Algorithmic Learning from Financial Predictions

Pongphat Taptagaporn

A thesis submitted to the Department of Mathematics of the London School of Economics for the degree of ***Doctor of Philosophy****, London, April 2017.*

"The advantage scientists bring into the game is not their mathematical or computational skills than their ability to think scientifically. They are less likely to accept an apparent winning strategy that might be a mere statistical fluke."

−James Harris Simons

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

Statement of conjoint work:

I confirm that some of the content of Chapter 4 of this thesis was presented as co-authored work with my supervisor, Dr Tugkan Batu, at the 27th International Conference on Algorithmic Learning Theory, published in the Lecture Notes in Artificial Intelligence, Volume 9925, pp 288-302, Springer 2016. Citation [BT16].

The contents of Chapter 5 of this thesis is also currently being reviewed for conference publication as co-authored work with Dr Tugkan Batu. Citation [BT17].

Abstract

We study how financial predictions can be used in learning algorithms for problems such as portfolio selection and derivatives pricing, from the perspective of minimizing regret; the worst-case loss (across all possible price paths) against some optimal benchmark model with superior information. Unlike most studies in financial mathematics, we do not make any underlying assumptions beyond the existence of such predictions, so our results are robust in the model-free sense.

This thesis consists of three main ideas:

1. Study a portfolio selection model that competes with an optimal static trading strategy (the best fixed strategy in hindsight) using predictions of the optimal portfolio allocation.

2. Study a portfolio selection model that competes (in probability) with an optimal dynamic trading strategy (the best greedy strategy in hindsight) using price predictions of each asset in the portfolio.

3. Derive robust derivative pricing bounds for vanilla options and various exotic derivatives based on price predictions of the underlying asset(s).

This work is focused on the mathematical analysis of these models, using techniques from theoretical algorithmic and statistical learning.

Acknowledgement

First and foremost I would like to sincerely thank Dr Tugkan Batu for the supervision throughout my PhD program. His guidance and constant support have been invaluable. I would also like to thank my second supervisor, Prof Martin Anthony, for his wisdom and feedback on my research. I am grateful to the examiners, Dr Yuri Kalnishkan and Dr Varun Kanade, for their useful comments on this thesis.

I would like to thank everyone in the Mathematics department at the LSE for providing a pleasant and stimulating atmosphere to carry out academic research. It has been enjoyable and satisfying.

Finally, I gratefully acknowledge funding from the LSE's PhD scholarship that has given me the opportunity to explore interesting research, and funding from the Mathematics department for attending overseas conferences, among other things.

Contents

Chapter 1

Introduction

Much of the studies in financial mathematics can be broadly categorised into the $\mathbb{Q}$ and $\mathbb{P}$ worlds. The $\mathbb{Q}$ world is in the realms of derivative pricing (calibrating the fair price of securities based on market variables), while the $\mathbb{P}$ world studies portfolio management problems (estimating the statistically derived probability distribution of asset prices and constructing efficient portfolios). In both cases, much of the work often rely on assumptions of the underlying market dynamics. In particular, these results typically make statements of the form:

"If variable X has dynamics Y, then ...."

Perhaps the most famous examples are the Merton portfolio [Mer71] and Black-Scholes option pricing [BS73], whose results depend on the underlying asset price evolving as Geometric Brownian Motion (GBM), among other things. While these have been widely adopted by practitioners and received much success from industry, there has also been much criticism about the inconsistencies of these underlying assumptions to the observed behaviour from the financial market. For example, the existence of volatility smile in the foreign exchange options market [Hul06].

This work aims to provide robust (model-free) approach to these problems in the $\mathbb{Q}$ and $\mathbb{P}$ worlds. In particular, we will provide algorithms for portfolio selection and derivatives pricing that will work regardless of how market dynamics behave. Instead of making assumptions on the underlying market dynamics, we assume the existence of predictions of various parameters to assist in the decision-making process, for example, predictions on the future asset price (often referred to as “alpha” in the financial industry) or optimal portfolio distribution.

Without making any assumptions on the underlying market dynamics, it is impossible to say anything meaningful about the average-case performance of these algorithms. Therefore, the performance of these models are analyzed relative to some optimal benchmark adversary (typically with access to superior information). The performance of such algorithms (relative to the optimal benchmark) would then depend on the quality of the predictions received.

1.1 Literature Review

A new field emerged in the 1990’s that uses game theory and machine learning to design portfolio selection models that performs competitively without making any conjecture on the future. The first known paper in this field by Thomas Cover [Cov91] introduced a portfolio model that makes decision on the portfolio distribution among assets purely based on current and past information; it does not assume any prediction mechanism. Most interestingly, he was able to prove a worst-case performance guarantee (across all possible price paths), as compared to the wealth of the best fixed (static) strategy in hindsight, without making any assumption on how the price must evolve. In particular, it was shown that a regret of

$$\max_{x^T} \left(\log S_T^* - \log \hat{S}_T \right) = O(\log T)$$

is attainable where S_T^* is the wealth obtained from the best constant-rebalanced portfolio (CRP) in hindsight, and $\hat{S}_T$ is the wealth obtained from the competing portfolio model (Cover’s universal portfolio), over T discrete time steps. The maximum difference in log-wealth is taken over all possible price paths x^T. Note that a CRP is defined as an investment strategy with the restriction that it must maintain a fixed proportion of wealth in each of the assets throughout all time steps, performing any required rebalancing as to maintain these proportions as the asset prices change. The best CRP is then the best of such strategy that maximizes wealth over all T time steps.

The general idea of Cover's algorithm is to take a weighted combination of portfolios according to some prior distribution (hence "universal"), and run them independently, performing any required rebalancing as to maintain the CRP assumption across each of these constituent portfolios. The key to proving the regret bound is that by using a weighted combination of portfolios, the algorithm would have picked a sufficient number of sample portfolios "nearby" the best CRP, which will then perform similarly. The bad performing portfolios would then die off and the better performing ones would dominate, resulting in a wealth that is bounded against the best CRP.

Since then, there has been much follow up work and extensions to Cover's original portfolio model. For example, some research [CO96, BS10, BS11, KACS15] extended Cover's universal portfolio to compete against a stronger benchmark using a concept of "side information". This is where the adversary reveals a side information (say, an integer between 1 and y) and the CRP restriction is applied on each state separately. In particular, there is now y different CRPs that may be used, depending on the side information in that particular time step. The benchmark in this case is the best set of y CRP's that achieves the highest wealth, given the observed sequence of side information. However, the regret bound of this model assumes that y is finite and does not grow with T, meaning that sublinear regret bound does not hold if the benchmark model uses a different portfolio in every time step, i.e., the side information never repeats.

One possible extension when studying portfolio models is to include the presence of market friction in the form of transaction cost, to mimic the behaviour of modern order-driven market. This concept was introduced in the context of universal portfolios by Blum and Kalai [BK99], where they charged a fixed percentage of commission as a proportion of the traded volume.

Hazan et al. [HK09] showed an alternative regret bound for Cover's universal portfolio of $O(\log Q)$, where Q is the quadratic variation of the underlying assets (similar to the notion of volatility). However, one can realistically expect Q to grow with T, hence still has the dependence on T nevertheless.

Recall that Cover's original portfolio model (as well as most research in this field) assumes trading in discrete time steps. Freund [Fre09] demonstrated the extension of another similar portfolio selection algorithm from Chaudhuri et al. [CFH09] in the continuous-time setting by modelling the stochastic price process as an Itô process.

More recent efforts [CYL$^+$12, RS13] incorporated predictions into online learning problems. These work look at the more general case of convex loss functions, as compared to the log-wealth in the portfolio setting. Some other variants of the universal portfolio can be found in [AH06, AHKS06, Cov96, GW12, HAK07, KW99, OC96, SL05]. Most of these models are based on the idea of taking a weighted combination of CRPs over the set of all possible portfolio vectors, as in the case of Cover's universal portfolio. Therefore, it was natural in these settings to then compare the wealth to the best fixed strategy in hindsight.

In 2006, DeMarzo et al. [DKM06] showed that the regret of portfolio selection algorithms naturally give rise to an upper bound for options price in the model-free sense, by replicating the payoff of an option by the returns of the best performing asset in hindsight. However, their bound still depends on the volatility of the underlying asset (through the quadratic variation Q), much like in the Black–Scholes framework. In particular, they showed

$$C(K, T) \leq \Theta(\sqrt{Q}),$$

where $C(K, T)$ denotes the price of a call option (with strike K at expiry T), without making any additional assumptions on the underlying price process.

Follow up work from Gofer et al. [GM11a, GM11b, Gof14] extended this result to price various exotic derivatives in the model-free sense, and Abernethy et al. [AFW12, ABFW13] showed that this option price bound converges to that of Black-Scholes in the limit; as each time step increment $\rightarrow 0$, analogous to the continuous time setting).

1.2 Structure of the Thesis

First we will introduce the required notations and preliminary background in Chapter 2 that will be used throughout the thesis. Thereafter, this thesis comprise of three main parts as described below.

Chapter 3: Static Trading Strategy

Extend the results of Cover's universal portfolio algorithm to account for transaction cost in the setting with multiple side information. We prove that logarithmic regret against the best CRP with side information is attainable, and provide an efficient approximation algorithm to compute such portfolio. We also examine the improvements that can be achieved by introducing the notion of predictions of the optimal portfolio distribution.

Chapter 4: Dynamic Trading Strategy

We look beyond the restriction of the CRP from Chapter 3 to design a portfolio selection algorithm that competes with a stronger benchmark, the best greedy portfolio, in the stochastic setting. To do this, we make use of price predictions and prove that small expected regret (and variance of regret) is attainable subject to the quality of such predictions. We also study the case of incorporating transaction cost, and show that sub-linear regret is not attainable in this setting with non-zero transaction costs. The computation of these portfolios will be shown to reduce to a linear program.

Chapter 5: Derivatives Pricing

Traditional option pricing model such as Black-Scholes assume that the underlying asset follows a GBM. Alternatively, we derive a robust upper bound on options price (that does not make any assumption on the underlying asset price process) using the expected regret bound from the predictive trading strategy from Chapter 4. We first show a bound for pricing vanilla options, then extend this to a number popular exotic derivatives.

Chapter 2

Preliminaries

This chapter will provide the preliminary definitions and ideas that will be used throughout the thesis.

Consider the scenario where we have m assets available for trading over T discrete time steps. Define

$$x_t = (x_t(1), \cdots, x_t(m)) \in \mathbb{R}_+^m$$

as a vector of price relatives (also known elsewhere as "returns vector") at time step t, that is, $x_t(i)$ is the ratio of the true market price of asset i at time t and time $t-1$. For example, if asset i did not change in price at time t then $x_t(i) = 1$. This will be defined for $1 \leq t \leq T \in \mathbb{N}$, and use x^t to denote the price path up to time t,

$$x^t := (x_1, \cdots, x_t).$$

Define a portfolio vector at time t as

$$b_t = (b_t(1), \cdots, b_t(m)) \in \mathcal{B} = \{b_t \in \mathbb{R}_+^m : \sum_{i=1}^{m} b_t(i) = 1\}$$

where $b_t(i)$ is the proportion of the portfolio's total wealth allocated to asset i at time t. From time step $t-1$ to t, if we invest using portfolio b_t in our trading strategy then our wealth will change by a factor of

$$b_t \cdot x_t,$$

i.e., the dot product of the two m-dimensional vectors, representing the change in value of the portfolio. Over T time steps, a trading strategy is specified by the sequence of portfolio vectors

$$b^T := (b_1, \cdots, b_T)$$

and the total wealth becomes[1]

$$S_T(b^T) = \prod_{t=1}^{T} b_t x_t.$$

Broadly speaking, S_T is the product of the wealth change across all time steps $t \in [T]$. Note that S_T has hidden dependency on x^T; we omit this for notational convenience. Typically we may need to re-distribute wealth between assets as to obtain the chosen portfolio vector for the next time step. We will call this re-distribution of wealth process "re-balancing".

