRL or Algorithm 6 along with Propositions 11 and 14 to obtain the following guarantee:

$$\sum_{t=1}^T \ell_t(\boldsymbol{w}_t) - \ell_t(\mathbf{u}) \leq \frac{1}{2}\eta^{-1}\|\mathbf{u}\|^2 + \frac{1}{2}\eta \sum_{t=1}^T (\|\nabla \ell_t(\boldsymbol{w}_t)\|^*)^2 = DG\sqrt{T}$$

where D is an upper bound on the *radius* of the set, G is an upper bound on the largest norm of the gradients of ℓ_t's, and we set $\eta = \frac{G\sqrt{T}}{D}$.

Note that

$$\boldsymbol{w}_{t+1} = \Pi_{\mathcal{K}}(\boldsymbol{w}_t - \eta \nabla \ell_t(\boldsymbol{w}_t)),$$

the algorithm of Zinkevich [13].

2.7.2 EG and Weighted Majority

Let $\mathcal{K}$ be the n-simplex. Define $\mathcal{R}(\boldsymbol{w}) = \sum_{i=1}^{n}(\boldsymbol{w}_i \log \boldsymbol{w}_i - \boldsymbol{w}_i)$. One can verify that $\nabla\mathcal{R}(\boldsymbol{w}) = \ln \boldsymbol{w}$ (where we take ln element-wise) and

$$D_{\mathcal{R}}(\boldsymbol{w}, \mathbf{y}) = \sum_{i=1}^{n} \boldsymbol{w}_i \ln \frac{\boldsymbol{w}_i}{\mathbf{y}_i} + (\mathbf{y}_i - \boldsymbol{w}_i).$$

Over the simplex, this divergence corresponds to the KL divergence. We claim that in this rare case, FTRL and Algorithm 6 coincide. Indeed, the update of FTRL (2.12) corresponds to

$$\boldsymbol{w}_{t+1} = \Pi_{\mathcal{R},\mathcal{K}} \tilde{\boldsymbol{w}}_{t+1} = \Pi_{\mathcal{R},\mathcal{K}}(\nabla\mathcal{R}^*(\ln(\tilde{\boldsymbol{w}}_t) - \ln\exp(\eta\nabla\ell_t(\boldsymbol{w}_t)))),$$

In particular, $\tilde{\boldsymbol{w}}_{t+1}(i) = \tilde{\boldsymbol{w}}_t(i)\exp(-\eta \mathbf{f}_t(i))$ (again, we abused notation by applying ln and exp to vectors element-wise.) It is easy to verify that the projection with respect to the relative entropy onto the simplex is equivalent to normalization. Equivalence of FTRL and Algorithm 6 come from the fact that updates are multiplicative and update-normalize-update-normalize is equivalent to update-update-normalize.

It is possible to verify that $\mathcal{R}$ is strongly convex with respect to $\|\cdot\|_1$ over the simplex. Hence, we conclude

$$\sum_{t=1}^{T} \ell_t(\boldsymbol{w}_t) - \ell_t(\mathbf{u}) \le \eta \sum_{t=1}^{T} \|\mathbf{f}_t\|_\infty^2 + \eta^{-1}(\mathcal{R}(\mathbf{u}) - \mathcal{R}(\tilde{\boldsymbol{w}}_1)).$$

Noting that $\mathcal{R}(\mathbf{u}) \le 0$ and $\mathcal{R}(\tilde{\boldsymbol{w}}_1) = -\log n$, we arrive at $2\sqrt{T\log n}$ as the bound on the regret. Here, we used $\tilde{\boldsymbol{w}}_1 = \mathbf{1}$, as $\nabla\mathcal{R}(\mathbf{1}) = 0$. A constant better than 2 can be achieved by a more careful analysis.

2.8 Time-varying learning rate

Suppose the learning rate η_t is varied throughout the game. Can this give us interesting guarantees that cannot be obtained with fixed η? In this section we will give a regret guarantee and a situation where such a method is strictly more powerful than fixing η ahead of time.

For simplicity, suppose that the problem is unconstrained, i.e. $\mathcal{K} = \mathbb{R}^n$. Let $\tilde{\boldsymbol{w}}_t$ be defined as in (2.2):

$$\boldsymbol{w}_{t+1} = \arg\min_{\boldsymbol{w}\in\mathcal{K}} \left[\sum_{s=1}^{t} \eta_s \ell_s(\boldsymbol{w}) + \mathcal{R}(\boldsymbol{w})\right].$$

The following lemma is an easy generalization of Lemma 1.

