1-1-2013

Predictable Sequences and Competing with Strategies

Wei Han
University of Pennsylvania, weihan.upenn@gmail.com

Predictable Sequences and Competing with Strategies

Abstract
First, we study online learning with an extended notion of regret, which is defined with respect to a set of strategies. We develop tools for analyzing the minimax rates and deriving efficient learning algorithms in this scenario. While the standard methods for minimizing the usual notion of regret fail, through our analysis we demonstrate the existence of regret-minimization methods that compete with such sets of strategies as: autoregressive algorithms, strategies based on statistical models, regularized least squares, and follow-the-regularized-leader strategies. In several cases, we also derive efficient learning algorithms.

Then we study how online linear optimization competes with strategies while benefiting from the predictable sequence. We analyze the minimax value of the online linear optimization problem and develop algorithms that take advantage of the predictable sequence and that guarantee performance compared to fixed actions. Later, we extend the story to a model selection problem on multiple predictable sequences. At the end, we re-analyze the problem from the perspective of dynamic regret.

Last, we study the relationship between Approximate Entropy and Shannon Entropy, and propose the adaptive Shannon Entropy approximation methods (e.g., Lempel-Ziv sliding window method) as an alternative approach to quantify the regularity of data. The new approach has the advantage of adaptively choosing the order of regularity.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Applied Mathematics

First Advisor
Alexander Rakhlin

Second Advisor
Abraham J. Wyner

Subject Categories
Applied Mathematics | Statistics and Probability

PREDICTABLE SEQUENCES AND COMPETING WITH STRATEGIES

Wei Han

A DISSERTATION

in

Applied Mathematics and Computational Science

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2013

Supervisor of Dissertation

Alexander Rakhlin
Assistant Professor of Statistics

Co-Supervisor of Dissertation

Abraham J. Wyner
Professor of Statistics

Graduate Group Chairperson

Charles L. Epstein
Thomas A. Scott Professor of Mathematics

Dissertation Committee:
Alexander Rakhlin, Assistant Professor of Statistics
Abraham J. Wyner, Professor of Statistics
Dean P. Foster, Professor of Statistics

Acknowledgements

It seems like yesterday that I arrived on the beautiful campus of the University of Pennsylvania, and it is hard to believe that four wonderful years have already passed. With memories everywhere across the campus, I owe thanks to many people in many ways.

I would like to express my deepest appreciation to all of my wonderful professors. Thanks to Professor Charles L. Epstein for accepting me to The Graduate Group in Applied Mathematics and Computational Science (AMCS). That marked the beginning of all the memories. Thanks also to Professor Abraham J. Wyner for introducing me to the field of statistics and to the Wharton Statistics Department and many interesting projects. I would also like to extend my appreciation to Professor Alexander Rakhlin for introducing me to the field of online learning, and for his support and encouragement throughout my research. Karthik Sridharan offered many brilliant and inspiring ideas, and contributed greatly to my research. Professor Dean Foster served on my thesis committee, and provided extremely valuable career advice. Many thanks, as well, to all of the great professors who instructed me during my twenty courses at Penn.

I could not have completed this program without the generous financial support of several professors and departments at the University of Pennsylvania. Thanks to Professor Charles L. Epstein, Professor Max Kelz, Professor Alexander Rakhlin, Department of Mathematics, AMCS and the Wharton Statistics Department.

Furthermore, I would like to express deep appreciation to all of the staff members of AMCS and the Wharton Statistics Department who helped me and made my life much easier. Thank you, Janet E. Burns, Audrey Masciocchi Caputo, Charlotte Merrick, Tanya Winder, Adam Greenberg, Carol Reich and Sarin Sieng.

In addition, thanks to all my friends, who took classes, shared offices and ideas, discussed problems, played bridge, travelled and tasted delicious food together.

Most especially, I would like to thank my family, my parents and my dear husband, for holding my hand and being so supportive.

ABSTRACT

PREDICTABLE SEQUENCES AND COMPETING WITH STRATEGIES

Wei Han

Alexander Rakhlin

Abraham J. Wyner

First, we study online learning with an extended notion of regret, which is defined with respect to a set of strategies. We develop tools for analyzing the minimax rates and deriving efficient learning algorithms in this scenario. While the standard methods for minimizing the usual notion of regret fail, through our analysis we demonstrate the existence of regret-minimization methods that compete with such sets of strategies as: autoregressive algorithms, strategies based on statistical models, regularized least squares, and follow-the-regularized-leader strategies. In several cases, we also derive efficient learning algorithms.

Then we study how online linear optimization competes with strategies while benefiting from the predictable sequence. We analyze the minimax value of the online linear optimization problem and develop algorithms that take advantage of the predictable sequence and that guarantee performance compared to fixed actions. Later, we extend the story to a model selection problem on multiple predictable sequences. At the end, we re-analyze the problem from the perspective of dynamic regret.

Last, we study the relationship between Approximate Entropy and Shannon

Entropy, and propose the adaptive Shannon Entropy approximation methods (e.g., Lempel-Ziv sliding window method) as an alternative approach to quantify the regularity of data. The new approach has the advantage of adaptively choosing the order of regularity.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Online learning is a subfield of machine learning. There are two key points to distinguish online learning from traditional machine learning algorithms and classical statistical methods. First, traditional machine learning algorithms and classical statistical methods usually have strong assumptions about the data generating mechanism. For example, data come in the independent and identically distributed (i.i.d.) fashion, or data are generated from fixed distributions. However, these assumptions are not true in many environments. Online learning aims to avoid these strong assumptions, and to make sure that the algorithm works well, regardless of the data generating mechanism. Second, online learning focuses on the environment where new data constantly arrive over time, and real-time decisions are made at every step. Therefore, online learning algorithms emphasize computational efficiency.

Let us illustrate these ideas by a stock trading example. There are many stocks in the stock market. In each period, we only have enough resources to purchase a subset of these stocks. To keep the illustration simple, let us assume that we only have limited money and can only buy one single stock every day. The potential profit to be made from each stock varies on each day. At the end of each day, we observe profits of all the stocks in the stock market. The process of buying stocks and observing profit is repeated day after day. Before buying stocks, we have collected the profit information from previous days. We are interested in designing a trading strategy that maximizes our profit from buying stocks.

How would classic statistical methods deal with this problem? A statistician might assume that there is a model for the profit to be gained from each stock, and previous observations are realizations from these distributions. Then, the statistician estimates the expectation of the profit of buying each stock on each day, and chooses the stock with the highest average of profit according to the history. If the data are indeed generated according to the assumed model, this algorithm will work well in the long run. But what if the assumption is inaccurate and the data do not follow this underlying structure? This strategy can perform badly in some scenarios. For example, suppose the profit of buying Stock A is alternatively 2 or 0, and the profit of buying Stock B is consistently 1. According to the previous model, we will randomly choose Stock A or Stock B on the odd days, and consistently choose Stock A on the even days. However, it is wiser to choose Stock B on all even days.

In the stock scenario, it is unreasonable to assume that the profits come from fixed distribution, and that the data generating process is highly non-stationary. Many factors influence the stock market. Also, other traders are constantly adjusting their trading strategies. This also affects stocks' profit.

Therefore, what if we know little about the data generating process? What if it is unreasonable to make too many assumptions about the data generating process? Online learning aims to solve such a problem. The key philosophy of online learning is to remain agnostic about how data are generated. Algorithms in online learning guarantee the performance even in the worst situation, and the performance measure does not depend on specific data structures.

Since we make no assumption about how the data are generated, then there can always be one situation that we make very little profit regardless of the algorithm. If we just keep our eyes on the profit, there is no hope to make any progress. So, it is critical to define the performance measure for online learning algorithms.

There is an algorithm that make is as much profit as that made from constantly choosing the best stock over time. In the long run, the difference between the average profit of this algorithm and the average profit of the best stock, which is defined as regret, goes to zero. Furthermore, the regret vanishes at the rate of $\sqrt{\log N/T}$, where N is the size of the stock market. Or, if we collect the opinions from N stock market experts, there exists an algorithm that earns almost the same profit as the best expert among these N experts. It is interesting to notice that the rate of convergence depends on the size of the stock set or the size of the expert set.

Regret is the key performance measure in online learning. Why should we only compete against fixed stocks / fixed experts? It only guarantees that the performance is as good as the best single stock / expert. If there is one stock / expert that performs well over this period of time, we are happy. But, what if none of these fixed stocks / expers performs well enough? If there is a strategy that can choose different stocks at different time based on revealed information, can we design an algorithm that performs as well as the best strategy?

Our research extends significantly the definition of regret, which is the most widely accepted performance measure in online leaning algorithms. Instead of competing with fixed actions, we are competing with strategies, which is defined as a set of functions that map from history to actions. As strategy is history-based, the regret competing with strategies provides a higher standard for performance measure of online learning algorithms.

The difficulty of competing with strategies is dealing with the exponentially growing size of the strategy set. Supposing that there are only finite stocks / experts, the number of strategies can increase exponentially as the duration of the period increases. If the number of stocks / experts is small, the regret competing with fixed actions / finite experts quickly converges to zero. However, if history-based strategies are considered, the size of the expert set N increases exponentially. For example, if we use history in the previous k steps, then there are 3^k strategies. The longer the history is, the larger the strategy set is. It is interesting to note that many strategies are similar to each other, although the number of potential strategies is quite large. It is beneficial to group strategies according to the similarity of strategies. If we can group strategies according to their similarities, then the regret reduces even more quickly.

In Chapter 3, we define several complexity measures of strategies that dramatically reduce the number of "effective strategies". Also, while earlier algorithms are only able to produce as much profit as fixed actions / finite experts, we derive novel algorithms that make as much profit as the best history-based strategy. Furthermore, we provide theoretical support for our algorithms.

Beyond our contribution to regret, we consider how to integrate outside information with the learning process. For example, there are seasonal factors that affect the stock market, and abnormal behaviors that occur around earnings announcements. In Chapter 4, the problem is formulated and efficient algorithms are derived to incorporate outside information.

Chapter 5 is about entropy estimation. Approximate Entropy, as an approximation of Kolmogrov-Sinai Entropy, is the widely accepted method to quantify the regularity in data, especially medical data. However, it quantifies the regularity only up to a predetermined order, while real data demand a much higher order. We demonstrate the connection between Approximate Entropy and Shannon Entropy. Based on that connection, we propose the adaptive Shannon Entropy approximation methods (e.g., Lempel-Ziv sliding window method) as an alternative approach to quantify the regularity of data. The new approach has the advantage of adaptively choosing the order of regularity to analyze based on the data. Later, we compare the results of the Lempel-Ziv sliding window method with Approximation Entropy on the electroencephalography (EEG) data to measure the depth of anesthesia. The Lempel-Ziv sliding window method yields more accurate results, especially for low entropy data.

Chapter 2

Online Learning

This chapter briefly describes the framework of online learning and related results in online learning. It follows closely to Alexander Rakhlin and Karthik Sridharan's lecture notes of STAT928: Statistical Learning Theory and Sequential Prediction [16].

2.1 Framework of Online Learning

Suppose both the set of the learner's choices $\mathcal{F}$ and the set of the adversary's choices $\mathcal{Z}$ are closed compact sets. The loss function $\ell : \mathcal{F} \times \mathcal{Z} \to \mathbb{R}$ measures the quality of the learner's choice. The learning framework is shown in the following table

Algorithm 1 Learning Framework

for $t = 1$ to T **do**
  We (The learner) pick(s) $f_t \in \mathcal{F}$
  The adversary picks $z_t \in \mathcal{Z}$
  We suffer loss $\ell(f_t, z_t)$
end for

$\mathcal{Z}$ can be also a set of experts. We, as the learner, choose to follow one expert's choice at each step, according to experts' historical performance. We can switch between experts, but can not mix experts' choice at one single step. After the whole learning process, we may regret and say "I should follow the best expert if I know the sequence in advance".

The difference between our cumulative loss and the cumulative loss of the best expert is used to evaluate the algorithm. The difference is defined as **Regret** and is formalized as

$$\mathbf{Reg}_T(\mathcal{F}) = \sum_{t=1}^{T} \ell(f_t, z_t) - \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell(f, z_t).$$

For the given pair $(\mathcal{F}, \mathcal{Z})$, the problem is called online learnable if there exists an algorithm that achieves sub-linear regret. One main problem in online learning is to analyze the learnability. Another problem is the construction of low-regret algorithms.

2.2 The Minimax Analysis

The minimax value of the online learning problem is defined as

$$\mathcal{V}_T(\mathcal{F}) = \inf_{p_1 \in \Delta(\mathcal{F})} \sup_{z_1 \in \mathcal{Z}} \mathbb{E}_{f_1 \sim p_1} \dots \inf_{p_T \in \Delta(\mathcal{F})} \sup_{z_T \in \mathcal{Z}} \mathbb{E}_{f_T \sim p_T} \left(\sum_{t=1}^{T} \ell(f_t, z_t) - \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell(f, z_t) \right),$$

where $\Delta(\mathcal{F})$ includes all distributions over the set $\mathcal{F}$. It is a powerful way to compress the process of online learning in one single formula.

The upper bound of $\mathcal{V}_T(\mathcal{F})$ guarantees the existence of learning algorithms that perform as well as the best one in the comparator set $\mathcal{F}$. The lower bound indicates that the adversary can cause that much damage regardless of learning algorithms. Also, the minimax regret gives us the access to analyze the learnability of more general online learning problem and to design learning algorithms.

Let us prepare several notations for further analysis. First, Rademacher random variables ϵ represents a fair coin flip, and it equals to 1 or -1 with probability 0.5. Sequences of Rademacher random variables are embedded in Sequential Rademacher Complexity [18] to capture the sequential nature of the online learning problem. Before we jump into the definition of Sequential Rademacher complexity, we first introduce the tree structure. A (complete binary) $\mathcal{Z}$-valued tree $\mathbf{z}$ of depth T is a collection of functions $\mathbf{z}_1, \dots, \mathbf{z}_T$ such that $\mathbf{z}_i : \{\pm 1\}^{i-1} \to \mathcal{Z}$ and $\mathbf{z}_1$ is a constant function. A sequence of i.i.d. Rademacher random variables $(\epsilon_1, \dots, \epsilon_T)$ defines a path in on the tree $\mathbf{z}$:

$$\mathbf{z}_1, \mathbf{z}_2(\epsilon_1), \mathbf{z}_3(\epsilon_1, \epsilon_2), \dots, \mathbf{z}_T(\epsilon_1, \dots, \epsilon_{T-1}).$$

Definition 2.2.1. The Sequential Rademacher Complexity of a function class $\mathcal{F}$ is defined as

$$\mathcal{R}_T^{\text{seq}}(\mathcal{F}) = \sup_{\mathbf{z}} \mathbb{E}_\epsilon \left[\sup_{f \in \mathcal{F}} \sum_{t=1}^{T} \epsilon_t \ell(f, \mathbf{z}_t(\epsilon_1, \dots, \epsilon_{t-1})) \right],$$

where the outer supremum is taken over all $\mathcal{Z}$-valued trees of depth T.

Furthermore, the value of the game $\mathcal{V}_T(\mathcal{F})$ is upper bounded by twice Sequential Rademacher Complexity $\mathcal{R}_T^{\text{seq}}(\mathcal{F})$.

Theorem 2.2.2. [18] *The minimax value of the online learning problem is bounded by*

$$\mathcal{V}_T(\mathcal{F}) \le 2\mathcal{R}_T^{seq}(\mathcal{F})$$

2.3 Algorithms in Online Learning

There are many interesting algorithms in online learning setting. I show two algorithms in this subsection, one is the Exponential Weights Algorithm and another is the Mirror Decent Algorithm.

2.3.1 Exponential Weights Algorithm

The Exponential Weights Algorithm focuses on the finite experts setting. For $t = 1, \dots, T$, the learner observes N difference choices $f_t^1, \dots, f_t^N \in \mathcal{F}$, chooses a distribution p_t in a $N-1$ simplex, picks one choice $f_t^{i_t}$ from these N choices according to the distribution p_t, observe z_t and suffers loss $\ell(f_t^{i_t}, z_t)$. The loss function $\ell : \mathcal{F} \times \mathcal{Z} \to \mathbb{R}$ is convex in its first argument and takes value in $[0,1]$. The goal is to minimize the regret defined as

$$\sum_{t=1}^{T} \mathop{\mathbb{E}}_{i_t \sim q_t} \ell(f_t^{i_t}, z_t) - \inf_{i \in \{1,\dots,N\}} \sum_{t=1}^{T} \ell(f_t^i, z_t).$$

The Exponential Weights Algorithm randomly chooses one expert according to the historical performance. The probability of choosing one expert reduces if its historical performance is worse than other experts, otherwise, the probability increases. The Exponential Weights Algorithm achieves the optimal convergence rate of regret $\mathcal{O}(\sqrt{T \ln N})$, where N is the size of the expert set.

Algorithm 2 Exponential Weights Algorithm

Initialize: $q_1 = (\frac{1}{N}, \dots, \frac{1}{N}), \eta = \sqrt{\frac{8 \ln N}{T}}$
for $t = 1$ to T **do**
  Sample $i_t \sim q_t$, and predict $f^{i_t} \in \mathcal{F}$
  Observe z_t and update

$$q_{t+1}(i) \propto q_t(i) e^{-\eta \ell(f^i, z_t)} \text{ for all } i \in \{1, \dots, N\}$$

end for

2.3.2 Mirror Descent Algorithm

Suppose both the set the learner's choices $\mathcal{F}$ and the outcome set $\mathcal{Z}$ are convex. At $t = 1, \dots, T$, the learner chooses $f_t \in \mathcal{F}$, observes $z_t \in \mathcal{Z}$ and suffers loss $\langle f_t, z_t \rangle$. The goal is to minimize the regret defined as

$$\sum_{t=1}^{T} \langle f_t, z_t \rangle - \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \langle f, z_t \rangle.$$

Let us prepare several definitions for the algorithm. First, a function $\mathcal{R}$ is σ-strongly convex over $\mathcal{F}$ with respect to $\|\cdot\|$ if

$$\mathcal{R}(a) \geq \mathcal{R}(b) + \langle \nabla \mathcal{R}(b), a - b \rangle + \frac{\sigma}{2} \|a - b\|^2$$

for all $a, b \in F$. $\nabla \mathcal{R}(b)$ can be replaced by any subgradient in $\partial \mathcal{R}(b)$ if $\mathcal{R}$ is non-differentiable. Then, the Bregman divergence $\mathcal{D}(a, b)$ with respect to the σ-strongly convex function $\mathcal{R}$ is defined as

$$\mathcal{D}_{\mathcal{R}}(a, b) = \mathcal{R}(a) - \mathcal{R}(b) - \langle \Delta \mathcal{R}(b), a - b \rangle$$

and the convex conjugate of function $\mathcal{R}$ is defined as

$$\mathcal{R}^{\star}(u) = \sup_{a} \langle u, a \rangle - \mathcal{R}(a).$$

Algorithm 3 Mirror Descent Algorithm

Input: $\mathcal{R}$ is σ-strongly convex with respect to $\|\cdot\|$, learning rate $\eta > 0$

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

$$f_{t+1} = \arg\min_{f \in \mathcal{F}} \langle f, z_t \rangle + \eta^{-1} \mathcal{D}_{\mathcal{R}}(f, f_t)$$

or, equivalently,

$$\tilde{f}_{t+1} = \nabla \mathcal{R}^{\star}(\nabla \mathcal{R}(f_t) - \eta z_t) \text{ and } f_{t+1} = \arg\min_{f \in \mathcal{F}} \mathcal{D}_{\mathcal{R}}(f, \tilde{f}_{t+1})$$

end for

Mirror Descent, as a general version of Gradient Descent, focuses on online convex optimization and is computational efficient. Figure 2.1 illustrates three steps of the Mirror Descent Algorithm. First, the input f_t in the primal space $\mathcal{D}$ is mapped to the dual space by $\nabla \mathcal{R}$. Then, the gradient descent step $-\eta z_t$ is done in the dual space $\mathcal{D}^{\star}$. At the end, the gradient descent update $\nabla \mathcal{R}(f_t) - \eta z_t$ is mapped back to the primal space $\mathcal{D}^{\star}$ by $\nabla \mathcal{R}^{\star}$.

2.4 The Algorithm Framework

Several learning algorithms developed in the later chapters are based on a principled way of deriving online learning algorithms from the minimax analysis. The local sequential Rademacher complexities and relaxation help us to obtain faster rates in online learning. Details about the algorithm framework are introduced in [15].

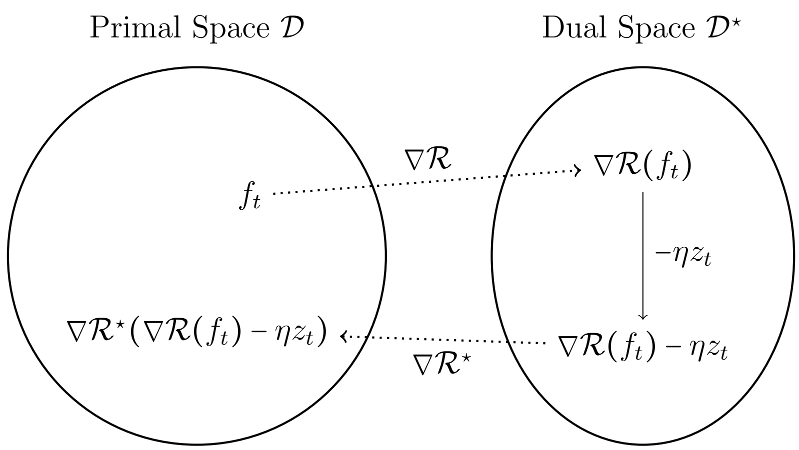


Figure 2.1: Mirror Descent

Chapter 3

Competing with Strategies

In this chapter, we study the problem of online learning with a notion of regret defined with respect to a set of strategies. We develop tools for analyzing the minimax rates and for deriving regret-minimization algorithms in this scenario. While the standard methods for minimizing the usual notion of regret fail, through our analysis we demonstrate existence of regret-minimization methods that compete with such sets of strategies as: autoregressive algorithms, strategies based on statistical models, regularized least squares, and follow the regularized leader strategies. In several cases we also derive efficient learning algorithms.

3.1 Introduction

The common criterion for evaluating an online learning algorithm is *regret*, that is the difference between the cumulative loss of the algorithm and the cumulative loss of the best fixed decision, chosen in hindsight. While much work has been done on understanding no-regret algorithms, such a definition of regret against a fixed decision often draws criticism: even if regret is small, the cumulative loss of a best *fixed* action can be large, thus rendering the result uninteresting. To address this problem, various generalizations of the regret notion have been proposed, including regret with respect to the cost of a "slowly changing" compound decision. While being a step in the right direction, such definitions are still "static" in the sense that the decision of each compound comparator per step does not depend on the sequence of realized outcomes.

Arguably, a more interesting (and more difficult to deal with) notion is that of performing as well as a set of *strategies* (or, *algorithms*). A strategy π is a sequence of functions π_t, for each time period t, mapping the observed outcomes to the next action. Of course, if the collection of such strategies is finite, we may disregard their dependence on the actual sequence and treat each strategy as a black box expert. This is precisely the reason the Multiplicative Weights and other expert algorithms gained such popularity. However, this "black box" approach is not always desirable since some measure of the "effective number" of experts must play a role in the complexity of the problem: experts that predict similarly should not count as two independent ones. But what is a notion of closeness of two strategies? Imagine that we would like to develop an algorithm that incurs loss comparable to that of the best of an infinite family of strategies. To obtain such a statement, one may try to discretize the space of strategies and invoke the black-box experts method. As we show in this chapter, such an approach will not always work. Instead, we present a theoretical framework for the analysis of "competing against strategies" and for algorithmic development, based on the ideas in [18, 15].

The strategies considered in this chapter are termed "simulatable experts" in [3]. The authors also distinguish *static* and *non-static* experts. In particular, for static experts and absolute loss, [2] were able to show that problem complexity is governed by the geometry of the class of static experts as captured by its i.i.d. Rademacher averages. For nonstatic experts, however, the authors note that "unfortunately we do not have a characterization of the minimax regret by an empirical process", due to the fact that the sequential nature of the online problems is at odds with the i.i.d.-based notions of classical empirical process theory. In recent years, however, a martingale generalization of empirical process theory has emerged, and these tools were shown to characterize learnability of online supervised learning, online convex optimization, and other scenarios [18, 1]. Yet, the machinery developed so far is not directly applicable to the case of general simulatable experts which can be viewed as mappings from an ever-growing set of histories to the space of actions. The goal of this chapter is precisely this: to extend the non-constructive as well as constructive techniques of [18, 15] to simulatable experts. We analyze a number of examples with the developed techniques, but we must admit that our work only scratches the surface. We can imagine further research developing methods that compete with interesting gradient descent methods (parametrized by step size choices), with Bayesian procedures (parametrized by choices of priors), and so on. We also note the connection to online algorithms, where one typically aims to prove a bound on the competitive ratio. Our results can be seen in that light as implying a competitive ratio of one.

We close the introduction with a high-level outlook, which builds on the ideas of [13]. Imagine we are faced with a sequence of data from a probabilistic source, such as a k-Markov model with unknown transition probabilities. A well developed statistical theory tells us how to estimate the parameter *under the assumption that the model is correct*. We may view an estimator as a *strategy* for predicting the next outcome. Suppose we have a set of possible models, with a good prediction strategy for each model. Now, let us lift the assumption that the sequence is generated by one of these models, and set the goal as that of performing as well as the best prediction strategy. In this case, if the observed sequence is indeed given by one of the models, our loss will be small because one of the strategies will perform well. If not, we still have a valid statement that does not rely on the fact that the model is "well specified". To illustrate the point, we will exhibit an example where we can compete with the set of all Bayesian strategies (parametrized by priors). We then obtain a statement that we perform as well as the best of them without assuming that the model is correct.

This chapter is organized as follows. In Section 3.2, we extend the minimax analysis of online learning problems to the case of competing with a set of strategies. In Section 3.3, we show that it is possible to compete with a set of autoregressive strategies, and that the usual online linear optimization algorithms do not attain the optimal bounds. We then derive an optimal and computationally efficient algorithm for one of the proposed regimes. In Section 3.4 we describe the general idea of competing with statistical models that use sufficient statistics, and demonstrate an example of competing with a set of strategies parametrized by priors. For this example, we derive an optimal and efficient randomized algorithm. In Section 3.5, we turn to the question of competing with regularized least squares algorithms indexed by the choice of a shift and a regularization parameter. In Section 3.6, we consider online linear optimization and show that it is possible to compete with Follow the Regularized Leader methods parametrized by a shift and by a step size schedule.

3.2 Minimax Regret and Sequential Rademacher Complexity

We consider the problem of online learning, or sequential prediction, that consists of T rounds. At each time $t = \{1, \dots, T\} \triangleq [T]$, the learner makes a prediction $f_t \in \mathcal{F}$ and observes an outcome $z_t \in \mathcal{Z}$, where $\mathcal{F}$ and $\mathcal{Z}$ are abstract sets of decisions and outcomes. Let us fix a loss function $\ell : \mathcal{F} \times \mathcal{Z} \mapsto \mathbb{R}$ that measures the quality of prediction. A *strategy* $\pi = (\pi_t)_{t=1}^T$ is a sequence of functions $\pi_t : \mathcal{Z}^{t-1} \mapsto \mathcal{F}$ mapping history of outcomes to a decision. Let Π denote a set of strategies. The regret with respect to Π is the difference between the cumulative loss of the player and the cumulative loss of the best strategy

$$\mathbf{Reg}_T = \sum_{t=1}^{T} \ell(f_t, z_t) - \inf_{\pi \in \Pi} \sum_{t=1}^{T} \ell(\pi_t(z_{1:t-1}), z_t).$$

where we use the notation $z_{1:k} \triangleq \{z_1, \dots, z_k\}$. We now define the value of the game against a set Π of strategies as

$$\mathcal{V}_T(\Pi) \triangleq \inf_{q_1 \in \mathcal{Q}} \sup_{z_1 \in \mathcal{Z}} \mathop{\mathbb{E}}_{f_1 \sim q_1} \dots \inf_{q_T \in \mathcal{Q}} \sup_{z_T \in \mathcal{Z}} \mathop{\mathbb{E}}_{f_T \sim q_T} [\mathbf{Reg}_T]$$

where $\mathcal{Q}$ and $\mathcal{P}$ are the sets of probability distributions on $\mathcal{F}$ and $\mathcal{Z}$, correspondingly. It was shown in [18] that one can derive non-constructive upper bounds on the value through a process of sequential symmetrization, and in [15] it was shown that these non-constructive bounds can be used as relaxations to derive an algorithm. This is the path we take in this chapter.

Let us describe an important variant of the above problem – that of *supervised learning*. Here, before making a real-valued prediction $\hat{y}_t$ on round t, the learner observes side information $x_t \in \mathcal{X}$. Simultaneously, the actual outcome $y_t \in \mathcal{Y}$ is chosen by Nature. A strategy can therefore depend on the history $x_{1:t-1}, y_{t-1}$ *and* the current x_t, and we write such strategies as $\pi_t(x_{1:t}, y_{1:t-1})$, with $\pi_t : \mathcal{X}^t \times \mathcal{Y}^{t-1} \mapsto \mathcal{Y}$. Fix some loss function $\ell(\hat{y}, y)$. The value $\mathcal{V}_T^S(\Pi)$ is then defined as

$$\sup_{x_1} \inf_{q_1 \in \Delta(\mathcal{Y})} \sup_{y_1 \in \mathcal{Y}} \mathop{\mathbb{E}}_{\hat{y}_1 \sim q_1} \ldots \sup_{x_T} \inf_{q_T \in \Delta(\mathcal{Y})} \sup_{y_T \in \mathcal{Y}} \mathop{\mathbb{E}}_{\hat{y}_T \sim q_T} \left[\sum_{t=1}^{T} \ell(\hat{y}_t, y_t) - \inf_{\pi \in \Pi} \sum_{t=1}^{T} \ell(\pi_t(x_{1:t}, y_{1:t-1}), y_t) \right]$$

To proceed, we need to define a notion of a tree. A $\mathcal{Z}$-valued tree $\mathbf{z}$ is a sequence of mappings $\{\mathbf{z}_1, \ldots, \mathbf{z}_T\}$ with $\mathbf{z}_t : \{\pm 1\}^{t-1} \mapsto \mathcal{Z}$. Throughout this chapter, $\epsilon_t \in \{\pm 1\}$ are i.i.d. Rademacher variables, and a realization of $\epsilon = (\epsilon_1, \ldots, \epsilon_T)$ defines a *path* on the tree, given by $\mathbf{z}_{1:t}(\epsilon) \triangleq (\mathbf{z}_1(\epsilon), \ldots, \mathbf{z}_t(\epsilon))$ for any $t \in [T]$. We write $\mathbf{z}_t(\epsilon)$ for $\mathbf{z}_t(\epsilon_{1:t-1})$. By convention, a sum $\sum_a^b = 0$ for $a > b$ and for simplicity assume that no loss is suffered on the first round.

Definition 3.2.1. Sequential Rademacher complexity of the set Π of strategies is defined as

$$\mathfrak{R}(\ell, \Pi) \triangleq \sup_{\mathbf{w}, \mathbf{z}} \mathbb{E}_\epsilon \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \epsilon_t \ell(\pi_t(\mathbf{w}_1(\epsilon), \ldots, \mathbf{w}_{t-1}(\epsilon)), \mathbf{z}_t(\epsilon)) \right] \tag{3.2.1}$$

where the supremum is over two $\mathcal{Z}$-valued trees $\mathbf{z}$ and $\mathbf{w}$ of depth T.

The $\mathbf{w}$ tree can be thought of as providing "history" while $\mathbf{z}$ providing "outcomes". We shall use these names throughout this chapter. The reader might notice that in the above definition, the outcomes and history are decoupled. We now state the main result:

Theorem 3.2.2. *The value of prediction problem with a set* Π *of strategies is upper bounded as*

$$\mathcal{V}_T(\Pi) \leq 2\mathfrak{R}(\ell, \Pi)$$

While the statement is visually similar to those in [18, 19], it does not follow from these works. Indeed, the proof needs to deal with the additional complications stemming from the dependence of strategies on the history. Further, we provide the proof for a more general case when sequences $z_1, \ldots, z_T$ are not arbitrary but need to satisfy constraints.

Proof. Let us prove a more general version of Theorem 3.2.2. The extra twist is that we allow constraints on the sequences $z_1, \ldots, z_T$ played by the adversary.

Specifically, the adversary at round t can only play x_t that satisfy constraint $C_t(z_1, \ldots, z_t) = 1$ where $(C_1, \ldots, C_T)$ is a predetermined sequence of constraints with $C_t : \mathcal{Z}^t \mapsto \{0, 1\}$. When each C_t is the function that is always 1 then we are in the setting of the theorem statement where we play an unconstrained/worst case adversary. However the proof here allows us to even analyze constrained adversaries which come in handy in many cases. Following [19], a *restriction* $\mathcal{P}_{1:T}$ on the adversary is a sequence $\mathcal{P}_1, \ldots, \mathcal{P}_T$ of mappings $\mathcal{P}_t : \mathcal{Z}^{t-1} \mapsto 2^{\mathcal{P}}$ such that $\mathcal{P}_t(z_{1:t-1})$ is a *convex* subset of $\mathcal{P}$ for any $z_{1:t-1} \in \mathcal{Z}^{t-1}$. In the present proof we will only consider *constrained adversaries*, where $\mathcal{P}_t = \Delta(\mathcal{C}_t(z_{1:t-1}))$ is the set of all distributions on the constrained subset

$$\mathcal{C}_t(z_{1:t-1}) \triangleq \{z \in \mathcal{Z} : C_t(z_1, \ldots, z_{t-1}, z) = 1\}.$$

defined at time t via a binary constraint $C_t : \mathcal{Z}^t \mapsto \{0, 1\}$. Notice that the set $\mathcal{C}_t(z_{1:t-1})$ is the subset of $\mathcal{Z}$ from which the adversary is allowed to pick instance z_t from given the history so far. It was shown in [19] that such constraints can model sequences with certain properties, such as slowly changing sequences, low-variance sequences, and so on. Let $\mathbf{C}$ be the set of $\mathcal{Z}$-valued trees $\mathbf{z}$ such that for every $\epsilon \in \{\pm 1\}^T$ and $t \in [T]$,

$$C_t(\mathbf{z}_1(\epsilon), \ldots, \mathbf{z}_t(\epsilon)) = 1,$$

that is, the set of trees such that the constraint is satisfied along any path. The statement we now prove is that the value of the prediction problem with respect to a set Π of strategies and against constrained adversaries (denoted by $\mathcal{V}_T(\Pi, \mathcal{C}_{1:T})$) is upper bounded by twice the sequential complexity

$$\sup_{\mathbf{w} \in \mathbf{C}, \mathbf{z}} \mathbb{E}_\epsilon \sup_{\pi \in \Pi} \sum_{t=1}^{T} \epsilon_t \ell(\pi_t(\mathbf{w}_1(\epsilon), \ldots, \mathbf{w}_{t-1}(\epsilon))), \mathbf{z}_t(\epsilon)) \qquad (3.2.2)$$

where it is crucial that the $\mathbf{w}$ tree ranges over trees that respect the constraints along all paths, while $\mathbf{z}$ is allowed to be an arbitrary $\mathcal{Z}$-valued tree. This fact that $\mathbf{w}$ respects the constraints is the only difference with the original statement of Theorem 3.2.2.