Similarly, for the trading strategies specified by $(\hat{b}_1, \ldots, \hat{b}_T)$ and $(b_1^*, \ldots, b_T^*)$, we will use $\hat{S}_T$ and S_T^*, respectively, to denote the wealth generated by the corresponding portfolios.

A constant-rebalanced portfolio (CRP) is defined as the subset of b^T with the additional constraint that the portfolio vector is the same throughout every time step, that is,

$$b_1 = \cdots = b_T.$$

Although the portfolio model investigated here has the restriction that all the wealth must be invested in one of the m assets (imposed by the condition that all of the individual asset wealths must sum up to one), this can be extended to a portfolio of $m + 1$ assets where the first m assets are as before, and the last one represents cash. Therefore, the returns x_t now has $m + 1$ dimension where the last element could represent risk-free interest rate (from the change in value of the riskless asset), analogous to much of the work in financial mathematics.

[1] The notations $b_t x_t$ is used as a short-hand for vector dot product, and we may occasionally refer to S_T instead of $S_T(b^T)$ for notational convenience.

2.1 Robust Performance Metric

There are a number of methods that have been used to measure performance of portfolio selection models in literature, most commonly wealth, or some risk-adjusted notion of wealth. For example, modern portfolio theory [Mar52] gives a framework to optimize the mean-variance of a portfolio.

Recall that our models do not make any assumption on the price movement of the assets. Of course, with no additional assumption we cannot have any guarantees regarding the future wealth. For example, if the price of all of the assets at a particular time decreases by 10%, our wealth will necessary decrease by 10%, regardless of the portfolio distribution.

Therefore, we can only compare the wealth obtained by the portfolio as compared to that of another portfolio. A common metric that has been used for this purpose is called *regret*.

To understand the notion of regret, first assume that we know the wealth S_T (over T time steps) of our portfolio model and the wealth S_T^* of the benchmark model (typically representing some notion of optimality). We wish to evaluate the growth-rate of wealth, denoted by W_T, such that $S_T = e^{W_T}$, and similarly W_T^* such that $S_T^* = e^{W_T^*}$. The worst-case difference (over all possible price paths) between these exponential growth is called *regret*, namely

$$R := \max_{x^T} \left(W_T^* - W_T \right) = \max_{x^T} \left(\log S_T^* - \log S_T \right).$$

This can also be viewed as the worst-case guarantee of the difference between the logarithmic wealth factors of the benchmark model and the competing portfolio model. Suppose now that we do not know the exact values of S_T and S_T^*, but may be able to derive some bounds on them. The problem of minimizing R is known as *regret minimization*. Intuitively, a smaller regret bound implies that our model is closer to the benchmark model (in the worst-case), thus closer to some notion of optimality.

Furthermore, to put the wealth ratio in context, we are generally interested in the exponential growth rate per time step. This can be written as

$$\frac{R}{T} = \frac{1}{T} \max_{x^T} \left(\log S_T^* - \log S_T \right).$$

We say that the portfolio model has sublinear regret if this value is $o(1)$ in T, or equivalently, $R = o(T)$. Intuitively, this means that the exponential growth rate of wealth of the competing portfolio model converges to that of the benchmark model, as the number of time steps grow large, $T \to \infty$.

For probabilistic portfolio selection models where the trading strategy depends on some random choices (for example, random predictions), the regret also becomes probabilistic. Then it seems natural to study the statistical properties of the regret such as expected regret

$$\mathbb{E}[R] := \mathbb{E}\left[\max_{x^T} \left(\log S_T^* - \log S_T\right)\right],$$

and the variance of regret

$$\mathrm{Var}[R] := \mathrm{Var}\left[\max_{x^T} \left(\log S_T^* - \log S_T\right)\right].$$

Academic studies in another related problem, the multi-armed bandit [BC12], had also considered the notion of pseudo-regret,

$$\bar{R} := \max_{x^T} \mathbb{E}\left[\log S_T^* - \log S_T\right],$$

While the notion of pseudo-regret is weaker than the expected regret with $\bar{R} \leq \mathbb{E}[R]$, bounds on the pseudo-regret imply bounds on the expected regret.

Much of the studies in online learning revolves around proving bounds for the regret (or its various properties), although often differ in context.

2.2 Transaction Cost

Any realistic trading strategy would have to consider the effect of transaction costs on its profitability. As seen across most financial exchanges, market makers, and brokers worldwide, the buying price of an asset is generally higher than its selling price due to bid-ask spread.

The concept of transaction costs was first introduced into the study of universal portfolios by Blum and Kalai [BK99], wherein their model charge a fixed percentage commission (of the traded size) on the purchase, but not on the sale, of assets. This is equivalent to charging commission on the purchase and sale of assets equally (as in modern limit-order markets), as the wealth from any asset we sold will have to be used to purchase another asset (whether it be kept in riskless cash, or another risky asset). We will use the same model here, though the choice of model doesn't significantly affect our results.

Given portfolio vectors $b_{t-1}, b_t \in \mathcal{B}$ and returns vector x_{t-1}, we want to re-balance from the vector

$$b'_{t-1} := b_{t-1} \cdot x_{t-1} \in \mathbb{R}^m$$

to

$$b_t \in \mathcal{B} \subset \mathbb{R}^m.$$

Given a transaction cost factor

$$c \in [0, 1]$$

indicating the proportion of cost to be paid from the value of assets purchased, the proportion of wealth retained after rebalancing can be expressed recursively as

$$\theta := \theta(b_{t-1}, b_t, x_{t-1}) = 1 - c \sum_{i:\beta_i > 0} \beta_i,$$

where

$$\beta_i = \theta b_t(i) - b_{t-1}(i) \cdot x_{t-1}(i) = \theta b_t(i) - b'_{t-1}(i)$$

indicates the quantity of asset i that needs to be bought or sold, depending on its sign. Intuitively, θ represents the proportion of the total wealth left after rebalancing. In the worst case, the market value of b' is at least $1 - c$ of the market value of b after rebalancing. In particular, rebalancing a portfolio will always retain at least $1 - c$ proportion of its wealth. Note that $c = 0$ means that no transaction cost is charges and hence can be ignored.

We denote by $\theta(b_{t-1}, b_t, x_{t-1})$ the multiplicative factor of decrease in wealth due to rebalancing from portfolio b_{t-1} (after observing the price change x_{t-1}) to portfolio b_t. Then, we can define the wealth of a portfolio model (with transaction cost) as

$$S_T = \prod_{t=1}^{T} b_t x_t \theta(b_{t-1}, b_t, x_{t-1}).$$

As a convention, we assume that there are no transaction costs associated with the initial positioning before the first time step: that is, $b_0 := b_1$, $x_0 = (1, \ldots, 1)$, and, thus, $\theta(b_0, b_1, x_0) = 1$.

Broadly speaking, S_T is the product of the wealth change across all time steps $t \in [T]$, where, at each step, we first pay a factor of $\theta(b_{t-1}, b_t, x_{t-1})$ in transaction cost for re-balancing b_{t-1} to b_t, and then experience a change $b_t x_t$ in wealth, once the price change is observed.

The transaction cost factor θ can be computed efficiently using either random sampling or a linear program. These will be demonstrated as part of the portfolio computation in Section 3.6 and Section 4.8, respectively.

Chapter 3

Static Trading Strategy

In this chapter, we present an extension of Cover's universal portfolio, incorporating the presence of transaction costs [BK99] in the setting with multiple discrete side information states [CO96]. We explore the case where we have a prediction mechanism that is able to indicate approximately the best portfolio distribution in some future time steps, and show that we are able to derive a portfolio selection algorithm that is competitive with the best static trading strategy.

First we define the static trading strategy that will be used as the benchmark model. Recall that a constant-rebalanced portfolio (CRP) is an investment strategy where at every time step invest its wealth according to some portfolio distribution, say b. Over T time steps, the wealth achieved by the CRP strategy then becomes

$$S_T(b) = \prod_{t=1}^{T} bx_t,$$

a special case of the (unrestricted) general strategy. We denote the best CRP in hindsight (over T time steps) by the best portfolio distribution b^* that maximizes precisely this wealth above. Formally,

$$b^* = \arg\max_{b \in \mathcal{B}} S_T(b),$$

and denote the corresponding wealth of b^* by

$$S_T^* = \max_{b \in \mathcal{B}} S_T(b).$$

3.1 Cover's Universal Portfolio

Now we will formally describe Cover's universal portfolio [Cov91]. Let μ be a prior (initial) distribution function over the space $\mathcal{B}$ of m assets with the standard condition that

$$\int_{\mathcal{B}} d\mu(b) = 1,$$

for $b \in \mathcal{B}$. A μ-weighted universal portfolio is, roughly speaking, a (measure-theoretic) weighted combination of many different CRPs, with the initial weighting rule according to the continuous prior distribution μ. Hence, the measure of wealth invested in the CRP denoted by b is $\mu(b)$, and so the wealth due to b up to time $t-1$ is $S_{t-1}(b)\mu(b)$, yielding a total wealth (across the whole spectrum of CRPs) of

$$\int_{\mathcal{B}} S_{t-1}(b)\, d\mu(b).$$

In Cover's definition of universal portfolio, the measure of each CRP are weighted by the amount of wealth they have historically generated. This is equivalent to picking many CRP's initially according to prior distribution μ, then letting them run independently for the entire game (from $t = 1$ to T). Conceptually, no re-balancing occurs across different CRP's, but they are each re-balanced independently of each other. Taking an integral over the continuous space of all portfolio vectors $\mathcal{B}$, we get the wealth distribution for the universal portfolio,

$$\int_{\mathcal{B}} b S_{t-1}(b)\, d\mu(b).$$

Dividing this by the total amount of wealth currently available at time step $t-1$ (so that the non-negative weights in the portfolio vector sums to 1), we get Cover's universal portfolio.

Definition 1 ([Cov91]) *The μ-weighted universal portfolio at time t is*

$$\hat{b}_t = \frac{\int_{\mathcal{B}} bS_{t-1}(b)\, d\mu(b)}{\int_{\mathcal{B}} S_{t-1}(b)\, d\mu(b)}.$$

This concept of increasing the wealth around the best performing portfolio is otherwise known as "experts" in the more general context. Note that this setting did not make any assumption on the asset prices, nor uses any concept of predictions. Thus it came as a surprise to many when Cover was able to prove a performance guarantee, although the regret is taken against a more restricted class of strategy, the best CRP in hindsight.

$$\max_{x^T} \left(\log S_T^* - \log \hat{S}_T \right) = O(\log T).$$

3.2 Side Information

A further extension to this idea is to loosen the constraints of the CRP to incorporate discrete state-space side information [CO96]. Formally, suppose each time step $t \in [T]$ has an associated label y_t known as side information, where each $y_t \in \mathcal{Y} = [k]$. In general, we expect the side information to be somewhat useful; a random sequence of side information is not of much use. For practical interpretation, it could represent a number of things such as market regime, technical indicators, trading signals, etc. Note that it is also not asset-specific (since we have one piece of information for each time step, but not for each asset), and y_t could possibly depend on x_t.