Lemma 17. *Suppose $\mathcal{K} = \mathbb{R}^n$, i.e. the problem is unconstrained. For any $\mathbf{u} \in \mathcal{K}$,*

$$\sum_{t=1}^{T} \ell_t(\tilde{\boldsymbol{w}}_t) - \ell_t(\mathbf{u}) \le \sum_{t=1}^{T} \eta_t^{-1}\left(D_{\Phi_t}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}) + D_{\Phi_{t-1}}(\mathbf{u}, \tilde{\boldsymbol{w}}_t) - D_{\Phi_t}(\mathbf{u}, \tilde{\boldsymbol{w}}_{t+1})\right)$$

The results of the previous sections are easily extended to varying η_t. For instance, as in Lemma 15 linearizing the functions ℓ_t by $\tilde{\ell}_t(\boldsymbol{w}) = \ell_t(\boldsymbol{w}_t) + \nabla\ell_t(\boldsymbol{w}_t)^\intercal(\boldsymbol{w} - \boldsymbol{w}_t)$, we observe that

$$\sum_{t=1}^{T} \tilde{\ell}_t(\tilde{\boldsymbol{w}}_t) - \tilde{\ell}_t(\mathbf{u}) \le \sum_{t=1}^{T} \eta_t^{-1}\left(D_{\mathcal{R}}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}) + D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_t) - D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_{t+1})\right) \tag{2.16}$$

Example: logarithmic regret for strongly-convex functions

Let us define a notion of strong convexity with respect to $\mathcal{R}$.

Definition 18. A function g over a convex set $\mathcal{K}$ is called σ-strongly convex with respect to $\mathcal{R}$ if

$$\forall \boldsymbol{w}, \mathbf{y} \in K \ , \ g(\boldsymbol{w}) \geq g(\mathbf{y}) + \nabla g(\mathbf{y})^{\intercal}(\boldsymbol{w} - \mathbf{y}) + \frac{\sigma}{2} D_{\mathcal{R}}(\boldsymbol{w}, \mathbf{y}).$$

It is clear that if ℓ_t is σ_t-strongly convex, then

$$\ell_t(\tilde{\boldsymbol{w}}_t) - \ell_t(\mathbf{u}) \leq \tilde{\ell}_t(\tilde{\boldsymbol{w}}_t) - \tilde{\ell}_t(\mathbf{u}) - \frac{\sigma_t}{2} D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_t).$$

Plugging into equation (2.16),

$$\begin{aligned}
\sum_{t=1}^{T} \ell_t(\tilde{\boldsymbol{w}}_t) - \ell_t(\mathbf{u}) &\leq \sum_{t=1}^{T} \left(\tilde{\ell}_t(\tilde{\boldsymbol{w}}_t) - \tilde{\ell}_t(\mathbf{u}) - \frac{\sigma_t}{2} D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_t) \right) \\
&\leq \sum_{t=1}^{T} \left(\eta_t^{-1} D_{\mathcal{R}}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}) + \eta_t^{-1} D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_t) - \eta_t^{-1} D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_{t+1}) - \frac{\sigma_t}{2} D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_t) \right) \\
&= \sum_{t=1}^{T} \eta_t^{-1} D_{\mathcal{R}}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}) + \sum_{t=1}^{T} \left((\eta_t^{-1} - \frac{\sigma_t}{2}) D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_t) - \eta_t^{-1} D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_{t+1}) \right) \\
&\leq \sum_{t=1}^{T} \eta_t^{-1} D_{\mathcal{R}}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1}) + \sum_{t=2}^{T} (\eta_t^{-1} - \frac{\sigma_t}{2} - \eta_{t-1}^{-1}) D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_t) + (\eta_1^{-1} - \frac{\sigma_1}{2}) D_{\mathcal{R}}(\mathbf{u}, \tilde{\boldsymbol{w}}_1)
\end{aligned}$$

Defining $\eta_t = \left(\frac{1}{2} \sum_{s=1}^{t} \sigma_s \right)^{-1}$, we obtain

$$\sum_{t=1}^{T} \ell_t(\tilde{\boldsymbol{w}}_t) - \ell_t(\mathbf{u}) \leq \sum_{t=1}^{T} \eta_t^{-1} D_{\mathcal{R}}(\tilde{\boldsymbol{w}}_t, \tilde{\boldsymbol{w}}_{t+1})$$

If $R(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|^2$,

$$\sum_{t=1}^{T} \ell_t(\tilde{\boldsymbol{w}}_t) - \ell_t(\mathbf{u}) \leq \frac{1}{2} \sum_{t=1}^{T} \eta_t^{-1} \|\eta_t \nabla \ell_t\|^2 \leq \sum_{t=1}^{T} \frac{G_t^2}{\sum_{s=1}^{t} \sigma_s},$$

which is $O(\log T)$ if all σ_t and G_t are constant. This recovers a lemma in Bartlett et al 2007.