For ease of notation we use $\langle\!\langle\ \rangle\!\rangle_{t=1}^{T}$ to denote repeated application of operators such has sup or inf. For instance, $\langle\!\langle \sup_{a_t \in A} \inf_{b_t \in B} \mathbb{E}_{r_t \sim P} \rangle\!\rangle_{t=1}^{T} [F(a_1, b_1, r_1, ..., a_T, b_T, r_T)]$ denotes $\sup_{a_1 \in A} \inf_{b_1 \in B} \mathbb{E}_{r_1 \sim P} \ldots \sup_{a_T \in A} \inf_{b_T \in B} \mathbb{E}_{r_T \sim P}[F(a_1, b_1, r_1, ..., a_T, b_T, r_T)]$.

The value of a prediction problem with respect to a set of strategies and against constrained adversaries can be written as :

$$\begin{aligned}
\mathcal{V}_T(\Pi, \mathcal{C}_{1:T}) &= \left\langle\!\!\left\langle \inf_{q_t \in \mathcal{Q}} \sup_{p_t \in \mathcal{P}_t(z_{1:t-1})} \mathop{\mathbb{E}}_{f_t \sim q_t, z_t \sim p_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \left[\sum_{t=1}^{T} \ell(f_t, z_t) - \inf_{\pi \in \Pi} \ell(\pi_t(z_{1:t-1}), z_t) \right] \\
&= \left\langle\!\!\left\langle \sup_{p_t \in \mathcal{P}_t(z_{1:t-1})} \mathop{\mathbb{E}}_{z_t \sim p_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \inf_{f_t \in \mathcal{F}} \mathbb{E}_{z'_t} \ell(f_t, z'_t) - \ell(\pi_t(z_{1:t-1}), z_t) \right] \\
&\le \left\langle\!\!\left\langle \sup_{p_t \in \mathcal{P}_t(z_{1:t-1})} \mathop{\mathbb{E}}_{z_t \sim p_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \mathbb{E}_{z'_t} \ell(\pi_t(z_{1:t-1}), z'_t) - \ell(\pi_t(z_{1:t-1}), z_t) \right] \\
&\le \left\langle\!\!\left\langle \sup_{p_t \in \mathcal{P}_t(z_{1:t-1})} \mathop{\mathbb{E}}_{z_t, z'_t \sim p_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \ell(\pi_t(z_{1:t-1}), z'_t) - \ell(\pi_t(z_{1:t-1}), z_t) \right]
\end{aligned}$$

Let us now define the "selector function" $\chi : \mathcal{Z} \times \mathcal{Z} \times \{\pm 1\} \mapsto \mathcal{Z}$ by

$$\chi(z, z', \epsilon) = \begin{cases} z' & \text{if } \epsilon = -1 \\ z & \text{if } \epsilon = 1 \end{cases}$$

In other words, χ_t selects between z_t and z'_t depending on the sign of ϵ. We will use the shorthand $\chi_t(\epsilon_t) \triangleq \chi(z_t, z'_t, \epsilon_t)$ and $\chi_{1:t}(\epsilon_{1:t}) \triangleq (\chi(z_1, z'_1, \epsilon_1), \ldots, \chi(z_t, z'_t, \epsilon_t))$. We can then re-write the last statement as

$$\begin{aligned}
\left\langle\!\!\left\langle \sup_{p_t \in \mathcal{P}_t(\chi_{1:t-1}(\epsilon_{1:t-1}))} \mathop{\mathbb{E}}_{z_t, z'_t \sim p_t} \mathop{\mathbb{E}}_{\epsilon_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi \in \Pi} \Big[\sum_{t=1}^{T} \epsilon_t \big(\ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), z'_t) \\
- \ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), z_t) \big) \Big] \qquad (3.2.3)
\end{aligned}$$

After z_t, z'_t and ϵ_t are revealed, $\chi_t(\epsilon_t)$ is fixed and can only be either z_t or z'_t. We can remove the dependency of $\chi_t(\epsilon)$ on ϵ, and replace $\chi_t(\epsilon)$ by y_t, which is either z_t or z'_t. Therefore, the last statement is upper bounded

$$\begin{aligned}
&\sup_{p_1 \in \mathcal{P}_1} \mathop{\mathbb{E}}_{z_1, z'_1 \sim p_1} \mathop{\mathbb{E}}_{\epsilon_1} \sup_{y_1 \in \{z_1, z'_1\}} \sup_{p_2 \in \mathcal{P}_2(y_1)} \mathop{\mathbb{E}}_{z_2, z'_2 \sim p_2} \mathop{\mathbb{E}}_{\epsilon_2} \sup_{y_2 \in \{z_2, z'_2\}} \cdots \sup_{y_{T-1} \in \{z_{T-1}, z'_{T-1}\}} \sup_{p_T \in \mathcal{P}_T(y_{1:T-1})} \mathop{\mathbb{E}}_{z_T, z'_T \sim p_T} \mathop{\mathbb{E}}_{\epsilon_T} \\
&\qquad \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \epsilon_t \left(\ell(\pi_t(y_{1:t-1}), z'_t) - \ell(\pi_t(y_{1:t-1}), z_t) \right) \right] \\
&= \sup_{z_1, z'_1 \in \mathcal{C}_1} \mathop{\mathbb{E}}_{\epsilon_1} \sup_{y_1 \in \{z_1, z'_1\}} \sup_{z_2, z'_2 \in \mathcal{C}_2(y_1)} \mathop{\mathbb{E}}_{\epsilon_2} \sup_{y_2 \in \{z_2, z'_2\}} \cdots \sup_{y_{T-1} \in \{z_{T-1}, z'_{T-1}\}} \sup_{z_T, z'_T \in \mathcal{C}_T(y_{1:T-1})} \mathop{\mathbb{E}}_{\epsilon_T} \\
&\qquad \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \epsilon_t \left(\ell(\pi_t(y_{1:t-1}), z'_t) - \ell(\pi_t(y_{1:t-1}), z_t) \right) \right]
\end{aligned}$$

Furthermore, as $\{z_t, z'_t\} \in \mathcal{C}_t(y_{1:t-1})$ and $y_t \in \{z_t, z'_t\}$, we can conclude that $y_t \in \mathcal{C}_t(y_{1:t-1})$. If we drop the constraint on z_t and z'_t, and loosen the constraint on y_t to be $y_t \in \mathcal{C}_t(y_{1:t-1})$, the last statement is upper bounded by

$$\begin{aligned}
&\sup_{z_1,z_1'\in\mathcal{Z}} \mathbb{E}_{\epsilon_1} \sup_{y_1\in\mathcal{C}_1} \sup_{z_2,z_2'\in\mathcal{Z}} \mathbb{E}_{\epsilon_2} \sup_{y_2\in\mathcal{C}_2(y_1)} \cdots \sup_{y_{T-1}\in\mathcal{C}_{T-1}(y_{1:T-2})} \sup_{z_T,z_T'\in\mathcal{Z}} \mathbb{E}_{\epsilon_T} \\
&\qquad \sup_{\pi\in\Pi}\left[\sum_{t=1}^{T}\epsilon_t\left(\ell(\pi_t(y_{1:t-1}),z_t') - \ell(\pi_t(y_{1:t-1}),z_t)\right)\right] \\
&= 2\sup_{z_1\in\mathcal{Z}} \mathbb{E}_{\epsilon_1} \sup_{y_1\in\mathcal{C}_1} \sup_{z_2\in\mathcal{Z}} \mathbb{E}_{\epsilon_2} \sup_{y_2\in\mathcal{C}_2(y_1)} \cdots \sup_{y_{T-1}\in\mathcal{C}_{T-1}(y_{1:T-2})} \sup_{z_T\in\mathcal{Z}} \mathbb{E}_{\epsilon_T} \sup_{\pi\in\Pi}\left[\sum_{t=1}^{T}\epsilon_t\ell(\pi_t(y_{1:t-1}),z_t)\right]
\end{aligned} \tag{3.2.4}$$

since the two terms obtaining by splitting the supremum are the same. Next, we replace y_t by w_{t+1} and add supremum over w_1 at the beginning. Since w_1 does not appear in the loss function, the last statement can be rewritten as

$$2\sup_{w_1\in\mathcal{Z}} \sup_{z_1\in\mathcal{Z}} \mathbb{E}_{\epsilon_1} \sup_{w_2\in\mathcal{C}_1} \sup_{z_2\in\mathcal{Z}} \mathbb{E}_{\epsilon_2} \cdots \sup_{w_T\in\mathcal{C}_T(w_{1:T-1})} \sup_{z_T\in\mathcal{Z}} \mathbb{E}_{\epsilon_T} \sup_{\pi\in\Pi}\left[\sum_{t=1}^{T}\epsilon_t\ell(\pi_t(w_{2:t}),z_t)\right]$$

Now, we exchange the order of suprema and expectation and also maintain the constrains,

$$2\sup_{\mathbf{w}\in\mathcal{C}'} \sup_{\mathbf{z}} \mathbb{E}_{\epsilon} \sup_{\pi\in\Pi}\left[\sum_{t=1}^{T}\epsilon_t\ell(\pi_t(\mathbf{w}_{2:t}(\epsilon)),\mathbf{z}_t(\epsilon))\right] = (*)$$

In this step, we passed to the tree notation. Importantly, tree $\mathbf{w}$ does not range over all trees, but can only be a join of two trees in set $\mathcal{C}$, i.e.

$$\mathcal{C}' = \left\{\mathbf{w} : \forall\epsilon_1, \mathbf{w}(\epsilon_1)\in\mathcal{C}\right\}$$

Define $\mathbf{w}^* = \mathbf{w}(-1)$ and $\mathbf{w}^{**} = \mathbf{w}(+1)$, we can expend the expectation in $(*)$ with respect to ϵ_1 of the above expression by

$$\begin{aligned}
&\sup_{\mathbf{w}^*\in\mathbf{C}} \sup_{\mathbf{z}} \mathbb{E}_{\epsilon_{2:T}} \sup_{\pi\in\Pi}\left[-\ell(\pi_1(\cdot),\mathbf{z}_1(\cdot)) + \sum_{t=2}^{T}\epsilon_t\ell(\pi_t(\mathbf{w}^*_{1:t-1}(\epsilon)),\mathbf{z}_t(\epsilon))\right] \\
&+ \sup_{\mathbf{w}^{**}\in\mathbf{C}} \sup_{\mathbf{z}} \mathbb{E}_{\epsilon_{2:T}} \sup_{\pi\in\Pi}\left[\ell(\pi_1(\cdot),\mathbf{z}_1(\cdot)) + \sum_{t=2}^{T}\epsilon_t\ell(\pi_t(\mathbf{w}^{**}_{1:t-1}(\epsilon)),\mathbf{z}_t(\epsilon))\right].
\end{aligned}$$

With the assumption that we do not suffer lose at the first round, which means $\ell(\pi_1(\cdot),\mathbf{z}_1(\cdot)) = 0$, we can see that both terms achieve the suprema with the same $\mathbf{w}^* = \mathbf{w}^{**}$. Therefore, the above expression can be rewrite as

$$\sup_{\mathbf{w}\in\mathbf{C}} \sup_{\mathbf{z}} \mathbb{E}_{\epsilon_{2:T}} \sup_{\pi\in\Pi}\left[\sum_{t=1}^{T}\epsilon_t\ell(\pi_t(\mathbf{w}_{1:t-1}(\epsilon)),\mathbf{z}_t(\epsilon))\right]$$

which is precisely (3.2.2). This concludes the proof of Theorem 3.2.2. □

As we show below, the sequential Rademacher complexity on the right-hand side allows us to analyze general non-static experts, thus addressing the question raised in [2]. As the first step, we can "erase" a Lipschitz loss function, leading to the sequential Rademacher complexity of Π without the loss and without the $\mathbf{z}$ tree:

$$\mathfrak{R}(\Pi) \triangleq \sup_{\mathbf{w}} \mathfrak{R}(\Pi, \mathbf{w}) \triangleq \sup_{\mathbf{w}} \mathbb{E}_\epsilon \sup_{\pi\in\Pi} \left[\sum_{t=1}^T \epsilon_t \pi_t(\mathbf{w}_{1:t-1}(\epsilon))\right]$$

For example, suppose $\mathcal{Z} = \{0,1\}$, the loss function is the indicator loss, and strategies have potentially dependence on the full history. Then one can verify that

$$\begin{aligned}
&\sup_{\mathbf{w},\mathbf{z}} \mathbb{E}_\epsilon \sup_{\pi\in\Pi^k} \left[\sum_{t=1}^T \epsilon_t \mathbf{1}\left\{\pi_t(\mathbf{w}_{1:t-1}(\epsilon)) \neq \mathbf{z}_t(\epsilon)\right\}\right] \\
&= \sup_{\mathbf{w},\mathbf{z}} \mathbb{E}_\epsilon \sup_{\pi\in\Pi^k} \left[\sum_{t=1}^T \epsilon_t \Big(\pi_t(\mathbf{w}_{1:t-1}(\epsilon))(1-2\mathbf{z}_t(\epsilon)) + \mathbf{z}_t(\epsilon)\Big)\right] = \mathfrak{R}(\Pi)
\end{aligned} \tag{3.2.5}$$

The same result holds when $\mathcal{F} = [0,1]$ and ℓ is the absolute loss. The process of "erasing the loss" (or, contraction) extends quite nicely to problems of supervised learning. Let us state the second main result:

Theorem 3.2.3. *Suppose the loss function $\ell : \mathcal{Y}\times\mathcal{Y} \mapsto \mathbb{R}$ is convex and L-Lipschitz in the first argument, and let $\mathcal{Y} = [-1,1]$. Then*

$$\mathcal{V}_T^S(\Pi) \leq 2L \sup_{\mathbf{x},\mathbf{y}} \mathop{\mathbb{E}}_{\epsilon} \sup_{\pi\in\Pi} \left[\sum_{t=1}^T \epsilon_t \pi_t(\mathbf{x}_{1:t}(\epsilon), \mathbf{y}_{1:t-1}(\epsilon))\right]$$

where $(\mathbf{x}_{1:t}(\epsilon), \mathbf{y}_{1:t-1}(\epsilon))$ naturally takes place of $\mathbf{w}_{1:t-1}(\epsilon)$ in Theorem 3.2.2. Further, if $\mathcal{Y} = [-1,1]$ and $\ell(\hat{y}, y) = |\hat{y} - y|$,

$$\mathcal{V}_T^S(\Pi) \geq \sup_{\mathbf{x}} \mathop{\mathbb{E}}_{\epsilon} \sup_{\pi\in\Pi} \left[\sum_{t=1}^T \epsilon_t \pi_t(\mathbf{x}_{1:t}(\epsilon), \epsilon_{1:t-1})\right].$$

Proof. By convexity of the loss,

$$\begin{aligned}
&\left\langle\!\!\left\langle \sup_{x_t\in\mathcal{X}} \inf_{q_t\in\Delta(\mathcal{Y})} \sup_{y_t\in\mathcal{Y}} \mathop{\mathbb{E}}_{\hat{y}_t\sim q_t} \right\rangle\!\!\right\rangle_{t=1}^T \left[\sum_{t=1}^T \ell(\hat{y}_t, y_t) - \inf_{\pi\in\Pi} \sum_{t=1}^T \ell(\pi_t(x_{1:t}, y_{1:t-1}), y_t)\right] \\
&\leq \left\langle\!\!\left\langle \sup_{x_t\in\mathcal{X}} \inf_{q_t\in\Delta(\mathcal{Y})} \sup_{y_t\in\mathcal{Y}} \mathop{\mathbb{E}}_{\hat{y}_t\sim q_t} \right\rangle\!\!\right\rangle_{t=1}^T \sup_{\pi\in\Pi} \left[\sum_{t=1}^T \ell'(\hat{y}_t, y_t)(\hat{y}_t - \pi_t(x_{1:t}, y_{1:t-1}))\right] \\
&\leq \left\langle\!\!\left\langle \sup_{x_t\in\mathcal{X}} \inf_{q_t\in\Delta(\mathcal{Y})} \sup_{y_t\in\mathcal{Y}} \mathop{\mathbb{E}}_{\hat{y}_t\sim q_t} \sup_{s_t\in[-L,L]} \right\rangle\!\!\right\rangle_{t=1}^T \sup_{\pi\in\Pi} \left[\sum_{t=1}^T s_t(\hat{y}_t - \pi_t(x_{1:t}, y_{1:t-1}))\right]
\end{aligned}$$

where in the last step we passed to an upper bound by allowing for the worst-case choice s_t of the derivative. We will often omit the range of the variables in our notation, and it is understood that s_t's range over $[-L, L]$, while $y_t, \hat{y}_t$ over $\mathcal{Y}$ and x_t's over $\mathcal{X}$. Now, by Jensen's inequality, we pass to an upper bound by exchanging $\mathbb{E}_{\hat{y}_t}$ and $\sup_{y_t \in \mathcal{Y}}$:

$$\left\langle\!\!\left\langle \sup_{x_t} \inf_{q_t \in \Delta(\mathcal{Y})} \mathbb{E}_{\hat{y}_t \sim q_t} \sup_{y_t} \sup_{s_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} s_t(\hat{y}_t - \pi_t(x_{1:t}, y_{1:t-1})) \right]$$
$$= \left\langle\!\!\left\langle \sup_{x_t} \inf_{\hat{y}_t \in \mathcal{Y}} \sup_{y_t, s_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} s_t(\hat{y}_t - \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

Consider the last step, assuming all the other variables fixed:

$$\sup_{x_T} \inf_{\hat{y}_T} \sup_{y_T, s_T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} s_t(\hat{y}_t - \pi_t(x_{1:t}, y_{1:t-1})) \right]$$
$$= \sup_{x_T} \inf_{\hat{y}_T} \sup_{p_T \in \Delta(\mathcal{Y} \times [-L, L])} \mathbb{E}_{(y_T, s_T) \sim p_T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} s_t(\hat{y}_t - \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

where the distribution p_T ranges over all distributions on $\mathcal{Y} \times [-L, L]$. Now observe that the function inside the infimum is convex in $\hat{y}_T$, and the function inside $\sup_{p_T}$ is linear in the distribution p_T. Hence, we can appeal to the minimax theorem, obtaining equality of the last expression to

$$\sup_{x_T} \sup_{p_T \in \Delta(\mathcal{Y} \times [-L, L])} \inf_{\hat{y}_T} \mathbb{E}_{(y_T, s_T) \sim p_T} \left[\sum_{t=1}^{T} s_t \hat{y}_t - \inf_{\pi \in \Pi} \sum_{t=1}^{T} s_t \pi_t(x_{1:t}, y_{1:t-1})) \right]$$
$$= \sum_{t=1}^{T-1} s_t \hat{y}_t + \sup_{x_T} \sup_{p_T} \inf_{\hat{y}_T} \mathbb{E}_{(y_T, s_T) \sim p_T} \left[s_T \hat{y}_T - \inf_{\pi \in \Pi} \sum_{t=1}^{T} s_t \pi_t(x_{1:t}, y_{1:t-1})) \right]$$
$$= \sum_{t=1}^{T-1} s_t \hat{y}_t + \sup_{x_T} \sup_{p_T} \left[\inf_{\hat{y}_T} \left(\mathbb{E}_{(y_T, s_T) \sim p_T} s_T \right) \hat{y}_T - \mathbb{E}_{(y_T, s_T) \sim p_T} \inf_{\pi \in \Pi} \sum_{t=1}^{T} s_t \pi_t(x_{1:t}, y_{1:t-1})) \right]$$
$$= \sum_{t=1}^{T-1} s_t \hat{y}_t + \sup_{x_T} \sup_{p_T} \mathbb{E}_{(y_T, s_T) \sim p_T} \left[\inf_{\hat{y}_T} \left(\mathbb{E}_{(y_T, s_T) \sim p_T} s_T \right) \hat{y}_T - \inf_{\pi \in \Pi} \sum_{t=1}^{T} s_t \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

We can now upper bound the choice of $\hat{y}_T$ by that given by π_T, yielding an upper bound

$$\sum_{t=1}^{T-1} s_t \hat{y}_t + \sup_{x_T, p_T} \mathbb{E}_{(y_T, s_T) \sim p_T} \sup_{\pi \in \Pi} \left[\inf_{\hat{y}_T} \left(\mathbb{E}_{(y_T, s_T) \sim p_T} s_T \right) \hat{y}_T - \sum_{t=1}^{T} s_t \pi_t(x_{1:t}, y_{1:t-1})) \right]$$
$$= \sum_{t=1}^{T-1} s_t \hat{y}_t + \sup_{x_T, p_T} \mathbb{E}_{\substack{(y_T, s_T) \\ \sim p_T}} \sup_{\pi \in \Pi} \left[\left(\mathbb{E}_{\substack{(y'_T, s'_T) \\ \sim p_T}} s'_T - s_T \right) \pi_T(x_{1:T}, y_{1:T-1}) - \sum_{t=1}^{T-1} s_t \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

It is not difficult to verify that this process can be repeated for $T-1$ and so on.

The resulting upper bound is therefore

$$
\begin{aligned}
\mathcal{V}_T^S(\Pi) &\leq \left\langle\!\!\left\langle \sup_{x_t, p_t} \mathop{\mathbb{E}}_{(y_t, s_t)\sim p_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi\in\Pi} \left[\sum_{t=1}^{T} \left(\mathop{\mathbb{E}}_{(y'_t, s'_t)\sim p_t} s'_t - s_t \right) \pi_t(x_{1:t}, y_{1:t-1}) \right] \\
&\leq \left\langle\!\!\left\langle \sup_{x_t, p_t} \mathop{\mathbb{E}}_{\substack{(y_t, s_t)\sim p_t \\ (y'_t, s'_t)\sim p_t}} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi\in\Pi} \left[\sum_{t=1}^{T} (s'_t - s_t)\, \pi_t(x_{1:t}, y_{1:t-1}) \right] \\
&= \left\langle\!\!\left\langle \sup_{x_t, p_t} \mathop{\mathbb{E}}_{\substack{(y_t, s_t)\sim p_t \\ (y'_t, s'_t)\sim p_t}} \mathop{\mathbb{E}}_{\epsilon_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi\in\Pi} \left[\sum_{t=1}^{T} \epsilon_t\, (s'_t - s_t)\, \pi_t(x_{1:t}, y_{1:t-1}) \right] \\
&\leq \left\langle\!\!\left\langle \sup_{x_t} \sup_{\substack{(y_t, s_t) \\ (y'_t, s'_t)}} \mathop{\mathbb{E}}_{\epsilon_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi\in\Pi} \left[\sum_{t=1}^{T} \epsilon_t\, (s'_t - s_t)\, \pi_t(x_{1:t}, y_{1:t-1}) \right] \\
&\leq \left\langle\!\!\left\langle \sup_{x_t, y_t} \sup_{s'_t, s_t} \mathop{\mathbb{E}}_{\epsilon_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi\in\Pi} \left[\sum_{t=1}^{T} \epsilon_t\, (s'_t - s_t)\, \pi_t(x_{1:t}, y_{1:t-1}) \right] \\
&\leq 2 \left\langle\!\!\left\langle \sup_{x_t, y_t} \sup_{s_t\in[-L, L]} \mathop{\mathbb{E}}_{\epsilon_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi\in\Pi} \left[\sum_{t=1}^{T} \epsilon_t s_t \pi_t(x_{1:t}, y_{1:t-1}) \right]
\end{aligned}
$$

Since the expression is convex in each s_t, we can replace the range of s_t by $\{-L, L\}$, or, equivalently,

$$
\mathcal{V}_T^S(\Pi) \leq 2L \left\langle\!\!\left\langle \sup_{x_t, y_t} \sup_{s_t\in\{-1,1\}} \mathop{\mathbb{E}}_{\epsilon_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi\in\Pi} \left[\sum_{t=1}^{T} \epsilon_t s_t \pi_t(x_{1:t}, y_{1:t-1}) \right] \tag{3.2.6}
$$

Now consider any arbitrary function $\psi : \{\pm 1\} \mapsto \mathbb{R}$, we have that

$$
\sup_{s\in\{\pm1\}} \mathbb{E}_\epsilon \left[\psi(s\cdot\epsilon) \right] = \sup_{s\in\{\pm1\}} \frac{1}{2} \left(\psi(+s) + \psi(-s) \right) = \frac{1}{2} \left(\psi(+1) + \psi(-1) \right) = \mathbb{E}_\epsilon \left[\psi(\epsilon) \right]
$$

Since in Equation (3.2.6), for each t, s_t and ϵ_t appear together as $\epsilon_t \cdot s_t$ using the above equation repeatedly, we conclude that

$$
\begin{aligned}
\mathcal{V}_T^S(\Pi) &\leq 2L \left\langle\!\!\left\langle \sup_{x_t, y_t} \mathop{\mathbb{E}}_{\epsilon_t} \right\rangle\!\!\right\rangle_{t=1}^{T} \sup_{\pi\in\Pi} \left[\sum_{t=1}^{T} \epsilon_t \pi_t(x_{1:t}, y_{1:t-1}) \right] \\
&= 2L \sup_{\mathbf{x}, \mathbf{y}} \mathop{\mathbb{E}}_{\epsilon} \sup_{\pi\in\Pi} \left[\sum_{t=1}^{T} \epsilon_t \pi_t(\mathbf{x}_{1:t}(\epsilon), \mathbf{y}_{1:t-1}(\epsilon)) \right]
\end{aligned}
$$

The lower bound is obtained by the same argument as in [18].

□

Let us present a few simple examples as a warm-up.

Example 3.2.4 (History-independent strategies). Let $\pi^f \in \Pi$ be constant history-independent strategies $\pi_1^f = \ldots = \pi_T^f = f \in \mathcal{F}$. Then (3.2.1) recovers the definition of sequential Rademacher complexity in [18].

Example 3.2.5 (Static experts). For static experts, each strategy π is a predetermined sequence of outcomes, and we may therefore associate each π with a vector in $\mathcal{Z}^T$. A direct consequence of Theorem 3.2.3 for any convex L-Lipschitz loss is that

$$\mathcal{V}(\Pi) \le 2L\mathbb{E}_\epsilon \sup_{\pi\in\Pi}\left[\sum_{t=1}^{T} \epsilon_t \pi_t\right]$$

which is simply the classical i.i.d. Rademacher averages. For the case of $\mathcal{F} = [0,1]$, $\mathcal{Z} = \{0,1\}$, and the absolute loss, this is the result of [2].

Example 3.2.6 (Finite-order Markov strategies). Let Π^k be a set of strategies that only depend on the k most recent outcomes to determine the next move. Theorem 3.2.2 implies that the value of the game is upper bounded as

$$\mathcal{V}(\Pi^k) \le 2\sup_{\mathbf{w},\mathbf{z}} \mathbb{E}_\epsilon \sup_{\pi\in\Pi^k}\left[\textstyle\sum_{t=1}^{T} \epsilon_t \ell(\pi_t(\mathbf{w}_{t-k}(\epsilon),\ldots,\mathbf{w}_{t-1}(\epsilon)),\mathbf{z}_t(\epsilon))\right]$$

Now, suppose that $\mathcal{Z}$ is a finite set, of cardinality s. Then there are effectively s^{s^k} strategies π. The bound on the sequential Rademacher complexity then scales as $\sqrt{2s^k \log(s) T}$, recovering the result of [7] (see [3, Cor. 8.2]).

In addition to providing an understanding of minimax regret against a set of strategies, sequential Rademacher complexity can serve as a starting point for algorithmic development. As shown in [15], any *admissible relaxation* can be used to define a succinct algorithm with a regret guarantee. For the setting of this chapter, this means the following. Let $\mathbf{Rel} : \mathcal{Z}^t \mapsto \mathbb{R}$, for each t, be a collection of functions satisfying two conditions:

$$\forall t,\ \inf_{q_t\in\mathcal{Q}} \sup_{z_t\in\mathcal{Z}} \left\{ \mathop{\mathbb{E}}_{f_t\sim q_t} \ell(f_t,z_t) + \mathbf{Rel}(z_{1:t}) \right\} \le \mathbf{Rel}(z_{1:t-1}),$$
$$\text{and} \quad -\inf_{\pi\in\Pi} \sum_{t=1}^{T} \ell(\pi_t(z_{1:t-1}),z_t) \le \mathbf{Rel}(z_{1:T})\ .$$

Then we say that *the relaxation is admissible.* It is then easy to show that regret of any algorithm that ensures above inequalities is bounded by $\mathbf{Rel}(\{\})$.

Theorem 3.2.7. *The conditional sequential Rademacher complexity with respect to* Π

$$\begin{aligned}&\mathfrak{R}(\ell,\Pi|z_1,\ldots,z_t)\\ &\quad \triangleq \sup_{\mathbf{z},\mathbf{w}} \mathop{\mathbb{E}}_{\epsilon_{t+1:T}} \sup_{\pi\in\Pi}\left[2\sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t},\mathbf{w}_{1:s-t-1}(\epsilon)),\mathbf{z}_{s-t}(\epsilon)) - \sum_{s=1}^{t} \ell(\pi_s(z_{1:s-1}),z_s)\right]\end{aligned}$$

is admissible.

Proof. Denote $L_t(\pi) = \sum_{s=1}^{t} \ell(\pi_s(z_{1:s-1}), z_s)$. The first step of the proof is an application of the minimax theorem (we assume the necessary conditions hold):

$$
\begin{aligned}
&\inf_{q_t \in \Delta(\mathcal{F})} \sup_{z_t \in \mathcal{Z}} \Bigg\{ \mathop{\mathbb{E}}_{f_t \sim q_t} \left[\ell(f_t, z_t) \right] \\
&\qquad + \sup_{\mathbf{z},\mathbf{w}} \mathop{\mathbb{E}}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \left[2 \sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_t(\pi) \right] \Bigg\} \\
&= \sup_{p_t \in \Delta(\mathcal{Z})} \inf_{f_t \in \mathcal{F}} \Bigg\{ \mathop{\mathbb{E}}_{z_t \sim p_t} \left[\ell(f_t, z_t) \right] \\
&\qquad + \mathop{\mathbb{E}}_{z_t \sim p_t} \sup_{\mathbf{z},\mathbf{w}} \mathop{\mathbb{E}}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \left[2 \sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_t(\pi) \right] \Bigg\}
\end{aligned}
$$

For any $p_t \in \Delta(\mathcal{Z})$, the infimum over f_t of the above expression is equal to

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{z_t \sim p_t} \sup_{\mathbf{z},\mathbf{w}} \mathop{\mathbb{E}}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[2 \sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \\
&\qquad\qquad + \inf_{f_t \in \mathcal{F}} \mathop{\mathbb{E}}_{z_t \sim p_t} \left[\ell(f_t, z_t) \right] - \ell(\pi_t(z_{1:t-1}), z_t) \Bigg] \\
&\le \mathop{\mathbb{E}}_{z_t \sim p_t} \sup_{\mathbf{z},\mathbf{w}} \mathop{\mathbb{E}}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[2 \sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \\
&\qquad\qquad + \mathop{\mathbb{E}}_{z_t \sim p_t} \left[\ell(\pi_t(z_{1:t-1}), z_t) \right] - \ell(\pi_t(z_{1:t-1}), z_t) \Bigg] \\
&\le \mathop{\mathbb{E}}_{z_t, z_t' \sim p_t} \sup_{\mathbf{z},\mathbf{w}} \mathop{\mathbb{E}}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[2 \sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \\
&\qquad\qquad + \ell(\pi_t(z_{1:t-1}), z_t') - \ell(\pi_t(z_{1:t-1}), z_t) \Bigg]
\end{aligned}
$$

We now argue that the independent z_t and z_t' have the same distribution p_t, and thus we can introduce a random sign ϵ_t. The above expression then equals to

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{z_t, z_t' \sim p_t} \mathop{\mathbb{E}}_{\epsilon_t} \sup_{\mathbf{z},\mathbf{w}} \mathop{\mathbb{E}}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[2 \sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) \\
&\qquad\qquad - L_{t-1}(\pi) + \epsilon_t(\ell(\pi_t(z_{1:t-1}), \chi_t(-\epsilon_t))) - \ell(\pi_t(z_{1:t-1}), \chi_t(\epsilon_t))) \Bigg] \\
&\le \mathop{\mathbb{E}}_{z_t, z_t' \sim p_t} \sup_{z'', z'''} \mathop{\mathbb{E}}_{\epsilon_t} \sup_{\mathbf{z},\mathbf{w}} \mathop{\mathbb{E}}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[2 \sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) \\
&\qquad\qquad - L_{t-1}(\pi) + \epsilon_t(\ell(\pi_t(z_{1:t-1}), z_t'') - \ell(\pi_t(z_{1:t-1}), z_t''')) \Bigg]
\end{aligned}
$$

Splitting the resulting expression into two parts, we arrive at the upper bound of

$$
\begin{aligned}
&2 \underset{z_t,z'_t\sim p_t}{\mathbb{E}} \sup_{z''} \underset{\epsilon_t}{\mathbb{E}} \sup_{\mathbf{z},\mathbf{w}} \underset{\epsilon_{t+1:T}}{\mathbb{E}} \sup_{\pi\in\Pi} \Bigg[\sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t-1},\chi_t(\epsilon_t),\mathbf{w}_{1:s-t-1}(\epsilon)),\mathbf{z}_{s-t}(\epsilon)) \\
&\qquad\qquad\qquad\qquad\qquad -\frac{1}{2}L_{t-1}(\pi) + \epsilon_t \ell(\pi_t(z_{1:t-1}),z''_t)\Bigg] \\
&\le \sup_{z,z',z''} \underset{\epsilon_t}{\mathbb{E}} \sup_{\mathbf{z},\mathbf{w}} \underset{\epsilon_{t+1:T}}{\mathbb{E}} \sup_{\pi\in\Pi} \Bigg[\sum_{s=t+1}^{T} 2\epsilon_s \ell(\pi_s((z_{1:t-1},\chi_t(\epsilon_t),\mathbf{w}_{1:s-t-1}(\epsilon)),\mathbf{z}_{s-t}(\epsilon)) \\
&\qquad\qquad\qquad\qquad\qquad -L_{t-1}(\pi) + \epsilon_t \ell(\pi_t(z_{1:t-1}),z''_t)\Bigg] \\
&\le \mathfrak{R}_T(\Pi|z_1,\dots,z_{t-1}).
\end{aligned}
$$

The first inequality is true as we upper bounded the expectation by the supremum. The last inequality is easy to verify, as we are effectively filling in a root z_t and z'_t for the two subtrees, for $\epsilon_t = +1$ and $\epsilon_t = -1$, respectively, and jointing the two trees with a $\varnothing$ root.