In this setting, the portfolio distribution $b \in \mathcal{B}^k$ includes k different portfolio vectors from $\mathcal{B}$, corresponding to each of the k side information states. Note that $b \times \cdots \times b \in \mathcal{B}^k$ is the Cartesian product of k identical portfolio vectors $b \in \mathcal{B}$. At time t, the trader will invest using the i^{th} portfolio distribution whenever the side information indicates $y_t = i$. We can similarly extend the notion of CRP to account for side information by having k different CRPs corresponding to each of the k side information states. Formally,

$$b^* = \arg \max_{(b_1, \cdots, b_k) \in \mathcal{B}^k} S_T(b_{y_1}, \cdots, b_{y_T}),$$

and can define S_T^* similarly. If we wish to further optimise the side information state, in other words, find the best sequence of discrete states $y = (y_1, \cdots, y_T)$ that maximizes wealth, this can be calculated as

$$y = \arg \max_{(y_1,\cdots,y_T)\in[k]^T} \max_{(b_1,\cdots,b_k)\in\mathcal{B}^k} S_T(b_{y_1}, \cdots, b_{y_T}).$$

The universal portfolio can be extended to incorporate side information as shown below.

Definition 2 ([CO96]) *The μ-weighted universal portfolio (with side information of k states) at time t is*

$$\hat{b}_t(y) = \frac{\int_{\mathcal{B}} bS_{t-1}(b|y)\, d\mu(b)}{\int_{\mathcal{B}} S_{t-1}(b|y)\, d\mu(b)},$$

where $S_i(b|y)$ is the wealth obtained by the CRP b_y along the subsequence $\{j \le t : y_j = y\}$, in other words, the contribution from state y based on past performance.

In this case, it was shown that the regret (against the best CRP with side information) can be bounded as $O(k \log T)$. This is problematic if k is large, for example if each time steps has unique side information then $k = T$, and the regret becomes at least linear in T. Therefore, we generally assume k to be relatively small compared to T, e.g., $k \in o(T)$.

3.3 Our Extension

Suppose we have the prediction $\tilde{b} \in \mathcal{B}^k$ that is a 'good' approximation of the best CRP (with side information) in hindsight $b^* \in \mathcal{B}^k$. (The specific definition of 'good' will be discussed in more details later). We could potentially make use of this information to give us a performance advantage. First we need some definitions. Given $b \in \mathcal{B}, \tilde{b} \in \mathcal{B}^k$ and $0 \le \epsilon \le 1$, define

$$\tilde{b}_\epsilon(b) = (1-\epsilon)\tilde{b} + \epsilon(b \times \cdots \times b) \in \mathcal{B}^k,$$

and thus

$$\tilde{b}_\epsilon(b, y) = (1-\epsilon)\tilde{b}(y) + \epsilon b \in \mathcal{B},$$

i.e., $\tilde{b}_\epsilon(b, y) \in \mathcal{B}$ is the portfolio vector of $\tilde{b}_\epsilon(b) \in \mathcal{B}^k$ when we observed state $y \in [k]$. Also, $\tilde{b}_\epsilon(b, y)$ contains the portfolio vector $b \in \mathcal{B}$ translated by some proportion of $\tilde{b}$, where $\tilde{b} \in \mathcal{B}^k$ depends on side information state y. Lastly, the ϵ parameter specifies the 'closeness' of $\tilde{b}_\epsilon(b)$ to $\tilde{b}$; this can be thought of as a free parameter in the strategy that specifies how closely to fit to the prediction, which can later be optimized. We can also extend the definition of $\tilde{b}_\epsilon(b)$ to contain such a set for all possible vectors b as

$$\tilde{b}_\epsilon = \{\tilde{b}_\epsilon(b) | b \in \mathcal{B}\} \subset \mathcal{B}^k.$$

This is analogous to the concept of ϵ-neighbourhood around $\tilde{b}$ in the field of Topology. Now consider the following portfolio model.

Definition 3 *Given the prediction $\tilde{b} \in \mathcal{B}^k$ and the closeness parameter $0 \leq \epsilon \leq 1$, the μ-weighted universal portfolio (with transaction costs and side information) at time t, depending on side information state $y \in [k]$, is specified by*

$$\hat{b}_t = \frac{\int_{\mathcal{B}} \tilde{b}_\epsilon(b, y) \cdot S_{t-1}(\tilde{b}_\epsilon(b)) \, d\mu(b)}{\int_{\mathcal{B}} S_{t-1}(\tilde{b}_\epsilon(b)) \, d\mu(b)}.$$

3.4 Interpretation of Our Model

Note that $\epsilon = 1$ is equivalent to not making use of the predicted portfolio distribution $\tilde{b}$, and likewise $\epsilon = 0$ is equivalent to investing entirely in the predicted portfolio distribution $\tilde{b}$. Therefore, ϵ can be thought of as the neighbourhood of the subspace of portfolio vectors around $\tilde{b}$ to sample from. Throughout this whole chapter, we will assume a uniform prior μ. We will also drop the $\mathcal{B}$ subscript in the integral, so this can be assumed, unless stated otherwise.

This definition of universal portfolio is a generalisation of the model with transaction cost, but is not necessarily a generalisation of the model with side information. This is because the universal portfolio vector is calculated in a different way; the model from Definition 2 assigns weighting to each b according to how much wealth it has previously generated while in that state (through the term $S_{t-1}(b|y)$), whereas this new model (Definition 3) considers its performance across all states (since $S_{t-1}(\tilde{b}_\epsilon(b))$ does not rely on the current side information state y). We made this modification as to give a more natural transition for combining the notion of side information with the presence of transaction costs.

Putting side information aside (assume $k = 1$), we can alternatively interpret the model from Definition 3 as taking a weighted average of $\tilde{b}$ and Cover's universal portfolio in each timestep, and rebalancing as to maintain this weighting. This is distinct from the simple model of placing ϵ of our wealth in $\tilde{b}$, the remaining wealth in Cover's universal portfolio at the initial timestep, and letting them run independently (without rebalancing between these two portfolios). In fact, this simple model would achieve only $1 - \epsilon$ proportion of the wealth obtained by Cover's universal portfolio (in asymptotics, as $T \to \infty$), as the wealth due to $\tilde{b}$ would die off relative to the best CRP, in the worst case; no one single CRP can compete (with sublinear regret) against the best CRP apart from the best CRP itself, hence the need for a universal portfolio algorithm. We will later show that the use of a prediction $\tilde{b}$ in Definition 3 can help improve the regret over that of Cover's model, in the generalised setting of side information and transaction cost.

The total wealth for the universal portfolio strategy $(\hat{b}_1, \cdots, \hat{b}_t)$, say $\hat{S}_t$, is defined as the integral of the wealth for each constituent portfolio $\tilde{b}_\epsilon(b)$, in particular,

$$\hat{S}_t = \int S_t(\tilde{b}_\epsilon(b))\, d\mu(b).$$

If we allow the transfer of wealth across portfolios, then the transaction costs could be reduced to achieve

$$\hat{S}_t \geq \int S_t(\tilde{b}_\epsilon(b))\, d\mu(b),$$

thereby accounting for additional cost savings. However, we will assume that no re-balancing occurs across portfolios, for simplicity.

3.5 Regret Bound

We will now provide the ideas needed to bound the wealth of the universal portfolio from Definition 3 against $\tilde{b}$, achieving sublinear regret. This will be done by proving that portfolios 'near' to each other perform similarly (Lemma 1), and that there are many portfolios that are 'near' each other (Lemma 2) by calculating the volume of its subspace.

We will then proceed to explore precisely how 'good' of an approximation $\tilde{b}$ needs to be with respect to b^*, resulting in bounding the wealth of the universal portfolio against b^*.

Lemma 1 *For $b \in \mathcal{B}, \tilde{b} \in \mathcal{B}^k$, $0 \leq \delta \leq 1$,*

$$\frac{S_T(\tilde{b}_\delta(b))}{S_T(\tilde{b})} \geq (1-\delta)^{(1+c)T}$$

Proof. Denote the side information state at time $t-1$ as y, and at time t as y' (y and y' are not necessarily distinct). Consider the portfolios $b \in \mathcal{B}, \tilde{b} \in \mathcal{B}^k$, and recall that by definition,

$$\tilde{b}_\delta(b, y) = (1-\delta)\tilde{b}(y) + \delta b,$$

$$\tilde{b}_\delta(b, y') = (1-\delta)\tilde{b}(y') + \delta b.$$

We will prove the desired bound by combining a series of bounds for each time step of the strategy. Figure 3.1 below demonstrates the single-time step workflow for the portfolios $\tilde{b}_\delta(b, y)$ and $\tilde{b}(y)$, respectively.

$$\tilde{b}_\delta(b, y) \longrightarrow \beta\tilde{b}(y) \longrightarrow \beta X \longrightarrow \kappa\beta\tilde{b}(y') \longrightarrow (1-c\delta)\kappa\beta\tilde{b}_\delta(b, y')$$

$$\tilde{b}(y) \longrightarrow X \longrightarrow \kappa\tilde{b}(y')$$

Figure 3.1: Proof workflow

We want to first compute the ratio of wealth in a single period between

$\tilde{b}_\delta(b)$ and $\tilde{b}$. To do this, we start with $\tilde{b}_\delta(b, y)$ which consists of some portion of $\tilde{b}(y)$ and b. We first pay some transaction cost to re-balance δb to get a new vector $\gamma\tilde{b}(y)$, where $\gamma \geq 0$ (in other words, the extra asset beyond $(1-\delta)\tilde{b}(y)$ cannot hurt), leaving us with $\beta\tilde{b}(y)$ with $\beta \geq 1-\delta$.

Now we move from time step $t-1$ to time step t and the price of the assets may change; the two portfolios $\tilde{b}(y)$ and $\beta\tilde{b}(y)$ becomes X and βX respectively. Note that $X \in \mathbb{R}_+^m$ but not necessarily $X \in \mathcal{B}$ as the entries may not sum up to 1. The ratio of wealth between portfolio βX and X is β (before re-balancing).

Next, each portfolio needs to perform re-balancing as to preserve the portfolio vectors by the definition of CRP. We re-balance X and βX to some proportion of $\tilde{b}(y')$ and $\tilde{b}_\delta(b, y')$, respectively. Suppose that in the former we get $\kappa\tilde{b}(y')$. For the latter, we first re-balance βX to $\kappa\beta\tilde{b}(y')$, then re-balancing this further to get $(1-c\delta)\kappa\beta\tilde{b}_\delta(b, y')$ (costing an additional $c\delta$ to re-balance to some proportion of b). Reading off directly from this, we get

$$\frac{\text{single time step wealth of } \tilde{b}_\delta(b)}{\text{single time step wealth of } \tilde{b}} \geq (1-c\delta)\beta \geq (1-c\delta)(1-\delta) \geq (1-\delta)^{(1+c)}$$

Taking the product of this over T time steps we get the desired result. $\square$

The above lemma can be interpreted as follows: the portfolio $\tilde{b}_\delta(b)$, which is defined to be near $\tilde{b}$ (in the sense that it consists of some proportion of $\tilde{b}$), will also have wealth similar to $\tilde{b}$ (in fact, no worse off than by a factor of $(1-\delta)^{(1+c)T}$ after T time steps). Now we need to show that there exists sufficiently many such portfolios, by analyzing the volume of the set of such portfolios, $\tilde{b}_\delta$ (or equivalently, measure).