2.9 Optimization versus Regret Minimization

In standard optimization, one is interested in finding $\boldsymbol{w}$ such that

$$g(\boldsymbol{w}) \leq \min_{\mathbf{u} \in \mathcal{K}} g(\mathbf{u}) + \epsilon,$$

for a convex function g. Efficiency of a particular optimization procedure is often measured as the number of calls to the oracle (e.g gradient information) required to find $\boldsymbol{w}$ to within ϵ.

It is natural to ask about the relation between optimization and regret minimization. One can view the latter as optimization of a changing objective, where the measure of success is calculated with respect to the whole interaction, not just the end objective. So far we showed that availability of a black-box for minimizing a regularized objective (FTRL) implies non-trivial bounds on the regret under some natural

assumptions. What about the other way around? Can one construct an optimization procedure given a regret-minimization black-box?

Suppose we have a procedure such that regret R_T grows sublinearly. Suppose we are interested in minimizing a function $g(\mathbf{u})$. To do so, we simply feed $g(\cdot)$ repeatedly to the regret-minimization black-box and obtain the following guaratee:

$$\sum_{t=1}^{T} g(\boldsymbol{w}_t) - \min_{\mathbf{u}\in\mathcal{K}} \sum_{t=1}^{T} g(\mathbf{u}) \leq R_T.$$

Using convexity of g and denoting $\boldsymbol{w}^\star = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{w}_t$ we see that

$$g(\boldsymbol{w}^\star) \leq \frac{1}{T}\sum_{t=1}^{T} g(\boldsymbol{w}_t) \leq \min_{\mathbf{u}\in\mathcal{K}} g(\mathbf{u}) + \frac{R_T}{T}.$$

To find out time complexity of the procedure, we set $R_T/T = \epsilon$ and solve for T. Thus, $O(\sqrt{T})$-type regret guarantees imply $O(1/\epsilon^2)$ convergence for the optimization procedure.

At a high level, we conclude that ability to optimize a function implies ability to achieve sublinear regret under some natural assumptions. Conversely, sublinear regret allows one to perform optimization. Of course, the rates of convergence, dependence on the dimension, and many other questions are not addressed in this simplistic reduction.

CHAPTER

THREE

FULL-INFORMATION: THE RANDOMIZATION APPROACH

The family of algorithms described in this section is called *Follow the Perturbed Leader*, as opposed to *Follow the Regularized Leader* of the previous section. Most of results of this section are adapted from [5] and the paper of Kalai & Vempala [7]. As a bibliographic remark we note that the randomization approach goes back to Hannan [6].

Recall Lemma 6, which holds for a convex $\mathcal{R}$ (not necessarily *strictly* convex). It says that the regret is constant if we know the future and play $\boldsymbol{w}_{t+1}$ instead of $\boldsymbol{w}_t$ on round t. While such algorithm is only hypothetical, the results of this section have the flavor of "showing that the randomized strategy is close to this hypothetical algorithm".

In the rest of the section, we will assume that the losses $\ell_t()$ are linear. Furthermore, we will take $\mathcal{R}(\boldsymbol{w}_t)$ to be linear and random: $\mathcal{R}(\boldsymbol{w}_t) = \mathbf{r}^\mathsf{T}\boldsymbol{w}_t$. Suppose $\mathbf{r}$ is drawn at the beginning of the game from the distribution f. In this section we study algorithms of the form Here, we aim to bound *expected* regret.

Algorithm 7: Follow the Perturbed Leader

$$\boldsymbol{w}_{t+1} = \arg\min_{\boldsymbol{w}\in\mathcal{K}} \left[\eta \sum_{s=1}^{t} \mathbf{f}_t^\mathsf{T}\boldsymbol{w} + \mathbf{r}^\mathsf{T}\boldsymbol{w} \right] \tag{3.1}$$

The above family of algorithms is called *Follow the Perturbed Leader*. In some sense, in expectation, the linear penalty $\mathbf{r}^\mathsf{T}\boldsymbol{w}$ acts as a convex regularizer.

Note that we cannot use the divergence techniques of the previous section because $\mathcal{R}$ is no longer strictly convex. However, we will exploit Eq. (2.9), which is derived without any convexity assumptions.

In the realm of randomized repeated games, we need to distinguish two types of adversaries. The first type is an *oblivious* adversary who fixes his moves before the game and does not see the instantiation of $\mathbf{r}$. The second type is an adaptive adversary, who can base his choices on our moves, making his decisions dependent on $\mathbf{r}$. For simplicity, we consider the first type throughout these lectures.