One can see that the proof of admissibility corresponds to one step minimax swap and symmetrization in the proof of [18]. In contrast, in the latter chapter, all T minimax swaps are performed at once, followed by T symmetrization steps. □

Conditional sequential Rademacher complexity can therefore be used as a starting point for possibly deriving computationally attractive algorithms, as shown throughout this chapter.

We may now define covering numbers for the set Π of strategies over the history trees. The development is a straightforward modification of the notions we developed in [18], where we replace "any tree $\mathbf{x}$" with a tree of histories $\mathbf{w}_{1:t-1}$.

Definition 3.2.8. A set V of $\mathbb{R}$-valued trees is an *α-cover* (with respect to ℓ_p) of a set of strategies Π on an $\mathcal{Z}^*$-valued history tree $\mathbf{w}$ if

$$
\forall \pi\in\Pi,\ \forall\epsilon\in\{\pm1\}^T,\quad \exists \mathbf{v}\in V \quad \text{s.t.} \quad \left(\tfrac{1}{T}\textstyle\sum_{t=1}^T |\pi_t(\mathbf{w}_{1:t-1}(\epsilon)) - \mathbf{v}_t(\epsilon)|^p\right)^{1/p} \le \alpha\ . \tag{3.2.7}
$$

An *α-covering number* $\mathcal{N}_p(\Pi,\mathbf{w},\alpha)$ is the size of the smallest α-cover.

For supervised learning, $(\mathbf{x}_{1:t}(\epsilon),\mathbf{y}_{1:t-1}(\epsilon))$ takes place of $\mathbf{w}_{1:t-1}(\epsilon)$. Now, for any history tree $\mathbf{w}$, sequential Rademacher averages of a class of $[-1,1]$-valued strategies Π satisfy

$$
\mathfrak{R}(\Pi,\mathbf{w}) \le \inf_{\alpha\ge0}\left\{\alpha T + \sqrt{2\log\mathcal{N}_1(\Pi,\mathbf{w},\alpha)T}\right\}
$$

and the Dudley entropy integral type bound also holds:

$$
\mathfrak{R}(\Pi,\mathbf{w}) \le \inf_{\alpha\ge0}\left\{4\alpha T + 12\sqrt{T}\int_\alpha^1 \sqrt{\log\ \mathcal{N}_2(\Pi,\mathbf{w},\delta)}\, d\delta\right\} \tag{3.2.8}
$$

In particular, this bound should be compared with Theorem 7 in [2], which employs a covering number in terms of a pointwise metric between strategies that requires closeness for *all histories and all time steps*. Second, the results of [2] for real-valued prediction require strategies to be bounded away from 0 and 1 by $\delta > 0$ and this restriction spoils the rates.

In the rest of this chapter, we show how the results of this section (a) yield proofs of existence of regret-minimization strategies with certain rates and (b) guide in the development of algorithms. For some of these examples, standard methods (such as Exponential Weights) come close to providing an optimal rate, while for others – fail miserably.

3.3 Competing with Autoregressive Strategies

In this section, we consider strategies that depend linearly on the past outcomes. To this end, we fix a set $\Theta \subset \mathbb{R}^k$, for some $k > 0$, and parametrize the set of strategies as

$$\Pi_\Theta = \left\{\pi^\theta : \pi^\theta_t(z_1, \dots, z_{t-1}) = \textstyle\sum_{i=0}^{k-1} \theta_{i+1} z_{t-k+i}, \quad \theta = (\theta_1, \dots, \theta_k) \in \Theta\right\}$$

For consistency of notation, we assume that the sequence of outcomes is padded with zeros for $t \le 0$. First, as an example where known methods *can* recover the correct rate, we consider the case of a constant look-back of size k. We then extend the study to cases where neither the regret behavior nor the algorithm is known in the literature, to the best of our knowledge.

3.3.1 Finite Look-Back

Suppose $\mathcal{Z} = \mathcal{F} \subset \mathbb{R}^d$ are ℓ_2 unit balls, the loss is $\ell(f, z) = \langle f, z\rangle$, and $\Theta \subset \mathbb{R}^k$ is also a unit ℓ_2 ball. Denoting by $\mathbf{W}_{(t-k:t-1)} = [\mathbf{w}_{t-k}(\epsilon), \dots, \mathbf{w}_{t-1}(\epsilon)]$ a matrix with columns in $\mathcal{Z}$,

$$\begin{aligned}
\mathfrak{R}(\ell, \Pi_\Theta) &= \sup_{\mathbf{w},\mathbf{z}} \mathbb{E}_\epsilon \sup_{\theta\in\Theta} \left[\sum_{t=1}^T \epsilon_t \left\langle \pi^\theta(\mathbf{w}_{t-k:t-1}(\epsilon)), \mathbf{z}_t(\epsilon)\right\rangle\right] && (3.3.1)\\
&= \sup_{\mathbf{w},\mathbf{z}} \mathbb{E}_\epsilon \sup_{\theta\in\Theta} \left[\sum_{t=1}^T \epsilon_t \mathbf{z}_t(\epsilon)^\top \mathbf{W}_{(t-k:t-1)} \cdot \theta\right] \\
&= \sup_{\mathbf{w},\mathbf{z}} \mathbb{E}_\epsilon \left\|\sum_{t=1}^T \epsilon_t \mathbf{z}_t(\epsilon)^\top \mathbf{W}_{(t-k:t-1)}\right\| \le \sqrt{kT} && (3.3.2)
\end{aligned}$$

In fact, this bound against all strategies parametrized by Θ is achieved by the gradient descent (GD) method with the simple update $\theta_{t+1} = \mathrm{Proj}_\Theta(\theta_t - \eta\, [z_{t-k}, \dots, z_{t-1}]^\top z_t)$ where Proj_Θ is the Euclidean projection onto the set Θ. This can be seen by writing the loss as

$$\langle [z_{t-k}, \dots, z_{t-1}] \cdot \theta_t, z_t\rangle = \langle \theta_t, [z_{t-k}, \dots, z_{t-1}]^\top z_t\rangle.$$

The regret of GD, $\sum_{t=1}^{T}\langle\theta_t, [z_{t-k},\ldots,z_{t-1}]^\intercal z_t\rangle - \inf_{\theta\in\Theta}\sum_{t=1}^{T}\langle\theta, [z_{t-k},\ldots,z_{t-1}]^\intercal z_t\rangle$, is precisely regret against strategies in Θ, and analysis of GD yields the rate in (3.3.1).

3.3.2 Full Dependence on History

The situation becomes less obvious when $k = T$ and strategies depend on the full history. The regret bound in (3.3.1) is vacuous, and the question is whether a better bound can be proved, under some additional assumptions on Θ. Can such a bound be achieved by GD?

For simplicity, consider the case of $\mathcal{F} = \mathcal{Z} = [-1,1]$, and assume that $\Theta = B_p(1) \subset \mathbb{R}^T$ is a unit ℓ_p ball, for some $p \geq 1$. Since $k = T$, it is easier to re-index the coordinates so that

$$\pi_t^\theta(z_{1:t-1}) = \textstyle\sum_{i=1}^{t-1}\theta_i z_i.$$

The sequential Rademacher complexity of the strategy class is

$$\begin{aligned}\mathfrak{R}(\ell,\Pi_\Theta) &= \sup_{\mathbf{w},\mathbf{z}}\mathbb{E}\sup_{\theta\in\Theta}\left[\sum_{t=1}^{T}\epsilon_t\pi^\theta(\mathbf{w}_{1:t-1}(\epsilon))\cdot\mathbf{z}_t(\epsilon)\right]\\ &= \sup_{\mathbf{w},\mathbf{z}}\mathbb{E}\sup_{\theta\in\Theta}\left[\sum_{t=1}^{T}\left(\sum_{i=1}^{t-1}\theta_i\mathbf{w}_i(\epsilon)\right)\epsilon_t\mathbf{z}_t(\epsilon)\right].\end{aligned}$$

Rearranging the terms, the last expression is equal to

$$\sup_{\mathbf{w},\mathbf{z}}\mathbb{E}\sup_{\theta\in\Theta}\left[\sum_{t=1}^{T-1}\theta_t\mathbf{w}_t(\epsilon)\cdot\left(\sum_{i=t+1}^{T}\epsilon_i\mathbf{z}_i(\epsilon)\right)\right] \leq \sup_{\mathbf{w},\mathbf{z}}\mathbb{E}\left[\|\mathbf{w}_{1:T-1}(\epsilon)\|_q\cdot\max_{1\leq t\leq T}\left|\sum_{i=t+1}^{T}\epsilon_i\mathbf{z}_i(\epsilon)\right|\right]$$

where q is the Hölder conjugate of p. Observe that

$$\begin{aligned}\sup_{\mathbf{z}}\mathbb{E}\sup_{1\leq t\leq T}\left|\sum_{i=t}^{T}\epsilon_i\mathbf{z}_i(\epsilon)\right| &\leq \sup_{\mathbf{z}}\mathbb{E}\left[\left|\sum_{i=1}^{T}\epsilon_i\mathbf{z}_i(\epsilon)\right| + \sup_{1\leq t\leq T}\left|\sum_{i=1}^{t-1}\epsilon_i\mathbf{z}_i(\epsilon)\right|\right]\\ &\leq 2\sup_{\mathbf{z}}\mathbb{E}\sup_{1\leq t\leq T}\left|\sum_{i=1}^{t}\epsilon_i\mathbf{z}_i(\epsilon)\right|\end{aligned}$$

Since $\{\epsilon_t\mathbf{z}_t(\epsilon) : t = 1,\ldots,T\}$ is a bounded martingale difference sequence, the last term is of the order of $\mathcal{O}(\sqrt{T})$. Now, suppose there is some $\beta > 0$ such that $\|\mathbf{w}_{1:T-1}(\epsilon)\|_q \leq T^\beta$ for all ϵ. This assumption can be implemented if we consider constrained adversaries, where such ℓ_q-bound is required to hold for any prefix $\mathbf{w}_{1:t}(\epsilon)$ of history (In Appendix, we prove Theorem 3.2.2 for the case of constrained sequences). Then $\mathfrak{R}(\ell,\Pi_\Theta) \leq C\cdot T^{\beta+1/2}$ for some constant C. We now compare the rate of convergence of sequential Rademacher and the rate of the mirror descent algorithm for different settings of q in Table 3.1. If $\|\theta\|_p \leq 1$ and $\|\mathbf{w}\|_q \leq T^\beta$ for $q \geq 2$, the convergence rate of mirror descent with Legendre function $F(\theta) = \frac{1}{2}\|\theta\|_p^2$ is $\sqrt{q-1}T^{\beta+1/2}$ (see [21]).

	Θ	$\mathbf{w}_{1:T}$	sequential Radem. rate	Mirror descent rate
	$B_1(1)$	$\lVert \mathbf{w}_{1:T-1} \rVert_\infty \le 1$	$\sqrt{T}$	$\sqrt{T \log T}$
$q \ge 2$	$B_p(1)$	$\lVert \mathbf{w}_{1:T-1} \rVert_q \le T^\beta$	$T^{\beta+1/2}$	$\sqrt{q-1}T^{\beta+1/2}$
	$B_2(1)$	$\lVert \mathbf{w}_{1:T-1} \rVert_2 \le T^\beta$	$T^{\beta+1/2}$	$T^{\beta+1/2}$
$1 \le q \le 2$	$B_p(1)$	$\lVert \mathbf{w}_{1:T-1} \rVert_q \le T^\beta$	$T^{\beta+1/2}$	$T^{\beta+1/q}$
	$B_\infty(1)$	$\lVert \mathbf{w}_{1:T-1} \rVert_1 \le T^\beta$	$T^{\beta+1/2}$	T

Table 3.1: Comparison of the rates of convergence (up to constant factors)

We observe that mirror descent, which is known to be optimal for online linear optimization, and which gives the correct rate for the case of *bounded* look-back strategies, in several regimes fails to yield the correct rate for more general linearly parametrized strategies. Even in the most basic regime where Θ is a unit ℓ_1 ball and the sequence of data is not constrained (other than $\mathcal{Z} = [-1,1]$), there is a gap of $\sqrt{\log T}$ between the Rademacher bound and the guarantee of mirror descent. Is there an algorithm that removes this factor?

3.3.3 Algorithms for $\Theta = B_1(1)$

For the example considered in the previous section, with $\mathcal{F} = \mathcal{Z} = [-1,1]$ and $\Theta = B_1(1)$, the conditional sequential Rademacher complexity of Theorem 3.2.7 becomes

$$\begin{aligned}\mathfrak{R}_T(\Pi|z_1,\dots,z_t) &= \sup_{\mathbf{z},\mathbf{w}} \mathop{\mathbb{E}}_{\epsilon_{t+1:T}} \sup_{\pi\in\Pi} \left[2\sum_{s=t+1}^{T} \epsilon_s \pi_s(z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)) \cdot \mathbf{z}_s(\epsilon) - \sum_{s=1}^{t} \pi_s(z_{1:s-1}) \cdot z_s \right] \\ &\le \sup_{\mathbf{w}} \mathop{\mathbb{E}}_{\epsilon_{t+1:T}} \sup_{\pi\in\Pi} \left[2\sum_{s=t+1}^{T} \epsilon_s \pi_s(z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)) - \sum_{s=1}^{t} z_s \pi_s(z_{1:s-1}) \right]\end{aligned}$$

where the $\mathbf{z}$ tree is "erased", as at the end of the proof of Theorem 3.2.3. Define $a_s(\epsilon) = 2\epsilon_s$ for $s > t$ and $-z_s$ otherwise; $b_i(\epsilon) = \mathbf{w}_i(\epsilon)$ for $i > t$ and z_i otherwise. We can then simply write

$$\begin{aligned}&\sup_{\mathbf{w}} \mathop{\mathbb{E}}_{\epsilon_{t+1:T}} \sup_{\theta\in\Theta} \left[\sum_{s=1}^{T} a_s(\epsilon) \sum_{i=1}^{s-1} \theta_i b_i(\epsilon) \right] \\ &= \sup_{\mathbf{w}} \mathop{\mathbb{E}}_{\epsilon_{t+1:T}} \sup_{\theta\in\Theta} \left[\sum_{s=1}^{T-1} \theta_s b_s(\epsilon) \sum_{i=s+1}^{T} a_i(\epsilon) \right] \\ &\le \mathop{\mathbb{E}}_{\epsilon_{t+1:T}} \max_{1\le s\le T} \left| \sum_{i=s}^{T} a_i(\epsilon) \right|\end{aligned}$$

which we may use as a relaxation:

Lemma 3.3.1. *Define $a_s^t(\epsilon) = 2\epsilon_s$ for $s > t$, and $-z_s$ otherwise. Then,*

$$\mathbf{Rel}(z_{1:t}) = \mathbb{E}_{\epsilon_{t+1:T}} \max_{1\le s\le T} \left| \textstyle\sum_{i=s}^{T} a_i^t(\epsilon) \right|$$

is an admissible relaxation.

Proof. The first step of the proof is an application of the minimax theorem (we assume the necessary conditions hold):

$$\begin{aligned}
&\inf_{q_t\in\Delta(\mathcal{F})}\sup_{z_t\in\mathcal{Z}}\left\{\mathop{\mathbb{E}}_{f_t\sim q_t} f_t\cdot z_t + \mathop{\mathbb{E}}_{\epsilon_{t+1:T}}\max_{1\le s\le T}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|\right\}\\
&=\sup_{p_t\in\Delta(\mathcal{Z})}\inf_{f_t\in\mathcal{F}}\left\{f_t\cdot\mathop{\mathbb{E}}_{z_t\sim p_t} z_t + \mathop{\mathbb{E}}_{z_t\sim p_t}\mathop{\mathbb{E}}_{\epsilon_{t+1:T}}\max_{1\le s\le T}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|\right\}
\end{aligned}$$

For any $p_t\in\Delta(\mathcal{Z})$, the infimum over f_t of the above expression is equal to

$$\begin{aligned}
&-\left|\mathop{\mathbb{E}}_{z_t\sim p_t} z_t\right| + \mathop{\mathbb{E}}_{z_t\sim p_t}\mathop{\mathbb{E}}_{\epsilon_{t+1:T}}\max\left\{\max_{s>t}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|,\max_{s\le t}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|\right\}\\
&\le\mathop{\mathbb{E}}_{z_t\sim p_t}\mathop{\mathbb{E}}_{\epsilon_{t+1:T}}\max\left\{\max_{s>t}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|,\max_{s\le t}\left|\sum_{i=s}^T a_i^t(\epsilon)+\mathop{\mathbb{E}}_{z'_t\sim p_t} z'_t\right|\right\}\\
&\le\mathop{\mathbb{E}}_{z_t,z'_t\sim p_t}\mathop{\mathbb{E}}_{\epsilon_{t+1:T}}\max\left\{\max_{s>t}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|,\max_{s\le t}\left|\sum_{i\ge s,i\ne t} a_i^t(\epsilon)+(z'_t-z_t)\right|\right\}
\end{aligned}$$

We now argue that the independent z_t and z'_t have the same distribution p_t, and thus we can introduce a random sign ϵ_t. The above expression then equals to

$$\begin{aligned}
&\mathop{\mathbb{E}}_{z_t,z'_t\sim p_t}\mathop{\mathbb{E}}_{\epsilon_{t:T}}\max\left\{\max_{s>t}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|,\max_{s\le t}\left|\sum_{i\ge s,i\ne t} a_i^t(\epsilon)+\epsilon_t(z'_t-z_t)\right|\right\}\\
&\le\mathop{\mathbb{E}}_{z_t\sim p_t}\mathop{\mathbb{E}}_{\epsilon_{t:T}}\max\left\{\max_{s>t}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|,\max_{s\le t}\left|\sum_{i\ge s,i\ne t} a_i^t(\epsilon)+2\epsilon_t z_t\right|\right\}
\end{aligned}$$

Now, the supremum over p_t is achieved at a delta distribution, yielding an upper bound

$$\begin{aligned}
&\sup_{z_t\in[-1,1]}\mathop{\mathbb{E}}_{\epsilon_{t:T}}\max\left\{\max_{s>t}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|,\max_{s\le t}\left|\sum_{i\ge s,i\ne t} a_i^t(\epsilon)+2\epsilon_t z_t\right|\right\}\\
&\le\mathop{\mathbb{E}}_{\epsilon_{t:T}}\max\left\{\max_{s>t}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|,\max_{s\le t}\left|\sum_{i\ge s,i\ne t} a_i^t(\epsilon)+2\epsilon_t\right|\right\}\\
&=\mathop{\mathbb{E}}_{\epsilon_{t:T}}\max_{1\le s\le T}\left|\sum_{i=s}^T a_i^{t-1}(\epsilon)\right|
\end{aligned}$$

□

With this relaxation, the following method attains $\mathcal{O}(\sqrt{T})$ regret: prediction at step t is

$$q_t=\operatorname*{argmin}_{q\in[-1,1]}\sup_{z_t\in\{\pm1\}}\left\{\mathop{\mathbb{E}}_{f_t\sim q} f_t\cdot z_t+\mathbb{E}_{\epsilon_{t+1:T}}\max_{1\le s\le T}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|\right\}$$

where the sup over $z_t \in [-1, 1]$ is achieved at $\{\pm 1\}$ due to convexity. Following [15], we can also derive randomized algorithms, which can be viewed as "randomized playout" generalizations of the Follow the Perturbed Leader algorithm.

Lemma 3.3.2. *Consider the randomized strategy where at round t we first draw $\epsilon_{t+1}, \ldots, \epsilon_T$ uniformly at random and then further draw our move f_t according to the distribution*

$$
\begin{aligned}
q_t(\epsilon) &= \underset{q \in [-1,1]}{\operatorname{argmin}} \ \sup_{z_t \in \{-1,1\}} \left\{ \mathbb{E}_{f_t \sim q} f_t \cdot z_t + \max_{1 \le s \le T} \left| \sum_{i=s}^T a_i^t(\epsilon) \right| \right\} \\
&= \tfrac{1}{2} \Big(\ \max\left\{ \max_{s=1,\ldots,t} \left| - \sum_{i=s}^{t-1} z_i + 1 + 2 \sum_{i=t+1}^T \epsilon_i \right| \ , \quad \max_{s=t+1,\ldots,T} \left| 2 \sum_{i=s}^T \epsilon_i \right| \right\} \\
&\qquad - \max\left\{ \max_{s=1,\ldots,t} \left| - \sum_{i=s}^{t-1} z_i - 1 + 2 \sum_{i=t+1}^T \epsilon_i \right| \ , \quad \max_{s=t+1,\ldots,T} \left| 2 \sum_{i=s}^T \epsilon_i \right| \right\} \Big)
\end{aligned}
$$

The expected regret of this randomized strategy is upper bounded by sequential Rademacher complexity: $\mathbb{E}[\mathbf{Reg}_T] \le 2\mathfrak{R}_T(\Pi)$, which was shown to be $\mathcal{O}(\sqrt{T})$ (see Table 3.1).

Proof. Let q_t be the randomized strategy where we draw $\epsilon_{t+1}, \ldots, \epsilon_T$ uniformly at random and pick

$$
q_t(\epsilon) = \underset{q \in [-1,1]}{\operatorname{argmin}} \ \sup_{z_t \in \{-1,1\}} \left\{ \underset{f_t \sim q}{\mathbb{E}} f_t \cdot z_t + \max_{1 \le s \le T} \left| \sum_{i=s}^T a_i^t(\epsilon) \right| \right\} \tag{3.3.3}
$$

Then,

$$
\begin{aligned}
& \sup_{z_t \in \{-1,1\}} \left\{ \underset{f_t \sim q_t}{\mathbb{E}} f_t \cdot z_t + \mathbb{E}_{\epsilon_{t+1:T}} \max_{1 \le s \le T} \left| \sum_{i=s}^T a_i^t(\epsilon) \right| \right\} \\
&= \sup_{z_t \in \{-1,1\}} \left\{ \mathbb{E}_{\epsilon_{t+1:T}} \underset{f_t \sim q_t(\epsilon)}{\mathbb{E}} f_t \cdot z_t + \mathbb{E}_{\epsilon_{t+1:T}} \max_{1 \le s \le T} \left| \sum_{i=s}^T a_i^t(\epsilon) \right| \right\} \\
&\le \mathbb{E}_{\epsilon_{t+1:T}} \left[\sup_{z_t} \left\{ \underset{f_t \sim q_t(\epsilon)}{\mathbb{E}} f_t \cdot z_t + \max_{1 \le s \le T} \left| \sum_{i=s}^T a_i^t(\epsilon) \right| \right\} \right] \\
&= \mathbb{E}_{\epsilon_{t+1:T}} \left[\inf_{q_t \in \Delta(\mathcal{F})} \sup_{z_t} \left\{ \underset{f_t \sim q_t}{\mathbb{E}} f_t \cdot z_t + \max_{1 \le s \le T} \left| \sum_{i=s}^T a_i^t(\epsilon) \right| \right\} \right]
\end{aligned}
$$

where the last step is due to the way we pick our predictor $f_t(\epsilon)$ given random draw of ϵ's in Equation (3.3.3). We now apply the minimax theorem, yielding the following upper bound on the term above:

$$
\mathbb{E}_{\epsilon_{t+1:T}} \left[\sup_{p_t \in \Delta(\mathcal{Z})} \inf_{f_t} \left\{ \underset{z_t \sim p_t}{\mathbb{E}} f_t \cdot z_t + \underset{z_t \sim p_t}{\mathbb{E}} \max_{1 \le s \le T} \left| \sum_{i=s}^T a_i^t(\epsilon) \right| \right\} \right]
$$

This expression can be re-written as

$$
\begin{aligned}
&\mathbb{E}_{\epsilon_{t+1:T}}\left[\sup_{p_t\in\Delta(\mathcal{Z})}\mathbb{E}_{z_t\sim p_t}\inf_{f_t}\left\{\mathbb{E}_{z'_t\sim p_t} f_t\cdot z'_t + \max_{1\le s\le T}\left|\sum_{i=s}^{T} a_i^t(\epsilon)\right|\right\}\right]\\
&\le \mathbb{E}_{\epsilon_{t+1:T}}\left[\sup_{p_t\in\Delta(\mathcal{Z})}\mathbb{E}_{z_t\sim p_t}\left\{-\left|\mathbb{E}_{z'_t\sim p_t} z'_t\right| + \max_{1\le s\le T}\left|\sum_{i=s}^{T} a_i^t(\epsilon)\right|\right\}\right]\\
&\le \mathbb{E}_{\epsilon_{t+1:T}}\left[\sup_{p_t\in\Delta(\mathcal{Z})}\mathbb{E}_{z_t\sim p_t}\max\left\{\max_{s\le t}\left|\sum_{i=s}^{T} a_i^t(\epsilon) + \mathbb{E}_{z'_t\sim p_t} z'_t\right|, \max_{s>t}\left|\sum_{i=s}^{T} a_i^t(\epsilon)\right|\right\}\right]\\
&\le \mathbb{E}_{\epsilon_{t+1:T}}\left[\sup_{p_t\in\Delta(\mathcal{Z})}\mathbb{E}_{z_t,z'_t\sim p_t}\max\left\{\max_{s\le t}\left|\sum_{i\ge s, i\ne t}^{T} a_i^t(\epsilon) + (z'_t - z_t)\right|, \max_{s>t}\left|\sum_{i=s}^{T} a_i^t(\epsilon)\right|\right\}\right]
\end{aligned}
$$

We now argue that the independent z_t and z'_t have the same distribution p_t, and thus we can introduce a random sign ϵ_t. The above expression then equals to

$$
\begin{aligned}
&\mathbb{E}_{\epsilon_{t+1:T}}\left[\sup_{p_t\in\Delta(\mathcal{Z})}\mathbb{E}_{z_t,z'_t\sim p_t}\mathbb{E}_{\epsilon_t}\max\left\{\max_{s\le t}\left|\sum_{i\ge s, i\ne t}^{T} a_i^t(\epsilon) + \epsilon_t(z'_t - z_t)\right|, \max_{s>t}\left|\sum_{i=s}^{T} a_i^t(\epsilon)\right|\right\}\right]\\
&\le \mathbb{E}_{\epsilon_{t+1:T}}\sup_{z_t\in\{-1,1\}}\mathbb{E}_{\epsilon_t}\max_{1\le s\le T}\left|\sum_{i=s}^{T} a_i^{t-1}(\epsilon)\right| = \mathbb{E}_{\epsilon_{t:T}}\max_{1\le s\le T}\left|\sum_{i=s}^{T} a_i^{t-1}(\epsilon)\right|
\end{aligned}
$$

□

The time consuming parts of the above randomized method are to draw $T-t$ random bits at round t and to calculate the partial sums. However, we may replace Rademacher random variables by Gaussian $\mathcal{N}(0,1)$ random variables and use known results on the distributions of extrema of a Brownian motion. To this end, define a Gaussian analogue of conditional sequential Rademacher complexity

$$
\begin{aligned}
&\mathcal{G}_T(\Pi|z_1,\dots,z_t)\\
&= \sup_{\mathbf{z},\mathbf{w}}\mathbb{E}_{\sigma_{t+1:T}}\sup_{\pi\in\Pi}\left[\sqrt{2\pi}\sum_{s=t+1}^{T}\sigma_s\ell(\pi_s((z_{1:t},\mathbf{w}_{1:s-t-1}(\epsilon)),\mathbf{z}_{s-t}(\epsilon)) - \sum_{s=1}^{t}\ell(\pi_s(z_{1:s-1}),z_s)\right]
\end{aligned}
$$

where $\sigma_t \sim \mathcal{N}(0,1)$, and $\epsilon = (\mathrm{sign}(\sigma_1),\dots,\mathrm{sign}(\sigma_T))$. For our example the $\mathcal{O}(\sqrt{T})$ bound can be shown for $\mathcal{G}_T(\Pi)$ by calculating the expectation of the maximum of Brownian motion. Proofs similar to Theorem 3.2.2 and Theorem 3.2.7 show that the conditional Gaussian complexity $\mathcal{G}_T(\Pi|z_1,\dots,z_t)$ is an upper bound on $\mathfrak{R}_T(\Pi|z_1,\dots,z_t)$ and is admissible.

Theorem 3.3.3. *The conditional sequential Rademacher complexity with respect to* Π

$$
\begin{aligned}
&\mathcal{G}_T(\ell,\Pi|z_1,\dots,z_t)\\
&\triangleq \sup_{\mathbf{z},\mathbf{w}}\mathbb{E}_{\sigma_{t+1:T}}\sup_{\pi\in\Pi}\left[\sqrt{2\pi}\sum_{s=t+1}^{T}\sigma_s\ell(\pi_s((z_{1:t},\mathbf{w}_{1:s-t-1}(\epsilon)),\mathbf{z}_{s-t}(\epsilon)) - \sum_{s=1}^{t}\ell(\pi_s(z_{1:s-1}),z_s)\right]
\end{aligned}
$$

is admissible.

Proof. Denote $L_t(\pi) = \sum_{s=1}^{t} \ell(\pi_s(z_{1:s-1}), z_s)$. Let $c = \mathbb{E}_\sigma |\sigma| = \sqrt{2/\pi}$. The first step of the proof is an application of the minimax theorem (we assume the necessary conditions hold):

$$\begin{aligned}
&\inf_{q_t \in \Delta(\mathcal{F})} \sup_{z_t \in \mathcal{Z}} \Bigg\{ \mathbb{E}_{f_t \sim q_t} [\ell(f_t, z_t)] \\
&\qquad + \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \left[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_t(\pi) \right] \Bigg\} \\
&= \sup_{p_t \in \Delta(\mathcal{Z})} \inf_{f_t \in \mathcal{F}} \Bigg\{ \mathbb{E}_{z_t \sim p_t} [\ell(f_t, z_t)] \\
&\qquad + \mathbb{E}_{z_t \sim p_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \left[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_t(\pi) \right] \Bigg\}
\end{aligned}$$

For any $p_t \in \Delta(\mathcal{Z})$, the infimum over f_t of the above expression is equal to

$$\begin{aligned}
&\mathbb{E}_{z_t \sim p_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \\
&\qquad\qquad + \inf_{f_t \in \mathcal{F}} \mathbb{E}_{z_t \sim p_t} [\ell(f_t, z_t)] - \ell(\pi_t(z_{1:t-1}), z_t) \Bigg] \\
&\le \mathbb{E}_{z_t \sim p_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \\
&\qquad\qquad + \mathbb{E}_{z_t \sim p_t} [\ell(\pi_t(z_{1:t-1}), z_t)] - \ell(\pi_t(z_{1:t-1}), z_t) \Bigg] \\
&\le \mathbb{E}_{z_t, z'_t \sim p_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \\
&\qquad\qquad + \ell(\pi_t(z_{1:t-1}), z'_t) - \ell(\pi_t(z_{1:t-1}), z_t) \Bigg]
\end{aligned}$$

We now argue that the independent z_t and z'_t have the same distribution p_t, and thus we can introduce a gaussian random variable σ_t and a random sign $\epsilon_t = \text{sign}(\sigma_t)$. The above expression then equals to

$$\begin{aligned}
&\mathbb{E}_{z_t, z'_t \sim p_t} \mathbb{E}_{\sigma_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \\
&\qquad\qquad + \epsilon_t (\ell(\pi_t(z_{1:t-1}), \chi_t(-\epsilon_t))) - \ell(\pi_t(z_{1:t-1}), \chi_t(\epsilon_t))) \Bigg] \\
&\le \mathbb{E}_{z_t, z'_t \sim p_t} \mathbb{E}_{\sigma_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \\
&\qquad\qquad + \epsilon_t \mathbb{E}_{\sigma_t} \left| \frac{\sigma_t}{c} \right| (\ell(\pi_t(z_{1:t-1}), \chi_t(-\epsilon_t))) - \ell(\pi_t(z_{1:t-1}), \chi_t(\epsilon_t))) \Bigg]
\end{aligned}$$

Put the expectation outside and use the fact $\epsilon_t|\sigma_t| = \sigma_t$, we get

$$\begin{aligned}
&\mathop{\mathbb{E}}_{z_t, z_t' \sim p_t} \mathop{\mathbb{E}}_{\sigma_t} \sup_{\mathbf{z},\mathbf{w}} \mathop{\mathbb{E}}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \\
&\qquad\qquad + \frac{\sigma_t}{c} (\ell(\pi_t(z_{1:t-1}), \chi_t(-\epsilon_t))) - \ell(\pi_t(z_{1:t-1}), \chi_t(\epsilon_t))) \Bigg] \\
&\leq \mathop{\mathbb{E}}_{z_t, z_t' \sim p_t} \sup_{z'', z'''} \mathop{\mathbb{E}}_{\sigma_t} \sup_{\mathbf{z},\mathbf{w}} \mathop{\mathbb{E}}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[\frac{2}{c} \sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) \\
&\qquad\qquad - L_{t-1}(\pi) + \frac{\sigma_t}{c} (\ell(\pi_t(z_{1:t-1}), z_t'') - \ell(\pi_t(z_{1:t-1}), z_t''')) \Bigg]
\end{aligned}$$

Splitting the resulting expression into two parts, we arrive at the upper bound of

$$\begin{aligned}
&2 \mathop{\mathbb{E}}_{z_t, z_t' \sim p_t} \sup_{z''} \mathop{\mathbb{E}}_{\sigma_t} \sup_{\mathbf{z},\mathbf{w}} \mathop{\mathbb{E}}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[\frac{1}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) \\
&\qquad\qquad - \frac{1}{2} L_{t-1}(\pi) + \frac{\sigma_t}{c} \ell(\pi_t(z_{1:t-1}), z_t'') \Bigg] \\
&\leq \sup_{z, z', z''} \mathop{\mathbb{E}}_{\sigma_t} \sup_{\mathbf{z},\mathbf{w}} \mathop{\mathbb{E}}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \Bigg[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \\
&\qquad\qquad + \frac{2\sigma_t}{c} \ell(\pi_t(z_{1:t-1}), z_t'') \Bigg] \\
&\leq \mathcal{G}_T(\ell, \Pi | z_1, \ldots, z_{t-1}).
\end{aligned}$$

□

Furthermore, the proof of Lemma 3.3.2 holds for Gaussian random variables, and gives the randomized algorithm as in Lemma 3.3.2 with ϵ_t replaced by σ_t. It is not difficult to see that we can keep track of the maximum and minimum of $\left\{- \sum_{i=s}^{t-1} z_i \right\}$ between rounds in $\mathcal{O}(1)$ time. We can then draw three random variables from the joint distribution of the maximum, the minimum and the endpoint of a Brownian Motion and calculate the prediction in $\mathcal{O}(1)$ time per round of the game (the joint distribution can be found in [11]). In conclusion, we have derived an algorithm that for the case of $\Theta = B_1(1)$, with time complexity of $\mathcal{O}(1)$ per round and the optimal regret bound of $\mathcal{O}(\sqrt{T})$. We leave it as an open question to develop efficient and optimal algorithms for the other settings in Table 3.1.