Lemma 2 *For $\tilde{b} \in \mathcal{B}^k$ and $0 \leq \delta \leq 1$,*

$$\mathrm{Vol}(\tilde{b}_\delta) = \delta^{k(m-1)}\mathrm{Vol}(\mathcal{B}^k)$$

Proof. The set $\mathcal{B}$ is convex (in fact, a simplex) and lies in an $m-1$ dimensional space, since the last component can be computed from the rest due to the condition that they must sum up to 1. So the set $\mathcal{B}^k$ (with side information states) lies in a space of $k(m-1)$ dimensions. Therefore,

$$\begin{aligned}
\mathrm{Vol}(\tilde{b}_\delta) &= \prod_{y=1}^{k} \mathrm{Vol}(\tilde{b}_\delta(y)) && (3.1)\\
&= \mathrm{Vol}(\tilde{b}_\delta(1))^k \\
&= \mathrm{Vol}(\{(1-\delta)\tilde{b}(y) + \delta b \mid b \in \mathcal{B}\})^k \\
&= \mathrm{Vol}(\{\delta b \mid b \in \mathcal{B}\})^k && (3.2)\\
&= \left(\delta^{m-1}\mathrm{Vol}(\mathcal{B})\right)^k \\
&= \delta^{k(m-1)}\mathrm{Vol}(\mathcal{B}^k)
\end{aligned}$$

where (3.1) is due to independence of the side information states, and (3.2) is because the set

$$\{(1-\delta)\tilde{b}(y) + \delta b \mid b \in \mathcal{B}\}$$

is the same as the set $\{\delta b \mid b \in \mathcal{B}\}$ shifted from the origin by $(1-\delta)\tilde{b}(y)$. $\square$

Now that we have the required ideas, we will use the previous lemmas to bound the wealth of the universal portfolio (from Definition 3) with respect to $\tilde{b}$. For notational convenience, we will use $\tilde{S}_T$ to denote $S_T(\tilde{b})$.

Lemma 3 *For $\tilde{b} \in \mathcal{B}^k$ and $0 \le \epsilon \le 1$,*

$$\frac{\hat{S}_T}{\tilde{S}_T} \ge \frac{1 + (1-\epsilon)^{(1+c)T+1}\left(\epsilon^{k(m-1)}\left((1+c)T+1\right)^{k(m-1)} - (1-\epsilon)^{k(m-1)}\right)}{\left((1+c)T+1\right)^{k(m-1)}}$$

Proof. We can compute their wealth ratio as

$$\begin{aligned}
\frac{\hat{S}_T}{\tilde{S}_T} &\ge \frac{\int S_T(\tilde{b}_\epsilon(b))\, d\mu(b)}{\tilde{S}_T}, \text{ by definition of wealth of } \hat{S}_T \\
&= \int \frac{S_T(\tilde{b}_\epsilon(b))}{\tilde{S}_T}\, d\mu(b), \text{ since } \tilde{S}_T \text{ is independent of } b \\
&= \int \frac{S_T(\tilde{b}_\delta(b'))}{\tilde{S}_T}\, d\mu(b),
\end{aligned}$$

where

$$\delta = \min\{\delta | \tilde{b}_\epsilon(b) \in \tilde{b}_\delta\}$$

and b' is chosen such that

$$\tilde{b}_\epsilon(b) = \tilde{b}_\delta(b').$$

We can further bound this integral as

$$\begin{aligned}
\frac{\hat{S}_T}{\tilde{S}_T} &\geq \int (1-\delta)^n \, d\mu(b), \text{ by Lemma 1, where } n = (1+c)T \\
&\geq \int_{(1-\epsilon)^n}^1 \Pr_{b \in \mathcal{B}^k} [(1-\delta)^n \geq z] \, dz, \qquad (3.3) \\
&= \int_{(1-\epsilon)^n}^1 \Pr_{b \in \mathcal{B}^k} [\delta \leq 1 - z^{1/n}] \, dz, \text{ by re-arranging the equation}
\end{aligned}$$

where (3.3) is due to the identity

$$\int v \, dx = \int_0^1 \Pr[v \geq z] \, dz$$

for non-negative random variable v. By Lemma 2,

$$\Pr_{b \in \mathcal{B}^k} [\delta \leq 1 - z^{1/n}] = \text{Vol}(\tilde{b}_{1-z^{1/n}}) = (1 - z^{1/n})^{k(m-1)} \text{Vol}(\mathcal{B}^k).$$

Define $r = k(m-1)$. Now the theorem reduces down to evaluating the integral

$$\int_{(1-\epsilon)^n}^1 (1 - z^{1/n})^r \, dz.$$

We can do this by using a change of variable $u = z^{1/n}$, then repeatedly applying integration by parts. Note that using this substitution we have $z = u^n$, hence $\frac{dz}{du} = nu^{n-1}$ and $dz = nu^{n-1}du$. Thus we get

$$\int_{(1-\epsilon)^n}^1 (1 - z^{1/n})^r \, dz = n \int_{1-\epsilon}^1 u^{n-1}(1-u)^r \, du$$

Define

$$F(r, n-1) = \int_{1-\epsilon}^1 u^{n-1}(1-u)^r \, du,$$

then the above reduces to evaluating $tF(r, n-1)$. Using integration by parts, we can obtain a recursive rule for when $r, n-1 \geq 1$ as

$$\begin{aligned} F(r,n-1) &= \int_{1-\epsilon}^{1} u^{n-1}(1-u)^r \, du \\ &= \frac{(1-u)^r u^n}{n}\Big|_{1-\epsilon}^{1} + \frac{r}{n}\int_{1-\epsilon}^{1} u^n (1-u)^{r-1} \, du \\ &= \frac{\epsilon^r(1-\epsilon)^n}{n} + \frac{r}{n}\int_{1-\epsilon}^{1} u^n(1-u)^{r-1} \, du \\ &= \frac{\epsilon^r(1-\epsilon)^n}{n} + \frac{r}{n}F(r-1,n), \end{aligned}$$

with the base case

$$F(0, n+r-1) = \int_{1-\epsilon}^{1} u^{n+r-1} \, du = \left[\frac{u^{n+r}}{n+r}\right]_{1-\epsilon}^{1} = \frac{1-(1-\epsilon)^{n+r}}{n+r}.$$

Notice that if $\epsilon = 0$ or 1, then

$$\frac{\epsilon^r(1-\epsilon)^n}{n} = 0.$$

Otherwise if $0 < \epsilon < 1$ then

$$\frac{\epsilon^r(1-\epsilon)^n}{n} > 0,$$

so this term can be omitted when finding a lower bound of the above function. Using the recursive rule, we can find a good lower bound for the desired integral. For simplicity, we will include the term

$$\frac{\epsilon^r(1-\epsilon)^n}{n}$$

only for the first recursion, otherwise the bound would get complicated very quickly. In particular,

$$nF(r,n-1) = \epsilon^r(1-\epsilon)^n + n\frac{r}{n}F(r-1,n),$$

and applying the recursion we get

$$n\frac{r}{n}F(r-1,n) \geq n\frac{r!(n-1)!}{(n+r-1)!}F(0,n+r-1).$$

Combining these we get

$$\begin{aligned}
nF(r,n-1) &\geq \epsilon^r(1-\epsilon)^n + n\frac{r!(n-1)!}{(n+r-1)!}F(0,n+r-1) \\
&= \epsilon^r(1-\epsilon)^n + \frac{r!n!}{(n+r-1)!}\frac{1-(1-\epsilon)^{n+r}}{n+r} \\
&= \epsilon^r(1-\epsilon)^n + (1-(1-\epsilon)^{n+r})\binom{n+r}{r}^{-1} \\
&\geq \epsilon^r(1-\epsilon)^n + \frac{1-(1-\epsilon)^{n+r}}{(n+1)^r} \\
&= \frac{(n+1)^r\epsilon^r(1-\epsilon)^n}{(n+1)^r} + \frac{1-(1-\epsilon)^{n+r}}{(n+1)^r} \\
&= \frac{1+(1-\epsilon)^n(\epsilon^r(n+1)^r-(1-\epsilon)^r)}{(n+1)^r}
\end{aligned}$$

Substituting back $n=(1+c)T$ and $r=k(m-1)$ gets the desired result. □

Recall that $\tilde{b}$ was defined previously as a 'good' approximation to b^*. Now we will formalise this notion of 'goodness' and give a sufficient condition such that there is also a sublinear difference between the log-wealth of the universal portfolio $\hat{b}$ and the best CRP b^*. First we define some notations which we will use later.

Definition 4 *For $\tilde{b}, b^* \in \mathcal{B}^k$, $\tilde{b}$ is α-close to b^* if for every $y \in [k]$,*

$$b^*(y) \geq (1-\alpha)\tilde{b}(y).$$

Note that we could have substituted b^* in place of $\tilde{b}$ in Lemma 3, and get a sublinear difference between the log-wealth of $\hat{b}$ and b^*. The problem with this is that, by our definition of universal portfolio, this would require us to know b^* in advanced. Instead, we assume that we have some idea of certain regions of b^*, but perhaps not its entire constituents. So let's assume that we are able to approximate $\tilde{b}$ such that it is α-close to b^*. The closeness parameter ϵ in the universal portfolio model is chosen according to α, as seen in the proof of the theorem below.

Theorem 4 *Suppose $\tilde{b}$ is α-close to b^*. Then for all values of m, k, c, α and $\epsilon \geq \alpha$, the regret against b^* is bounded as*

$$\max_{x^T} \left(\log S_T^* - \log \hat{S}_T \right) \leq k(m-1) \log[(1+c)T + 1] + \zeta,$$

where the remainder ζ can be written as

$$\zeta = k(m-1) \log \left[\frac{1-\alpha}{1-\alpha/\epsilon} \right]$$
$$- \log \left[1 + (1-\epsilon)^{(1+c)T+1} \left(\epsilon^{k(m-1)} \left((1+c)T+1 \right)^{k(m-1)} - (1-\epsilon)^{k(m-1)} \right) \right].$$

Proof. Recall that in our definition of universal portfolio (Definition 3), we took an integral over all $b \in \mathcal{B}$. However, the actual portfolio vectors used inside the integral are from the set $\tilde{b}_\epsilon$. If we can show that

$$b^*_\lambda \subseteq \tilde{b}_\epsilon$$

for some $0 \leq \lambda \leq 1$, this is equivalent to saying that the universal portfolio model with $\tilde{b} := b^*$ is contained in the model with $\tilde{b} := \tilde{b}$ (and we know that the model with $\tilde{b} := b^*$ has sublinear regret between b^* and $\hat{b}$, by Lemma 3). Furthermore, the ratio of the volumes of these two sets is a constant independent of n, since

$$\mathrm{Vol}(b^*_\lambda)/\mathrm{Vol}(\tilde{b}_\epsilon) = (\lambda/\epsilon)^{k(m-1)},$$

implying that

$$\tilde{S}_T \geq (\lambda/\epsilon)^{k(m-1)} S_T^*.$$

Now we will show that $b^*_\lambda \subseteq \tilde{b}_\epsilon$. Suppose $z \in b^*_\lambda$, then

$$z \geq (1-\lambda) b^*$$

by definition. Furthermore, by the closeness condition (Definition 4),

$$z \geq (1-\lambda)(1-\alpha)\tilde{b}.$$

Let's choose λ so that

$$(1-\lambda)(1-\alpha) = (1-\epsilon),$$

implying that

$$z \geq (1-\epsilon)\tilde{b}$$

and hence $z \in \tilde{b}_\epsilon$. This choice of lambda is always possible, and between 0 and 1 when $\epsilon \geq \alpha$. In summary,

$$\frac{S_T^*}{\hat{S}_T} \leq (\epsilon/\lambda)^{k(m-1)} \frac{\tilde{S}_T}{\hat{S}_T} = \left(\frac{1-\alpha}{1-\alpha/\epsilon}\right)^{k(m-1)} \frac{\tilde{S}_T}{\hat{S}_T},$$

since $(1-\lambda)(1-\alpha) = (1-\epsilon)$, and Lemma 3 gives the bound of $\frac{\tilde{S}_T}{\hat{S}_T}$. $\square$

Notice that for $\epsilon = 1$ (equivalent to ignoring the prediction $\tilde{b}$ altogether), we have that $\zeta = 0$, thus getting the same regret bound in a generalised case of side information [CO96] and transaction cost [BK99].