Theorem 19. *Suppose $\mathbf{f}_t \in \mathcal{R}_+^n$, $\mathcal{K} \in \mathcal{R}_+^n$, and $f(\mathbf{r})$ has support in $\mathcal{R}_+^n$. Then, for any $\mathbf{u} \in \mathcal{K}$,*

$$\mathbb{E}\left[\sum_{t=1}^{T} \mathbf{f}_t^\mathsf{T}\boldsymbol{w}_t - \sum_{t=1}^{T} \mathbf{f}_t^\mathsf{T}\mathbf{u}\right] \le \sum_{t=1}^{T} \mathbf{f}_t^\mathsf{T}\left[\int_{\{\mathbf{r}:f(\mathbf{r})\ge f(\mathbf{r}-\eta\mathbf{f}_t)\}} \boldsymbol{w}_t f(\mathbf{r})d\mathbf{r}\right] + \eta^{-1}\mathbb{E}\sup_{\boldsymbol{w}\in\mathcal{K}} \mathbf{r}^\mathsf{T}\boldsymbol{w}$$

Proof. Let $\boldsymbol{w}^* = \arg\min_{\boldsymbol{w}\in\mathcal{K}} \sum_{t=1}^T \mathbf{f}_t^\intercal \boldsymbol{w}$. By Eq. (2.9),

$$\sum_{t=1}^T \mathbf{f}_t^\intercal \boldsymbol{w}_t - \sum_{t=1}^T \mathbf{f}_t^\intercal \boldsymbol{w}^* \le \sum_{t=1}^T \mathbf{f}_t^\intercal (\boldsymbol{w}_t - \boldsymbol{w}_{t+1}) + \eta^{-1}(\mathbf{r}^\intercal \boldsymbol{w}^* - \mathbf{r}^\intercal \boldsymbol{w}_1).$$

Taking expectations and dropping the negative term,

$$\mathbb{E}\left[\sum_{t=1}^T \mathbf{f}_t^\intercal \boldsymbol{w}_t - \sum_{t=1}^T \mathbf{f}_t^\intercal \boldsymbol{w}^*\right] \le \mathbb{E}\left[\sum_{t=1}^T \mathbf{f}_t^\intercal (\boldsymbol{w}_t - \boldsymbol{w}_{t+1})\right] + \eta^{-1}\mathbb{E}\sup_{\boldsymbol{w}\in\mathcal{K}} \mathbf{r}^\intercal \boldsymbol{w}.$$

Now,

$$\mathbb{E}\mathbf{f}_t^\intercal \boldsymbol{w}_t = \int \mathbf{f}_t^\intercal \arg\min_{\boldsymbol{w}\in\mathcal{K}} \left[(\eta \sum_{s=1}^{t-1} \mathbf{f}_s + \mathbf{r})^\intercal \boldsymbol{w}\right] f(\mathbf{r}) d\mathbf{r}$$

and

$$\begin{aligned}\mathbb{E}\mathbf{f}_t^\intercal \boldsymbol{w}_{t+1} &= \int \mathbf{f}_t^\intercal \arg\min_{\boldsymbol{w}\in\mathcal{K}} \left[(\eta \sum_{s=1}^{t} \mathbf{f}_s + \mathbf{r})^\intercal \boldsymbol{w}\right] f(\mathbf{r}) d\mathbf{r} \\ &= \int \mathbf{f}_t^\intercal \arg\min_{\boldsymbol{w}\in\mathcal{K}} \left[(\eta \sum_{s=1}^{t-1} \mathbf{f}_s + \mathbf{r}')^\intercal \boldsymbol{w}\right] f(\mathbf{r}' - \eta\mathbf{f}_t) d\mathbf{r}'\end{aligned}$$

where we made a substitution of variables $\mathbf{r}' = \mathbf{r} + \mathbf{f}_t$. Combining,

$$\begin{aligned}\mathbb{E}\mathbf{f}_t^\intercal (\boldsymbol{w}_t - \boldsymbol{w}_{t+1}) &= \int \mathbf{f}_t^\intercal \arg\min_{\boldsymbol{w}\in\mathcal{K}} \left[(\eta \sum_{s=1}^{t-1} \mathbf{f}_s + \mathbf{r})^\intercal \boldsymbol{w}\right] [f(\mathbf{r}) - f(\mathbf{r} - \eta\mathbf{f}_t)]\, d\mathbf{r} \\ &= \int \mathbf{f}_t^\intercal \boldsymbol{w}_t\, [f(\mathbf{r}) - f(\mathbf{r} - \eta\mathbf{f}_t)]\, d\mathbf{r} \\ &= \mathbf{f}_t^\intercal \int \boldsymbol{w}_t\, [f(\mathbf{r}) - f(\mathbf{r} - \eta\mathbf{f}_t)]\, d\mathbf{r} \\ &\le \mathbf{f}_t^\intercal \int_{\{\mathbf{r}: f(\mathbf{r}) \ge f(\mathbf{r}-\eta\mathbf{f}_t)\}} \boldsymbol{w}_t f(\mathbf{r}) d\mathbf{r}\end{aligned}$$

where the last inequality follows by the positivity assumption on the vectors. □

Next, we prove an analogous theorem where we relax the restriction to the positive orthant.