3.4 Competing with Statistical Models

In this section we consider competing with a set of strategies that arise from statistical models. For example, for the case of Bayesian models, strategies are parametrized by the choice of a prior. Regret bounds with respect to a set of such methods can be thought of as a robustness statement: we are aiming to perform as well as the strategy with the best choice of a prior. We start this section with a general setup that needs further investigation.

3.4.1 Compression and Sufficient Statistics

Assume that strategies in Π have a particular form: they all work with a "sufficient statistic", or, more loosely, *compression* of the past data. Suppose "sufficient statistics" can take values in some set Γ. Fix a set $\bar{\Pi}$ of mappings $\bar{\pi} : \Gamma \mapsto \mathcal{F}$. We assume that all the strategies in Π are of the form $\pi_t(z_1, \ldots, z_{t-1}) = \bar{\pi}(\gamma(z_1, \ldots, z_{t-1}))$ for some $\bar{\pi} \in \bar{\Pi}$ and $\gamma : \mathcal{Z}^* \mapsto \Gamma$. Such a bottleneck Γ can arise due to a finite memory or finite precision, but can also arise if the strategies in Π are actually solutions to a statistical problem. If we assume a certain stochastic source for the data, we may estimate the parameters of the model, and there is often a natural set of sufficient statistics associated with it. If we collect all such solutions to stochastic models in a set Π, we may compete with all these strategies as long as Γ is not too large and the dependence of estimators on these sufficient statistics is smooth. With the notation introduced in this paper, we need to study the sequential Rademacher complexity for strategies Π, which can be upper bounded by the complexity of $\bar{\Pi}$ on Γ-valued trees:

$$\mathfrak{R}(\Pi) \leq \sup_{\mathbf{g},\mathbf{z}} \mathbb{E}_\epsilon \sup_{\bar{\pi}\in\bar{\Pi}} \left[\sum_{t=1}^{T} \epsilon_t \ell(\bar{\pi}(\mathbf{g}_t(\epsilon)), \mathbf{z}_t(\epsilon)) \right]$$

This complexity corresponds to our intuition that with sufficient statistics the dependence on the ever-growing history can be replaced with the dependence on a *summary* of the data. Next, we consider one particular case of this general idea, and refer to [8] for more details on these types of bounds.

3.4.2 Bernoulli Model with a Beta Prior

Suppose the data $z_t \in \{0,1\}$ is generated according to Bernoulli distribution with parameter p, and the prior on $p \in [0,1]$ is $p \sim Beta(\alpha, \beta)$. Given the data $\{z_1, \ldots, z_{t-1}\}$, the maximum a posteriori (MAP) estimator of p is $\hat{p} = (\sum_{i=1}^{t-1} z_i + \alpha - 1)/(t-1+\alpha+\beta-2)$. We now consider the problem of competing with $\Pi = \{\pi^{\alpha,\beta} : \alpha > 1, \beta \in (1, C_\beta]\}$ for some C_β, where each $\pi^{\alpha,\beta}$ predicts the corresponding MAP value for the next round:

$$\pi_t^{\alpha,\beta}(z_1, \ldots, z_{t-1}) = \left(\textstyle\sum_{i=1}^{t-1} z_i + \alpha - 1\right)/(t - 1 + \alpha + \beta - 2) \ .$$

Let us consider the absolute loss, which is equivalent to probability of a mistake of the randomized prediction[1] with bias $\pi_t^{\alpha,\beta}$. Thus, the loss of a strategy $\pi^{\alpha,\beta}$ on round

[1] Alternatively, we can consider strategies that predict according to $\mathbf{1}\{\hat{p} \geq 1/2\}$, which better matches the choice of an absolute loss. However, in this situation, an experts algorithm on an appropriate discretization attains the bound.

t is $\left|\pi_t^{\alpha,\beta}(z_{1:t-1}) - z_t\right|$. Using Theorem 3.2.2 and the argument in (3.2.5) to erase the outcome tree, we conclude that there exists a regret minimization algorithm against the set Π which attains regret of at most

$$2\sup_{\mathbf{w}} \mathbb{E}_\epsilon \sup_{\alpha,\beta}\left[\sum_{t=1}^T \epsilon_t \frac{\sum_{i=1}^{t-1}\mathbf{w}_i(\epsilon)+\alpha-1}{t-1+\alpha+\beta-2}\right] .$$

To analyze the rate exhibited by this upper bound, construct a new tree with $\mathbf{g}_1(\epsilon) = 1$ and $\mathbf{g}_t(\epsilon) = \frac{\sum_{i=1}^{t-1}\mathbf{w}_i(\epsilon)+\alpha-1}{t+\alpha-2} \in [0,1]$ for $t \geq 2$. With this notation, we can simply re-write the last expression as twice

$$\sup_{\mathbf{g}} \mathbb{E}_\epsilon \sup_{\alpha,\beta}\left[\sum_{t=1}^T \epsilon_t \mathbf{g}_t(\epsilon)\frac{t+\alpha-2}{t+\alpha+\beta-3}\right]$$

The supremum ranges over all $[0,1]$-valued trees $\mathbf{g}$, but we can pass to the supremum over all $[-1,1]$-valued trees (thus making the value larger). We then observe that the supremum is achieved at a $\{\pm1\}$-valued tree $\mathbf{g}$, which can then be erased as in the end of the proof of Theorem 3.2.3 (roughly speaking, it amounts to renaming ϵ_t into $\epsilon_t\mathbf{g}_t(\epsilon_{1:t-1})$). We obtain an upper bound

$$\mathfrak{R}(\Pi) \leq \mathbb{E}_\epsilon \sup_{\alpha,\beta}\sum_{t=1}^T \frac{\epsilon_t(t+\alpha-2)}{t+\alpha+\beta-3} \leq \mathbb{E}_\epsilon\left|\sum_{t=1}^T \epsilon_t\right| + \mathbb{E}_\epsilon \sup_{\alpha,\beta}\left|\sum_{t=1}^T \frac{\epsilon_t(\beta-1)}{t+\alpha+\beta-3}\right| = (\sqrt{C_\beta}+1)\sqrt{T} \tag{3.4.1}$$

where we used Cauchy-Schwartz inequality for the second term. We note that an experts algorithm would require a discretization that depends on T and will yield a regret bound of order $O(\sqrt{T\log T})$. It is therefore interesting to find an algorithm that avoids the discretization and obtains this regret. To this end, we take the derived upper bound on the sequential Rademacher complexity and prove that it is an admissible relaxation.

Lemma 3.4.1. *The relaxation*

$$\mathbf{Rel}(z_{1:t}) = \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\alpha,\beta}\left[2\sum_{s=t+1}^T \epsilon_s \cdot \frac{s+\alpha-2}{s+\alpha+\beta-3} - \sum_{s=1}^t \left|\frac{\sum_{i=1}^{s-1} z_i}{s+\alpha+\beta-3} - z_s\right|\right]$$

is admissible.

Proof. Denote

$$L_t(\alpha,\beta) = \sum_{s=1}^t \left|\frac{\frac{\sum_{i=1}^{s-1} z_i}{s+\alpha-2}}{1+\frac{\beta-1}{s+\alpha-2}} - z_s\right| .$$

The first step of the proof is an application of the minimax theorem:

$$\begin{aligned}
&\inf_{q_t\in\Delta(\mathcal{F})}\sup_{z_t\in\mathcal{Z}}\left\{\mathop{\mathbb{E}}_{f_t\sim q_t}|f_t-z_t| + \mathop{\mathbb{E}}_{\epsilon_{t+1:T}}\sup_{\alpha,\beta}\left[2\sum_{s=t+1}^T \epsilon_s\cdot\frac{1}{1+\frac{\beta-1}{s+\alpha-2}} - L_t(\alpha,\beta)\right]\right\}\\
&= \sup_{p_t\in\Delta(\mathcal{Z})}\inf_{f_t\in\mathcal{F}}\left\{\mathop{\mathbb{E}}_{z_t\sim p_t}|f_t-z_t| + \mathop{\mathbb{E}}_{z_t\sim p_t}\mathop{\mathbb{E}}_{\epsilon_{t+1:T}}\sup_{\alpha,\beta}\left[2\sum_{s=t+1}^T \epsilon_s\cdot\frac{1}{1+\frac{\beta-1}{s+\alpha-2}} - L_t(\alpha,\beta)\right]\right\}
\end{aligned}$$

For any $p_t \in \Delta(\mathcal{Z})$, the infimum over f_t of the above expression is equal to

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{z_t\sim p_t} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\alpha,\beta} \left[2\sum_{s=t+1}^{T} \frac{\epsilon_s}{1+\frac{\beta-1}{s+\alpha-2}} - L_{t-1}(\alpha,\beta) + \inf_{f_t\in\mathcal{F}} \mathop{\mathbb{E}}_{z_t\sim p_t} |f_t - z_t| - \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1+\frac{\beta-1}{t+\alpha-2}} - z_t \right| \right] \\
&\le \mathop{\mathbb{E}}_{z_t\sim p_t} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\alpha,\beta} \left[2\sum_{s=t+1}^{T} \frac{\epsilon_s}{1+\frac{\beta-1}{s+\alpha-2}} - L_{t-1}(\alpha,\beta) + \mathop{\mathbb{E}}_{z'_t\sim p_t} \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1+\frac{\beta-1}{t+\alpha-2}} - z'_t \right| - \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1+\frac{\beta-1}{t+\alpha-2}} - z_t \right| \right] \\
&\le \mathop{\mathbb{E}}_{z_t, z'_t\sim p_t} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\alpha,\beta} \left[2\sum_{s=t+1}^{T} \frac{\epsilon_s}{1+\frac{\beta-1}{s+\alpha-2}} - L_{t-1}(\alpha,\beta) + \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1+\frac{\beta-1}{t+\alpha-2}} - z'_t \right| - \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1+\frac{\beta-1}{t+\alpha-2}} - z_t \right| \right]
\end{aligned}
$$

We now argue that the independent z_t and z'_t have the same distribution p_t, and thus we can introduce a random sign ϵ_t. The above expression then equals to

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{z_t, z'_t\sim p_t} \mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[2\sum_{s=t+1}^{T} \frac{\epsilon_s}{1+\frac{\beta-1}{s+\alpha-2}} - L_{t-1}(\alpha,\beta) + \epsilon_t \left(\left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1+\frac{\beta-1}{t+\alpha-2}} - z'_t \right| - \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1+\frac{\beta-1}{t+\alpha-2}} - z_t \right| \right) \right] \\
&\le \sup_{\substack{z_t, z'_t \\ \in\mathcal{Z}}} \mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[2\sum_{s=t+1}^{T} \frac{\epsilon_s}{1+\frac{\beta-1}{s+\alpha-2}} - L_{t-1}(\alpha,\beta) + \epsilon_t \left(\left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1+\frac{\beta-1}{t+\alpha-2}} - z'_t \right| - \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1+\frac{\beta-1}{t+\alpha-2}} - z_t \right| \right) \right]
\end{aligned}
$$

where we upper bounded the expectation by the supremum. Splitting the resulting expression into two parts, we arrive at the upper bound of

$$
\begin{aligned}
&2\sup_{z_t\in\mathcal{Z}} \mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[\sum_{s=t+1}^{T} \frac{\epsilon_s}{1+\frac{\beta-1}{s+\alpha-2}} - \frac{1}{2} L_{t-1}(\alpha,\beta) + \epsilon_t \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1+\frac{\beta-1}{t+\alpha-2}} - z_t \right| \right] \\
&= 2\sup_{z_t\in\mathcal{Z}} \mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[\sum_{s=t+1}^{T} \frac{\epsilon_s}{1+\frac{\beta-1}{s+\alpha-2}} - \frac{1}{2} L_{t-1}(\alpha,\beta) + \epsilon_t \cdot \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1+\frac{\beta-1}{t+\alpha-2}} (1-2z_t) - \epsilon_t z_t \right] \\
&= 2\mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[\sum_{s=t+1}^{T} \frac{\epsilon_s}{1+\frac{\beta-1}{s+\alpha-2}} - \frac{1}{2} L_{t-1}(\alpha,\beta) + \epsilon_t \cdot \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1+\frac{\beta-1}{t+\alpha-2}} \right]
\end{aligned}
$$

where the last step is due to the fact that for any $z_t \in \{0,1\}$, $\epsilon_t(1-2z_t)$ has the same distribution as ϵ_t. We then proceed to upper bound

$$
\begin{aligned}
&2\sup_{p} \mathbb{E}_{a\sim p} \mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[\sum_{s=t+1}^{T} \frac{\epsilon_s}{1+\frac{\beta-1}{s+\alpha-2}} - \frac{1}{2} L_{t-1}(\alpha,\beta) + \epsilon_t \cdot \frac{a}{1+\frac{\beta-1}{t+\alpha-2}} \right] \\
&\le 2\sup_{a\in\{\pm 1\}} \mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[\sum_{s=t+1}^{T} \frac{\epsilon_s}{1+\frac{\beta-1}{s+\alpha-2}} - \frac{1}{2} L_{t-1}(\alpha,\beta) + \epsilon_t \cdot \frac{a}{1+\frac{\beta-1}{t+\alpha-2}} \right] \\
&\le 2\mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[\sum_{s=t}^{T} \frac{\epsilon_s}{1+\frac{\beta-1}{s+\alpha-2}} - \frac{1}{2} L_{t-1}(\alpha,\beta) \right]
\end{aligned}
$$

The initial condition is trivially satisfied as

$$\mathbf{Rel}(z_{1:T}) = -\inf_{\alpha,\beta} \sum_{s=1}^{T} \left| \frac{\frac{\sum_{i=1}^{s-1} z_i}{s+\alpha-2}}{1 + \frac{\beta-1}{s+\alpha-2}} - z_s \right|$$

□

Given that this relaxation is admissible, we have a guarantee that the following algorithm attains the rate $(\sqrt{C_\beta} + 1)\sqrt{T}$ given in (3.4.1):

$$q_t = \arg\min_{q\in[0,1]} \max_{z_t\in\{0,1\}} \Big\{ \mathbb{E}_{f\sim q}|f - z_t| \\ + \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\alpha,\beta} \left[2\sum_{s=t+1}^{T} \epsilon_s \cdot \frac{s+\alpha-2}{s+\alpha+\beta-3} - \sum_{s=1}^{t} \left| \frac{\sum_{i=1}^{s-1} z_i}{s+\alpha+\beta-3} - z_s \right| \right] \Big\}$$

In fact, q_t can be written as

$$q_t = \frac{1}{2}\Bigg\{ \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\alpha,\beta} \left[2\sum_{s=t+1}^{T} \epsilon_s \cdot \frac{s+\alpha-2}{s+\alpha+\beta-3} - \sum_{s=1}^{t-1} (1-2z_s) \cdot \frac{\sum_{i=1}^{s-1} z_i}{s+\alpha+\beta-3} + \frac{\sum_{i=1}^{t-1} z_i}{t+\alpha+\beta-3} \right] \\ - \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\alpha,\beta} \left[2\sum_{s=t+1}^{T} \epsilon_s \cdot \frac{s+\alpha-2}{s+\alpha+\beta-3} - \sum_{s=1}^{t-1} (1-2z_s) \cdot \frac{\sum_{i=1}^{s-1} z_i}{s+\alpha+\beta-3} - \frac{\sum_{i=1}^{t-1} z_i}{t+\alpha+\beta-3} \right] \Bigg\}$$

For a given realization of random signs, the supremum is an optimization of a sum of linear fractional functions of two variables. Such an optimization can be carried out in time $O(T\log T)$ (see [4]). To deal with the expectation over random signs, one may either average over many realizations or use the random playout idea and only draw one sequence. Such an algorithm is admissible for the above relaxation, obtains the $O(\sqrt{T})$ bound, and runs in $O(T\log T)$ time per step. We leave it as an open problem whether a more efficient algorithm with $O(\sqrt{T})$ regret exists.

3.5 Competing with Regularized Least Squares

Consider the supervised learning problem with $\mathcal{Y} = [-1,1]$ and some set $\mathcal{X}$. Consider the Regularized Least Squares (RLS) strategies, parametrized by a regularization parameter λ and a shift w_0. That is, given data $(x_1,y_1),\dots,(x_t,y_t)$, the strategy solves

$$\arg\min_w \textstyle\sum_{i=1}^t (y_i - \langle x_i, w\rangle)^2 + \lambda\|w - w_0\|^2 \ .$$

For a given pair λ and w_0, the solution is

$$w_{t+1}^{\lambda,w_0} = w_0 + (X^\top X + \lambda I)^{-1} X^\top Y,$$

where $X \in \mathbb{R}^{t\times d}$ and $Y \in \mathbb{R}^{t\times 1}$ are the usual matrix representations of the data $x_{1:t}, y_{1:t}$. We would like to compete against a set of such RLS strategies which make prediction $\left\langle w_{t-1}^{\lambda,w_0}, x_t \right\rangle$, given side information x_t. Since the outcomes are in $[-1,1]$, without loss of generality we clip the predictions of strategies to this interval, thus making our regret minimization goal only harder. To this end, let $c(a) = a$ if $a \in [-1,1]$ and $c(a) = \mathrm{sign}(a)$ for $|a| > 1$. Thus, given side-information $x_t \in \mathcal{X}$, the prediction of strategies in $\Pi = \{\pi^{\lambda,w_0} : \lambda \geq \lambda_{\min} > 0, \|w_0\|_2 \leq 1\}$ is simply the clipped product

$$\pi_t^{\lambda,w_0}(x_{1:t}, y_{1:t-1}) = c\left(\left\langle w_{t-1}^{\lambda,w_0}, x_t \right\rangle\right) .$$

Let us take the squared loss function $\ell(\hat{y}, y) = (\hat{y} - y)^2$.

Lemma 3.5.1. *For the set Π of strategies defined above, the minimax regret of competing against Regularized Least Squares strategies is*

$$\mathcal{V}_T(\Pi) \leq c\sqrt{T \log(T\lambda_{\min}^{-1})}$$

for an absolute constant c.

Proof. Given an $\mathcal{X}$-valued tree $\mathbf{x}$ and a $\mathcal{Y}$-valued tree $\mathbf{y}$, let us write $\mathbf{X}_t(\epsilon)$ for the matrix consisting of $(\mathbf{x}_1(\epsilon), \ldots, \mathbf{x}_{t-1}(\epsilon))$ and $\mathbf{Y}_t$ for the vector $(\mathbf{y}_1(\epsilon), \ldots, \mathbf{y}_{t-1}(\epsilon))$. By Theorem 3.2.3, the minimax regret is bounded by

$$\begin{aligned}
&4 \sup_{\mathbf{x},\mathbf{y}} \mathbb{E}_\epsilon \sup_{\pi^{\lambda,w_0} \in \Pi} \left[\sum_{t=1}^T \epsilon_t \pi_t^{\lambda,w_0}(\mathbf{x}_{1:t}(\epsilon), \mathbf{y}_{1:t-1}(\epsilon)) \right] \\
&= 4 \sup_{\mathbf{x},\mathbf{y}} \mathbb{E}_\epsilon \sup_{\lambda,w_0} \left[\sum_{t=1}^T \epsilon_t c\left(\left\langle (\mathbf{X}_t(\epsilon)^\intercal \mathbf{X}_t(\epsilon) + \lambda I)^{-1} \mathbf{X}_t(\epsilon)^\intercal \mathbf{Y}_t(\epsilon), \mathbf{x}_t(\epsilon) \right\rangle + \langle w_0, \mathbf{x}_t(\epsilon) \rangle \right) \right]
\end{aligned}$$

Since the output of the clipped strategies in Π is between -1 and 1, the Dudley integral gives an upper bound

$$\mathfrak{R}(\Pi, (\mathbf{x},\mathbf{y})) \leq \inf_{\alpha \geq 0} \left\{ 4\alpha T + 12\sqrt{T} \int_\alpha^1 \sqrt{\log\ \mathcal{N}_2(\Pi, (\mathbf{x},\mathbf{y}), \delta)}\, d\delta \right\}$$

Define the set of strategies before clipping:

$$\Pi' = \left\{ \pi' : \pi'_t(x_{1:t}, y_{1:t-1}) = \left\langle w_0 + (X^\intercal X + \lambda I)^{-1} X^\intercal Y, x_t \right\rangle, \|w_0\| \leq 1,\ \lambda > \lambda_{\min} \right\}$$

If V is a δ-cover of Π' on $(\mathbf{x},\mathbf{y})$, then V is also an δ-cover of Π as $|c(x) - c(x')| \leq |x - y|$. Therefore, for any $(\mathbf{x},\mathbf{y})$,

$$\mathcal{N}_2(\Pi, (\mathbf{x},\mathbf{y}), \delta) \leq \mathcal{N}_2(\Pi', (\mathbf{x},\mathbf{y}), \delta)$$

and

$$\mathfrak{R}(\Pi, (\mathbf{x},\mathbf{y})) \leq \inf_{\alpha \geq 0} \left\{ 4\alpha T + 12\sqrt{T} \int_\alpha^1 \sqrt{\log\ \mathcal{N}_2(\Pi', (\mathbf{x},\mathbf{y}), \delta)}\, d\delta \right\} .$$

If W is a $\delta/2$-cover of the set of strategies $\Pi_{w_0} = \{\langle w_0, \mathbf{x}_t(\epsilon)\rangle : \|w_0\| \le 1\}$ on a tree $\mathbf{x}$, and Λ is a $\delta/2$-cover of the set of strategies

$$\Pi_\lambda = \left\{\pi : \pi_t(x_{1:t}, y_{1:t-1}) = \left\langle (X^\top X + \lambda I)^{-1} X^\top Y, x_t \right\rangle : \lambda > \lambda_{\min}\right\}$$

then $W \times \Lambda$ is an δ-cover of Π'. Therefore,

$$\mathcal{N}_2(\Pi', (\mathbf{x},\mathbf{y}), \delta) \le \mathcal{N}_2(\Pi_{w_0}, (\mathbf{x},\mathbf{y}), \delta/2) \times \mathcal{N}_2(\Pi_\lambda, (\mathbf{x},\mathbf{y}), \delta/2).$$

Hence,

$$\begin{aligned}
&\mathfrak{R}(\Pi, (\mathbf{x},\mathbf{y})) \\
&\quad \le \inf_{\alpha \ge 0}\left\{4\alpha T + 12\sqrt{T}\int_\alpha^1 \sqrt{\log\ \mathcal{N}_2(\Pi_{w_0}, (\mathbf{x},\mathbf{y}), \delta/2) + \log\ \mathcal{N}_2(\Pi_\lambda, (\mathbf{x},\mathbf{y}), \delta/2)}\, d\delta\right\} \\
&\quad \le \inf_{\alpha \ge 0}\left\{4\alpha T + 12\sqrt{T}\int_\alpha^1 \sqrt{\log\ \mathcal{N}_2(\Pi_{w_0}, (\mathbf{x},\mathbf{y}), \delta/2)}\, d\delta\right\} \\
&\quad + 12\sqrt{T}\int_0^1 \sqrt{\log\ \mathcal{N}_2(\Pi_\lambda, (\mathbf{x},\mathbf{y}), \delta/2)}\, d\delta
\end{aligned}$$

The first term is the Dudley integral of the set of static strategies Π_{w_0} given by $w_0 \in B_2(1)$, and it is exactly the complexity studied in [18] where it is shown to be $O(\sqrt{T\log(T)})$. We now provide a bound on the covering number for the second term. It is easy to verify that the following identity holds

$$(X^\top X + \lambda_2 I_d)^{-1} - (X^\top X + \lambda_1 I_d)^{-1} = (\lambda_1 - \lambda_2)(X^\top X + \lambda_1 I_d)^{-1}(X^\top X + \lambda_2 I_d)^{-1}$$

by right- and left-multiplying both sides by $(X^\top X + \lambda_2 I_d)$ and $(X^\top X + \lambda_1 I_d)$, respectively. Let $\lambda_1, \lambda_2 > 0$. Then, assuming that $\|x_t\|_2 \le 1$ and $y_t \in [-1,1]$ for all t,

$$\begin{aligned}
&\left\|(X_t X + \lambda_2 I_d)^{-1} X^\top Y - (X^\top X + \lambda_1 I_d)^{-1} X^\top Y\right\|_2 \\
&= |\lambda_2 - \lambda_1| \left\|(X^\top X + \lambda_1 I_d)^{-1}(X^\top X + \lambda_2 I_d)^{-1} X^\top Y\right\|_2 \\
&\le |\lambda_2 - \lambda_1| \frac{1}{\lambda_1\lambda_2} \|X^\top Y\|_2 \le \left|\lambda_1^{-1} - \lambda_2^{-1}\right| t
\end{aligned}$$

Hence, for $|\lambda_1^{-1} - \lambda_2^{-1}| \le \delta/T$, we have $\|(X^\top X + \lambda_2 I_d)^{-1} X^\top Y - (X^\top X + \lambda_1 I_d)^{-1} X^\top Y\|_2 \le \delta$, and thus the discretization of λ^{-1} on $(0, \lambda_{\min}^{-1}]$ gives an ℓ_∞-cover, and the size of the cover at scale δ is $\lambda_{\min}^{-1} T \delta^{-1}$. The Dudley entropy integral yields the bound of,

$$\mathfrak{R}(\Pi, (\mathbf{x},\mathbf{y})) \le 12\sqrt{T}\int_0^1 \sqrt{\log(2T\lambda_{\min}^{-1}\delta^{-1})}\, d\delta \le 12\sqrt{T}\left(1 + \sqrt{\log(2T\lambda_{\min}^{-1})}\right).$$

This concludes the proof. $\square$

Observe that $\lambda_{\min}^{-1}$ enters only logarithmically, which allows us to set, for instance, $\lambda_{\min} = 1/T$. Finally, we mention that the set of strategies includes $\lambda = \infty$. This setting corresponds to a static strategy $\pi_t^{\lambda,w_0}(x_{1:t}, y_{1:t-1}) = \langle w_0, x_t \rangle$ and regret against such a static family parametrized by $w_0 \in B_2(1)$ is exactly the objective of online linear regression [22]. Lemma 3.5.1 thus shows that it is possible to have vanishing regret with respect to a much larger set of strategies. It is an interesting open question of whether one can develop an efficient algorithm with the above regret guarantee.

3.6 Competing with Follow the Regularized Leader Strategies

Consider the problem of online linear optimization with the loss function $\ell(f_t, x_t) = \langle f_t, z_t \rangle$ for $f_t \in \mathcal{F}$, $z_t \in \mathcal{Z}$. For simplicity, assume that $\mathcal{F} = \mathcal{Z} = B_2(1)$. An algorithm commonly used for online linear and online convex optimization problems is the Follow the Regularized Leader (FTRL) algorithm. We now consider competing with a family of FTRL algorithms $\pi^{w_0,\lambda}$ indexed by $w_0 \in \{w : \|w\| \le 1\}$ and $\lambda \in \Lambda$ where Λ is a family of functions $\lambda : \mathbb{R}^+ \times [T] \mapsto \mathbb{R}^+$ specifying a schedule for the choice of regularization parameters. Specifically we consider strategies $\pi^{w_0,\lambda}$ such that $\pi_t^{w_0,\lambda}(z_1, \ldots, z_{t-1}) = w_t$ where

$$w_t = w_0 + \underset{w:\|w\|\le 1}{\operatorname{argmin}} \left\{ \textstyle\sum_{i=1}^{t-1} \langle w, z_i \rangle + \frac{1}{2}\lambda \left(\left\| \sum_{i=1}^{t-1} z_i \right\|, t \right) \|w\|^2 \right\} \tag{3.6.1}$$

This can be written in closed form as

$$w_t = w_0 - (\sum_{i=1}^{t-1} z_i) / \max\left\{ \lambda\left(\left\| \sum_{i=1}^{t-1} z_i \right\|, t \right), \left\| \sum_{i=1}^{t-1} z_i \right\| \right\}.$$

Lemma 3.6.1. *For a given class Λ of functions indicating choices of the regularization parameters, define a class Γ of functions on $[0,1] \times [1/T, 1]$ specified by*

$$\Gamma = \left\{ \gamma : \forall b \in [1/T, 1], a \in [0,1], \gamma(a,b) = \min\left\{ \frac{a/(b-1)}{\lambda(a/(b-1), 1/b)}, 1 \right\}, \lambda \in \Lambda \right\}$$

Then the value of the online learning game competing against FTRL strategies given by Equation 3.6.1 is bounded as

$$\mathcal{V}_T(\Pi_\Lambda) \le 4\sqrt{T} + 2\,\mathcal{R}_T(\Gamma)$$

where $\mathcal{R}_T(\Gamma)$ is the sequential Rademacher complexity [18] of Γ.

Proof. Using Theorem 3.2.2,

$$
\begin{aligned}
&\mathcal{V}_T(\Pi_\Lambda) \le 2\mathfrak{R}(\ell, \Pi_\Lambda) \\
&= 2 \sup_{\mathbf{z},\mathbf{z}'} \mathbb{E}_\epsilon \sup_{\mathbf{w}_0 : \|\mathbf{w}_0\| \le 1, \lambda \in \Lambda} \left[\sum_{t=1}^{T} \epsilon_t \left\langle \mathbf{w}_0 - \frac{\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)}{\max\left\{\lambda\left(\left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|, t\right), \left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|\right\}}, \mathbf{z}'_t(\epsilon) \right\rangle \right]
\end{aligned}
$$

which we can upper bound by splitting the supremum into two:

$$
\begin{aligned}
&2 \sup_{\mathbf{z}'} \mathbb{E}_\epsilon \sup_{\mathbf{w}_0 : \|\mathbf{w}_0\| \le 1} \left[\sum_{t=1}^{T} \epsilon_t \left\langle \mathbf{w}_0, \mathbf{z}'_t(\epsilon) \right\rangle \right] \\
&\quad + 2 \sup_{\mathbf{z},\mathbf{z}'} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \epsilon_t \left\langle \frac{\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)}{\max\left\{\lambda\left(\left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|, t\right), \left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|\right\}}, \mathbf{z}'_t(\epsilon) \right\rangle \right]
\end{aligned}
$$

The first term is simply

$$
2 \sup_{\mathbf{z}'} \mathbb{E}_\epsilon \left\| \sum_{t=1}^{T} \epsilon_t \mathbf{z}'_t(\epsilon) \right\| \le 2\sqrt{T}.
$$

The second term can be written as

$$
\begin{aligned}
&2 \sup_{\mathbf{z},\mathbf{z}'} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \epsilon_t \left\langle \frac{\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)}{\left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|}, \mathbf{z}'_t(\epsilon) \right\rangle \frac{\left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|}{\max\left\{\lambda\left(\left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|, t\right), \left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|\right\}} \right] \\
&\le 2 \sup_{\mathbf{z}} \sup_{\mathbf{s}} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \epsilon_t \mathbf{s}_t(\epsilon) \frac{\left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|}{\max\left\{\lambda\left(\left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|, t\right), \left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|\right\}} \right]
\end{aligned}
$$

and the tree $\mathbf{s}$ can be erased (see end of the proof of Theorem 3.2.3), yielding an upper bound

$$
\begin{aligned}
&2 \sup_{\mathbf{z}} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \frac{\epsilon_t \left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|}{\max\left\{\lambda\left(\left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|, t\right), \left\|\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)\right\|\right\}} \right] \\
&\le 2 \sup_{\mathbf{a}} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \frac{\epsilon_t \mathbf{a}_t(\epsilon)}{\max\left\{\lambda\left(\mathbf{a}_t(\epsilon), t\right), \mathbf{a}_t(\epsilon)\right\}} \right] \\
&\le 2 \sup_{\mathbf{a}} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \frac{\epsilon_t}{\max\left\{\frac{\lambda(\mathbf{a}_t(\epsilon), t)}{\mathbf{a}_t(\epsilon)}, 1\right\}} \right] \\
&= 2 \sup_{\mathbf{a}} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \epsilon_t \min\left\{ \frac{\mathbf{a}_t(\epsilon)}{\lambda\left(\mathbf{a}_t(\epsilon), t\right)}, 1 \right\} \right] \\
&= 2 \sup_{\mathbf{b}} \mathbb{E}_\epsilon \sup_{\gamma \in \Gamma} \left[\sum_{t=1}^{T} \epsilon_t \gamma\left(\mathbf{b}_t(\epsilon), 1/t\right) \right] \\
&\le 2\, \mathcal{R}_T(\Gamma)
\end{aligned}
$$

where in the above $\mathbf{a}$ is a $\mathbb{R}^+$-valued tree such that $\mathbf{a}_t : \{\pm 1\}^{t-1} \mapsto [0, t-1]$, $\mathbf{b}$ is a $[1/T, 1]$-value tree and

$$\Gamma = \left\{ \gamma : \forall b \in [1/T, 1], a \in [0,1], \gamma(a,b) = \min \left\{ \frac{a/(b-1)}{\lambda(a/(b-1), 1/b)}, 1 \right\}, \lambda \in \Lambda \right\}.$$

□

Notice that if $|\Lambda| < \infty$ then the second term is bounded as $\mathcal{R}_T(\Gamma) \le \sqrt{T \log |\Lambda|}$. However, we may compete with an infinite set of step-size rules. Indeed, each $\gamma \in \Gamma$ is a function $[0,1]^2 \mapsto [0,1]$. Hence, even if one considers Γ to be the set of all 1-Lipschitz functions (Lipschitz w.r.t., say, ℓ_∞ norm), it holds that $\mathcal{R}_T(\Gamma) \le 2\sqrt{T \log T}$. We conclude that it is possible to compete with set of FTRL strategies that pick any w_0 in unit ball as starting point and further use for regularization parameter schedule any $\lambda : \mathbb{R}^2 \mapsto \mathbb{R}$ that is such that $\frac{a/(b-1)}{\lambda(a/(b-1),1/b)}$ is a 1-Lipchitz function for every $a, b \in [1/T, 1]$.