If α is sufficiently small (equivalent to having very good prediction) depending on the other parameters, it is possible to choose ϵ such that $\zeta < 0$, thus yielding an improved regret bound over previous results, in the generalised setting. For example, in the case where $k = 1$ (no side information), $m = 2$ (two asset case), $c = 0$ (no transaction cost), and suppose $\alpha = 0.01$. Now if we choose $\epsilon = 0.1$ then one can verify that $\zeta < 0$ for $1 \leq T \leq 32$. On the other hand, choosing $\epsilon = 0.05$ gives $\zeta < 0$ for $6 \leq T \leq 41$, etc.

3.6 Portfolio Computation

The definition of universal portfolios rely on the ability to pick continuous quantities of each portfolio vector from the set $\mathcal{B}^k$. Practically this is hard to achieve, but we could approximate it by picking enough samples from the prior distribution μ, and invest an amount of wealth equally between all of them. Chebyshev's inequality will guarantee that we can get sufficiently close to the mesure-theoretic definition of universal portfolio, with high probability, although the sample complexity will grow in the same rate as the ratio in Theorem 4, as we shall see later.

Figure 3.2 below outlines this randomized approximation scheme to compute the universal portfolio from Definition 3. It is an extension and generalization of the randomized approximation scheme that was briefly mentioned

Input variables : $(\mu, y, \tilde{b}, \epsilon, \gamma, \delta, R)$
μ is the prior distribution on $\mathcal{B}$ for initial portfolio vectors
$y = (y_1, \cdots, y_T)$ are the side information states
$\tilde{b}$ is an approximation of b^* and ϵ is the closeness parameter
R is the best known upper bound for $S_T^*/\hat{S}_T$

1. Define $N := (R-1)/\gamma^2\delta$.
2. Pick the sample set $B = \{b_1, \cdots, b_N\}$ i.i.d. where each $b_i \in \mathcal{B}$, according to the prior distribution μ.
3. Initialize $X_i \leftarrow 1$, where each X_i keeps track of the wealth of b_i.
4. Run them independently as follows:
   For time steps $t \in [T]$ and samples $i \in [N]$,
   (a) Invest X_i of the wealth according to portfolio vector $\tilde{b}_\epsilon(b_i, y_t)$.
   (b) Update the wealth of that sample, $X_i \leftarrow S_t\left(\tilde{b}_\epsilon(b_i)\right)$.

Output : $\check{S}_T := \frac{1}{N}\sum_{i=1}^{N} X_i$

Figure 3.2: Randomized approximation

by Blum and Kalai [BK99] which approximated Cover's universal portfolio (without side information).

It will later become clear as to the particular choice of N in the Figure 3.2. The input variables γ and δ in the algorithm above will directly control the accuracy of the approximation. In particular, the algorithm in Figure 3.2 achieves a wealth of at least $1-\gamma$ times as large as the universal portfolio with probability $1-\delta$. To prove this, we will make use of the one-sided version of Chebyshev's inequality (otherwise known as Chebyshev-Cantelli) as in the lemma below.

Lemma 5 *For a random variable X and for any $a > 0$,*

$$\Pr[X \leq \mathbb{E}[X] - a] \leq \frac{\mathrm{Var}[X]}{\mathrm{Var}[X] + a^2}.$$

We can now proceed to prove the bound on the approximation scheme, in addition to calculating its sample complexity.

Theorem 6 *With probability at least* $1-\delta$*, the approximation from Figure 3.2 achieves a wealth* $\check{S}_T$ *of at least* $1-\gamma$ *times as large as* $\hat{S}_T$*. In other words,*

$$\Pr\left[\frac{\check{S}_T}{\hat{S}_T} > 1-\gamma\right] \geq 1-\delta.$$

This requires sample complexity

$$O\left(\frac{((1+c)T)^{k(m-1)}}{\gamma^2\delta}\right).$$

Proof. As in the algorithm, let X_i represent the wealth derived from the i^{th} sample b_i, where the b_i's are chosen i.i.d. according to μ,

$$X_i = S_T\left(\tilde{b}_\epsilon(b_i)\right).$$

For some $\gamma, \delta \in [0,1]$, let

$$N = (R-1)/\gamma^2\delta$$

where R is an upper bound for $S_T^*/\hat{S}_T$. Define

$$X = \frac{1}{N}\sum_{i=1}^{N} X_i.$$

Then by linearity of expectation and variance,

$$\mathbb{E}[X] = \mathbb{E}[X_i]$$

and

$$\mathrm{Var}[X] = \mathrm{Var}[X_i]/N.$$

We want to prove that

$$\Pr\left[\frac{\check{S}_T}{\hat{S}_T} > 1-\gamma\right] \geq 1-\delta,$$

which is equivalent to

$$\Pr\left[\frac{X}{\mathbb{E}[X]} \leq 1-\gamma\right] \leq \delta,$$

since $\check{S}_T = X$ and $\hat{S}_T = \mathbb{E}[X_i] = \mathbb{E}[X]$. Re-arranging this and applying

Lemma 5 we get

$$\Pr\left[X \le \mathbb{E}[X] - \gamma\mathbb{E}[X]\right] \le \frac{\mathrm{Var}[X]}{\mathrm{Var}[X] + (\gamma\mathbb{E}[X])^2} = \frac{\mathrm{Var}[X_i]}{\mathrm{Var}[X_i] + N\left(\gamma\mathbb{E}[X]\right)^2},$$

where the last equality is due to $\mathrm{Var}[X] = \mathrm{Var}[X_i]/N$. Therefore it remains to show that

$$\frac{\mathrm{Var}[X_i]}{\mathrm{Var}[X_i] + N\left(\gamma\mathbb{E}[X]\right)^2} \le \delta,$$

which is equivalent to

$$\frac{\mathrm{Var}[X_i](1-\delta)}{\delta\left(\gamma\mathbb{E}[X]\right)^2} \le N,$$

since we assume $\mathbb{E}[X] > 0$. We can bound this as

$$\begin{aligned}
\frac{\mathrm{Var}[X_i](1-\delta)}{\delta\left(\gamma\mathbb{E}[X]\right)^2} &= \frac{\mathrm{Var}[X_i](1-\delta)}{\delta\gamma^2\mathbb{E}[X_i]^2}, \text{ since } \mathbb{E}[X] = \mathbb{E}[X_i] \\
&\le \frac{\mathrm{Var}[X_i]}{\delta\gamma^2\mathbb{E}[X_i]^2}, \text{ since } \delta \ge 0 \\
&= \frac{\mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2}{\delta\gamma^2\mathbb{E}[X_i]^2}, \text{ by definition of variance} \\
&= \frac{1}{\gamma^2\delta}\left(\frac{\mathbb{E}[X_i^2]}{\mathbb{E}[X_i]^2} - 1\right).
\end{aligned}$$

Using the fact that $\mathbb{E}[X_i] = \hat{S}_T$ and $\mathbb{E}[X_i^2] \le \check{S}_T\hat{S}_T$, we get

$$\begin{aligned}
\frac{1}{\gamma^2\delta}\left(\frac{\mathbb{E}[X_i^2]}{\mathbb{E}[X_i]^2} - 1\right) &\le \frac{1}{\gamma^2\delta}\left(\frac{\check{S}_T\hat{S}_T}{\hat{S}_T^2} - 1\right) \\
&= \frac{1}{\gamma^2\delta}\left(\frac{\check{S}_T}{\hat{S}_T} - 1\right) \\
&\le \frac{1}{\gamma^2\delta}\left(\frac{S_T^*}{\hat{S}_T} - 1\right), \text{ since } \check{S}_T \le S_T^* \text{ (the best CRP)} \\
&\le \frac{R-1}{\gamma^2\delta}, \text{ since } S_T^*/\hat{S}_T \le R \text{ by definition} \\
&= N.
\end{aligned}$$

To get a bound on the sample complexity, we need a bound on R. Theorem 4 implies that

$$R \in O\left(((1+c)T)^{k(m-1)}\right).$$

Using this we obtain the sample complexity

$$N = \frac{R-1}{\gamma^2\delta} \in O\left(\frac{((1+c)T)^{k(m-1)}}{\gamma^2\delta}\right).$$

$\square$

When the number of states k and number of assets m are fixed (independent of T), the sample complexity from the theorem above is polynomial in T, although potentially of high order (depending on k, m). On the other hand, the sample complexity is exponential in m and k.

Kalai and Vempala [KV02] showed a more efficient randomized approximation scheme for the classical Cover's universal portfolio (without side information or transaction costs). They derived a random walk sampling algorithm that yields sample complexity that is polynomial in m and of low-order polynomial in T (something like $O(T^3 \log^2 T)$). Therefore it may be possible to extend their technique to get a more efficient method than the one presented here. It should be noted though that this is unlikely to be straightforward; Kalai and Vempala [KV02] mentioned that they do not know of a way to derive an implementation with the presence of transaction costs ($c > 0$), and furthermore it must account for multiple side information states (possibly $k \geq 2$) and prediction inputs as in our new model.

Chapter 4

Dynamic Trading Strategy

In this chapter, we introduced a counterpart to the portfolio selection model from the last chapter that balances the reward from rebalancing the portfolio (based on information received from a price prediction) against the transaction cost incurred, and find an optimal point in between as to maximise cost adjusted wealth.

The ideas in this chapter will go beyond the restriction imposed by the CRP, and instead, we devise a model that is competitive with the best greedy portfolio in a stochastic setting: one that makes the optimal decision as if it knows the next time step's price. To do this, we suppose that our model has access to a price prediction $\tilde{x}_t$ (of the next time step, $t+1$) that follows some probability distribution $\tilde{x}_t \sim \mathcal{D}_t(x_t)$, where x_t is the later observed price change (in the sense that x_t is revealed by the adversary after the player draws the prediction and chooses his portfolio). In this model, we quantify the precise relationship between the expected regret and the accuracy of such predictions. Note that we allow the prediction accuracy to vary over time, as reflected by the dependence of $\mathcal{D}_t$ on the current time step t. We demonstrate that for certain probability distributions $\mathcal{D}_t$,

$$\mathbb{E}_{\tilde{x}_t \sim \mathcal{D}_t(x_t)} \left[\max_{x^T} \left(\log S_T^* - \log \hat{S}_T \right) \right] = o(T)$$

is attainable, subject to some restrictions on the accuracy of $\tilde{x}_t$'s: namely, that the integral of the tail probabilities (of mis-estimation) must converge to zero as t grows. Intuitively, this is equivalent to improving our predictions through learning from past outcomes, and the requirement is that the model must be learning at a rate fast enough as to satisfy a certain sufficient condition which we will later prove.