Theorem 20. *For any* $\mathbf{u} \in \mathcal{K}$,

$$\mathbb{E}\sum_{t=1}^T \mathbf{f}_t^\intercal \boldsymbol{w}_t \le \sup_{\mathbf{r},t} \frac{f(\mathbf{r})}{f(\mathbf{r} - \eta\mathbf{f}_t)} \left[\sum_{t=1}^T \mathbf{f}_t^\intercal \mathbf{u} + \eta^{-1}\mathbb{E}\sup_{\boldsymbol{w}\in\mathcal{K}} \mathbf{r}^\intercal \boldsymbol{w} + \eta^{-1}\mathbb{E}\sup_{\boldsymbol{w}\in\mathcal{K}} -\mathbf{r}^\intercal \boldsymbol{w}\right].$$

Proof.

$$\begin{aligned}\mathbb{E}\mathbf{f}_t^\intercal \boldsymbol{w}_t &= \int \mathbf{f}_t^\intercal \arg\min_{\boldsymbol{w}\in\mathcal{K}} \left[\left(\eta \sum_{s=1}^{t-1} \mathbf{f}_s + \mathbf{r}\right)^\intercal \boldsymbol{w}\right] f(\mathbf{r}) d\mathbf{r} \\ &\le \sup_{\mathbf{r},t} \frac{f(\mathbf{r})}{f(\mathbf{r} - \eta\mathbf{f}_t)} \int \mathbf{f}_t^\intercal \arg\min_{\boldsymbol{w}\in\mathcal{K}} \left[\left(\eta \sum_{s=1}^{t} \mathbf{f}_s + \mathbf{r}\right)^\intercal \boldsymbol{w}\right] f(\mathbf{r}) d\mathbf{r} \\ &= \sup_{\mathbf{r},t} \frac{f(\mathbf{r})}{f(\mathbf{r} - \eta\mathbf{f}_t)} \mathbb{E}\mathbf{f}_t^\intercal \boldsymbol{w}_{t+1}\end{aligned}$$

The result follows from Eq. (2.9) similarly to the previous theorem. □

3.1 Example: Experts Setting (application of Theorem 19)

Suppose $\mathcal{K}$ is the N-simplex. Draw $\mathbf{r} \sim \text{Unif}([0,1]^N)$. Suppose $\mathbf{f}_t \in [0,1]^N$ are the losses of experts at time t. Applying Theorem 19,

$$\begin{aligned}
\mathbb{E}R_T &\leq \sum_{t=1}^{T} \mathbf{f}_t^{\mathsf{T}} \left[\int_{\{\mathbf{r}: f(\mathbf{r}) \geq f(\mathbf{r}-\eta\mathbf{f}_t)\}} \boldsymbol{w}_t f(\mathbf{r}) d\mathbf{r} \right] + \eta^{-1} \mathbb{E} \sup_{\boldsymbol{w} \in \mathcal{K}} \mathbf{r}^t r \boldsymbol{w} \\
&\leq \sum_{t=1}^{T} \int_{\{\mathbf{r}: f(\mathbf{r}) \geq f(\mathbf{r}-\eta\mathbf{f}_t)\}} f(\mathbf{r}) d\mathbf{r} + \eta^{-1} \\
&\leq \sum_{t=1}^{T} \text{Vol}(\{\mathbf{r} : \exists i \text{ s.t. } \mathbf{r}[i] - \eta \mathbf{f}_t[i] < 0\}) + \eta^{-1} \\
&\leq \sum_{t=1}^{T} \eta \sum_{i=1}^{N} \mathbf{f}_t[i] + \eta^{-1} \\
&\leq T\eta N + \eta^{-1} \\
&= 2\sqrt{TN}
\end{aligned}$$

if we set $\eta = \frac{1}{\sqrt{TN}}$. The upper bounds on the volume above hold because the set $\text{Vol}(\{\mathbf{r} : \exists i \text{ s.t. } \mathbf{r}[i] - \eta \mathbf{f}_t[i] < 0\})$ can be covered by N slabs with dimensions $\eta \mathbf{f}_t[i] \times 1 \times 1 \ldots \times 1$.