Beyond the finite and Lipschitz cases shown above, it would be interesting to analyze richer families of step size schedules, and possibly derive efficient algorithms.

Chapter 4

Predictable Sequences and Competing with Strategies

In this chapter, we study how online linear optimization competes with strategies, while benefiting from the predictable sequence. We analyze the minimax value of the online learning optimization problem and develop an algorithm that minimizes the regret. Then, we extend the online optimization problem to one with multiple predictable sequences. We derive efficient regret-minimizing algorithms for two cases: (a) finite number of predictable processes, and (b) infinite predictable processes with only one optimal strategy. Last, we re-analyze the online linear optimization problem using dynamic regret.

4.1 Introduction

Let us first restate the notations for the online learning problem that consists of T rounds. At each time $t \in \{1, \dots, T\} \triangleq [T]$, the learner predicts $f_t \in \mathcal{F}$ and observes an outcome $z_t \in \mathcal{Z}$, where $\mathcal{F}$ and $\mathcal{Z}$ are sets of decisions and outcomes. The learner suffers a loss of $\ell(f_t, z_t)$ at the t-th round, where $\ell : \mathcal{F} \times \mathcal{Z} \to \mathbb{R}$ measures the quality of prediction. A strategy $\pi = (\pi_t)_{t=1}^T$ is a sequence of functions $\pi_t : \mathcal{Z}^{t-1} \to \mathcal{F}$, and Π is a set of strategies. The regret competing with the strategy set Π is defined as

$$\mathbf{Reg}_T(\Pi) = \sum_{t=1}^T \ell(f_t, z_t) - \inf_{\pi \in \Pi} \sum_{t=1}^T \ell(\pi_t(z_{1:t-1}), z_t),$$

where $z_{1:t} = (z_1, \dots, z_t)$ and the minimax value of this prediction problem is defined as

$$\mathcal{V}_T(\Pi) \triangleq \inf_{p_1 \in \Delta(\mathcal{F})} \sup_{z_1 \in \mathcal{Z}} \mathop{\mathbb{E}}_{f_1 \sim p_1} \dots \inf_{p_T \in \Delta(\mathcal{F})} \sup_{z_T \in \mathcal{Z}} \mathop{\mathbb{E}}_{f_T \sim p_T} \mathbf{Reg}_T(\Pi),$$

where $\Delta(\mathcal{F})$ is the set of all probability distributions on $\mathcal{F}$.

During the learning process, the learner may receive outside information. For example, it is reasonable to expect high revenue during Thanksgiving season, or to expect stock market fluctuations on an earning announcement day. The outside information can also be history-based. We embed the outside information as the predictable sequence, and the true outcome is separated from the predictable sequence by adversarial noise. It can also be roughly described as in [17]:

$$\text{outcome} = \text{predictable sequence} \ + \ \text{adversarial noise}.$$

A predictable sequence $M = (M_t)_{t=1}^T$ is a sequence of functions $M_t : \mathcal{Z}^{t-1} \to \mathcal{Z}$. If the outcome is guaranteed to be the same as the predictable process, then the optimal strategy $\pi^M = (\pi_t^M)_{t=1}^T$ is a sequence of functions $\pi_t^M : \mathcal{Z}^{t-1} \to \mathcal{Z}$, which satisfies

$$\pi_t^M(z_{1:t-1}) = \arg\min_{f \in \mathcal{F}} \ell(f, M_t(z_{1:t-1})).$$

[17] presents methods for online optimization problem that take advantage of the predictable process. However, the regret analysis in [17] is competing with fixed actions. The learner benefits from the predictable process, and it puts fixed actions at a disadvantage. In this chapter, we restudy the predictable process discussed in [17], but use the regret competing with strategies.

We define

$$\Pi_{\mathcal{F}}^M = \{\Pi_{\alpha,f}^M = (\alpha f + (1-\alpha)\pi_t^M)_{t=1}^T, \alpha \in [0,1], f \in \mathcal{F}\}$$

as the strategy set. The strategy set $\Pi_{\mathcal{F}}^M$ contains all fixed actions $\{\Pi_{1,f}^M\}_{f \in \mathcal{F}} = \{(f)_{t=1}^T\}_{f \in \mathcal{F}}$ and also the optimal strategy $\Pi_{0,f}^M = (\pi_t^M)_{t=1}^T$. Fixed actions $\{(f)_{t=1}^T\}_{f \in \mathcal{F}}$ are included to prevent from getting hurt by unreliable/unstable predictable sequence. It is usually more difficult to compete with $\Pi_{\mathcal{F}}^M$, than to compete with the union of fixed actions $\{(f)_{t=1}^T\}_{f \in \mathcal{F}}$ and the optimal strategy $(\pi_t^M)_{t=1}^T$.

If the loss function $\ell : \mathcal{F} \times \mathcal{Z} \to \mathbb{R}$ is linear as $\ell(f,z) = \langle f, z \rangle$, competing with $\Pi_{\mathcal{F}}^M$ is equivalent to competing with the union of fixed actions $\{(f)_{t=1}^T\}_{f \in \mathcal{F}}$ and the optimal strategy $(\pi_t^M)_{t=1}^T$. The equivalence of the infima of the cumulative losses leads to the equivalence of the regrets over two different strategy sets. We compare the infima of the cumulative losses over two different strategy sets in the following equations:

$$\begin{aligned}
&\inf_{f \in \mathcal{F}, \alpha \in [0,1]} \sum_{t=1}^T \langle \alpha f + (1-\alpha)\pi_t^M(z_{1:t-1}), z_t \rangle \\
&= \inf_{\alpha \in [0,1]} \inf_{f \in \mathcal{F}} \left[\alpha \sum_{t=1}^T \langle f, z_t \rangle + (1-\alpha) \sum_{t=1}^T \langle \pi_t^M(z_{1:t-1}), z_t \rangle \right] \\
&= \inf_{\alpha \in [0,1]} \left[\alpha \inf_{f \in \mathcal{F}} \sum_{t=1}^T \langle f, z_t \rangle + (1-\alpha) \sum_{t=1}^T \langle \pi_t^M(z_{1:t-1}), z_t \rangle \right] \\
&= \min \left\{ \sum_{t=1}^T \langle \pi_t^M(z_{1:t-1}), z_t \rangle, \inf_{f \in \mathcal{F}} \sum_{t=1}^T \langle f, z_t \rangle \right\}.
\end{aligned}$$

The first equality holds because of the linearity of the loss function ℓ with respect to the first argument. The second equality holds because f only appears in the term $\sum_{t=1}^{T}\langle f, z_t\rangle$. And, the last equality holds because $\alpha \inf_{f\in\mathcal{F}} \sum_{t=1}^{T}\langle f, z_t\rangle + (1-\alpha)\sum_{t=1}^{T}\langle \pi_t^M(z_{1:t-1}), z_t\rangle$ is linear with respect to α, and the infimum of a linear function over a convex set $[0,1]$ is always at the boundary $\{0,1\}$.

The regret competing with the strategy set Π_F^M is defined as

$$\mathbf{Reg}_T(\Pi_{\mathcal{F}}^M) = \sum_{t=1}^{T}\ell(f_t, z_t) - \inf_{f\in\mathcal{F},\alpha\in[0,1]}\sum_{t=1}^{T}\ell(\alpha f + (1-\alpha)\pi_t^M(z_{1:t-1}), z_t),$$

and the minimax value of the online linear optimization problem competing with the strategy set Π_F^M is

$$\mathcal{V}_T(\Pi_{\mathcal{F}}^M) = \inf_{p_1\in\Delta(\mathcal{F})}\sup_{z_1\in\mathcal{Z}_1}\mathop{\mathbb{E}}_{f_1\sim p_1}\dots\inf_{p_T\in\Delta(\mathcal{F})}\sup_{z_T\in\mathcal{Z}_T(z_{1:T-1})}\mathop{\mathbb{E}}_{f_T\sim p_T}$$
$$\left[\sum_{t=1}^{T}\ell(f_t, z_t) - \inf_{f\in\mathcal{F},\alpha\in[0,1]}\sum_{t=1}^{T}\ell(\alpha f + (1-\alpha)\pi_t^M(z_{1:t-1}), z_t)\right]$$

This chapter is organized as follows. In Section 4.2, we analyze the minimax value of online optimization that takes advantage of the predictable process M and competes with the strategy set $\Pi_{\mathcal{F}}^M$. Then, we apply the minimax analysis on the online linear optimization problem. In Section 4.3, we show one online linear optimization algorithm that achieves the minimax value. The algorithm benefits from the given predictable sequence M_t, and also avoids being hurt if the outcome deviates from the predictable sequence. In Section 4.4, we consider the environment with multiple predictable processes. Specifically, we derive algorithms on two cases, (a) finite number of predictable processes and (b) infinite number of predictable sequences and only one optimal strategy. In Section 4.5, we view the whole problem from the viewpoint of dynamic regret [9].

4.2 Minimax Regret

In this chapter, we focus on linear loss $\ell(f,z) = \langle f, z\rangle$, where $\mathcal{F}$ is a unit ball with respect to $\|\cdot\|$ and $\mathcal{Z}$ is a unit ball with respect to the dual norm $\|\cdot\|_\star$. If the learner receives the information that the coming outcome z_t is guaranteed to be $M_t(z_{1:t-1})$, the optimal strategy is to predict

$$\pi_t^M(z_{1:t-1}) = \inf_{f_t\in\mathcal{F}}\langle f_t, M_t(z_{1:t-1})\rangle.$$

If $\|\cdot\|$ is a Euclidean norm, the optimal strategy is to predict

$$\pi_t^M(z_{1:t-1}) = -\frac{M_t(z_{1:t-1})}{\|M_t(z_{1:t-1})\|}.$$

If the learner receives the information that the coming outcome z_t is close to $M_t(z_{1:t-1})$, it is still wise to adjust the prediction according to the predictable sequence $M_t(z_{1:t-1})$.

Since the learner has the extra information M_t, it is a significant advantage compared to fixed actions. Therefore, instead of competing with fixed actions, we use regret competing with strategies. The regret competing with the strategy set $\Pi_{\mathcal{F}}^M$ is defined as

$$\mathbf{Reg}_T^\ell(\Pi_{\mathcal{F}}^M) = \sum_{t=1}^T \langle f_t, z_t\rangle - \inf_{f\in\mathcal{F},\alpha\in[0,1]} \sum_{t=1}^T \langle \alpha f + (1-\alpha)\pi_t^M(z_{1:t-1}), z_t\rangle,$$

where the comparator term is the smaller one between the cumulative loss of the best fixed action $\inf_{f\in\mathcal{F}} \sum_{t=1}^T \langle f, z_t\rangle$ and the cumulative loss of the optimal strategy $\sum_{t=1}^T \langle \pi_t^M(z_{1:t-1}), z_t\rangle$. Further, we define the minimax value of the online linear optimization problem competing with the strategy set Π_F^M as

$$\mathcal{V}_T^\ell(\Pi_{\mathcal{F}}^M) = \inf_{p_1\in\Delta(\mathcal{F})} \sup_{z_1\in\mathcal{Z}_1} \mathop{\mathbb{E}}_{f_1\sim p_1} \dots \inf_{p_T\in\Delta(\mathcal{F})} \sup_{z_T\in\mathcal{Z}_T(z_{1:T-1})} \mathop{\mathbb{E}}_{f_T\sim p_T} \left[\sum_{t=1}^T \langle f_t, z_t\rangle - \inf_{f\in\mathcal{F},\alpha\in[0,1]} \sum_{t=1}^T \langle \alpha f + (1-\alpha)\pi_t^M(z_{1:t-1}), z_t\rangle\right]$$

Theorem 4.2.1. *Suppose $\mathcal{F}$ is a unit ball with respect to the norm $\|\cdot\|$, $\mathcal{Z}$ is unit balls with respect to the dual norm $\|\cdot\|_\star$ and the loss function $\ell(f,z)$ is linear $\ell(f,z) = \langle f, z\rangle$. If the outcome $z_t \in \mathcal{Z}$ is always σ_t-close to the predictable sequence M_t, i.e.*

$$\|z_t - M_t(z_{1:t-1})\|_\star \le \sigma_t,$$

for $t = 1,\dots,T$, then the outcome set at time t can be defined as $\mathcal{Z}_t(z_{1:t-1}) = \{z \in \mathcal{Z} : \|z - M_t(z_{1:t-1})\|_\star \le \sigma_t\}$. Therefore, the minimax value of the online linear optimization problem with respect to the strategy set $\Pi_{\mathcal{F}}^M$ is upper bounded by

$$\mathcal{V}_T^\ell(\Pi_{\mathcal{F}}^M) \le c_0 \sqrt{\sum_{t=1}^T \sigma_t^2}$$

where the constant c_0 depends on the smoothness of the norm $\|\cdot\|$. If the norm $\|\cdot\|$ is L_p-norm ($p > 2$) in Euclidean space, the constant c_0 is $2 + \frac{1}{\sqrt{p-1}}$.

Before we prove Theorem 4.2.1, let us first prove a more general version of Theorem 4.2.1 with general loss functions.

Lemma 4.2.2. *Suppose the prediction set $\mathcal{F}$ is a unit ball with respect to the norm $\|\cdot\|$, the outcome set $\mathcal{Z}$ is a unit ball with respect to the dual norm $\|\cdot\|_\star$ and the* *loss function is $\ell : \mathcal{F} \times \mathcal{Z} \to \mathbb{R}$. If the outcome $z_t \in \mathcal{Z}$ is σ_t-close to the predictable sequence M_t, i.e.*

$$\|z_t - M_t(z_{1:t-1})\|_\star \le \sigma_t,$$

for $t = 1, \dots, T$, then the outcome set at time t is defined as $\mathcal{Z}_t(z_{1:t-1}) = \{z \in \mathcal{Z} : \|z - M_t(z_{1:t-1})\|_\star \le \sigma_t\}$. Therefore, the minimax value of the prediction problem competing with the strategy set Π and constrained adversaries is upper bounded by

$$\begin{aligned}
&\mathcal{V}_T(\Pi) \\
&\triangleq \inf_{p_1 \in \Delta(\mathcal{F})} \sup_{z_1 \in \mathcal{Z}_1} \mathbb{E}_{f_1 \sim p_1} \dots \inf_{p_T \in \Delta(\mathcal{F})} \sup_{z_T \in \mathcal{Z}_T(z_{1:T-1})} \mathbb{E}_{f_T \sim p_T} \left[\sum_{t=1}^T \ell(f_t, z_t) - \inf_{\pi \in \Pi} \sum_{t=1}^T \ell(\pi_t(z_{1:t-1}), z_t) \right] \\
&\le 2 \sup_{\mathbf{w} \in \mathbf{C}} \sup_{\mathbf{z}} \mathbb{E}_{\epsilon} \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t \left(\ell(\pi_t(\mathbf{w}_{1:t-1}(\epsilon)), \mathbf{z}_t(\epsilon)) - \ell(\pi_t(\mathbf{w}_{1:t-1}(\epsilon)), M_t(\mathbf{w}_{1:t-1}(\epsilon))) \right) \right]
\end{aligned}$$

where $\mathbf{w}$ and $\mathbf{z}$ are $\mathcal{Z}$-valued trees. $\mathbf{C}$ is the set of $\mathcal{Z}$-valued trees $\mathbf{z}$ such that all paths are σ_t-close to the predictable sequence M, i.e., for every $\epsilon \in \{\pm 1\}^T$ and $t \in [T]$, $\|\mathbf{z}_t(\epsilon) - M_t(\mathbf{z}_{1:t-1}(\epsilon))\|_\star \le \sigma_t$.

Proof. According to (3.2.3) in Theorem 3.2.2, the minimax value of the online linear optimization problem with respect to the strategy set Π and constrained adversaries is upper bounded by

$$\begin{aligned}
\mathcal{V}_T(\Pi) \le \sup_{p_1 \in \mathcal{P}_1(\cdot)} \mathbb{E}_{z_1, z_1' \sim p_1} \mathbb{E}_{\epsilon_1} \cdots \sup_{p_T \in \mathcal{P}_T(\chi_{1:T-1}(\epsilon_{1:T-1}))} \mathbb{E}_{z_T, z_T' \sim p_T} \mathbb{E}_{\epsilon_T} \\
\sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t \left(\ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), z_t') - \ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), z_t) \right) \right],
\end{aligned}$$

where $\mathcal{P}_t(\chi_{1:t-1}(\epsilon_{1:t-1}))$ is a simplex on the set $\mathcal{Z}_t(\chi_{1:t-1}(\epsilon_{1:t-1}))$. The "selector function" $\chi : \mathcal{Z} \times \mathcal{Z} \times \{\pm 1\} \to \mathcal{Z}$ is defined as $\chi(z_t, z_t', +1) = z_t$ and $\chi(z_t, z_t', -1) = z_t'$. The selector function selects between z_t and z_t' depending on the third argument. When the context is clear, we use $\chi_t(\epsilon)$ to represent $\chi(z_t, z_t', \epsilon)$ for simplicity. Next, we add and subtract corresponding loss terms with M_t to the last statement. The minimax value of the prediction problem is upper bounded by

$$\begin{aligned}
\mathcal{V}_T(\Pi) \le \sup_{p_1 \in \mathcal{P}_1} \mathbb{E}_{z_1, z_1' \sim p_1} \mathbb{E}_{\epsilon_1} \sup_{p_2 \in \mathcal{P}_2(\chi_1(\epsilon_1))} \mathbb{E}_{z_2, z_2' \sim p_2} \mathbb{E}_{\epsilon_2} \cdots \sup_{p_T \in \mathcal{P}_T(\chi_{1:T-1}(\epsilon_{1:T-1}))} \mathbb{E}_{z_T, z_T' \sim p_T} \mathbb{E}_{\epsilon_T} \\
\sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t \left(\ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), z_t') - \ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), M_t(\chi_{1:t-1}(\epsilon_{1:t-1}))) \right) \right. \\
\left. + \sum_{t=1}^T \epsilon_t \left(\ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), M_t(\chi_{1:t-1}(\epsilon_{1:t-1}))) - \ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), z_t) \right) \right]
\end{aligned}$$

After z_t, z'_t and ϵ_t are revealed, $\chi_t(\epsilon_t)$ is fixed and can only be either z_t or z'_t. We remove the dependency of $\chi_t(\epsilon_t)$ on ϵ_t, and replace $\chi_t(\epsilon_t)$ by y_t, which is either z_t or z'_t. Therefore, the minimax value of the prediction problem is upper bounded by

$$\begin{aligned}
\mathcal{V}_T(\Pi) \le \sup_{p_1\in\mathcal{P}_1} \mathop{\mathbb{E}}_{z_1,z'_1\sim p_1} \mathop{\mathbb{E}}_{\epsilon_1} \sup_{y_1\in\{z_1,z'_1\}} \sup_{p_2\in\mathcal{P}_2(y_1)} \mathop{\mathbb{E}}_{z_2,z'_2\sim p_2} \mathop{\mathbb{E}}_{\epsilon_2} \cdots \sup_{y_{T-1}\in\{z_{T-1},z'_{T-1}\}} \sup_{p_T\in\mathcal{P}_T(y_{1:T-1})} \mathop{\mathbb{E}}_{z_T,z'_T\sim p_T} \mathop{\mathbb{E}}_{\epsilon_T} \\
\sup_{\pi\in\Pi}\Bigg[\sum_{t=1}^{T} \epsilon_t \left(\ell(\pi_t(y_{1:t-1}), z'_t) - \ell(\pi_t(y_{1:t-1}), M_t(y_{1:t-1}))\right) \\
+ \sum_{t=1}^{T} \epsilon_t \left(\ell(\pi_t(y_{1:t-1})), M_t(y_{1:t-1})) - \ell(\pi_t(y_{1:t-1}), z_t)\right)\Bigg] \\
\le \sup_{z_1,z'_1\in\mathcal{Z}_1} \mathop{\mathbb{E}}_{\epsilon_1} \sup_{y_1\in\{z_1,z'_1\}} \sup_{z_2,z'_2\in\mathcal{Z}_2(y_1)} \mathop{\mathbb{E}}_{\epsilon_2} \sup_{y_2\in\{z_2,z'_2\}} \cdots \sup_{y_{T-1}\in\{z_{T-1},z'_{T-1}\}} \sup_{z_T,z'_T\in\mathcal{Z}_T(y_{1:T-1})} \mathop{\mathbb{E}}_{\epsilon_T} \\
\sup_{\pi\in\Pi}\Bigg[\sum_{t=1}^{T} \epsilon_t \left(\ell(\pi_t(y_{1:t-1}), z'_t) - \ell(\pi_t(y_{1:t-1}), M_t(y_{1:t-1}))\right) \\
+ \sum_{t=1}^{T} \epsilon_t \left(\ell(\pi_t(y_{1:t-1})), M_t(y_{1:t-1})) - \ell(\pi_t(y_{1:t-1}), z_t)\right)\Bigg]
\end{aligned}$$

Furthermore, as $z_t, z'_t \in \mathcal{Z}_t(y_{1:t-1})$ and $y_t \in \{z_t, z'_t\}$, then $y_t \in \mathcal{Z}_t(y_{1:t-1})$ is true. If we drop the constraints on z_t and z'_t, and loosen the constraint on y_t to be $y_t \in \mathcal{Z}_t(y_{1:t-1})$, the minimax value is upper bounded by

$$\begin{aligned}
\mathcal{V}_T(\Pi) \le \sup_{z_1,z'_1\in\mathcal{Z}} \mathop{\mathbb{E}}_{\epsilon_1} \sup_{y_1\in\mathcal{Z}_1} \sup_{z_2,z'_2\in\mathcal{Z}} \mathop{\mathbb{E}}_{\epsilon_2} \sup_{y_2\in\mathcal{Z}_2(y_1)} \cdots \sup_{y_{T-1}\in\mathcal{Z}_{T-1}(y_{1:T-2})} \sup_{z_T,z'_T\in\mathcal{Z}} \mathop{\mathbb{E}}_{\epsilon_T} \\
\sup_{\pi\in\Pi}\Bigg[\sum_{t=1}^{T} \epsilon_t \left(\ell(\pi_t(y_{1:t-1}), z'_t) - \ell(\pi_t(y_{1:t-1}), M_t(y_{1:t-1}))\right) \\
+ \sum_{t=1}^{T} \epsilon_t \left(\ell(\pi_t(y_{1:t-1})), M_t(y_{1:t-1})) - \ell(\pi_t(y_{1:t-1}), z_t)\right)\Bigg] \\
\le 2 \sup_{z_1\in\mathcal{Z}} \mathop{\mathbb{E}}_{\epsilon_1} \sup_{y_1\in\mathcal{Z}_1} \sup_{z_2\in\mathcal{Z}} \mathop{\mathbb{E}}_{\epsilon_2} \sup_{y_2\in\mathcal{Z}_2(y_1)} \cdots \sup_{y_{T-1}\in\mathcal{Z}_{T-1}(y_{1:T-2})} \sup_{z_T\in\mathcal{Z}} \mathop{\mathbb{E}}_{\epsilon_T} \\
\sup_{\pi\in\Pi}\left[\sum_{t=1}^{T} \epsilon_t \left(\ell(\pi_t(y_{1:t-1}), z_t) - \ell(\pi_t(y_{1:t-1}), M_t(y_{1:t-1}))\right)\right]
\end{aligned} \tag{4.2.1}$$

since the two terms obtaining by splitting the supremum are the same. If we define a new strategy π' and a new loss function L as

$$L(\pi'_t(y_{1:t-1}), z_t) \triangleq \ell(\pi_t(y_{1:t-1}), z_t) - \ell(\pi_t(y_{1:t-1}), M_t(y_{1:t-1})),$$

then (4.2.1) matches the format of (3.2.4). Therefore, steps of Theorem 3.2.2 after (3.2.4) go through and the minimax value of the prediction problem is upper bounded by

$$\begin{aligned}&\mathcal{V}_T(\Pi)\\&\leq 2\sup_{\mathbf{w}\in\mathbf{C}}\sup_{\mathbf{z}}\mathbb{E}_{\epsilon}\sup_{\pi\in\Pi}\left[\sum_{t=1}^{T}\epsilon_t\left(\ell(\pi_t(\mathbf{w}_{1:t-1}(\epsilon)),\mathbf{z}_t(\epsilon))-\ell(\pi_t(\mathbf{w}_{1:t-1}(\epsilon)),M_t(\mathbf{w}_{1:t-1}(\epsilon)))\right)\right],\end{aligned}$$

where $\mathbf{w}$ and $\mathbf{z}$ are $\mathcal{Z}$-valued trees. $\mathbf{C}$ is a set of $\mathcal{Z}$-valued trees and all paths are σ_t-close to the predictable sequence. □

Now, let us prove three propositions to prepare for the proof of Theorem 4.2.1.

Proposition 4.2.3 (a simplified version of Example 13 in Section 12.1 of [16]). *If* $\mathbf{x}$ *is a* $\mathbb{R}$*-valued tree, then*

$$\left(\mathbb{E}_{\epsilon}\left|\sum_{t=1}^{T}\epsilon_t\mathbf{x}_t(\epsilon)\right|\right)^2\leq\mathbb{E}_{\epsilon}\sum_{t=1}^{T}|\mathbf{x}_s(\epsilon)|^2$$

Proof. If $\mathbf{x}$ is a $\mathbb{R}$-valued tree, then

$$\left(\mathbb{E}_{\epsilon}\left|\sum_{t=1}^{T}\epsilon_t\mathbf{x}_t(\epsilon)\right|\right)^2\leq\mathbb{E}_{\epsilon}\left|\sum_{t=1}^{T}\epsilon_t\mathbf{x}_t(\epsilon)\right|^2=\mathbb{E}_{\epsilon}\sum_{s,t=1}^{T}\epsilon_s\epsilon_t\mathbf{x}_s(\epsilon)\mathbf{x}_t(\epsilon)=\mathbb{E}_{\epsilon}\sum_{t=1}^{T}|\mathbf{x}_s(\epsilon)|^2$$

where the first inequality holds because of Jensen's inequality $f(\mathbb{E}X)\leq\mathbb{E}(fX)$, which holds for $\mathbb{R}$-valued random variable X and convex function $f(\cdot)$, and the second equality holds by expanding the squared term, and the third equality holds as the $\mathbb{E}_{\epsilon}\,\epsilon_s\epsilon_t\mathbf{x}_s(\epsilon)\mathbf{x}_t(\epsilon)=0$ when $s\neq t$. □

Let us fill in some background on convex duality. A differentiable function $\Psi:\mathcal{F}\to\mathbb{R}$ is σ-strongly convex with respect to $\|\cdot\|$, that is

$$\forall f_1,f_2\in\mathcal{F},\Psi(f_1)-\Psi(f_2)\geq\langle f_1-f_2,\nabla\Psi(f_2)\rangle+\frac{\sigma}{2}\|f_1-f_2\|^2.$$

Define $\Psi^\star:\mathcal{Z}\to\mathbb{R}$ as the Fenchel conjugate of Ψ, that is

$$\Psi^\star(z)\triangleq\sup_{f\in\mathcal{F}}\langle f,z\rangle-\Psi(f).$$

Then, the definition of the Fenchel conjugate shows the Fenchel-Young inequality, i.e.,

$$\forall f\in\mathcal{F},\forall z\in\mathcal{Z},\langle f,z\rangle\leq\Psi(f)+\Psi^\star(z),$$

and Lemma 15 in Appendix A.4 of [20] shows that

$$\forall z_1,z_2\in\mathcal{Z},\Psi^\star(z_1)-\Psi^\star(z_2)\leq\langle\nabla\Psi^\star(z_2),z_1-z_2\rangle+\frac{1}{2\sigma}\|z_1-z_2\|_\star^2.\tag{4.2.2}$$

Proposition 4.2.4 (a modification of Lemma 2 in [10]). *Let $Z_i \in \mathcal{Z}$ be mean zero independent random vectors, then*

$$\mathbb{E}\left[\Psi^\star\left(\sum_{t=1}^{T} Z_t\right)\right] \le \frac{\sum_{t=1}^{T}\mathbb{E}\,\|Z_t\|_\star^2}{2\sigma}.$$

Proof. 4.2.2 shows

$$\Psi^\star\left(\sum_{s=1}^{t+1} Z_s\right) - \Psi^\star\left(\sum_{s=1}^{t} Z_s\right) \le \left\langle \nabla\Psi^\star\left(\sum_{s=1}^{t} Z_s\right), Z_{t+1}\right\rangle + \frac{1}{2\sigma}\|Z_{t+1}\|_\star^2.$$

Taking expectation with respect to $Z_1, \ldots, Z_t, Z_{t+1}$ and noting $\mathbb{E}\,Z_{t+1} = 0$, then

$$\mathbb{E}\,\Psi^\star\left(\sum_{s=1}^{t+1} Z_s\right) - \mathbb{E}\,\Psi^\star\left(\sum_{s=1}^{t} Z_s\right) \le \left\langle \mathbb{E}\,\nabla\Psi^\star\left(\sum_{s=1}^{t} Z_s\right), \mathbb{E}\,Z_{t+1}\right\rangle + \frac{\mathbb{E}\,\|Z_{t+1}\|_\star^2}{2\sigma} = \frac{\mathbb{E}\,\|Z_{t+1}\|_\star^2}{2\sigma}.$$

Summing the last statement from $t = 0$ to $t = T - 1$, we have

$$\mathbb{E}\,\Psi^\star\left(\sum_{t=1}^{T} Z_t\right) \le \frac{\sum_{t=1}^{T}\mathbb{E}\,\|Z_t\|_\star^2}{2\sigma}.$$

□

Proposition 4.2.5 (a modification of Proposition 12 in [19]). *If $\mathcal{F}$ is a unit ball with respect to $\|\cdot\|$, then*

$$\mathbb{E}_{\epsilon}\sup_{f\in\mathcal{F}}\left\langle f, \sum_{t=1}^{T}\epsilon_t\mathbf{x}_t(\epsilon)\right\rangle \le c\sqrt{\sum_{t=1}^{T}\mathbb{E}_{\epsilon}\,\|\mathbf{x}_t(\epsilon)\|_\star^2},$$

where the constant c depends on the smoothness of the norm $\|\cdot\|$. If the norm $\|\cdot\|$ is L_p-norm ($p > 2$) in Euclidean space, the constant c is $\frac{1}{2\sqrt{p-1}}$.