4.1 Greedy Portfolio

At time $t \in [T]$, suppose our model has access to a prediction such that it follows some probability distribution with respect to the later observed price change: that is, $\tilde{x}_t \sim \mathcal{D}_t(x_t)$. Note that the distribution $\mathcal{D}_t$ may depend on the current time step t (hence, the subscript) and x_t, possibly hiding further dependencies on additional parameters such as volatility. Based on this prediction, we can compute a portfolio distribution as to optimize the wealth.

Definition 5 *For each $t \in [T]$, given a predicted price-change $\tilde{x}_t$ of the observed price change x_t such that $\tilde{x}_t \sim \mathcal{D}_t(x_t)$ for some probability distribution $\mathcal{D}_t$, the portfolio distribution at time t is specified by*

$$\hat{b}_t := \arg\max_{b \in \mathcal{B}} b\tilde{x}_t \theta(\hat{b}_{t-1}, b, x_{t-1}).$$

The benchmark model, which we call the optimal greedy portfolio, is defined similarly as, for each time t,

$$b_t^* = \arg\max_{b \in \mathcal{B}} b x_t \theta(b_{t-1}^*, b, x_{t-1}).$$

Note that the above models considers the tradeoff between the transaction cost of shifting to a "better" portfolio against the expected benefit of doing such a rebalancing given the prediction or actual outcome, respectively. In the case where the optimization yields multiple solutions, we canonically choose the one with the least transaction costs. This will be made more precise in Section 4.8.

In the rest of the chapter, we investigate how close the wealth of the above portfolio model is to the optimal greedy portfolio. As a measure of performance, we consider the expected regret $\mathbb{E}[R]$. Namely,

$$\mathbb{E}_{\tilde{x}_t \sim \mathcal{D}_t(x_t)} \left[\max_{x^T} \left(\log S_T^* - \log \hat{S}_T \right) \right].$$

This can be interpreted as enumerating through all possible price predictions $\tilde{x}^T$ and choosing the outcome of price sequence x^T that maximises regret for each choice of $\tilde{x}^T$. Each of these choices of $\tilde{x}^T$ occurs with some probability depending on x^T and $\mathcal{D}_t$ for $t \in [T]$, and we take the expectation over these probabilities.

Note that a bound on the expected-regret will also imply a bound (of the same order, up to some constant) for the regret with high probability, by Markov's inequality. In particular, for all $k > 0$,

$$\Pr\left[R \leq k\mathbb{E}[R]\right] \geq 1 - \frac{1}{k}.$$

4.2 Impossibility without Predictions

We assume the existence of predictions because otherwise it is impossible to achieve sublinear expected regret, as demonstrated in the below theorem.

Theorem 7 *If there are no predictions, in particular, no restriction on how $\tilde{x}_t$ relates to x_t, then sublinear regret is not achievable,*

$$\mathbb{E}[R] = \Omega(T).$$

Proof. Consider the two asset case, $m = 2$. Suppose the probabilistic trader invests according to the portfolio distribution $b \in \mathcal{B} \subset \mathbb{R}^2$ with probability $f_t(b)$ at time $t \in [T]$. Then for each $b \in \mathcal{B} \subset \mathbb{R}^2$, the adversary chooses a returns vector $g_t(b) := x_t$ at time $t \in [T]$ as

$$g_t(b) = \begin{cases} (0,1) & \text{if } b(1) \geq \frac{1}{2}, \\ (1,0) & \text{otherwise.} \end{cases}$$

The first condition is equivalent to $b(1) \geq b(2)$ as they must sum up to one. The equivalent best greedy portfolio in hindsight at time $t \in [T]$ is

$$h_t(b) = \begin{cases} (0,1) & \text{if } g(b) = (0,1), \\ (1,0) & \text{if } g(b) = (1,0). \end{cases}$$

This will yield a single-time step expected wealth for the trader at time $t \in [T]$ of

$$\begin{aligned}
\int_{\mathcal{B}} f_t(b) \cdot g_t(b)\, df_t(b) &= \int_{b(1)\geq\frac{1}{2}} f_t(b) \cdot g_t(b)\, df_t(b) + \int_{b(1)<\frac{1}{2}} f_t(b) \cdot g_t(b)\, df_t(b) \\
&= \int_{b(1)\geq\frac{1}{2}} f_t(b) \cdot (0,1)\, df_t(b) + \int_{b(1)<\frac{1}{2}} f_t(b) \cdot (1,0)\, df_t(b) \\
&\leq \int_{b(1)\geq\frac{1}{2}} \left(\frac{1}{2}, \frac{1}{2}\right) \cdot (0,1)\, df_t(b) + \int_{b(1)<\frac{1}{2}} \left(\frac{1}{2}, \frac{1}{2}\right) \cdot (1,0)\, df_t(b) \\
&= \frac{1}{2},
\end{aligned}$$

where last inequality is due to the fact that the best portfolio within the constraints that maximises wealth is $\left(\frac{1}{2}, \frac{1}{2}\right)$ in both cases. The single-time step wealth of the best greedy portfolio at time $t \in [T]$ is

$$\begin{aligned}
\int_{\mathcal{B}} h_t(b) \cdot g_t(b)\, df_t(b) &= \int_{b(1)\geq\frac{1}{2}} h_t(b) \cdot g_t(b)\, df_t(b) + \int_{b(1)<\frac{1}{2}} h_t(b) \cdot g_t(b)\, df_t(b) \\
&= \int_{b(1)\geq\frac{1}{2}} (0,1) \cdot (0,1)\, df_t(b) + \int_{b(1)<\frac{1}{2}} (1,0) \cdot (1,0)\, df_t(b) \\
&= 1.
\end{aligned}$$

Therefore, the single-time step expected wealtlh ratio is at least 2. Combining this over T time steps give at least a linear expected regret. □

Therefore, the key ingredients to achieving a good performance (against the optimal greedy portfolio) are accurate price predictions.

4.3 Expected Regret Bound

We analyze the expected regret $\mathbb{E}[R]$, where the choice of portfolio distributions depend directly on the random predictions $\tilde{x}_t \sim \mathcal{D}_t(x_t)$ and x_t is chosen adversarially, for each $t \in [T]$. The theorem below gives an upper bound on the expected regret of the strategy from Definition 5 against the optimal greedy portfolio as a function of the distributions $\mathcal{D}_t$ of predictions in each time step, in the presence of transaction costs.

Theorem 8 *The expected regret of the portfolio strategy from Definition 5 can be bounded from above as*

$$\mathbb{E}[R] \leq \gamma + 2\sum_{t=1}^{T}\int_{0}^{\infty} \Pr_{\tilde{x}_t \sim \mathcal{D}_t(x_t)} [\tilde{x}_t \notin (e^{-z}x_t, e^{z}x_t)]\, dz\,,[1]$$

where γ accounts for the regret arising from the positioning error of the portfolio and is defined as

$$\gamma = -\sum_{t=1}^{T} \mathbb{E}\left[\log \frac{\theta(\hat{b}_{t-1}, b_t^*, x_{t-1})}{\theta(b_{t-1}^*, b_t^*, x_{t-1})}\right].$$

Proof. We fix some time t and consider the ratio of the single-time-step wealth change of our portfolio to that of the benchmark at time t in order to bound the regret arising from that time step. The regret associated with the time step t has two sources: positioning error of the current portfolio that results in transaction costs and inaccurate price predictions. We define

$$\rho_t = \frac{\theta(\hat{b}_{t-1}, b_t^*, x_{t-1})}{\theta(b_{t-1}^*, b_t^*, x_{t-1})}$$

to capture the regret arising from the positioning error of the portfolio at time step t: for example, when b_{t-1}^* was in a better position than $\hat{b}_{t-1}$ to minimise transaction costs when rebalancing at time t. Now, suppose that[2]

$$(1-\delta)x_t \preceq \tilde{x}_t \preceq (1-\delta)^{-1}x_t,$$

at time step t, for some δ such that $0 \leq \delta < 1$. Then, for any $\hat{b}_t, b_t^*, \hat{b}_{t-1}, b_{t-1}^* \in \mathcal{B}$, we have the following bound on the ratio of the single-time-step wealths:

$$\begin{aligned}
\frac{\hat{b}_t x_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}{b_t^* x_t \theta(b_{t-1}^*, b_t^*, x_{t-1})} &\geq (1-\delta)\frac{\hat{b}_t \tilde{x}_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}{b_t^* x_t \theta(b_{t-1}^*, b_t^*, x_{t-1})} && (4.1)\\
&\geq (1-\delta)^2\frac{\hat{b}_t \tilde{x}_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}{b_t^* \tilde{x}_t \theta(b_{t-1}^*, b_t^*, x_{t-1})} && (4.2)\\
&\geq (1-\delta)^2 \rho_t. && (4.3)
\end{aligned}$$

[1]The notations $\notin$ is applied as a component-wise conjunction, and e^{-z}, e^{z} multiplies on to each element of x_t, since $x_t, \tilde{x}_t$ are multidimensional.

[2]The notations $\preceq$, $\succeq$, $\prec$, and $\succ$ denote component-wise vector inequalities.

In the above, (4.1) is due to

$$x_t \succeq (1-\delta)\tilde{x}_t,$$

(4.2) is due to

$$\tilde{x}_t \succeq (1-\delta)x_t,$$

and (4.3) is due to the fact that

$$\begin{aligned}\hat{b}_t\tilde{x}_t\theta(\hat{b}_{t-1},\hat{b}_t,x_{t-1}) &\leq b_t^*\tilde{x}_t\theta(\hat{b}_{t-1},b_t^*,x_{t-1})\\ &= \rho_t b_t^*\tilde{x}_t\theta(b_{t-1}^*,b_t^*,x_{t-1}),\end{aligned}$$

as $\hat{b}_t$ was chosen to maximise its single-time-step wealth by Definition 5. For each time step $t \in [T]$, we define deviation δ_t of x_t and $\tilde{x}_t$ as

$$\delta_t := \min\{\delta \geq 0 \mid (1-\delta)x_t \preceq \tilde{x}_t \preceq (1-\delta)^{-1}x_t\}.$$

Intuitively, this is the deviation of the predicted price change from the observed price change. We can now calculate the expected regret as follows.

$$\begin{aligned}\mathbb{E}[R] &= \mathbb{E}\left[\max_{x^T}\log\left(\frac{S_T^*}{\hat{S}_T}\right)\right]\\ &= \mathbb{E}\left[\max_{x^T}\log\left(\prod_{t=1}^{T}\frac{b_t^*x_t\theta(b_{t-1}^*,b_t^*,x_{t-1})}{\hat{b}_tx_t\theta(\hat{b}_{t-1},\hat{b}_t,x_{t-1})}\right)\right]\\ &\leq \mathbb{E}\left[\log\left(\prod_{t=1}^{T}(1-\delta_t)^{-2}\rho_t^{-1}\right)\right] && (4.4)\\ &\leq \sum_{t=1}^{T}2\mathbb{E}\left[-\log(1-\delta_t)\right]-\mathbb{E}\left[\log\rho_t\right], && (4.5)\end{aligned}$$

where (4.4) is by the inequality from (4.3), and (4.5) follows from linearity of expectation. We now will now use

$$\gamma = -\sum_{t=1}^{T}\mathbb{E}[\log\rho_t]$$

to denote the "positioning error," and continue our analysis of the first term on the right hand side of the inequality.