Note that the optimal dependence on N is logarithmic and achieved by exponential weights (regularization with entropy). In fact, using Theorem 20 achieves the optimal dependence. This is exhibited in the next Example.

3.2 Example: Experts setting (application of Theorem 20)

As in the previous example, suppose $\mathcal{K}$ is the N-simplex and $\mathbf{f}_t \in [0,1]^N$ are the losses of experts at time t. However, we now set the distribution of $\mathbf{r}$ to be a two-sided exponential: $f(\mathbf{r}) = (1/2)^N \exp(-\|\mathbf{r}\|_1)$.

First, one can argue (see [5], page 77) that it suffices to prove the result for $\|\mathbf{f}_t\|_1 \leq 1$: it is advantageous for the adversary to make only one expert incur loss at each round.

Now,

$$\frac{f(\mathbf{r})}{f(\mathbf{r} - \eta \mathbf{f}_t)} = \exp(-\|\mathbf{r}\| + \|\mathbf{r} - \eta \mathbf{f}_t\|_1) \leq \exp(\eta \|\mathbf{f}_t\|_1) \leq \exp(\eta).$$

Furthermore,

$$\begin{aligned}
\mathbb{E} \sup_{\boldsymbol{w}} \mathbf{r}^{\mathsf{T}} \boldsymbol{w} + \mathbb{E} \sup_{\boldsymbol{w}} -\mathbf{r}^{\mathsf{T}} \boldsymbol{w} &\leq 2\mathbb{E} \sup_{\boldsymbol{w}} \mathbf{r}^{\mathsf{T}} \boldsymbol{w} \\
&= 2 \int_0^{\infty} \Pr \left[\max_{i=1\ldots N} \mathbf{r}[i] > u \right] du \\
&\leq 2\nu + \frac{2N}{\eta} e^{-\eta\nu} \quad \text{for any } \nu > 0 \\
&= \frac{2(1 + \ln N)}{\eta} \quad \text{for } \nu = \frac{\ln N}{\eta}.
\end{aligned}$$

From Theorem 20 we obtain, for any $\mathbf{u} \in \mathcal{K}$

$$\mathbb{E} \sum_{t=1}^{T} \mathbf{f}_t^{\mathsf{T}} \boldsymbol{w}_t \leq e^{\eta} \left(\sum_{t=1}^{T} \mathbf{f}_t^{\mathsf{T}} \mathbf{u} + \frac{2(1 + \ln N)}{\eta} \right).$$

Note that $e^\eta \le 1 + (e-1)\eta$ for $\eta \in [0,1]$. Hence,

$$\mathbb{E}\sum_{t=1}^{T} \mathbf{f}_t^\intercal \boldsymbol{w}_t \le \sum_{t=1}^{T} \mathbf{f}_t^\intercal \mathbf{u} + (e-1)\eta \sum_{t=1}^{T} \mathbf{f}_t^\intercal \mathbf{u} + \frac{2(1+\ln N)}{\eta} \le \sum_{t=1}^{T} \mathbf{f}_t^\intercal \mathbf{u} + (e-1)\eta T + \frac{2(1+\ln N)}{\eta}.$$

Optimization over η yields

$$\mathbb{E}\sum_{t=1}^{T} \mathbf{f}_t^\intercal \boldsymbol{w}_t \le \sum_{t=1}^{T} \mathbf{f}_t^\intercal \mathbf{u} + 2\sqrt{2(e-1)(1+\ln N)T}$$

for $\eta = \sqrt{\frac{2(1+\ln N)}{(e-1)T}}$.

3.3 Example: Online Shortest Path

In this setting, there is a fixed DAG with source u and sink v. At each time step, the player picks a path from u to v and the opponent reveals the cost of each edge. The loss is the cost of the chosen path.

This problem can be viewed as an instance of OCO. Indeed, we associate each path with some $\boldsymbol{w}_t \in \{0,1\}^{|E|}$, where $|E|$ is the number of edges. Hence, the set of all valid paths is $\mathcal{K} \subseteq \{0,1\}^{|E|}$. The adversary picks delays $\mathbf{f}_t \in \mathcal{R}_+^{|E|}$ and the loss of the player is $\mathbf{f}_t^\intercal \boldsymbol{w}_t$.