Proof. By linearity and Fenchel-Young inequality,

$$\mathbb{E}_{\epsilon}\sup_{f\in\mathcal{F}}\left\langle f, \sum_{t=1}^{T}\epsilon_t\mathbf{x}_t(\epsilon)\right\rangle = \frac{1}{\lambda}\mathbb{E}_{\epsilon}\sup_{f\in\mathcal{F}}\left\langle f, \lambda\sum_{t=1}^{T}\epsilon_t\mathbf{x}_t(\epsilon)\right\rangle \le \frac{1}{\lambda}\mathbb{E}_{\epsilon}\left[\sup_{f\in\mathcal{F}}\Psi(f) + \Psi^\star\left(\lambda\sum_{t=1}^{T}\epsilon_t\mathbf{x}_t(\epsilon)\right)\right]$$

Using Proposition 4.2.4, we have

$$\begin{aligned}
\mathbb{E}_{\epsilon}\sup_{f\in\mathcal{F}}\left\langle f, \sum_{t=1}^{T}\epsilon_t\mathbf{x}_t(\epsilon)\right\rangle &\le \frac{1}{\lambda}\sup_{f\in\mathcal{F}}\Psi(f) + \frac{1}{\lambda}\mathbb{E}_{\epsilon}\left[\Psi^\star\left(\lambda\sum_{t=1}^{T}\epsilon_t\mathbf{x}_t(\epsilon)\right)\right] \\
&\le \frac{1}{\lambda}\sup_{f\in\mathcal{F}}\Psi(f) + \frac{1}{\lambda}\frac{\sum_{t=1}^{T}\mathbb{E}_{\epsilon}\,\|\lambda\epsilon_t\mathbf{x}_t(\epsilon)\|_\star^2}{2\sigma} \\
&= \frac{1}{\lambda}\sup_{f\in\mathcal{F}}\Psi(f) + \frac{\lambda\sum_{t=1}^{T}\mathbb{E}_{\epsilon}\,\|\epsilon_t\mathbf{x}_t(\epsilon)\|_\star^2}{2\sigma} \\
&= \frac{1}{\lambda}\sup_{f\in\mathcal{F}}\Psi(f) + \frac{\lambda\sum_{t=1}^{T}\mathbb{E}_{\epsilon}\,\|\mathbf{x}_t(\epsilon)\|_\star^2}{2\sigma}
\end{aligned}$$

If $\lambda = \sqrt{\frac{2\sigma \sup_{f\in\mathcal{F}} \Psi(f)}{\sum_{t=1}^T \mathbb{E}_\epsilon \|\mathbf{x}_t(\epsilon)\|_\star^2}}$, then

$$\mathbb{E}_\epsilon \sup_{f\in\mathcal{F}} \left\langle f, \sum_{t=1}^T \epsilon_t \mathbf{x}_t(\epsilon) \right\rangle \le \sqrt{\frac{\sup_{f\in\mathcal{F}} \Psi(f)}{2\sigma}} \cdot \sqrt{\sum_{t=1}^T \mathbb{E}_\epsilon \|\mathbf{x}_t(\epsilon)\|_\star^2}.$$

If the norm $\|\cdot\|$ is L_p-norm ($p > 2$) in Euclidean space, then the function $\Psi(\cdot) = \frac{1}{2}\|\cdot\|_p^2$ is $(p-1)$-strongly convex. Therefore,

$$\frac{\sup_{f\in\mathcal{F}} \Psi(f)}{2\sigma} = \frac{1/2}{2(p-1)} = \frac{1}{4(p-1)}$$

and

$$\mathbb{E}_\epsilon \sup_{f\in\mathcal{F}} \left\langle f, \sum_{t=1}^T \epsilon_t \mathbf{x}_t(\epsilon) \right\rangle \le \frac{1}{2\sqrt{p-1}} \cdot \sqrt{\sum_{t=1}^T \mathbb{E}_\epsilon \|\mathbf{x}_t(\epsilon)\|_\star^2}.$$

□

Now, let us apply Lemma 4.2.2 and the last three propositions to prove Theorem 4.2.1.

Proof. of Theorem 4.2.1. Suppose the loss function is linear $\ell(f,z) = \langle f, z\rangle$ and the strategy set is $\Pi_{\mathcal{F}}^M$. Using Lemma 4.2.2, we conclude that

$$\mathcal{V}_T^\ell(\Pi_{\mathcal{F}}^M) \le 2 \sup_{\mathbf{w}\in\mathbf{C}} \sup_{\mathbf{z}} \mathbb{E}_\epsilon \sup_{\pi\in\Pi_{\mathcal{F}}^M} \left[\sum_{t=1}^T \epsilon_t \langle \pi_t(\mathbf{w}_{1:t-1}(\epsilon)), \mathbf{z}_t(\epsilon) - M_t(\mathbf{w}_{1:t-1}(\epsilon))\rangle \right].$$

Then,

$$\mathcal{V}_T^\ell(\Pi_{\mathcal{F}}^M) \le 2 \sup_{\mathbf{w}\in\mathbf{C}} \sup_{\mathbf{z}} \mathbb{E}_\epsilon \sup_{\substack{f\in\mathcal{F}\\ \alpha\in[0,1]}} \left[\sum_{t=1}^T \epsilon_t \langle \alpha f + (1-\alpha)\pi_t^M(\mathbf{w}_{1:t-1}(\epsilon)), \mathbf{z}_t(\epsilon) - M_t(\mathbf{w}_{1:t-1}(\epsilon))\rangle \right]$$

according to the definition of the strategy set $\Pi_{\mathcal{F}}^M$. By the linearity of the loss function, we have

$$
\begin{aligned}
\mathcal{V}_T^{\ell}(\Pi_{\mathcal{F}}^M) \le & 2\sup_{\mathbf{w}\in\mathbf{C}}\sup_{\mathbf{z}}\mathop{\mathbb{E}}_{\epsilon}\sup_{\alpha\in[0,1]}\sup_{f\in\mathcal{F}}\Bigg[\alpha\sum_{t=1}^{T}\epsilon_t\langle f,\mathbf{z}_t(\epsilon)-M_t(\mathbf{w}_{1:t-1}(\epsilon))\rangle \\
& \qquad +(1-\alpha)\sum_{t=1}^{T}\epsilon_t\langle \pi_t^M(\mathbf{w}_{1:t-1}(\epsilon)),\mathbf{z}_t(\epsilon)-M_t(\mathbf{w}_{1:t-1}(\epsilon))\rangle\Bigg] \\
= & 2\sup_{\mathbf{w}\in\mathbf{C}}\sup_{\mathbf{z}}\mathop{\mathbb{E}}_{\epsilon}\sup_{\alpha\in[0,1]}\Bigg[\alpha\sup_{f\in\mathcal{F}}\sum_{t=1}^{T}\epsilon_t\langle f,\mathbf{z}_t(\epsilon)-M_t(\mathbf{w}_{1:t-1}(\epsilon))\rangle \\
& \qquad +(1-\alpha)\sum_{t=1}^{T}\epsilon_t\pi_t^M(\mathbf{w}_{1:t-1}(\epsilon)),\mathbf{z}_t(\epsilon)-M_t(\mathbf{w}_{1:t-1}(\epsilon))\rangle\Bigg] \\
= & 2\sup_{\mathbf{w}\in\mathbf{C}}\sup_{\mathbf{z}}\mathop{\mathbb{E}}_{\epsilon}\max\Bigg\{\sup_{f\in\mathcal{F}}\sum_{t=1}^{T}\epsilon_t\langle f,\mathbf{z}_t(\epsilon)-M_t(\mathbf{w}_{1:t-1}(\epsilon))\rangle, \\
& \qquad \sum_{t=1}^{T}\epsilon_t\langle \pi_t^M(\mathbf{w}_{1:t-1}(\epsilon)),\mathbf{z}_t(\epsilon)-M_t(\mathbf{w}_{1:t-1}(\epsilon))\rangle\Bigg\}
\end{aligned}
$$

The last equality holds because of the definition of dual norm. Then, the minimax value of the prediction problem is upper bounded by

$$
\begin{aligned}
\mathcal{V}_T^{\ell}(\Pi_{\mathcal{F}}^M) \le & 2\sup_{\mathbf{w}\in\mathbf{C}}\sup_{\mathbf{z}}\mathop{\mathbb{E}}_{\epsilon}\sup_{f\in\mathcal{F}}\sum_{t=1}^{T}\epsilon_t\langle f,\mathbf{z}_t(\epsilon)-M_t(\mathbf{w}_{1:t-1}(\epsilon))\rangle \\
& + 2\sup_{\mathbf{w}\in\mathbf{C}}\sup_{\mathbf{z}}\mathop{\mathbb{E}}_{\epsilon}\left|\sum_{t=1}^{T}\epsilon_t\langle \pi_t^M(\mathbf{w}_{1:t-1}(\epsilon)),\mathbf{z}_t(\epsilon)-M_t(\mathbf{w}_{1:t-1}(\epsilon))\rangle\right|.
\end{aligned}
$$

According to Proposition 4.2.3, we have

$$
\begin{aligned}
& \sup_{\mathbf{w}\in\mathbf{C}}\sup_{\mathbf{z}}\mathop{\mathbb{E}}_{\epsilon}\left|\sum_{t=1}^{T}\epsilon_t\langle \pi_t^M(\mathbf{w}_{1:t-1}(\epsilon)),\mathbf{z}_t(\epsilon)-M_t(\mathbf{w}_{1:t-1}(\epsilon))\rangle\right| \\
& \quad \le \sup_{\mathbf{w}\in\mathbf{C}}\sup_{\mathbf{z}}\left(\mathop{\mathbb{E}}_{\epsilon}\sum_{t=1}^{T}\left|\epsilon_t\langle \pi_t^M(\mathbf{w}_{1:t-1}(\epsilon)),\mathbf{z}_t(\epsilon)-M_t(\mathbf{w}_{1:t-1}(\epsilon))\rangle\right|^2\right)^{1/2} \le \sqrt{\sum_{t=1}^{T}\sigma_t^2}
\end{aligned}
$$

According to Proposition 4.2.5, we have

$$
\begin{aligned}
& \sup_{\mathbf{w}\in\mathbf{C}}\sup_{\mathbf{z}}\mathop{\mathbb{E}}_{\epsilon}\sup_{f\in\mathcal{F}}\sum_{t=1}^{T}\epsilon_t\langle f,\mathbf{z}_t(\epsilon)-M_t(\mathbf{w}_{1:t-1}(\epsilon))\rangle \\
& \quad \le \sup_{\mathbf{w}\in\mathbf{C}}\sup_{\mathbf{z}} c\sqrt{\sum_{t=1}^{T}\mathop{\mathbb{E}}_{\epsilon}\|\mathbf{z}_t(\epsilon)-M_t(\mathbf{w}_{1:t-1}(\epsilon))\|_\star^2} \le c\sqrt{\sum_{t=1}^{T}\sigma_t^2}
\end{aligned}
$$

where the constant c depends on the smoothness of the norm $\|\cdot\|$. Therefore,

$$\mathcal{V}_T^{\ell}(\Pi_{\mathcal{F}}^M) \le 2(1+c)\sqrt{\sum_{t=1}^{T}\sigma_t^2}.$$

If the norm $\|\cdot\|$ is L_p-norm ($p>2$) in Euclidean space, then the constant c is $\frac{1}{2\sqrt{p-1}}$ and

$$\mathcal{V}_T^{\ell}(\Pi_{\mathcal{F}}^M) \le (2+\frac{1}{\sqrt{p-1}})\sqrt{\sum_{t=1}^{T}\sigma_t^2}.$$

□

4.3 Algorithm

Theorem 4.2.1 indicates the existence of better online optimization methods if the predictable sequence is available. In this section, we present an algorithm that obtains an upper bound that matches the minimax value in Theorem 4.2.1. Our algorithm mainly combines the Exponential Weights Algorithm and the Optimistic Mirror Descent Algorithm. The Optimistic Mirror Descent Algorithm is introduced in [17], and we show the Optimistic Mirror Descent Algorithm and several related results in Subsection 4.3.1.

4.3.1 Optimistic Mirror Descent Algorithm

Algorithm 4 Optimistic Mirror Descent Algorithm

Input: $\mathcal{R}$ is a 1- strongly convex function with respect to $\|\cdot\|$, learning rate $\eta>0$
Initialize: $f_1 = g_1 = \arg\min_{g\in\mathcal{F}}\mathcal{R}(g)$
for $t=1$ to T **do**
predict f_t and update

- $g_{t+1} = \arg\min_{g\in\mathcal{F}} \eta\langle g, z_t\rangle + D_{\mathcal{R}}(g, g_t)$
- $f_{t+1} = \arg\min_{f\in\mathcal{F}} \eta\langle f, M_{t+1}\rangle + D_{\mathcal{R}}(f, g_{t+1})$

end for

Lemma 4.3.1 (the same as Lemma 2 in [17]). *Let $\mathcal{F}$ be a convex set in a Banach space $\mathcal{B}$ and $\mathcal{X}$ be a convex set in the dual space $\mathcal{B}^\star$. Let $\mathcal{R}:\mathcal{B}\to\mathbb{R}$ be a 1-strongly* *convex function on $\mathcal{F}$ with respect to $\|\cdot\|$. For any strategy of Nature, the Optimistic Mirror Descent Algorithm yields, for any $f^\star \in \mathcal{F}$,*

$$\sum_{t=1}^{T}\langle f_t, z_t\rangle - \sum_{t=1}^{T}\langle f^\star, z_t\rangle \le \eta^{-1}R_{\max}^2 + \frac{\eta}{2}\sum_{t=1}^{T}\|z_t - M_t\|_\star^2$$

where $R_{\max}^2 = \max_{f\in\mathcal{F}}\mathcal{R}(f) - \min_{f\in\mathcal{F}}\mathcal{R}(f)$.

Suppose $\mathcal{F} \in \mathbb{R}^d$ is the probability simplex and $\mathcal{Z}$ is ℓ_∞ ball. If

$$\mathcal{R}(w) = \sum_{t=1}^{d} w(i)\log w(i) - 1$$

for $w \in \mathcal{F}$, then the Optimistic Mirror Descent Algorithm (Algorithm 4) is simplify to the Optimistic Exponential Weights Algorithm (Algorithm 5).

Algorithm 5 Optimistic Exponential Weights Algorithm

Input: learning rate $\eta > 0$
Initialize: $w_1 = v_1 = (\frac{1}{d}, \dots, \frac{1}{d})$
for $t = 1$ to T **do**
predict w_t and update

- $v_{t+1}(i) \propto \exp\{-\eta \sum_{s=1}^{t} z_s(i)\}$
- $w_{t+1}(i) \propto \exp\{-\eta \sum_{s=1}^{t} z_s(i) - \eta M_{t+1}(i)\}$

end for

Lemma 3 in [17] shows that the Optimistic Mirror Descent achieves a regret bound in terms of local norms

$$\|v\|_w = \sqrt{v^T\nabla^2\mathcal{R}(w)v}, \text{ for } v \in \mathcal{F}$$

and

$$\|z\|_w^\star = \sqrt{z^T\nabla^2\mathcal{R}(w)^{-1}z}, \text{ for } z \in \mathcal{Z},$$

where the Hessian $\nabla^2\mathcal{R}(w)$ is $\mathrm{diag}(w(1)^{-1}, \dots, w(d)^{-1})$.

Lemma 4.3.2 (the same as Lemma 3 in [17])**.** *The Optimistic Mirror Descent on the probability simplex enjoys, for any $w^\star \in \mathcal{F}$,*

$$\sum_{t=1}^{T}\langle w_t - w^\star, z_t\rangle \le 2\eta\sum_{t=1}^{T}(\|z_t - M_t\|_{w_t}^\star)^2 + \frac{\log d}{\eta}$$

as long as $\eta\|z_t - M_t\|_\infty \le 1/4$ at each step.

Algorithm 6 Main Algorithm

Input: $\mathcal{R}$ is a 1- strongly convex function with respect to $\|\cdot\|$, learning rate $\eta_1 > 0$ and $\eta_2 \in (0, \frac{1}{4}]$

Initialize: $f_1 = g_1 = \arg\min_{f \in \mathcal{F}} \mathcal{R}(f)$ and $\alpha_1 = \frac{1}{2}$

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

predict $h_t = \alpha_t f_t + (1 - \alpha_t)\pi_t^M$

observe z_t, suffer loss $\langle h_t, z_t \rangle$, update

- $g_{t+1} = \arg\min_{g \in \mathcal{F}} \eta_1 \langle g, z_t \rangle + D_{\mathcal{R}}(g, g_t)$
- $f_{t+1} = \arg\min_{f \in \mathcal{F}} \eta_1 \langle f, M_{t+1} \rangle + D_{\mathcal{R}}(f, g_{t+1})$
- $\alpha_{t+1} = \frac{\exp\left(-\eta_2 \sum_{s=1}^{t} \langle f_s, z_s \rangle - \eta_2 \langle f_{t+1}, M_{t+1} \rangle\right)}{\exp\left(-\eta_2 \sum_{s=1}^{t} \langle f_s, z_s \rangle - \eta_2 \langle f_{t+1}, M_{t+1} \rangle\right) + \exp\left(-\eta_2 \sum_{s=1}^{t} \langle \pi_s^M, z_s \rangle - \eta_2 \langle \pi_{t+1}^M, M_{t+1} \rangle\right)}$

end for

4.3.2 Main Algorithm

Algorithm 6 combines the Exponential Weights Algorithm and the Optimistic Mirror Descent Algorithm [17]. The update step in the Optimistic Mirror Descent is exactly the f_t update in Algorithm 6. According to Lemma 4.3.1, the cumulative loss of the Optimistic Mirror Descent Algorithm is upper bounded by the cumulative loss of the best fixed action plus a measure of closeness between the outcome and the predictable sequence. It guarantees the low regret competing with fixed actions $\{f \in \mathcal{F}\}$.

As we are competing with both fixed actions $\{f \in \mathcal{F}\}$ and the optimal strategy π^M, the Optimistic Exponential Weights Algorithm balances between fixed actions and the optimal strategy. The optimal strategy π^M and the Optimistic Mirror Descent Algorithm $(f_t)_{t=1}^T$ are viewed as two experts. The parameter α, the update of the Optimistic Exponential Weights Algorithm, tunes the weights on two experts based on the historical performance and also the estimated further performance according to the predictable sequence M_t.

At step t, we assign weight α_t to f_t, and weight $1-\alpha_t$ to π_t^M. Then the weighted loss of $\langle f_t, z_t \rangle$ and $\langle \pi_t^M, z_t \rangle$ is

$$\alpha_t \langle f_t, z_t \rangle + (1 - \alpha_t)\langle \pi_t^M, z_t \rangle.$$

This process can also be viewed as an online linear optimization problem, the protocol is equivalent to (a) predict a vector $(\alpha_t, 1 - \alpha_t)$ in a simplex, (b) receive the predictable vector $(\langle f_t, M_t \rangle, \langle \pi_t^M, M_t \rangle)$, (c) observe the loss

$$\langle (\alpha_t, 1 - \alpha_t), (\langle f_t, z_t \rangle, \langle \pi_t^M, z_t \rangle) \rangle.$$

This loss equals to the weighted loss of $\langle f_t, z_t\rangle$ and $\langle \pi_t^M, z_t\rangle$. This change of viewpoint helps us use the result in Lemma 5.

Algorithm 6 mainly follows the setting of the Optimistic Mirror Descent Algorithm in Algorithm 5. $\mathcal{R}$ is a 1-strongly convex function with respect to a norm $\|\cdot\|$, and $D_{\mathcal{R}}(\cdot,\cdot)$ denotes the Bregman divergence with respect to $\mathcal{R}$. $\|\cdot\|_\star$ is dual to $\|\cdot\|$. Also, for all algorithms presented in this section, M_1 is assumed to be 0 without loss of generality. With the assumption that we receive the predictable sequence $\{M_t\}_t$ from outside, we do not write the dependence of $\{M_t\}_t$ on the past explicitly.

Theorem 4.3.3. *Let $\mathcal{F}$ be a unit ball with respect to $\|\cdot\|$, $\mathcal{Z}$ be a unit ball with respect to the dual norm $\|\cdot\|_\star$ and $\mathcal{R}:\mathcal{F}\to\mathbb{R}$ be a 1-strongly convex function on $\mathcal{F}$ with respect to $\|\cdot\|$. For any sequence $z_1,\dots,z_T\in\mathcal{Z}$, $\alpha^\star\in[0,1]$ and $f^\star\in\mathcal{F}$, Algorithm 6 yields*

$$\sum_{t=1}^{T}\langle h_t, z_t\rangle - \sum_{t=1}^{T}\langle \alpha^\star f^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle \le \left(2\eta_2 + \frac{\eta_1}{2}\right)\sum_{t=1}^{T}\|z_t - M_t\|_\star^2 + \frac{1}{\eta_1}R_{\max}^2 + \frac{1}{\eta_2},$$

where $R_{\max}^2 = \max_{f\in\mathcal{F}}\mathcal{R}(f) - \min_{f\in\mathcal{F}}\mathcal{R}(f)$, $\eta_1>0$ and $\eta_2\in(0,1/4]$.

Proof. If the Optimistic Mirror Descent Algorithm $(f_t)_{t=1}^T$ and the optimal strategy Π^M are viewed as two experts, $\langle f_t, z_t\rangle$ and $\langle \pi_t^M, z_t\rangle$ are the corresponding losses at step t. Then, there are three equivalent ways to represent the regret,

$$\begin{aligned}
&\sum_{t=1}^{T}\langle h_t, z_t\rangle - \sum_{t=1}^{T}\langle \alpha^\star f^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle \\
&= \sum_{t=1}^{T}\alpha_t\langle f_t, z_t\rangle + (1-\alpha_t)\langle \pi_t^M, z_t\rangle - \sum_{t=1}^{T}\alpha^\star\langle f_t, z_t\rangle + (1-\alpha^\star)\langle \pi_t^M, z_t\rangle \\
&= \sum_{t=1}^{T}\langle(\alpha_t, 1-\alpha_t), (\langle f_t, z_t\rangle, \langle \pi_t^M, z_t\rangle)\rangle - \sum_{t=1}^{T}\langle(\alpha^\star, 1-\alpha^\star), (\langle f_t, z_t\rangle, \langle \pi_t^M, z_t\rangle)\rangle
\end{aligned}$$

The first equality holds by replacing h_t by $\alpha_t f_t + (1-\alpha_t)\pi_t^M$, and the second equality holds by reorganizing the inner product. The last formula matches the format in Lemma 4.3.2.

At time t, we predict $w_t = (\alpha_t, 1-\alpha_t)$, and suffer loss

$$\langle(\alpha_t, 1-\alpha_t), (\langle f_t, z_t\rangle, \langle \pi_t^M, z_t\rangle)\rangle.$$

According to Lemma 4.3.2 (the dimension of the simplex $d=1$), the Optimistic Exponential Weights Algorithm with learning rate η_2 enjoys, for any $\alpha^\star\in[0,1]$,

$$\sum_{t=1}^{T}\langle(\alpha_t, 1-\alpha_t), (\langle f_t, z_t\rangle, \langle \pi_t^M, z_t\rangle)\rangle - \sum_{t=1}^{T}\langle(\alpha^\star, 1-\alpha^\star), (\langle f_t, z_t\rangle, \langle \pi_t^M, z_t\rangle)\rangle$$
$$\leq 2\eta_2 \sum_{t=1}^{T}\left[\|(\langle f_t, z_t\rangle, \langle \pi_t^M, z_t\rangle) - (\langle f_t, M_t\rangle, \langle \pi_t^M, M_t\rangle)\|^\star_{(\alpha_t, 1-\alpha_t)}\right]^2 + \frac{1}{\eta_2}$$
$$= 2\eta_2 \sum_{t=1}^{T}\left[\|(\langle f_t, z_t - M_t\rangle, \langle \pi_t^M, z_t - M_t\rangle)\|^\star_{(\alpha_t, 1-\alpha_t)}\right]^2 + \frac{1}{\eta_2} := A(\eta_2).$$

as long as $\eta_2\|(\langle f_t, z_t - M_t\rangle, \langle \pi_t^M, z_t - M_t\rangle)\|_\infty \leq 1/4$ at each step. This constraint is always true as long as $\eta_2 \leq 1/4$. As $h_t = \alpha_t f_t + (1-\alpha_t)\pi_t^M$, the inequation above is equivalent to

$$\sum_{t=1}^{T}\langle h_t, z_t\rangle = \sum_{t=1}^{T}\langle \alpha_t f_t + (1-\alpha_t)\pi_t^M, z_t\rangle \leq \alpha^\star \sum_{t=1}^{T}\langle f_t, z_t\rangle + (1-\alpha^\star)\sum_{t=1}^{T}\langle \pi_t^M, z_t\rangle + A(\eta_2). \tag{4.3.1}$$

On the other hand, f_t is the update of the Optimistic Mirror Descent Algorithm with learning rate η_1. According to Lemma 4.3.2, for any $f^\star \in \mathcal{F}$

$$\sum_{t=1}^{T}\langle f_t, z_t\rangle - \sum_{t=1}^{T}\langle f^\star, z_t\rangle \leq \eta_1^{-1}R_{\max}^2 + \frac{\eta_1}{2}\sum_{t=1}^{T}(\|z_t - M_t\|_\star)^2 := B(\eta_1). \tag{4.3.2}$$

If inequalities (4.3.1) and (4.3.1) are combined, for any $f^\star \in \mathcal{F}$ and $\alpha^\star \in [0,1]$,

$$\sum_{t=1}^{T}\langle h_t, z_t\rangle \leq \alpha^\star\left[\sum_{t=1}^{T}\langle f^\star, z_t\rangle + B(\eta_1)\right] + (1-\alpha^\star)\sum_{t=1}^{T}\langle \pi_t^M, z_t\rangle + A(\eta_2)$$
$$= \sum_{t=1}^{T}\langle \alpha^\star f^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle + A(\eta_2) + \alpha^\star B(\eta_1)$$
$$\leq \sum_{t=1}^{T}\langle \alpha^\star f^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle + A(\eta_2) + B(\eta_1)$$

Then,

$$\sum_{t=1}^{T}\langle h_t, z_t\rangle - \sum_{t=1}^{T}\langle \alpha^\star f^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle \leq A(\eta_2) + B(\eta_1)$$
$$= A(\eta_2) + \eta_1^{-1}R_{\max}^2 + \frac{\eta_1}{2}\sum_{t=1}^{T}(\|z_t - M_t\|_\star)^2$$

As local norm $\|z\|_w^\star$ is defined as $\sqrt{z^T \nabla^2 \mathcal{R}(w)^{-1} z}$ and the Hessian $\nabla^2 \mathcal{R}(w)$ is $\mathrm{diag}(w(1)^{-1}, \ldots, w(d)^{-1})$, then

$$
\begin{aligned}
A(\eta_2) &= 2\eta_2 \sum_{t=1}^{T} \left[\|(\langle f_t, z_t - M_t\rangle, \langle \pi_t^M, z_t - M_t\rangle)\|_{(\alpha_t, 1-\alpha_t)}^\star \right]^2 + \frac{1}{\eta_2} \\
&= 2\eta_2 \sum_{t=1}^{T} \left[\alpha_t \left(\langle f_t, z_t - M_t\rangle\right)^2 + (1-\alpha_t) \left(\langle \pi_t^M, z_t - M_t\rangle\right)^2 \right] + \frac{1}{\eta_2} \\
&\leq 2\eta_2 \sum_{t=1}^{T} \left[\alpha_t \|f_t\|^2 \|z_t - M_t\|_\star^2 + (1-\alpha_t) \|\pi_t^M\|^2 \|z_t - M_t\|_\star^2 \right] + \frac{1}{\eta_2} \\
&\leq 2\eta_2 \sum_{t=1}^{T} \|z_t - M_t\|_\star^2 + \frac{1}{\eta_2}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\sum_{t=1}^{T} \langle h_t, z_t\rangle - \sum_{t=1}^{T} \langle \alpha^\star f^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle \\
&\leq 2\eta_2 \sum_{t=1}^{T} \|z_t - M_t\|_\star^2 + \frac{1}{\eta_2} + \frac{\eta_1}{2} \sum_{t=1}^{T} \|z_t - M_t\|_\star^2 + \frac{1}{\eta_1} R_{\max}^2
\end{aligned}
$$

□

If we know $\sum_{t=1}^T \|z_t - M_t\|_\star^2$ ahead of time, the regret bound $\mathcal{O}(\sqrt{\sum_{t=1}^T \|z_t - M_t\|_\star^2})$ is achieved by choosing $\eta_1 = \left(\sum_{t=1}^T \|z_t - M_t\|_\star^2 / 2R_{\max}\right)^{-1/2}$ and $\eta_2 = \left(2\sum_{t=1}^T \|z_t - M_t\|_\star^2\right)^{-1/2}$. Moreover, the standard doubling trick helps us to obtain the convergence rate even $\sum_{t=1}^T \|z_t - M_t\|_\star^2$ is unknown in advance.

Lemma 4.3.4. *Divide the learning problem into phases, with a constant learning rate $\lambda_i = \lambda_0 2^{-i}$ throughout the i-th phase, for some $\lambda_0 > 0$. Define for $i \geq 1$*

$$
s_{i+1} = \min\left\{ \tau : \frac{5\lambda_i}{2} \sum_{t=s_i}^{\tau} \|z_t - M_t\|_\star^2 > \frac{1}{\lambda_i}(1 + R_{\max}^2) \right\}
$$

to be the start of the phase $i+1$, and $s_1 = 1$. Let N be the last phase and let $S_{N+1} = T+1$. Then, Algorithm 6 with time-varying learning parameters η_1 and η_2 yields

$$
\sum_{t=1}^{T} \langle h_t, z_t\rangle - \sum_{t=1}^{T} \langle \alpha^\star f^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle \leq C_M \sqrt{\sum_{t=1}^{T} \|z_t - M_t\|_\star^2},
$$

where the constant C_M is problem dependent.

Proof. Let $\eta = \eta_1 = \eta_2$, then Theorem 4.3.3 shows that Algorithm 6 yields

$$\sum_{t=1}^{T}\langle h_t, z_t\rangle - \sum_{t=1}^{T}\langle \alpha^\star f^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle \le \frac{5\eta}{2}\sum_{t=1}^{T}\|z_t - M_t\|_\star^2 + \frac{1}{\eta}(1+R_{\max}^2).$$

Without loss of generality, assume $N > 1$. Then,

$$\begin{aligned}
&\sum_{t=1}^{T}\langle h_t, z_t\rangle - \sum_{t=1}^{T}\langle \alpha^\star f^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle \\
&\le \sum_{k=1}^{N}\left[\sum_{t=s_k}^{s_{k+1}-1}\langle h_t, z_t\rangle - \sum_{t=s_k}^{s_{k+1}-1}\langle \alpha^\star f^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle\right] \\
&\le \sum_{k=1}^{N}\left[\frac{5\lambda_k}{2}\sum_{t=s_k}^{s_{k+1}-1}\|z_t - M_t\|_\star^2 + \frac{1}{\lambda_k}(1+R_{\max}^2)\right] \\
&\le 2\sum_{k=1}^{N}\frac{1}{\lambda_k}(1+R_{\max}^2),
\end{aligned}$$

where the last inequality holds because of the definition of s_k. Also observe that

$$\frac{5\lambda_{N-1}}{2}\sum_{t=s_{N-1}}^{s_N}\|z_t - M_t\|_\star^2 > \frac{1}{\lambda_{N-1}}(1+R_{\max}^2),$$

which implies

$$\lambda_0^{-1}2^N = \lambda_N^{-1} = 2\lambda_{N-1}^{-1} \le \sqrt{\frac{10\sum_{t=s_{N-1}}^{s_N}\|z_t - M_t\|_\star^2}{1+R_{\max}^2}} \le \sqrt{\frac{10\sum_{t=1}^{T}\|z_t - M_t\|_\star^2}{1+R_{\max}^2}}$$

Hence, the regret is upper bounded by

$$2(1+R_{\max}^2)\sum_{k=1}^{N}\lambda_k^{-1} = 2(1+R_{\max}^2)\lambda_0^{-1}2^N\sum_{k=1}^{N}2^{k-N} \le 4\sqrt{10(1+R_{\max}^2)\sum_{t=1}^{T}\|z_t - M_t\|_\star^2}.$$

For general version of the Doubling Trick, details are shown in Appendix B of [17]. □

4.4 Learning the Predictable Processes

Algorithm 6 yields regret of $\mathcal{O}(\sqrt{\sum_{t=1}^{T}\|z_t - M_t\|_\star^2})$ if we have access to the predictable sequence $(M_t)_{t\ge1}$. It shows that we can benefit from the predictable sequence. Sometimes, there are multiple predictable sequences $\mathcal{M} = \{(M_t)_{t\ge1}\}$, instead of one single predictable sequence. We know that the outcome is close to one sequence in the set $\mathcal{M} = \{(M_t)_{t\ge1}\}$, but have no idea which one it is. This idea is formalized as:

$$\exists M^\star \in \mathcal{M}, \text{ such that } \|z_t - M_t^\star(z_{1:t-1})\|_\star \le \sigma_t, \ \forall t \in [T].$$

It is interesting to consider the problem of choosing the best predictable sequence from the predictable sequence set $\mathcal{M} = \{(M_t)_{t\ge1}\}$. In the online learning language, it can be expressed as competing with all of the optimal strategies π^M for all $M \in \mathcal{M}$, and also all of the fixed actions $\{f \in \mathcal{F}\}$. To formalize the idea, we want to compete with

$$\min\left\{\inf_{f\in\mathcal{F}} \sum_{t=1}^T \langle f, z_t\rangle, \inf_{M\in\mathcal{M}} \sum_{t=1}^T \langle \pi_t^M, z_t\rangle\right\},$$

which is equivalent to

$$\inf_{\substack{f\in\mathcal{F}\\ M\in\mathcal{M}\\ \alpha\in[0,1]}} \sum_{t=1}^T \langle \alpha f + (1-\alpha)\pi_t^M, z_t\rangle,$$

and the new format makes it easier to apply the result in Lemma 4.3.4.

In this section, we explore the model selection problem in two cases. The first case is when the size of the predictable processes is finite, and we handle this case by adding one more layer of the Exponential Weights Algorithm on top of Algorithm 6. In each step, we first apply Algorithm 6 to each predictable sequence $M \in \mathcal{M}$, and then aggregate the output by the Exponential Weights Algorithm.

The second case is when there is only one optimal strategy, even though there are infinite predictable sequences. One example is the online linear optimization. If these infinite predictable sequences have the same direction, the optimal strategy is the same even through the magnitude of these predictable sequences may vary widely. We solve this case by combining the Exponential Weights Algorithm, Optimistic Mirror Descent and Gradient Descent algorithm.

4.4.1 Finite Predictable Processes

Let $\mathcal{F}$ be a unit ball with respect to $\|\cdot\|$, and $\mathcal{Z}$ be a unit ball with respect to the dual norm $\|\cdot\|_\star$. Suppose $\mathcal{M}$ is a set of predictable processes, and the outcome $z_1, \ldots, z_t \in \mathcal{Z}$ is close to one of the learning process $\{(M_t)_{t\ge1}\}$, i.e., there exists $M \in \mathcal{M}$ such that $\|z_t - M_t(z_{1:t-1})\|_\star \le \sigma_t$ for all $t \in \{1, \ldots, T\}$.