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{E}\left[-\log(1-\delta_t)\right] &= \sum_{t=1}^{T} \int_0^\infty \Pr_{\tilde{x}_t}[-\log(1-\delta_t) \geq z]\, dz \\
&= \sum_{t=1}^{T} \int_0^\infty \Pr_{\tilde{x}_t}[1-\delta_t \leq e^{-z}]\, dz, \\
&= \sum_{t=1}^{T} \int_0^\infty 1 - \Pr_{\tilde{x}_t}[1-\delta_t > e^{-z}]\, dz, \\
&= \sum_{t=1}^{T} \int_0^\infty 1 - \Pr_{\tilde{x}_t}[e^{-z}x_t \prec \tilde{x}_t \prec e^{z}x_t]\, dz,
\end{aligned}$$

where the last line above is obtained from applying the definition of δ_t, giving us the bound on expected regret. □

Note that γ from Theorem 8 captures the positioning error of our model arising from transaction costs. Hence, in the absence of transaction costs (that is, when $c = 0$), we have that $\gamma = 0$. In fact, we later prove in Section 4.4 that, in general, $\gamma = \Omega(T)$ for non-zero transaction costs (that is, when $c > 0$), by showing that there exists a sequence x^T that yields an expected regret at least linear in T.

We also observe that $\gamma = 0$ in the weaker case when x_t is a random variable that is independent of x_{t-1} (hence, also independent of b^*_{t-1} and $\hat{b}_{t-1}$), for all time steps $t \in [T]$, whereas Theorem 8 is stronger as it makes no assumption on how x_t are chosen. This is because

$$\mathbb{E}\left[\log\theta(b^*_{t-1}, b^*_t, x_{t-1})\right] = \mathbb{E}\left[\log\theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})\right],$$

intuitively meaning that the random choice of x_t and $\tilde{x}_t$ are just as likely be favourable to b^*_{t-1} as it is to $\hat{b}_{t-1}$. For example, suppose that we define

$$\tilde{x}_t = (1, \ldots, 1)$$

and x_t is drawn from some log-normal distribution with mean $\tilde{x}_t$. Then, this is equivalent to assuming that the returns x_t follows a Geometric Brownian

Motion and that the current price is the best prediction of the next time step's price; similar to the assumption underlying much of the work in financial mathematics.

Finally, setting γ aside, the result above gives us a good intuition on what the expected regret looks like. Namely, in each time step the regret can be thought of to be no larger than the sum of an integral of the tail probabilities. Having a small expected regret then hinges on efficiently bounding these tail probabilities.

4.4 Non-zero Transaction Cost

We will now show that for any class of non-trivial distributions $\mathcal{D}_t$, the expected-regret bound above will not be sublinear for non-zero transaction cost (in effect, showing that γ is necessarily linear in T, for any $c > 0$). This is because there exists a sequence of returns x_t for $t \in [T]$ that will favour b_t^*'s position, hence, yielding a large enough regret.

Here, we define a *non-trivial distribution* as one where the preimage of the cumulative distribution function is non-empty at some value inside a constant interval around $\frac{1}{2}$. Note that any class of continuous distributions satisfies this criteria.

Theorem 9 *Given non-trivial $\mathcal{D}_t$ for all $t \in [T]$, when $c > 0$,*

$$\mathbb{E}[R] = \Omega(T).$$

Proof. To prove that the expected regret is not necessarily sublinear in the case of non-zero transaction cost, it is enough to come up with a sequence of x_t that breaks this sub-linearity. Therefore, we will give a way to construct such x_t for each $t \in [T]$ in the two-asset case ($m = 2$), where b_t^* and $\hat{b}_t$ will always take the values of either $(0, 1)$ or $(1, 0)$ by our construction of the re-balancing scheme from Section 4.8.

First we will describe the proof concept qualitatively, then subsequently provide the precise details. We will choose x_t in such a way that, when our portfolio disagrees with that of the competing benchmark, it will force some loss by making our portfolio rebalance (due to holding the asset with inferior next timestep return). On the other hand, when our portfolio agrees with that of the competing benchmark, we will choose an x_t as to result in equal probability of rebalancing or staying put, depending on the random variable $\tilde{x}_t \sim \mathcal{D}_t(x_t)$. This would imply that, in the long run, our portfolio disagrees with the competing benchmark in a fixed proportion of the timesteps. Summing up the fixed loss in these timesteps yield linear regret.

Now we give the formal proof. For time step t, assume that $\hat{b}_{t-1} = (0, 1)$, without loss of generality, with b^*_{t-1} is $(0, 1)$ or $(1, 0)$. We will calculate the single-time-step loss

$$\frac{b^*_t x_t \theta(b^*_{t-1}, b^*_t, x_{t-1})}{\hat{b}_t x_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}$$

in these two cases separately.

State 1 (Different) $b^*_{t-1} = (1, 0)$

The adversary chooses $x_t = (1, 1 - c)$, resulting in a single-time-step loss of $\frac{1}{1-c}$, regardless of the choice $\tilde{x}_t \sim \mathcal{D}_t(x_t)$.

State 2 (Same) $b^*_{t-1} = (0, 1)$

The adversary chooses $x_t = (\xi_t, 1)$, where ξ_t is chosen such that

$$\Pr_{\tilde{x}_t \sim \mathcal{D}_t((\xi_t, 1))} \left[\frac{\tilde{x}_t(1)}{\tilde{x}_t(2)} > \frac{1}{1-c} \right] = \frac{1}{2}.$$

Intuitively, this is the choice of price relative vector where the portfolio model (as represented by $\hat{b}_t$) has equal probabilities of shifting or staying put. This implies that $\Pr_{\tilde{x}_t \sim \mathcal{D}_t(x_t)}[\hat{b}_t = b^*_t] = \frac{1}{2}$, and the single-time-step loss may be as small as 1 in this case. Note that this choice of ξ_t exists if the preimage of the CDF of $\mathcal{D}_t$ at $\frac{1}{2}$ is non-empty. One can easily extend this proof to cases where the preimage of the CDF is non-empty at some value inside a constant interval around $\frac{1}{2}$.

With this information, we can model the dynamics of the portfolio as a Markov chain with these two states (Different and Same). The transition probability matrix of that Markov chain, assuming worst-case, i.e., the lowest probability of staying in "different", is

$$\begin{array}{c c} & \begin{array}{cc} \text{Different} & \text{Same} \end{array} \\ \begin{array}{c} \text{Different} \\ \text{Same} \end{array} & \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \end{array}$$

which implies a limiting distribution $\pi = (\frac{1}{3}, \frac{2}{3})$. Using this, the expected regret (over all possible x_t) can be lower-bounded by the linear expected regret (over the particular choice of x_t, as described above).

$$\begin{aligned}
\mathbb{E}[R] &= \mathbb{E}\left[\max_{x^T} \log\left(\frac{S_T^*}{\hat{S}_T}\right)\right] \\
&\geq \mathbb{E}\left[\log\left(\frac{S_T^*}{\hat{S}_T}\right)\right] \\
&= \sum_{t=1}^{T} \mathbb{E}\left[\log \frac{b_t^* x_t \theta(b_{t-1}^*, b_t^*, x_{t-1})}{\hat{b}_t x_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}\right] \\
&= -\frac{1}{3} \sum_{t=1}^{T} \log(1-c) = \Theta(T),
\end{aligned}$$

where the last line follows from the fact that the portfolio needs to shift all its wealth in one third of the steps in the long run (due to the limiting distribution of the Markov chain above), each of which incurs a loss factor of $1 - c$. □

So now we have established that we cannot hope for sublinear expected regret in the presence of transaction costs, no matter the choice of $\mathcal{D}_t$ (as long as it is non-trivial).

4.5 Variance of Regret Bound

We can now prove a bound on the variance of regret, using much of the ideas from the proof of the bound on expected regret in Theorem 8.

Theorem 10 *The variance of regret of our portfolio strategy from Definition 5 can be bounded from above as*

$$\mathrm{Var}[R] \leq \eta + 4 \sum_{t=1}^{T} \int_0^{\infty} \Pr_{\tilde{x}_t \sim \mathcal{D}_t(x_t)} [\tilde{x}_t \notin (e^{-\sqrt{z}} x_t, e^{\sqrt{z}} x_t)] \, dz \,,$$

where η accounts for the variance in the regret arising from the positioning error and the covariance of the single-time-step wealth ratios, defined as

$$\begin{aligned}\eta = &- \sum_{t=1}^{T} \mathrm{Var}\left[\log \frac{\theta(\hat{b}_{t-1}, b_t^*, x_{t-1})}{\theta(b_{t-1}^*, b_t^*, x_{t-1})}\right] \\ &+ \sum_{t=1}^{T} \sum_{j \neq t} \mathrm{Cov}\left[\frac{b_t^* x_t \theta(b_{t-1}^*, b_t^*, x_{t-1})}{\hat{b}_t x_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}, \frac{b_j^* x_j \theta(b_{j-1}^*, b_j^*, x_{j-1})}{\hat{b}_j x_j \theta(\hat{b}_{j-1}, \hat{b}_j, x_{j-1})}\right].\end{aligned}$$

Proof.

$$\begin{aligned}\mathrm{Var}[R] &= \mathrm{Var}\left[\max_{x^T} \log\left(\frac{S_T^*}{\hat{S}_T}\right)\right] \\ &= \mathrm{Var}\left[\max_{x^T} \log\left(\prod_{t=1}^{T} \frac{b_t^* x_t \theta(b_{t-1}^*, b_t^*, x_{t-1})}{\hat{b}_t x_t \theta(\hat{b}_{t-1}, \hat{b}_t, x_{t-1})}\right)\right] \\ &\leq \mathrm{Var}\left[\log\left(\prod_{t=1}^{T} (1-\delta_t)^{-2} \rho_t^{-1}\right)\right] \\ &\leq \eta + 4 \sum_{t=1}^{T} \mathrm{Var}\left[-\log(1-\delta_t)\right],\end{aligned}$$

where η is the term representing the positioning errors and covariance terms, as described in the theorem statement. We continue to simplify the remaining part of the equation, making use of the inequality $\mathrm{Var}[R] \leq \mathbb{E}[R^2]$. Thus, we get

$$\begin{aligned}\sum_{t=1}^{T} \mathrm{Var}\left[-\log(1-\delta_t)\right] &\leq \sum_{t=1}^{T} \mathbb{E}\left[(-\log(1-\delta_t))^2\right] \\ &= \sum_{t=1}^{T} \int_0^{\infty} \Pr_{\tilde{x}_t}[-\log(1-\delta_t) \geq \sqrt{z}]\, dz \\ &= \sum_{t=1}^{T} \int_0^{\infty} \Pr_{\tilde{x}_t}[1-\delta_t \leq e^{-\sqrt{z}}]\, dz, \\ &= \sum_{t=1}^{T} \int_0^{\infty} 1 - \Pr_{\tilde{x}_t}[1-\delta_t > e^{-\sqrt{z}}]\, dz, \\ &= \sum_{t=1}^{T} \int_0^{\infty} 1 - \Pr_{\tilde{x}_t}[e^{-\sqrt{z}} x_t \prec \tilde{x}_t \prec e^{\sqrt{z}} x_t]\, dz,\end{aligned}$$

where the last line above is obtained from applying the definition of δ_t (as defined in the proof of Theorem 8), giving us the desired result. $\square$

Similarly to the case for expected regret discussed in Section 4.4, we also have that $\eta = 0$ in the zero-transaction cost scenario (that is, $c = 0$) or x_t is independently distributed from x_{t-1} for $t \in [T]$. This can be proven in a similar way to Theorem 9.