To use the Follow the Perturbed Leader methodology, we draw $\mathbf{r} \sim \text{Unif}([0,1]^{|E|})$. We suppose $\mathbf{f}_t \in [0,1]^{|E|}$ and the length of the longest path (number of edges) is ξ. Then, applying Theorem 19,

$$\begin{aligned}
\mathbf{f}_t \int_{\{\mathbf{r}: f(\mathbf{r}) \ge f(\mathbf{r}-\eta\mathbf{f}_t)\}} \boldsymbol{w}_t f(\mathbf{r}) d\mathbf{r} &\le \|\mathbf{f}_t\|_\infty \| \int_{\{\mathbf{r}: f(\mathbf{r}) \ge f(\mathbf{r}-\eta\mathbf{f}_t)\}} \boldsymbol{w}_t f(\mathbf{r}) d\mathbf{r}\|_1 \\
&\le \xi \int_{\{\mathbf{r}: f(\mathbf{r}) \ge f(\mathbf{r}-\eta\mathbf{f}_t)\}} f(\mathbf{r}) d\mathbf{r} \\
&\le \xi |E| \eta
\end{aligned}$$

Hence,

$$\mathbb{E}R_T \le \eta^{-1}\xi + \xi|E|\eta T = 2\xi\sqrt{|E|T}$$

with $\eta = \frac{1}{\sqrt{|E|T}}$. Again, logarithmic dependence on $|E|$ can be achieved by applying Theorem 20.

CHAPTER

FOUR

BANDIT PROBLEMS

In the bandit setting, only partial feedback $\ell_t(\boldsymbol{w}_t)$ is observed at step t. The problem can be called "0th-order sequential optimization". At the moment of writing we do not have interesting results for general losses ℓ_t beyond the linear ones. Hence, throughout this section we assume $\ell_t(\mathbf{u}) = \mathbf{f}_t^\intercal \mathbf{u}$. Most of the results in this section are taken from the recent paper [1].

What can we do if we only observe $\mathbf{f}_t^\intercal \boldsymbol{w}_t$ as our feedback? We cannot appeal to FTRL or Algorithm 6, as we do not have $\mathbf{f}_t$'s in our possession. A natural idea is to estimate $\mathbf{f}_t$; however, we have only one "poke" at the adversary to do so. Can we indeed estimate $\mathbf{f}_t$ from one sample? This turns out to be possible to do with a randomized algorithm, but the variance of such estimate will be large. Moreover, the precise manner in which we estimate the slope $\mathbf{f}_t$ will depend on the local geometry of the set $\mathcal{K}$ at $\boldsymbol{w}_t$. As we show, we can still use FTRL for bandit optimization, but it is necessary to prove a different bound on the regret, one which involves "local" norms.

It can be shown that for bandit optimization, a special regularizer is required: its Hessian should increase as $1/d$ or $1/d^2$ in terms of distance d to the boundary of the set (or anything in between these regimes). Integrating the requirement, we observe that the first condition implies entropy-like behavior while the second behaves like log of the distance to the boundary. While entropy can be easily defined for the simplex, for general convex sets $\mathcal{K}$ we will opt for the second type. Indeed, the regularizer which will work for our purposes is the self-concordant barrier, which exists for any convex set $\mathcal{K}$ and can be efficiently computed for many natural bodies.

First, define the following local Euclidean norm

$$\|\mathbf{u}\|_{\mathbf{z}} = \sqrt{\mathbf{u}^\intercal \nabla^2 \mathcal{R}(\mathbf{z}) \mathbf{u}}$$

with respect to a convex, twice-differentiable $\mathcal{R}$. Recall that $\Phi_t = \eta \sum_{s=1}^t \mathbf{f}_s + \mathcal{R}$.

We start with some results from the theory of Interior Point Methods. The *Newton decrement* for Φ_t is

$$\lambda(\boldsymbol{w}, \Phi_t) := \|\nabla \Phi_t(\boldsymbol{w})\|_{\boldsymbol{w}}^* = \|\nabla^2 \Phi_t(\boldsymbol{w})^{-1} \nabla \Phi_t(\boldsymbol{w})\|_{\boldsymbol{w}}.$$

and note that since $\mathcal{R}$ is self-concordant then so is Φ_t. The above quantity is intended to measure roughly how far your current point is from the global optimum:

Theorem 21 (From [8])**.** *For any self-concordant function g with $\boldsymbol{w}^* = \arg\min g$, whenever $\lambda(\boldsymbol{w}, g) < 1/2$, we have*

$$\|\boldsymbol{w} - \boldsymbol{w}^*\|_{\boldsymbol{w}} \leq 2\lambda(\boldsymbol{w}, g)$$

where the local norm $\|\cdot\|_{\boldsymbol{w}}$ is defined with respect to g, i.e. $\|\mathbf{y}\|_{\boldsymbol{w}} := \sqrt{\mathbf{y}^\intercal (\nabla^2 g(\boldsymbol{w}))\mathbf{y}}$.