Algorithm 7 chooses the optimal predictable sequence adaptively with three layers. On the bottom layer, the Optimistic Mirror Descent Algorithm updates f_t^M for every predictable sequence M. It guarantees the low-regret performance competing with fixed actions. On the second layer, the Exponential Weights Algorithm updates h_t^M for every M. This step balances the fixed actions and the optimal strategy with respect to M. According to Algorithm 6 in Section 3, these two layers produce updates that have regret bound $\mathcal{O}(\sqrt{\sum_{t=1}^T \|z_t - M_t\|_\star^2})$ when competing with optimal strategies π^M and fixed actions.

Algorithm 7 Finite Predictable Processes

Input: $\mathcal{R}$ 1-strongly convex with respect to $\|\cdot\|$, learning rate $\eta > 0$
Initialize: for every $M \in \mathcal{M}$, $q_1(M) = \frac{1}{|\mathcal{M}|}$, learning rate $\eta_1^M > 0$ and $\eta_2^M > 0$, $f_1^M = g_1^M = \arg\min_{f \in \mathcal{F}} \mathcal{R}(f)$ and $\alpha_1^M = \frac{1}{2}$
for $t = 1$ to T **do**
predict $g_t = \sum_{M \in \mathcal{M}} q_t(M) h_t^M$
Observe z_t and for every $M \in \mathcal{M}$

- $h_t^M = \alpha_t^M f_t^M + (1 - \alpha_t^M)\pi_t^M$
- $g_{t+1}^M = \arg\min_{g \in \mathcal{F}} \eta_1 \langle g, z_t \rangle + D_{\mathcal{R}}(g, g_t)$
- $f_{t+1}^M = \arg\min_{f \in \mathcal{F}} \eta_1 \langle f, M_{t+1} \rangle + D_{\mathcal{R}}(f, g_{t+1})$
- $\alpha_{t+1}^M = \frac{\exp\left(-\eta_2^M \sum_{s=1}^t \langle f_s^M, z_s \rangle - \eta_2^M \langle f_{t+1}^M, M_{t+1} \rangle\right)}{\exp\left(-\eta_2^M \sum_{s=1}^t \langle f_s^M, z_s \rangle - \eta_2^M \langle f_{t+1}^M, M_{t+1} \rangle\right) + \exp\left(-\eta_2^M \sum_{s=1}^t \langle \pi_s^M, z_s \rangle - \eta_2^M \langle \pi_{t+1}^M, M_{t+1} \rangle\right)}$

$q_{t+1}(M) \propto q_t(M) e^{-(\langle h_t^M, z_t \rangle + \|z_t\|_\star)/2}$
end for

The top layer g_t is the update of the Exponential Weights Algorithm that adaptively chooses the optimal predictable sequence. The multiple factor for weights update in this step is

$$\exp\{-(\langle h_t^M, z_t \rangle + \|z_t\|_\star)/2\},$$

instead of the exponential of the loss directly. It is because the particular regret bound for the Exponential Weights Algorithm ([3], Theorem 2.4, Corollary 2.4) is needed for the top layer, and it requires the loss per step to be a number between 0 and 1. The update g_t adaptively adjusts the weights on the middle layer updates $\{h_t^M\}_{M \in \mathcal{M}}$, and the optimal predictable sequence is adaptively selected.

Theorem 4.4.1. *Let $\mathcal{F}$ be a unit ball with respect to $\|\cdot\|$, and $\mathcal{Z}$ be a unit ball with respect to the dual norm $\|\cdot\|_\star$. Suppose $\mathcal{M}$ is a set of predictable processes, the cumulative loss of Algorithm 7 satisfies*

$$\sum_{t=1}^T \langle g_t, z_t \rangle \leq \widetilde{L_T} + \sqrt{\left(\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_\star\right) \ln|\mathcal{M}|} + \ln|\mathcal{M}|,$$

where

$$
\begin{aligned}
\widetilde{L_T} &\triangleq \min_{M\in\mathcal{M}} \sum_{t=1}^{T} \langle h_t^M, z_t\rangle \\
&\le \min_{M\in\mathcal{M}} \left\{ \inf_{f\in\mathcal{F},\alpha\in[0,1]} \sum_{t=1}^{T} \langle \alpha f + (1-\alpha)\pi_t^M, z_t\rangle + C_M \sqrt{\sum_{t=1}^{T} \|z_t - M_t\|_\star^2} \right\}.
\end{aligned}
$$

To interpret the result, we need to understand $\widetilde{L_T}$ and $\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_\star$. First, the range of $\widetilde{L_T}$ is $[-\sum_{t=1}^T \|z_t\|_\star, \sum_{t=1}^T \|z_t\|_\star]$, so the range of $\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_\star$ is $[0, 2\sum_{t=1}^T \|z_t\|_\star]$. If $\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_\star$ goes sub-linearly, we can benefit from the predictable sequence.

Also, according to the definition of $\widetilde{L_T}$, $\widetilde{L_T}$ is upper bounded by

$$
\begin{aligned}
\widetilde{L_T} \le \min \Bigg\{ & \min_{M\in\mathcal{M}} \left\{ \sum_{t=1}^{T} \langle \pi_t^M, z_t\rangle + C_M \sqrt{\sum_{t=1}^{T} \|z_t - M_t\|_\star^2} \right\}, \\
& \inf_{f\in\mathcal{F}} \sum_{t=1}^{T} \langle f, z_t\rangle + \min_{M\in\mathcal{M}} \left\{ C_M \sqrt{\sum_{t=1}^{T} \|z_t - M_t\|_\star^2} \right\} \Bigg\}.
\end{aligned}
$$

The first term captures the performances of predictable sequences and corresponding strategies $\{\pi^M\}_{M\in\mathcal{M}}$, and the second term captures the performance of fixed actions $\{f \in \mathcal{F}\}$. The result shows the regret continuously depends on the distance between the optimal predictable sequence and the outcome, i.e.,

$$
\min_{M\in\mathcal{M}} \left\{ C_M \sqrt{\sum_{t=1}^{T} \|z_t - M_T\|_\star^2} \right\}.
$$

We provide three corollaries to better understand Theorem 7, especially the $\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_\star$ term and $L^\star$ term. The first corollary interprets the $\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_2$ term when both $\mathcal{F}$ and $\mathcal{Z}$ are unit balls with respect to $\|\cdot\|_2$. The other two corollaries focus on two extreme cases. In the optimal case, the algorithm picks the optimal predictable sequence automatically without extra price, and it recovers the Halving algorithm. At the same time, the algorithm guarantees that we still perform as well as the best fixed action in term of T in any situation.

Proof. of Theorem 4.4.1. According to Lemma 4.3.4, the cumulative loss of h_t^M satisfies

$$
\sum_{t=1}^{T} \langle h_t^M, z_t\rangle \le \inf_{f\in\mathcal{F},\alpha\in[0,1]} \sum_{t=1}^{T} \langle \alpha f + (1-\alpha)\pi_t^M, z_t\rangle + C_M \sqrt{\sum_{t=1}^{T} \|z_t - M_t\|_\star^2}.
$$

Then, $\{h_t^M\}_{M\in\mathcal{M}}$ are treated as $|\mathcal{M}|$ experts, and corresponding losses are $(\langle h_t^M, z_t\rangle + \|z_t\|_\star)/2$. The Exponential Weights Algorithm (Theorem 2.4, [3]) with parameter η yields

$$\sum_{t=1}^T \frac{\langle g_t, z_t\rangle + \|z_t\|_\star}{2} \le \frac{\eta \min_{M\in\mathcal{M}} \sum_{t=1}^T \frac{\langle h_t^M, z_t\rangle + \|z_t\|_\star}{2} + \ln|\mathcal{M}|}{1-e^{-\eta}} = \frac{\eta \frac{\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_\star}{2} + \ln|\mathcal{M}|}{1-e^{-\eta}}.$$

The last equality holds as $\widetilde{L_T}$ is defined as $\min_{M\in\mathcal{M}}\sum_{t=1}^T \langle h_t^M, z_t\rangle$. According to Corollary 2.4 of [3], we set the parameter η to be $\ln\left(1 + \sqrt{4\ln|\mathcal{M}|/(\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_\star)}\right)$. Then, we have

$$\sum_{t=1}^T \frac{\langle g_t, z_t\rangle + \|z_t\|_\star}{2} - \frac{\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_\star}{2} \le \sqrt{2\frac{\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_\star}{2}\ln|\mathcal{M}|} + \ln|\mathcal{M}|.$$

By canceling the common term $\sum_{t=1}^T \|z_t\|_\star/2$ and moving $\widetilde{L_T}$ to the right-hand side, the last statement is equivalent to

$$\sum_{t=1}^T \langle g_t, z_t\rangle \le \widetilde{L_T} + \sqrt{\left(\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_\star\right)\ln|\mathcal{M}|} + \ln|\mathcal{M}|.$$

□

Corollary 4.4.2. *If both $\mathcal{F}$ and $\mathcal{Z}$ are unit balls with respect to $\|\cdot\|_2$, we have*

$$\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_2 \le \min_{M\in\mathcal{M}}\left(2\sum_{t=1}^T \|z_t - M_t\|_2 + C_M\sqrt{\sum_{t=1}^T \|z_t - M_t\|_2^2}\right).$$

If there exists one $M \in \mathcal{M}$ such that $\|z_t - M_t\|_2 \le \sigma_t$, then

$$\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_2 \le 2\sum_{t=1}^T \sigma_t + C\sqrt{\sum_{t=1}^T \sigma_t^2}.$$

Proof. According to the definition of $\widetilde{L_T}$, we have

$$\begin{aligned}
\widetilde{L_T} &\le \min_{M\in\mathcal{M}}\left\{\inf_{f\in\mathcal{F},\alpha\in[0,1]} \sum_{t=1}^T \langle \alpha f + (1-\alpha)\pi_t^M, z_t\rangle + C_M\sqrt{\sum_{t=1}^T \|z_t - M_t\|_\star^2}\right\} \\
&\le \min_{M\in\mathcal{M}}\left\{\sum_{t=1}^T \langle \pi_t^M, z_t\rangle + C_M\sqrt{\sum_{t=1}^T \|z_t - M_t\|_\star^2}\right\}
\end{aligned}$$

the last inequality holds by setting α to be 0. Now, let us focus on the $\sum_{t=1}^{T}\langle \pi_t^M, z_t\rangle$ term. If both $\mathcal{F}$ and $\mathcal{Z}$ are unit balls with respect to $\|\cdot\|_2$, we have

$$\langle \pi_t^M, z_t\rangle + \|z_t\|_2 = -\langle \frac{M_t}{\|M_t\|_2}, z_t\rangle + \|z_t\|_2 = \|z_t\|_2\left(1 - \langle \frac{M_t}{\|M_t\|_2}, \frac{z_t}{\|z_t\|_2}\rangle\right).$$

Define the inner product between the normalized z_t and M_t as α, we have

$$\langle \pi_t^M, z_t\rangle + \|z_t\|_2 \le 2\|z_t - M_t\|_2$$

as

- if $\langle M_t, z_t\rangle \ge 0$, i.e., $\alpha \in [0,1]$

$$\begin{aligned}
&\|z_t\|_2^2(1-\alpha^2) - \|z_t - M_t\|_2^2 \\
&= \|z_t\|_2^2(1-\alpha^2) - (\|z_t\|_2^2 - 2\|z_t\|_2\|M_t\|_2\alpha + \|M_t\|_2^2) \\
&= -(\|z_t\|_2^2\alpha^2 - 2\|z_t\|_2\|M_t\|_2\alpha + \|M_t\|_2^2) \\
&= -(\|z_t\|_2\alpha - \|M_t\|_2)^2 \le 0
\end{aligned}$$

  So,

$$\langle \pi_t^M, z_t\rangle + \|z_t\|_2 = \|z_t\|_2(1-\alpha) \le \|z_t\|_2\sqrt{1-\alpha^2} \le \|z_t - M_t\|_2$$

- if $\langle M_t, z_t\rangle < 0$, i.e., $\alpha \in [-1,0)$

$$\langle \pi_t^M, z_t\rangle + \|z_t\|_2 = (1-\alpha)\|z_t\|_2 \le 2\|z_t\|_2$$

  and

$$\|z_t\|_2^2 \le \|z_t\|_2^2 - 2\alpha\|z_t\|_2\|M_t\|_2 + \|M_t\|_2^2 = \|z_t - M_t\|_2^2$$

So,

$$\sum_{t=1}^{T}\langle \pi_t^M, z_t\rangle + \sum_{t=1}^{T}\|z_t\|_2 \le 2\sum_{t=1}^{T}\|z_t - M_t\|_2$$

and

$$\widetilde{L_T} + \sum_{t=1}^{T}\|z_t\|_2 \le \min_{M\in\mathcal{M}}\left(2\sum_{t=1}^{T}\|z_t - M_t\|_2 + C_M\sqrt{\sum_{t=1}^{T}\|z_t - M_t\|_2^2}\right).$$

$\square$

Corollary 4.4.3. *Suppose $\mathcal{F}$ and $\mathcal{Z}$ are in Euclidean space, then*

$$\sum_{t=1}^{T}\langle g_t, z_t\rangle \le \min\left\{\inf_{f\in\mathcal{F}}\sum_{t=1}^{T}\langle f, z_t\rangle, \inf_{M\in\mathcal{M}}\sum_{t=1}^{T}\langle \pi_t^M, z_t\rangle\right\} + \ln|\mathcal{M}|.$$

if there exists one $M^\star \in \mathcal{M}$ such that $z_t = M_t^\star$ $\forall t$.

It recovers the result of the Halving algorithm. We make at most $\ln|\mathcal{M}|$ mistakes, and suffer at most loss 1 for each mistake.

Proof. If there exists one $M^\star \in \mathcal{M}$ such that $z_t = M_t^\star$ $\forall t$, then the optimal strategy with respect to $M^\star$ is

$$\pi_t^{M^\star} = \arg\min_{f\in\mathcal{F}} \langle f, M_t\rangle = -\frac{z_t}{\|z_t\|}$$

and

$$\langle \pi_t^{M^\star}, M_t\rangle = -\|z_t\|_\star.$$

If we pick $\alpha = 0$ and $M = M^\star$,

$$\begin{aligned}\widetilde{L_T} &\le \min_{M\in\mathcal{M}} \left\{ \inf_{f\in\mathcal{F},\alpha\in[0,1]} \sum_{t=1}^T \langle \alpha f + (1-\alpha)\pi_t^M, z_t\rangle + C_M \sqrt{\sum_{t=1}^T \|z_t - M_t\|_\star^2} \right\} \\ &\le \sum_{t=1}^T \langle \pi_t^{M^\star}, z_t\rangle + C_{M^\star}\sqrt{\sum_{t=1}^T \|z_t - M_t^\star\|_\star^2} = \sum_{t=1}^T \langle \pi_t^{M^\star}, z_t\rangle = -\sum_{t=1}^T \|z_t\|_\star.\end{aligned}$$

Apply $\widetilde{L_T} \le \sum_{t=1}^T \langle \pi_t^{M^\star}, z_t\rangle$ and $\widetilde{L_T} + \sum_{t=1}^T \|z_t\|_\star = 0$ to Theorem 4.4.1, we have

$$\sum_{t=1}^T \langle g_t, z_t\rangle \le \sum_{t=1}^T \langle \pi_t^{M^\star}, z_t\rangle + \ln|\mathcal{M}| = \inf_{M\in\mathcal{M}} \sum_{t=1}^T \langle \pi_t^M, z_t\rangle + \ln|\mathcal{M}|.$$

It shows that the optimal predictable sequence is automatically picked, and the extra cumulative loss compared to follow the best predictable sequence is at most $\ln|\mathcal{M}|$. At the same time, as

$$\sum_{t=1}^T \langle g_t, z_t\rangle \le -\sum_{t=1}^T \|z_t\|_\star + \ln|\mathcal{M}| \le -\left\|\sum_{t=1}^T z_t\right\|_\star + \ln|\mathcal{M}| = \inf_{f\in\mathcal{F}} \sum_{t=1}^T \langle f, z_t\rangle + \ln|\mathcal{M}|,$$

this algorithm also guarantees that the cumulative loss is upper bounded by the cumulative loss of the best fixed action plus logarithm of the size of the predictable sequence set. □

Corollary 4.4.4. *If $\mathcal{F}$ and $\mathcal{Z}$ are unit balls with respect to ℓ_2-norm $\|\cdot\|_2$, then the cumulative loss of Algorithm 6 competing with fixed actions $\{f \in \mathcal{F}\}$ satisfies*

$$\begin{aligned}&\sum_{t=1}^T \langle g_t, z_t\rangle - \inf_{f\in\mathcal{F}} \sum_{t=1}^T \langle f, z_t\rangle \\ &\le \min_{M\in\mathcal{M}} C_M \sqrt{\sum_{t=1}^T \|z_t - M_t\|_2^2} \\ &\quad + \sqrt{\min_{M\in\mathcal{M}} \left(2\sum_{t=1}^T \|z_t - M_t\|_2 + C_M \sqrt{\sum_{t=1}^T \|z_t - M_t\|_2^2} \right) \ln|\mathcal{M}| + \ln|\mathcal{M}|}\end{aligned}$$

Proof. If $\mathcal{F}$ and $\mathcal{Z}$ are unit balls with respect to ℓ_2-norm $\|\cdot\|_2$, we have

$$\widetilde{L_T} + \sum_{t=1}^{T} \|z_t\|_2 \leq \min_{M \in \mathcal{M}} \left(2 \sum_{t=1}^{T} \|z_t - M_t\|_2 + C_M \sqrt{\sum_{t=1}^{T} \|z_t - M_t\|_2^2} \right).$$

If we only consider the regret competing with fixed action $\{f \in \mathcal{F}\}$ and pick $\alpha = 1$, we have

$$\begin{aligned}
\widetilde{L_T} &\leq \min_{M \in \mathcal{M}} \left\{ \inf_{f \in \mathcal{F}, \alpha \in [0,1]} \sum_{t=1}^{T} \langle \alpha f + (1-\alpha)\pi_t^M, z_t \rangle + C_M \sqrt{\sum_{t=1}^{T} \|z_t - M_t\|_\star^2} \right\} \\
&\leq \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \langle f, z_t \rangle + \min_{M \in \mathcal{M}} C_M \sqrt{\sum_{t=1}^{T} \|z_t - M_t\|_2^2}.
\end{aligned}$$

Applying previous two inequalities to Theorem 4.4.1, we have

$$\begin{aligned}
&\sum_{t=1}^{T} \langle g_t, z_t \rangle - \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \langle f, z_t \rangle \\
&\leq \min_{M \in \mathcal{M}} C_M \sqrt{\sum_{t=1}^{T} \|z_t - M_t\|_2^2} \\
&\quad + \sqrt{\min_{M \in \mathcal{M}} \left(2 \sum_{t=1}^{T} \|z_t - M_t\|_2 + C_M \sqrt{\sum_{t=1}^{T} \|z_t - M_t\|_2^2} \right) \ln |\mathcal{M}| + \ln |\mathcal{M}|}
\end{aligned}$$

□

If we put no constraint on the outcome, the regret competing with fixed actions is $c\sqrt{T \ln |\mathcal{M}| + \ln |\mathcal{M}|}$. Compared to the $\mathcal{O}(\sqrt{T})$ regret, we perform much better in the optimal case, and in the worst case, we do not suffer more in term of T.

4.4.2 Infinite Predictable Sequences, Only One Optimal Strategy

Let us consider another special case, where there are infinite predictable sequences with only one optimal strategy. Suppose $\mathcal{F}$ is a unit ball with respect to $\|\cdot\|$, and $\mathcal{Z}$ is a unit ball with respect to the dual norm $\|\cdot\|_\star$ and the loss function is linear $\ell(f, z) = \langle f, z \rangle$.

If we know the direction of the next move, but are not sure about the magnitude. There are potentially infinite predictable sequences based on different magnitudes.

However, there is only one optimal strategy as the loss function is linear. The optimal strategy is to choose the opposite direction with maximum magnitude. Mathematically, the sets of predictable sequence is $\{\lambda M\}_{\lambda\in(0,1]}$, where M is one predictable sequence based on history and $\|M\|_\star = 1$ and the optimal strategy is to predict $\pi_t^M = -M_t/\|M_t\|$ at time t.

In this section, we show an algorithm such that we benefit from these infinite predictable sequences and it also guarantees our performance even if the outcome is away from predictable sequences.

Algorithm 8 Infinite Predictable Sequences, Only One Optimal Strategy

---

Input: learning rate $\eta_0 > 0$, $\eta_1 > 0$ and $\eta_2 > 0$, $\mathcal{R}$ self-concordant barrier

Initialize $\lambda_1 = 0$, $f_1 = \arg\min_{f\in\mathcal{F}} \mathcal{R}(f)$ and $\alpha_1 = \frac{1}{2}$

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

predict $h_t = \alpha_t f_t + (1-\alpha_t)\pi_t^M$, observe z_t, suffer loss $\langle h_t, z_t\rangle$, update

- $\lambda_{t+1} = \Pi_{[0,1]}\left(\lambda_t - \eta_0 \frac{\partial}{\partial\lambda}\Big|_{\lambda=\lambda_t} \|z_t - \lambda M_t\|_\star^2\right)$
- $\widetilde{M}_{t+1} = \lambda_{t+1} M_{t+1}$
- $g_{t+1} = \arg\min_{g\in\mathcal{F}} \eta_1\langle g, z_t\rangle + D_{\mathcal{R}}(g, g_t)$
- $f_{t+1} = \arg\min_{f\in\mathcal{F}} \eta_1\langle f, \widetilde{M}_{t+1}\rangle + D_{\mathcal{R}}(f, g_{t+1})$
- $\alpha_{t+1} = \frac{\exp\left(-\eta_2\sum_{s=1}^t\langle f_s, z_s\rangle - \eta_2\langle f_{t+1}, \widetilde{M}_{t+1}\rangle\right)}{\exp\left(-\eta_2\sum_{s=1}^t\langle f_s, z_s\rangle - \eta_2\langle f_{t+1}, \widetilde{M}_{t+1}\rangle\right) + \exp\left(-\eta_2\sum_{s=1}^t\langle \pi_s^M, z_s\rangle - \eta_2\langle \pi_{t+1}^M, \widetilde{M}_{t+1}\rangle\right)}$

end for

It is also a three-layer algorithm. We first update the magnitude λ_t by gradient descent to get the optimal predictable precess $\widetilde{M}_t = \lambda_t M_t$. Then, we update f_t by the Optimistic Mirror Descent Algorithm with $\widetilde{M}_t$. Finally, we use the Exponential Weights Algorithm to balance between the optimal strategy π^M and fixed actions $\{f \in \mathcal{F}\}$.

Theorem 4.4.5. *For any $\alpha^\star \in (0,1]$ and $f^\star \in \mathcal{F}$, Algorithm 8 satisfies*

$$\sum_{t=1}^T\langle h_t, z_t\rangle - \sum_{t=1}^T\langle \alpha^\star f^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle \le C_1\sqrt{\inf_{\lambda\in(0,1]}\sum_{t=1}^T\|z_t - \lambda M_t\|_\star^2 + C_2\sqrt{T}}$$

Proof. π_t^M is the optimal strategy no matter how λ is chosen. According to Lemma 4.3.4, we have

$$\sum_{t=1}^T\langle h_t, z_t\rangle - \sum_{t=1}^T\langle \alpha^\star f^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle \le C_1\sqrt{\sum_{t=1}^T\|z_t - \widetilde{M}_t\|_\star^2}.$$

As λ_t is the update of gradient descent with the goal of minimizing the regret

$$\sum_{t=1}^{T} \|z_t - \lambda_t M_t\|_\star^2 - \inf_{\lambda \in (0,1]} \sum_{t=1}^{T} \|z_t - \lambda M_t\|_\star^2,$$

then

$$\sum_{t=1}^{T} \|z_t - \widetilde{M_t}\|_\star^2 \le \inf_{\lambda \in (0,1]} \sum_{t=1}^{T} \|z_t - \lambda M_t\|_\star^2 + C_2\sqrt{T}$$

due to the optimality of gradient descent and the convexity of the function $\|z_t - \lambda M_t\|_\star^2$ with respect to λ. □

Furthermore, we can build history dependent structure on the magnitude, instead of fixed magnitude $\lambda \in (0,1]$. Then, we apply techniques from "Competing with Strategies", and get more interesting results.

Also, we improve the result by taking the advantage of the convexity of the special loss function $\ell(\lambda, (M,z)) = \|z - \lambda M\|_\star^2$.

If $\|\cdot\|_\star$ is ℓ_2 norm

Theorem 4.4.6. *If $\lambda_t^\star = (\sum_{s=1}^{t} \langle M_s, z_s\rangle)/t$ and $\hat{\lambda}_t = \lambda_{t-1}^\star$, then*

$$\sum_{t=1}^{T} \|z_t - \lambda_t M_t\|_2^2 - \inf_{\lambda \in (0,1]} \sum_{t=1}^{T} \|z_t - \lambda M_t\|_2^2 \le \sum_{t=1}^{T} \frac{8}{t} \le 8(1 + \ln T)$$

Proof. This proof is similar to the proof of Theorem 3.1 in the Section 3.2 of [3]. Define loss function as $\ell(\lambda, (M_t, z_t))$, it is easy to verify that

$$\hat{\lambda}_t = \arg\min_{\lambda \in (0,1]} \sum_{s=1}^{t-1} \ell(\lambda, (M_s, z_s))$$

and

$$\lambda_t^\star = \arg\min_{\lambda \in (0,1]} \sum_{s=1}^{t-1} \ell(\lambda, (M_s, z_s)).$$

Then, for any $(M, z) \in \mathcal{Z} \times \mathcal{Z}$

$$\ell(\hat{\lambda}_t, (M,z)) - \ell(\lambda_t^\star, (M,z)) = \|z - \hat{\lambda}_t M\|^2 - \|z - \lambda_t^\star M\|^2 \le 4|\hat{\lambda}_t - \lambda_t^\star| = 4|\hat{\lambda}_t - \hat{\lambda}_{t-1}|$$
$$= 4\left|\frac{\sum_{s=1}^{t-1}\langle M_s, z_s\rangle}{t-1} - \frac{\sum_{s=1}^{t}\langle M_s, z_s\rangle}{t}\right| = 4\left|\frac{1}{t} \cdot \frac{\sum_{s=1}^{t-1}\langle M_s, z_s\rangle}{t-1} - \frac{\langle M_t, z_t\rangle}{t}\right| \le \frac{8}{t}$$

Both inequalities hold because the set $\mathcal{Z}$ is bounded. Therefore,

$$\sum_{t=1}^{T} \|z_t - \lambda_t M_t\|^2 - \inf_{\lambda \in (0,1]} \sum_{t=1}^{T} \|z_t - \lambda M_t\|^2 \le \sum_{t=1}^{T} \frac{8}{t} \le 8(1 + \ln T)$$

□

This theorem leads to the much lower regret of Algorithm 8 as

$$\sum_{t=1}^{T}\langle h_t, z_t\rangle - \sum_{t=1}^{T}\langle \alpha^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle \le c\sqrt{\inf_{\lambda\in(0,1]}\sum_{t=1}^{T}\|z_t - \lambda M_t\|_2^2 + 8(1+\ln T)}.$$

If $\|\cdot\|_\star$ is ℓ_p-norm in $\mathbb{R}^N$ space, where $p \in (1,2)$

Define $\ell(\lambda,(M,x)) = \|x - \lambda M\|_p^2$, it is easy to verify that

$$\frac{\partial}{\partial\lambda}\ell(\lambda,(M,x)) = 2\|\lambda M - x\|_p^{2-p}\sum_{i=1}^{N}\frac{M_{(i)}(\lambda M_{(i)} - z_{(i)})}{|\lambda M_{(i)} - z_{(i)}|^{p-2}} \in [-2,2]$$

and

$$\frac{\partial^2}{\partial\lambda^2}\ell(\lambda,(M,x))$$
$$= \frac{2(2-p)}{\|\lambda M - x\|_p}\left[\sum_{i=1}^{N}\frac{M_{(i)}(\lambda M_{(i)} - z_{(i)})}{|\lambda M_{(i)} - z_{(i)}|^{2-p}}\right]^2 + 2(p-1)\|\lambda M - x\|_p^{2-p}\sum_{i=1}^{N}\frac{M_{(i)}^2}{|\lambda M_{(i)} - z_{(i)}|^{2-p}} \ge 0$$

Theorem 4.4.7. *If $\frac{\partial^2}{\partial\lambda^2}\ell(\lambda,(M,x))$ is bounded from below by a constant $C > 0$, then the regret of the-best-expert algorithm, i.e. $\lambda_t = \arg\min_{\lambda\in(0,1]}\sum_{s=1}^{t-1}\|z_s - \lambda M_s\|_p^2$, satisfies*

$$\sum_{t=1}^{T}\|z_t - \lambda_t M_t\|_p^2 - \inf_{\lambda\in(0,1]}\sum_{t=1}^{T}\|z_t - \lambda M_t\|_p^2 \le \frac{16}{C}(1+\ln T)$$

Proof. It is a special case of Theorem 3.1 in [3]. To prove, let us check these assumptions one by one.

1. $\ell(\cdot,(M,x))$ is convex in λ and bounded (the original assumption is "ℓ takes values in [0,1]", while boundedness condition is enough for the proof.).
2. $\ell(\cdot,(M,x))$ is Lipschitz in its first argument, with constant $B = 2$.
3. $\ell(\cdot,(M,x))$ is twice differentiable and $\frac{\partial^2}{\partial\lambda^2}\ell(\lambda,(M,x))$ is bounded from below by a constant $C > 0$.
4. define $\lambda_t^\star$ as the solution of $\nabla\Psi_t(\lambda) = 0$, where $\Psi_t(\lambda) = \frac{1}{t}\sum_{s=1}^{t}\ell(\lambda,(M_s,z_s))$.

After all these assumptions are checked, we have

$$\sum_{t=1}^{T}\|z_t - \lambda_t M_t\|_p^2 - \inf_{\lambda\in(0,1]}\sum_{t=1}^{T}\|z_t - \lambda M_t\|_p^2 \le \frac{16}{C}(1+\ln T)$$

if we choose $\lambda_t = \arg\min_{\lambda\in(0,1]}\sum_{s=1}^{t-1}\|z_s - \lambda M_s\|_p^2$. □

This theorem also leads to the much lower regret of Algorithm 8 as

$$\sum_{t=1}^{T}\langle h_t, z_t\rangle - \sum_{t=1}^{T}\langle \alpha^\star + (1-\alpha^\star)\pi_t^M, z_t\rangle \le c\sqrt{\inf_{\lambda\in(0,1]}\sum_{t=1}^{T}\|z_t - \lambda M_t\|_p^2 + \frac{16}{c}(1+\ln T)}.$$

4.5 Dynamic Regret

If the optimal strategy evolutes close to a dynamic process $\Phi_t : \mathcal{F} \to \mathcal{F}$, [9] presents Dynamic Mirror Descent which combines online optimization with the dynamic process. The cumulative loss of Dynamic Mirror Descent comparing to the cumulative loss of any $(\theta_1, \theta_2, \ldots, \theta_T) \in \mathcal{F}^T$ is upper bounded by

$$\mathcal{O}\left(\sqrt{T}\left[1 + \sum_{t=1}^{T} \|\theta_{t+1} - \Phi_{t+1}(\theta_t)\|\right]\right),$$

where $\sum_{t=1}^{T} \|\theta_{t+1} - \Phi_{t+1}(\theta_t)\|$ measures how the competitor $(\theta_1, \theta_2, \ldots, \theta_T)$ is away from the dynamic process $(\Phi_1(\cdot), \Phi_2(\theta_1), \cdots, \Phi_T(\theta_{T-1}))$.

We improve the results by adding the predictable sequence. For a given loss function, the predictable sequence and the optimal strategy appear as pairs. We view the dynamic process $(\Phi_1(\cdot), \Phi_2(\theta_1), \cdots, \Phi_T(\theta_{T-1}))$ as the optimal strategy, and define the paired predictable sequence as $(M_1(\cdot), M_2(z_1), \ldots, M_T(z_{1:T-1}))$. Our goal is to benefit from both the optimal strategy and the predictable sequence, and to produce low regret online learning algorithm.

At time t, we predict $f_t \in \mathcal{F}$, observe outcome $z_t \in \mathcal{Z}$ and suffer loss $\langle f_t, z_t \rangle$. Algorithm 9 incorporates a dynamical model, denoted by $\Phi_t : \mathcal{F} \to \mathcal{F}$, and the paired predictable sequence $(M_1(\cdot), M_2(z_1), \ldots, M_T(z_{1:T-1}))$. For regret, we define the comparator sequence in term of $(\theta_1, \ldots, \theta_T) \in \mathcal{F}^T$, and Algorithm 9 admits a regret bound of the form

$$\mathcal{O}\left(\sqrt{\sum_{t=1}^{T} \|z_t - M_t\|_\star^2}\left[1 + \sum_{t=1}^{T} \|\theta_{t+1} - \Phi_{t+1}(\theta_t)\|\right]\right).$$

4.5.1 Dynamic Model with Predictable Sequence

Algorithm 9 combines the Optimistic Mirror Descent [17] and the Dynamic Mirror Descent [9]. The update of Dynamic Mirror Descent g_t reduces the regret by embedding the dynamic model / the optimal strategy. At the same time, the update of Optimistic Mirror Descent f_t takes advantage of the predictable sequence. Algorithm 9 and Algorithm 6 yield similar results, but from different viewpoints. Algorithm 9 starts from the optimal strategy $\Phi_t(\cdot)$, and finds the paired predictable sequence $M_t(\cdot)$. However, Algorithm 6 is in the opposite direction. It starts from the predictable sequence $M_t(\cdot)$, and finds the optimal strategy $\pi_t(\cdot)$.

Theorem 4.5.1. *Let $\mathcal{F}$ be a convex set in a Banach space $\mathcal{B}$ and $\mathcal{X}$ be convex sets in the dual space $\mathcal{B}^\star$. Let $\mathcal{R} : \mathcal{B} \to \mathbb{R}$ be a σ-strongly convex function on $\mathcal{F}$ with respect to $\|\cdot\|$. Let $\Phi_t : \mathcal{F} \to \mathcal{F}$ be a dynamical model such that $\Delta_{\Phi_t} \triangleq \max_{f,f} \mathcal{D}_{\mathcal{R}}(\Phi_t(f), \Phi_t(f')) - \mathcal{D}_{\mathcal{R}}(f, f') \le 0$ for $t = 1, 2, \ldots$. For any strategy of Nature*

Algorithm 9 Dynamic Model with Predictable Sequence

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

predict f_t, observe z_t, suffer loss $\langle f_t, z_t\rangle$, update

- $g_{t+1} = \arg\min_{g\in\mathcal{F}} \eta\langle g, z_t\rangle + \mathcal{D}_{\mathcal{R}}(g, \Phi_t(g_t))$
- $f_{t+1} = \arg\min_{f\in\mathcal{F}} \eta\langle f, M_{t+1}\rangle + \mathcal{D}_{\mathcal{R}}(f, \Phi_{t+1}(g_{t+1}))$

end for

and any $(\theta_1, \dots, \theta_T) \in \mathcal{F}^T$, *Algorithm 9 yields*

$$\sum_{t=1}^T \langle f_t, z_t\rangle \le \sum_{t=1}^T \langle \theta_t, z_t\rangle + \frac{\eta}{2\sigma}\sum_{t=1}^T \|z_t - M_t\|_\star^2 + \frac{1}{\eta}\left[D_{\max} + 4M\sum_{t=1}^T \|\theta_{t+1} - \Phi_{t+1}(\theta_t)\|\right]$$

where $D_{\max} = \max_{f,f'\in\mathcal{F}} \mathcal{D}_{\mathcal{R}}(f, f')$ *and* $M = \frac{1}{2}\max_{f\in\mathcal{F}} \|\nabla\mathcal{R}(f)\|$.

Proof. For any $\theta_t \in \mathcal{F}$,

$$\langle f_t, z_t\rangle - \langle\theta_t, z_t\rangle = \underbrace{\langle f_t - g_{t+1}, z_t - M_t\rangle}_{A_t} + \underbrace{\langle f_t - g_{t+1}, M_t\rangle}_{B_t} + \underbrace{\langle g_{t+1} - \theta_t, z_t\rangle}_{C_t}$$

First, according to Hölder's inequality,

$$A_t \le \|f_t - g_{t+1}\|\|z_t - M_t\|_\star \le \frac{\sigma}{2\eta}\|f_t - g_{t+1}\|^2 + \frac{\eta}{2\sigma}\|z_t - M_t\|_\star^2.$$

Proposition 18 in [5] shows that

$$\langle f - f_1, z\rangle \le \mathcal{D}_{\mathcal{R}}(f_1, f_0) - \mathcal{D}_{\mathcal{R}}(f_1, f) - \mathcal{D}_{\mathcal{R}}(f, f_0).$$

if $z \in \mathcal{Z}$, $f = \arg\min_{f\in\mathcal{F}}(\langle f, z\rangle + \mathcal{D}_{\mathcal{R}}(f, f_0))$ and $f_1 \in \mathcal{F}$. We apply this proposition to term B_t and C_t. Due to the optimality of $f_t = \arg\min_{f\in\mathcal{F}} \eta\langle f, M_t\rangle + \mathcal{D}_{\mathcal{R}}(f, \Phi_t(g_t))$, we have

$$B_t \le \frac{1}{\eta}\left[\mathcal{D}_{\mathcal{R}}(g_{t+1}, \Phi_t(g_t)) - \mathcal{D}_{\mathcal{R}}(g_{t+1}, f_t) - \mathcal{D}_{\mathcal{R}}(f_t, \Phi_t(g_t))\right].$$

Due to the optimality of $g_{t+1} = \arg\min_{g\in\mathcal{F}} \eta\langle g, z_t\rangle + \mathcal{D}_{\mathcal{R}}(g, \Phi_t(g_t))$, we have

$$C_t \le \frac{1}{\eta}\left[\mathcal{D}_{\mathcal{R}}(\theta_t, \Phi_t(g_t)) - \mathcal{D}_{\mathcal{R}}(\theta_t, g_{t+1}) - \mathcal{D}_{\mathcal{R}}(g_{t+1}, \Phi_t(g_t))\right].$$

Use the previous three inequalities about A_t, B_t and C_t, we conclude

$$
\begin{aligned}
&\langle f_t, z_t\rangle - \langle \theta_t, z_t\rangle \\
&\le \frac{\sigma}{2\eta}\|f_t - g_{t+1}\|^2 + \frac{\eta}{2\sigma}\|z_t - M_t\|_\star^2 \\
&\qquad\qquad + \frac{1}{\eta}\left[\mathcal{D}_\mathcal{R}(\theta_t, \Phi_t(g_t)) - \mathcal{D}_\mathcal{R}(g_{t+1}, f_t) - \mathcal{D}_\mathcal{R}(f_t, \Phi_t(g_t)) - \mathcal{D}_\mathcal{R}(\theta_t, g_{t+1})\right] \\
&\le \frac{\sigma}{2\eta}\|f_t - g_{t+1}\|^2 + \frac{\eta}{2\sigma}\|z_t - M_t\|_\star^2 + \frac{1}{\eta}\left[\mathcal{D}_\mathcal{R}(\theta_t, \Phi_t(g_t)) - \mathcal{D}_\mathcal{R}(g_{t+1}, f_t) - \mathcal{D}_\mathcal{R}(\theta_t, g_{t+1})\right],
\end{aligned}
$$

where the first inequality holds by canceling the $\mathcal{D}_\mathcal{R}(g_{t+1}, \Phi_t(g_t))$ term, and the second inequality holds by removing the negative term $-\mathcal{D}_\mathcal{R}(f_t, \Phi(g_t))$. By the strong convexity of $\mathcal{R}$, $\mathcal{D}_\mathcal{R}(g_{t+1}, f_t) \ge \frac{\sigma}{2}\|g_{t+1} - f_t\|^2$. Then, the term $\frac{\sigma}{2\eta}\|f_t - g_{t+1}\|^2$ and the term $-\frac{1}{\eta}\mathcal{D}_\mathcal{R}(g_{t+1}, f_t)$ are canceled. Thus,

$$
\langle f_t, z_t\rangle - \langle f_t, \theta_t\rangle \le \frac{\eta}{2\sigma}\|z_t - M_t\|_\star^2 + \frac{1}{\eta}\underbrace{\left[\mathcal{D}_\mathcal{R}(\theta_t, \Phi_t(g_t)) - \mathcal{D}_\mathcal{R}(\theta_t, g_{t+1})\right]}_{E_t}.
$$

E_t can be decomposed into sum of three parts

$$
\begin{aligned}
E_t = &\underbrace{\mathcal{D}_\mathcal{R}(\theta_t, \Phi_t(g_t)) - \mathcal{D}_\mathcal{R}(\theta_{t+1}, \Phi_{t+1}(g_{t+1}))}_{E_{t,0}} \\
&+ \underbrace{\mathcal{D}_\mathcal{R}(\theta_{t+1}, \Phi_{t+1}(g_{t+1})) - \mathcal{D}_\mathcal{R}(\Phi_{t+1}(\theta_t), \Phi_{t+1}(g_{t+1}))}_{E_{t,1}} \\
&+ \underbrace{\mathcal{D}_\mathcal{R}(\Phi_{t+1}(\theta_t), \Phi_{t+1}(g_{t+1})) - \mathcal{D}_\mathcal{R}(\theta_t, g_{t+1})}_{E_{t,2}},
\end{aligned}
$$

where

$$
E_{t,2} \le \Delta_{\Phi_{t+1}} \le 0.
$$

As

$$
\begin{aligned}
&\mathcal{D}_\mathcal{R}(a, b) - \mathcal{D}_\mathcal{R}(c, b) \\
&= \mathcal{R}(a) - \mathcal{R}(b) - \langle \nabla\mathcal{R}(b), a - b\rangle - \mathcal{R}(c) + \mathcal{R}(b) + \langle \nabla\mathcal{R}(b), c - b\rangle \\
&= \mathcal{R}(a) - \mathcal{R}(c) - \langle \nabla\mathcal{R}(b), a - c\rangle \le \langle \nabla\mathcal{R}(a) - \nabla\mathcal{R}(b), a - c\rangle,
\end{aligned}
$$

we have

$$
E_{t,1} \le \langle \nabla\mathcal{R}(\theta_{t+1}) - \nabla\mathcal{R}(\Phi_{t+1}(g_{t+1})), \theta_{t+1} - \Phi_{t+1}(\theta_t)\rangle \le 4M\|\theta_{t+1} - \Phi_{t+1}(\theta_t)\|.
$$

Therefore,

$$
E_t \le D_t - D_{t+1} + 4M\|\theta_{t+1} - \Phi_{t+1}(\theta_t)\|,
$$

where $D_t = \mathcal{D}_{\mathcal{R}}(\theta_t, \Phi_t(g_t))$ and

$$\langle f_t, z_t \rangle - \langle \theta_t, z_t \rangle \le \frac{\eta}{2\sigma} \| z_t - M_t \|_\star^2 + \frac{1}{\eta} \left[D_t - D_{t+1} + 4M \| \theta_{t+1} - \Phi_{t+1}(\theta_t) \| \right].$$

Summing over $t = 1, \ldots, T$, we have

$$\sum_{t=1}^{T} \langle f_t, z_t \rangle - \sum_{t=1}^{T} \langle \theta_t, z_t \rangle \le \frac{\eta}{2\sigma} \sum_{t=1}^{T} \| z_t - M_t \|_\star^2 + \frac{1}{\eta} \left[D_{\max} + 4M \sum_{t=1}^{T} \| \theta_{t+1} - \Phi_{t+1}(\theta_t) \| \right].$$

$\square$

4.5.2 Kalman Filter

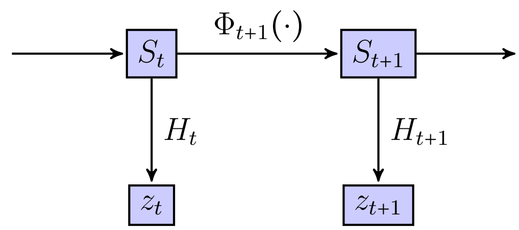


Figure 4.1: Kalman Filter Scheme

To demonstrate the Dynamic Mirror Descent with predictable sequence method, we consider the dynamic system with assumptions that (a) the underlying states $S_1, S_2, \ldots$ evolves as $S_t = \Phi_t(S_{t-1}) + w_t$, where $\Phi_t(S) = F_t S_{t-1} + B_t u_t$ and process noise $w_t \sim \mathcal{N}(0, Q_t)$, and (b) the observation z_t depends on the current state S_t as $z_t = H_t S_t + v_t$, where noise $v_t \sim \mathcal{N}(0, R_t)$.

If the model is accurate, the Kalman filter [23], which predicts

$$\hat{S}_{t|t-1} = \Phi_t(\hat{S}_{t-1|t-1}), \ P_{t|t-1} = F_t P_{t-1|t-1} F_t^T + Q_t$$

and updates

$$K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R_t)^{-1}, \ \hat{S}_{t|t} = \hat{S}_{t|t-1} + K_t(z_t - H_t \hat{S}_{t|t-1}), \ P_{t|t} = (I - K_t H_t) P_{t|t-1},$$

acts as an efficient and statistically optimal estimator of underlying states.

If these assumptions of underlying states and observations do not hold, we combine the Kalman filter with our framework. We choose the linear loss function $\langle f_t, z_t \rangle$ on our prediction of the underlying states f_t and the true observation z_t. As the optimal estimation of the coming observation z_{t+1} is $H_{t+1} \hat{S}_{t+1|t}$ after observing $z_1, \ldots, z_t$, it leads to the update equations

$$g_{t+1} = \arg\min_{g \in \mathcal{F}} \ \eta \langle g, z_t \rangle + \mathcal{D}_{\mathcal{R}}(g, \Phi_t(g_t))$$

and

$$f_{t+1} = \arg\min_{f \in \mathcal{F}} \eta \langle f, H_{t+1} \hat{S}_{t+1|t} \rangle + \mathcal{D}_{\mathcal{R}}(f, \Phi_{t+1}(g_{t+1})).$$

According to Theorem 4.5.1, Algorithm 9 yields

$$\sum_{t=1}^{T} \langle f_t, z_t \rangle \le \sum_{t=1}^{T} \langle \theta_t, z_t \rangle + \frac{\eta}{2\sigma} \sum_{t=1}^{T} \| z_t - H_t \hat{S}_{t|t-1} \|_\star^2 + \frac{1}{\eta} \left[D_{\max} + 4M \sum_{t=1}^{T} \| \theta_{t+1} - \Phi_{t+1}(\theta_t) \| \right]$$

where $D_{\max} = \max_{f, f' \in \mathcal{F}} \mathcal{D}_{\mathcal{R}}(f, f')$ and $M = \frac{1}{2} \max_{f \in \mathcal{F}} \| \nabla \mathcal{R}(f) \|$. If the model is accurate, the expected value of the square of the magnitude of the vector,

$$\mathbb{E} \left[\| z_t - H_t \hat{S}_{t|t-1} \|^2 \right],$$

is the trace of the covariance matrix $H_t P_{t|t-1} H_t^T + R_t$, i.e. the sum of the trace of matrix $H_t P_{t|t-1} H_t^T$ and the trace of matrix R_t.

Chapter 5

Shannon Entropy Over Approximate Entropy: an adaptive regularity measure

Approximate Entropy, as an approximation of Kolmogrov-Sinai Entropy, is the widely accepted method to quantify the regularity in data, especially medical data. However, it quantifies the regularity only up to predetermined order, while real data demand a much higher order. In this chapter, we demonstrate the connection between Approximate Entropy and Shannon Entropy. Based on that, we propose the adaptive Shannon Entropy approximation methods (e.g., Lempel-Ziv sliding window method) as an alternative approach to quantify the regularity of data. The new approach has the advantage of adaptively choosing the order of regularity to analyze based on the data. Later, we compare the results of Lempel-Ziv sliding window method with Approximation Entropy on the electroencephalography (EEG) data to measure the depth of anesthesia. The Lempel-Ziv sliding window method yields more accurate results, especially for low entropy data.

5.1 Introduction

Processed electroencephalographic (EEG) data is commonly used to quantify the depth of anesthetic hypnosis both clinically and in a laboratory setting, collapsing the complex patterns of the raw EEG signal into one (or a few) useful score(s). These scores help to achieve clinical goals of adequate hypnosis as well as to evaluate the results of experimental intervention on depth of anesthesia. Especially, Approximate Entropy [14] is a widely accepted method to quantify the regularity in EEG data.

Approximate Entropy arises as an approximation of Kolmogrov-Sinai Entropy, and [14] shows that the limit of Approximate Entropy is Kolmogrov-Sinai Entropy on the condition that (a) we have enough data, (b) the order of the regularity m

goes to infinity and (c) threshold r goes to zero. However, we never have the chance to have enough data and only compute one combination of parameters m and r on real data. It is an interesting problem to choose the order of the regularity m.

Instead of deriving the adaptive entropy estimation algorithm directly, we make connection between Approximate Entropy and Shannon Entropy. Then, we borrow the adaptive Shannon Entropy estimator and show the advantage comparing to Approximate Entropy.

5.2 Approximate Entropy

We start by defining Approximate Entropy.

Definition 5.2.1. Given a time series $u_1, u_2, \dots, u_N$, fix the order of the regularity $m \in \mathbb{N}^+$ and threshold $r > 0$, and form a sequence of vectors $x_i \in \mathbb{R}^m$, defined by

$$x_i = (u_i, u_{i+1}, \dots, u_{i+m-1}).$$

For each i, $1 \le i \le N - m + 1$ define

$$C_i^m(r) = \frac{\#\{j : d(x_i, x_j) \le r\}}{N - m + 1}.$$

We define $d(x_i, x_j)$, the distance between two vectors x_i and x_j as

$$d(x_i, x_j) = \max_{k=1,2,\dots,m} |u_{i+k-1} - u_{j+k-1}|.$$

Define

$$\Phi^m(r) = \frac{\sum_{i=1}^{N-m+1} \log C_i^m(r)}{N - m + 1},$$

then the Approximate Entropy is

$$\mathrm{ApEn}(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r).$$

Heuristically, ApEn quantifies the (logarithmic) likelihood that runs of patterns that are close remain close on next incremental comparisons.

Example 5.2.2. Suppose the time series is a repeat of sequence $(1,2,3)$, i.e. $u = (1,2,3,1,2,3,1,2,3,\dots)$, we first form a sequence of vectors

$$x_1 = (1,2,3), x_2 = (2,3,1), x_3 = (3,1,2), \dots \text{ when } m = 3$$

and

$$x_1 = (1,2), x_2 = (2,3), x_3 = (3,1), \dots \text{ when } m = 2.$$

Then, for threshold $r = 0$, define the percentage of the each vector happens in the sequence of vectors as

$$C_1^3 = \frac{1}{3}, C_2^3 = \frac{1}{3}, C_3^3 = \frac{1}{3}, \dots$$

and

$$C_1^2 = \frac{1}{3}, C_2^2 = \frac{1}{3}, C_3^2 = \frac{1}{3}, \ldots.$$

Furthermore, define the Φ function, which has a

$$\Phi^2(0) = \log\frac{1}{3}, \Phi^3(0) = \log\frac{1}{3}$$

Finally, we get

$$\mathrm{ApEn}(2, 0, N) = 0$$

5.3 Shannon Entropy and Entropy Rate

Here, we present several useful definitions related to entropy [6].

Definition 5.3.1. The Shannon Entropy of a random variable X with a probability mass function $p(x)$ is defined by

$$H(X) = -\sum_x p(x)\log_2 p(x)$$

we use logarithms to base 2 and without specific note, log in this article is to base 2.

The entropy of X can be interpreted as the expected value of the random variable $\log\frac{1}{p(X)}$, where X is drawn according to probability mass function $p(x)$. Thus,

$$H(X) = \mathbb{E}\log\frac{1}{p(X)}$$

Definition 5.3.2. The entropy rate of a stochastic process $\{X_i, i = 1, 2, \ldots\}$ is defined by

$$H(\mathcal{X}) = \lim_{n\to\infty}\frac{1}{n}H(X_1, X_2, \ldots, X_n)$$

where the limit exists.

If we have a sequence of n random variables, entropy rate measures how the entropy of the sequence grows with n. For example,

Example 5.3.3. $X_1, X_2, \ldots$ are i.i.d. random variables, then

$$H(\mathcal{X}) = \lim_{n\to\infty}\frac{H(X_1, X_2, \ldots, X_n)}{n} = \lim_{n\to\infty}\frac{nH(X_1)}{n} = H(X_1)$$

which is what one will expect for the entropy rate per symbol.

Example 5.3.4. The entropy rate of a m^{th} order Markov Chain is $H(X_{m+1}|X_1^m)$.

we can also define a related quantity for entropy rate

$$H'(\mathcal{X}) = \lim_{n\to\infty}H(X_n|X_{n-1}, \ldots, X_1)$$

when the limit exists. This is the conditional entropy of the last random variable given the past. The previous one is the per symbol entropy of the n random variables.

5.4 Approximate Entropy is equvilent to Conditional Entropy

In this section, we make the connection between Approximate Entropy and Shannon Entropy. First, we consider the case when the date are discrete.

5.4.1 For Discrete Data

Let us assume that the data are discrete, $u_i \in \mathcal{X} = \{1, 2, \ldots, |\mathcal{X}|\}$. There are maximum $|\mathcal{X}|^m$ different possible sequences with the length m, and we compact them in a set $\mathcal{X}^m$. If we define the frequency function $\hat{p}(X_1^m)$ on the set of X_1^m, then

$$C_i^m(r) = \hat{p}(X_1^m = x_i).$$

Therefore,

$$\begin{aligned}
\Phi^m(r) &= \sum_{i=1}^{N-m+1} \frac{\log C_i^m(r)}{N-m+1} \\
&= \sum_{i=1}^{N-m+1} \frac{\log \hat{p}(X_1^m = x_i)}{N-m+1} \\
&= \hat{\mathbb{E}}_{X_1^m} \log \hat{p}(X_1^m) = -\hat{H}(X_1^m)
\end{aligned}$$

Finally, we get

$$\mathrm{ApEn}(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r) = \hat{H}(X_1^{m+1}) - \hat{H}(X_1^m) = \hat{H}(X_{m+1}|X_1^m)$$

It is the conditional entropy of one variable given the previous m variables. It tells us how well we can predict one variable given m preceding variables. Let us present two examples to illustrate it.

Example 5.4.1. Consider the case that elements in the sequence are independent and identical distributed. Then,

$$\lim_{N \to +\infty} ApEn(m, r, N) = \lim_{N \to +\infty} \hat{H}(X_{m+1}|X_1^m) = H(X_1)$$

i.e. the limit of ApEn as N goes to infinity is the Shannon Entropy of the sequence. At the same time, it can also be interpreted as the entropy rate of the sequence.

Example 5.4.2. Consider the case that the sequence is a m^{th} order Markov chain. The limit of Approximate Entropy is indeed the entropy rate of the sequence. It is also true for lower order Markov Chain.

5.4.2 For Continuous Data

Theorem 5.4.3. *For an i.i.d. process with the density function $f(x)$, for any m,*

$$ApEn(m,r) = -\int f(y)\log\left(\int_{z=y-r}^{y+r} f(z)dz\right)dy$$

Proof. Given a time-series $u(1), u(2), \dots, u(N)$, fix m (positive integer) and r (positive real), and form a sequence of vectors $x(i)$ in $\mathcal{R}^m$, define by $x(i) = [u(i), u(i+1), \dots, u(i+m-1)]$. For each i, $1 \le i \le N-m+1$,

$$\begin{aligned}C_i^m(r) =& \Pr(x(j) : d[x(i), x(j)] \le r)\\ =& \Pr(u(j) : |u(i), u(j)| \le r) \cdot \Pr(u(j) : |u(i+1), u(j+1)| \le r)\\ &\cdots \Pr(u(j) : |u(i+m-1), u(j+m-1)| \le r)\\ =& \int_{u(i)-r}^{u(i)+r} f(z)dz \cdot \int_{u(i+1)-r}^{u(i+1)+r} f(z)dz \cdots \int_{u(i+m-1)-r}^{u(i+m-1)+r} f(z)dz\end{aligned}$$

Then, get the limit of $\Phi^m(r)$

$$\begin{aligned}\lim_{n\to\infty} \Phi^m(r) &= \lim_{n\to\infty} (N-m+1)^{-1} \sum_{i=1}^{N-m+1} \log C_i^m(r)\\ &= \lim_{n\to\infty} (N-m+1)^{-1} \sum_{i=1}^{N-m+1} \sum_{k=1}^{m} \log \int_{u(i+k-1)-r}^{u(i+k-1)+r} f(z)dz\\ &= \sum_{k=1}^{m} \mathbf{E}_{u^\star} \log \int_{u^\star-r}^{u^\star+r} f(z)dz\\ &= m \int f(y)\log\left(\int_{y-r}^{y+r} f(z)dz\right)dy\end{aligned}$$

Finally, the Approximation Entropy is

$$\mathrm{ApEn}(m,r,N) = \Phi^m(r) - \Phi^{m+1}(r) = -\int f(y)\log\left(\int_{y-r}^{y+r} f(z)dz\right)dy.$$

□

By the mean value theorem,

$$\mathrm{ApEn}(m,r) = -\int f(y)\log\left(2rf(y^\star)\right)dy$$

for $y-r < y^\star < y+r$ for all y. Therefore,

$$\mathrm{ApEn}(m,r) + \log(2r) \approx -\int f(y)\log f(y)dy.$$

For a general case, we can find that sequences with length m follows a probability density function p_m and sequences with length $m+1$ follows a probability density function p_{m+1}.

Theorem 5.4.4.

$$ApEn(m,r,N) = \Phi^m(r) - \Phi^{m+1}(r) \approx H(X_{m+1}|X_1^m) - \log(2r)$$

Proof. We assume the joint distribution of length m sequence is p_m, first,

$$\begin{aligned} C_i^m(r) &= \int_{u(i)-r}^{u(i)+r} \int_{u(i+1)-r}^{u(i+1)+r} \cdots \int_{u(i+m-1)-r}^{u(i+m-1)+r} p_m(y_i, y_{i+1}, \ldots, y_{i+m-1}) dy_{i+m-1} \ldots dy_{i+1} dy_i \\ &= (2r)^m p_m(y_i^\star, y_{i+1}^\star, \ldots, y_{i+m-1}^\star) \\ &\approx (2r)^m p_m(y_i, y_{i+1}, \ldots, y_{i+m-1}) \end{aligned}$$

Then,

$$\begin{aligned} \Phi^m(r) &= \mathbf{E}_i \log C_i^m(r) \\ &= \int\int \cdots \int p_m(y_1, y_2, \cdots, y_m) \log\left[(2r)^m p_m(y_i, y_{i+1}, \ldots, y_{i+m-1})\right] dy_1 dy_2 \ldots dy_m \\ &= \int\int \cdots \int p_m(y_1, y_2, \cdots, y_m) \left[m \log 2r + \log p_m(y_i, y_{i+1}, \ldots, y_{i+m-1})\right] dy_1 dy_2 \ldots dy_m \\ &= m \log 2r - H(X_1^m) \end{aligned}$$

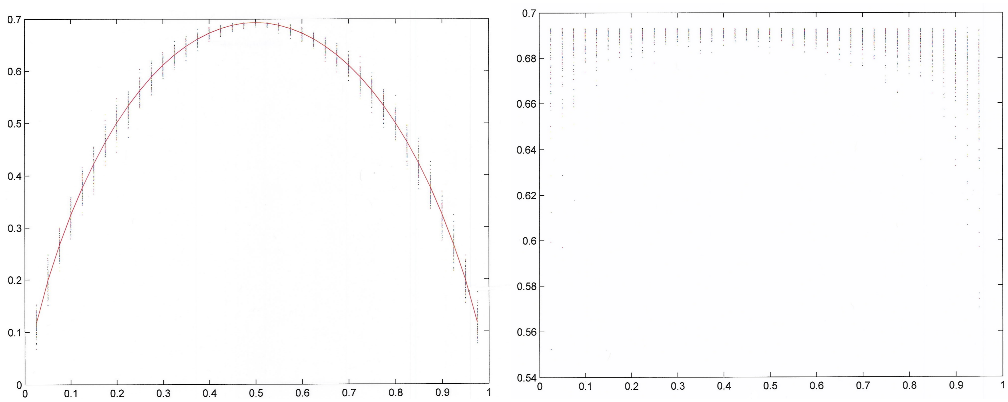
Finally, we get that Approximation Entropy is

$$\begin{aligned} \text{ApEn}(m,r,N) &= \Phi^m(r) - \Phi^{m+1}(r) \\ &\approx \left[m \log 2r - H(X_1^m)\right] - \left[(m+1)\log(2r) - H(X_1^{m+1})\right] \\ &= H(X_1^{m+1}) - H(X_1^m) - \log 2r \\ &= H(X_{m+1}|X_1^m) - \log 2r \end{aligned}$$

□

For Markov Chain with less than or equal to m^{th} order, Approximate Entropy is the entropy rate minus a constant and acts perfect as a measure for regularity. However, Approximate Entropy can not capture the regularity of higher order. We will show it by the following examples and set parameter $m = 2$. Figure 5.1a shows the second order Markov model where $X_{n+1} = X_n \oplus X_{n-1} \oplus e$ and the noise $e = 1$ w.p. p and $e = 0$ w.p. $1-p$. We can get the entropy rate $H(X_{n+1}|X_n, X_{n-1}) = H(e) = H(p)$ and the Approximate Entropy captures the regularity precisely. However, Figure 5.1b shows the third order Markov model where $X_{n+1} = X_n \oplus X_{n-1} \oplus X_{n-2} \oplus e$ and the noise $e = 1$ w.p. p and $e = 0$ w.p. $1-p$. We can get the entropy rate $H(X_{n+1}|X_{1:n}) = H(p)$ and the Approximate Entropy can not capture the regularity.

So far, we demonstrate the connection between Approximate Entropy and Shannon Entropy. Approximate Entropy quantifies the regularity only up to the prefixed order, while real data may carry much more. Shannon Entropy approximation methods (e.g., Lempel-Ziv sliding window method [12]) is an alternative approach to quantify the regularity of data.



(a) Second Order Markov Model (b) Third Order Markov Model

Figure 5.1: Approximate Entropy of Markov Model

5.5 Adaptive Entropy Estimation

We take Lempel-Ziv sliding window method as an example. Define L_i as the length of the shortest substring $X_{i:i+k-1}$ starting at position i that does not appear as a contiguous substring of the previous string. Use the follow term to quantify the regularity

$$\left(\frac{1}{n}\sum_{i=2}^{n}\frac{L_i}{\log i}\right)^{-1}$$

It adaptively chooses the order of regularity to analyze based on the data, and we show the comparison in Figure 5.2.

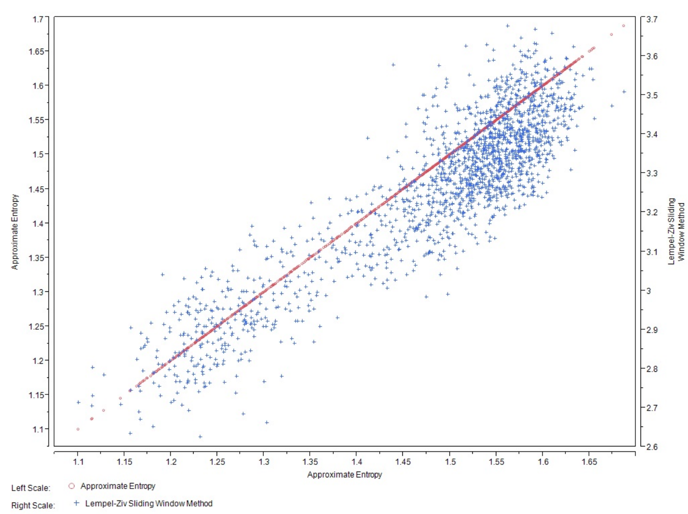


Figure 5.2: Approximate Entropy v.s. Lempel Ziv

Bibliography

[1] S. Ben-David, D. Pál, and S. Shalev-Shwartz. Agnostic online learning. In *Proceedings of the 22nd Annual Conference on Learning Theory*, 2009.

[2] N. Cesa-Bianchi and G. Lugosi. On prediction of individual sequences. *The Annals of Statistics*, 27(6):1865–1895, 1999.

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

[4] D.Z. Chen, O. Daescu, Y. Dai, N. Katoh, X. Wu, and J. Xu. Efficient algorithms and implementations for optimizing the sum of linear fractional functions, with applications. *Journal of Combinatorial Optimization*, 9(1):69–90, 2005.

[5] C. Chiang, T. Yang, C. Lee, M. Mahdavi, C. Lu, R. Jin, and S. Zhu. Online optimization with gradual variations. In *Proceedings of the 25th Annual Conference on Learning Theory*, 2012.

[6] T.M. Cover and J.A. Thomas. *Elements of Information Theory.* Wiley-Interscience, New York, NY, 1991.

[7] M. Feder, N. Merhav, and M. Gutman. Universal prediction of individual sequences. *IEEE Transactions on Information Theory*, 38(4):1258–1270, 1992.

[8] D.P. Foster, A. Rakhlin, K. Sridharan, and A. Tewari. Complexity-based approach to calibration with checking rules. In *Proceedings of the 24th Annual Conference on Learning Theory*, 2011.

[9] E.C. Hall and R.M. Willett. Online optimization in dynamic environments. 2013.

[10] S.M. Kakade, K. Sridharan, and A. Tewari. On the complexity of linear prediction: Risk bounds, margin bounds, and regularization. In *Advances in Neural Information Processing Systems*, pages 793–800, 2008.

[11] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer-Verlag, Berlin, 2nd edition, 1991.

[12] I. Kontoyiannis, P.H. Algoet, Yu.M. Suhov, and A.J. Wyner. Nonparametric entropy estimation for stationary processes and random fields, with applications to english text. *IEEE Transactions on Information Theory*, pages 1319–1327, 1998.

[13] N. Merhav and M. Feder. Universal prediction. *IEEE Transactions on Information Theory*, 44:2124–2147, 1998.

[14] S.M. Pincus, I.M. Gladstone, and R.A. Ehrenkranz. A regularity statistic for medical data analysis. *Journal of Clinical Monitoring and Computing*, pages 335–345, 1991.

[15] A. Rakhlin, O. Shamir, and K. Sridharan. Relax and randomize : From value to algorithms. In *Advances in Neural Information Processing Systems*, 2012.

[16] A. Rakhlin and K. Sridharan. `http://www-stat.wharton.upenn.edu/~rakhlin/courses/stat928/stat928_notes.pdf`, 2012.

[17] A. Rakhlin and K. Sridharan. Online learning with predictable sequences. In *Proceedings of the 26th Annual Conference on Learning Theory*, 2013.

[18] A. Rakhlin, K. Sridharan, and A. Tewari. Online learning: Random averages, combinatorial parameters, and learnability. In *Advances in Neural Information Processing Systems*, pages 1984–1992, 2010.

[19] A. Rakhlin, K. Sridharan, and A. Tewari. Online learning: Stochastic, constrained, and smoothed adversaries. In *Advances in Neural Information Processing Systems*, pages 1764–1772, 2011.

[20] S. Shalev-Shwartz. *Online Learning: Theory, Algorithms, and Applications*. PhD thesis, The Hebrew University of Jerusalem, 2007.

[21] N. Srebro, K. Sridharan, and A. Tewari. On the universality of online mirror descent. In *Advances in Neural Information Processing Systems*, 2011.

[22] V. Vovk. Competitive on-line linear regression. In *Advances in Neural Information Processing Systems*, pages 364–370, 1998.

[23] G. Welch and G. Bishop. An introduction to the kalman filter. 1995.