Let us introduce the following shorthand: $\|\mathbf{z}\|_t := \|\mathbf{z}\|_{\boldsymbol{w}_t}$ for the norm is defined with respect to $\boldsymbol{w}_t$. As $\boldsymbol{w}_{t+1}$ minimizes Φ_t and $\nabla^2 \Phi_t = \nabla^2 \mathcal{R}$, we immediately obtain

$$\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|_t \leq 2\lambda(\boldsymbol{w}_t, \Phi_t) = 2\eta \|\mathbf{f}_t\|_t^*$$

The last equality holds because, as is easy to check, $\nabla\Phi_t(\boldsymbol{w}_t) = \eta\mathbf{f}_t$. Applying Hölder's Inequality to Corollary 7, we have

Proposition 22. *Suppose for all $t \in \{1 \dots T\}$ we have $\eta\|\mathbf{f}_t\|_t^* \leq \frac{1}{2}$. Then FTRL with a self-concordant barrier $\mathcal{R}$ satisfies*

$$R_T(\mathbf{u}) \leq 2\eta \sum_{t=1}^{T} [\|\mathbf{f}_t\|_t^*]^2 + \eta^{-1}(\mathcal{R}(\mathbf{u}) - \mathcal{R}(\boldsymbol{w}_1)).$$

The result should be compared to Propositions 11 and 14 in Chapter 2.

We can prove a similar result in terms of local norms for $\mathcal{K}$ being a simplex and $\mathcal{R}$ the entropy function. First notice that $\nabla^2\mathcal{R}(\boldsymbol{w}) = \mathrm{diag}(\boldsymbol{w}[1]^{-1}, \dots, \boldsymbol{w}[n]^{-1})$, and that $1 - e^{-x} \leq x$ for all real x. Next,

$$\|\boldsymbol{w}_t - \tilde{\boldsymbol{w}}_{t+1}\|_t = \sqrt{\sum_{i=1}^{n} (\boldsymbol{w}_t[i] - \tilde{\boldsymbol{w}}_{t+1}[i])^2 / \boldsymbol{w}_t[i]} = \sqrt{\sum_{i=1}^{n} \boldsymbol{w}_t[i](1 - e^{-\eta\mathbf{f}_t[i]})^2} \leq \eta\sqrt{\sum_{i=1}^{n} \boldsymbol{w}_t[i]\mathbf{f}_t[i]^2} = \eta\|\mathbf{f}_t\|_t^*.$$

By Hölder's Inequality, in conjunction with Lemma 13,

$$R_T(\mathbf{u}) \leq \sum_{t=1}^{T} \|\mathbf{f}_t\|_t^* \|\boldsymbol{w}_t - \tilde{\boldsymbol{w}}_{t+1}\|_t + \eta^{-1}\mathcal{R}(\mathbf{u}) \leq \eta \sum_{t=1}^{T} (\|\mathbf{f}_t\|_t^*)^2 + \eta^{-1}\mathcal{R}(\mathbf{u}).$$

Proposition 23. *The exponential weights algorithm (i.e. FTRL with entropic regularization) enjoys the following bound:*

$$R_T(\mathbf{u}) \leq \eta \sum_{t=1}^{T} [\|\mathbf{f}_t\|_t^*]^2 + \eta^{-1}\mathcal{R}(\mathbf{u}).$$

[TO BE CONTINUED]

CHAPTER

FIVE

MINIMAX RESULTS AND LOWER BOUNDS

CHAPTER

SIX

VARIANCE BOUNDS

CHAPTER

SEVEN

STOCHASTIC APPROXIMATION

BIBLIOGRAPHY

[1] J. Abernethy and A. Rakhlin. On high probability bounds for bandit problems, 2009. In submission.

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

[3] Olivier Bousquet. Some recent advances in machine learning. Talk, October 2006. Available at http://ml.typepad.com/Talks/iwfhr.pdf.

[4] Y. Censor and S. A. Zenios. *Parallel Optimization: Theory, Algorithms, and Applications*. Oxford University Press, 1997.

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

[6] James Hannan. Approximation to bayes risk in repeated play. *In M. Dresher, A. W. Tucker, and P. Wolfe, editors, Contributions to the Theory of Games, volume III*, pages 97–139, 1957.

[7] Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005.

[8] A. Nemirovski and M. Todd. Interior-point methods for optimization. *Acta Numerica*, pages 191—234, 2008.

[9] Shai Shalev-Shwartz. *Online Learning: Theory, Algorithms, and Applications*. PhD thesis, Hebrew University, 2007.

[10] Shai Shalev-Shwartz and Yoram Singer. A primal-dual perspective of online learning algorithms. *Mach. Learn.*, 69(2-3):115–142, 2007.

[11] A. N. Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet Math. Dokl.*, 4:1035–1038, 1963.

[12] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-posed Problems*. W. H. Winston, Washington, D.C., 1977.

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