4.6 General $\mathcal{D}_t$

Herein we will assume that $c = 0$, as Theorem 9 shows that we cannot hope for sublinear expected regret in the presence of transaction costs.

Let $\mathcal{D}_t$ be parametrised by two parameters, μ_t (mean) and σ_t (standard deviation). We will look only at log-returns (rather than absolute returns); this is quite a standard notion in financial mathematics for a number of reasons [BS73, KS91, Mer71]. In particular, we will say that the log-predicted returns ($\ln \tilde{x}_t$) are distributed around the mean (defined as the log-observed returns, $\ln x_t$) with some standard deviation σ_t. Formally,

$$\ln \tilde{x}_t \sim \mathcal{D}_{\ln x_t, \sigma_t^2}$$

for some distribution $\mathcal{D}$, or simply

$$\tilde{x}_t \sim \ln \mathcal{D}_{\ln x_t, \sigma_t^2}$$

for short-hand. As the portfolio vectors are multi-dimensional, we denote (for notational convenience)

$$\sigma_t := (\sigma_t, \cdots, \sigma_t) \in \mathbb{R}_+^m,$$

and apply the logarithm and distribution element-wise: that is,

$$\ln x_t = \ln \left(x_t(1), \cdots, x_t(m) \right) = \left(\ln x_t(1), \cdots, \ln x_t(m) \right),$$

and, thus,

$$\ln \mathcal{D}_{\ln x_t, \sigma_t^2} = \ln \mathcal{D}_{\ln x_t(1), \sigma_t^2} \times \cdots \times \ln \mathcal{D}_{\ln x_t(m), \sigma_t^2}.$$

Note that applying Chebyshev's inequality

$$Pr(|x - \mu| \geq z) \leq \sigma_t^2 / z^2$$

gives the following guarantee for some generalised distributions $\mathcal{D}_t$:

$$\begin{aligned}
\mathbb{E}[R] &\leq 2 \sum_{t=1}^{T} \int_0^\infty \Pr_{\tilde{x}_t \sim \mathcal{D}_t(x_t)} [\tilde{x}_t \notin (e^{-z} x_t, e^z x_t)] \, dz \\
&\leq 2 \sum_{t=1}^{T} \int_0^{\sigma_t} 1 \, dz + \int_{\sigma_t}^\infty \frac{\sigma_t^2}{z^2} \, dz = 4 \sum_{t=1}^{T} \sigma_t,
\end{aligned}$$

where the last inequality is due to Chebyshev's and trivially bounding the probability by 1 when $z < \sigma_t$. Therefore, a necessary and sufficient condition for sublinear expected regret is $\sigma_t \to 0$ as $t \to \infty$, which means that one can only hope for sublinear expected regret if the trader learn the predictions over time. We will show in Section 4.7 that we can substantially improve on the above factor of 4 for particular cases of $\mathcal{D}_t$.

Note that Chebyshev's inequality is unable to achieve a reasonable bound on the variance of regret since

$$\begin{aligned}\mathrm{Var}[R] &\leq 4\sum_{t=1}^{T}\int_0^\infty \Pr_{\tilde{x}_t\sim\mathcal{D}_t(x_t)}[\tilde{x}_t \notin (e^{-\sqrt{z}}x_t, e^{\sqrt{z}}x_t)]\, dz \\ &\leq 4\sum_{t=1}^{T}\int_0^\infty \frac{\sigma_t^2}{z}\, dz = 4\sum_{t=1}^{T}\sigma_t^2\int_0^\infty \frac{1}{z}\, dz,\end{aligned}$$

and the integral of $\frac{1}{z}$ is not convergent; this is the well-known harmonic series. However, finite bounds on the variance of regret (based on σ_t's) can be derived for particular distributions $\mathcal{D}_t$, as will be shown in Section 4.7.

4.7 Special Cases of $\mathcal{D}_t$

Given the above results are for a generically distributed $\tilde{x}_t \sim \mathcal{D}_t(x_t)$, we will now look at some particular examples of $\mathcal{D}_t$ and compute the required quality of prediction in order to achieve sublinear expected regret and variance of regret. The table below gives a summary of the upper bounds obtained for particular distributions $\mathcal{D}_t$.

	$\mathbb{E}[R]$	$\mathrm{Var}[R]$
Log-uniform	$\sqrt{3}\sum\sigma_t$	$4\sum\sigma_t^2$
Log-linear	$2\sqrt{\frac{2}{3}}\sum\sigma_t$	$4\sum\sigma_t^2$
Log-normal	$2.1\sum\sigma_t$	$7.5\sum\sigma_t^2$

4.7.1 Log-Uniformly Distributed Predictions

Suppose that $\tilde{x}_t \sim \ln\mathcal{U}_{\ln x_t,\sigma_t^2}$, where $\mathcal{U}$ is the uniform distribution on the log-returns between the range $[-\sqrt{3}\sigma_t, \sqrt{3}\sigma_t]$ with standard deviation σ_t and the following probability density function

$$f(y) = \begin{cases} \frac{1}{2\sqrt{3}\sigma_t} & \text{if } 0 \leq |y - \ln x_t| \leq \sqrt{3}\sigma_t, \\ 0 & \text{otherwise.}\end{cases}$$

In this case,

$$\Pr_{\tilde{x}_t \sim \ln \mathcal{U}_{\ln x_t, \sigma_t^2}}[\tilde{x}_t \notin (e^{-z}x_t, e^z x_t)] = \begin{cases} 1 - \frac{z}{\sqrt{3}\sigma_t} & \text{if } 0 \le z \le \sqrt{3}\sigma_t, \\ 0 & \text{if } z > \sqrt{3}\sigma_t, \end{cases}$$

$$\Pr_{\tilde{x}_t \sim \ln \mathcal{U}_{\ln x_t, \sigma_t^2}}[\tilde{x}_t \notin (e^{-\sqrt{z}}x_t, e^{\sqrt{z}} x_t)] = \begin{cases} 1 - \frac{\sqrt{z}}{\sqrt{3}\sigma_t} & \text{if } 0 \le z \le 3\sigma_t^2, \\ 0 & \text{if } z > 3\sigma_t^2. \end{cases}$$

Therefore, applying Theorem 8 and Theorem 10 yields

$$\mathbb{E}[R] \le 2 \sum_{t=1}^{T} \int_0^{\sqrt{3}\sigma_t} \left(1 - \frac{z}{\sqrt{3}\sigma_t}\right) dz = \sqrt{3} \sum_{t=1}^{T} \sigma_t,$$

$$\mathrm{Var}[R] \le 4 \sum_{t=1}^{T} \int_0^{3\sigma_t^2} \left(1 - \frac{\sqrt{z}}{\sqrt{3}\sigma_t}\right) dz = 4 \sum_{t=1}^{T} \sigma_t^2.$$

4.7.2 Log-Linearly Distributed Predictions

Suppose that $\tilde{x}_t \sim \ln \mathcal{L}_{\ln x_t, \sigma_t^2}$, where $\mathcal{L}$ is the linearly-decreasing distribution with largest density at the mean, $\ln x_t$, and with standard deviation σ_t. More precisely, it has the following probability density function

$$f(y) = \begin{cases} \frac{1}{\sqrt{6}\sigma_t} - \frac{|y - \ln x_t|}{6\sigma_t^2} & \text{if } 0 \le |y - \ln x_t| \le \sqrt{6}\sigma_t, \\ 0 & \text{otherwise.} \end{cases}$$

In this case,

$$\Pr_{\tilde{x}_t \sim \ln \mathcal{L}_{\ln x_t, \sigma_t^2}}[\tilde{x}_t \notin (e^{-z}x_t, e^z x_t)] = \begin{cases} 1 - 2\frac{z}{\sqrt{6}\sigma_t} + \frac{z^2}{6\sigma_t^2} & \text{if } 0 \le z \le \sqrt{6}\sigma_t, \\ 0 & \text{if } z > \sqrt{6}\sigma_t, \end{cases}$$

$$\Pr_{\tilde{x}_t \sim \ln \mathcal{L}_{\ln x_t, \sigma_t^2}}[\tilde{x}_t \notin (e^{-\sqrt{z}}x_t, e^{\sqrt{z}} x_t)] = \begin{cases} 1 - 2\frac{\sqrt{z}}{\sqrt{6}\sigma_t} + \frac{z}{6\sigma_t^2} & \text{if } 0 \le z \le 6\sigma_t^2, \\ 0 & \text{if } z > 6\sigma_t^2. \end{cases}$$

Therefore, applying Theorem 8 and Theorem 10 yields

$$\mathbb{E}[R] \leq 2\sum_{t=1}^{T} \int_{0}^{\sqrt{6}\sigma_t} \left(1 - 2\frac{z}{\sqrt{6}\sigma_t} + \frac{z^2}{6\sigma_t^2}\right) dz = 2\sqrt{\frac{2}{3}} \sum_{t=1}^{T} \sigma_t,$$

$$\mathrm{Var}[R] \leq 4\sum_{t=1}^{T} \int_{0}^{6\sigma_t^2} \left(1 - 2\frac{\sqrt{z}}{\sqrt{6}\sigma_t} + \frac{z}{6\sigma_t^2}\right) dz = 4\sum_{t=1}^{T} \sigma_t^2.$$

4.7.3 Log-Normally Distributed Predictions

We will now look at the particular case when $\mathcal{D}_t$ is log-normally distributed (analogous to Geometric Brownian Motion). Suppose that $\tilde{x}_t \sim \ln \mathcal{N}_{\ln x_t, \sigma_t^2}$, then

$$\mathbb{E}[R] \leq 4\sum_{t=1}^{T} \int_{0}^{\infty} \Pr_{y \sim \mathcal{N}_{0,1}} [y > z/\sigma_t] \, dz.$$

To achieve a sublinear expected regret then depends on the ability to obtain an appropriate sequence of predictions with σ_t such that

$$\frac{1}{T} \sum_{t=1}^{T} \int_{0}^{\infty} \Pr_{y \sim \mathcal{N}_{0,1}} [y > z/\sigma_t] \, dz \to 0,$$

as $T \to \infty$. This has a very natural interpretation; the above condition can be viewed as an integral over the tail probabilities of the standard normal distribution, where the size of the tail is determined by σ_t. Similarly, the variance of regret in this case can be bounded as

$$\mathrm{Var}[R] \leq 8\sum_{t=1}^{T} \int_{0}^{\infty} \Pr_{y \sim \mathcal{N}_{0,1}} [y > \sqrt{z}/\sigma_t] \, dz.$$

Next we give a way to simplify the bound on the expected regret in terms of σ_t's. Recall that the probability density function of the standard normal distribution $\mathcal{N}_{0,1}$ is

$$f(x) = \frac{1}{\sqrt{2\pi}} \, e^{-x^2/2} \, .$$

In order to bound the tail probability in the expected-regret expression above, we will use the following function $\hat{f}$ to upper bound f:

$$\hat{f}(x) = \begin{cases} 1 - \frac{x^2}{2} + \frac{x^4}{8} & \text{for } 0 \le x < 1, \\ e^{-x/2} & \text{for } x \ge 1. \end{cases}$$

The first part of the function $\hat{f}$ is the Maclaurin series for $e^{-x^2/2}$ with three terms and is an upper bound on $e^{-x^2/2}$. Hence, note that, for all $x \ge 0$, $f(x) \le \hat{f}(x)/\sqrt{2\pi}$. The graph in Figure 4.1 shows the relationship between $f(x)$ (without the scaling factor $\frac{1}{\sqrt{2\pi}}$) and the components of $\hat{f}$.

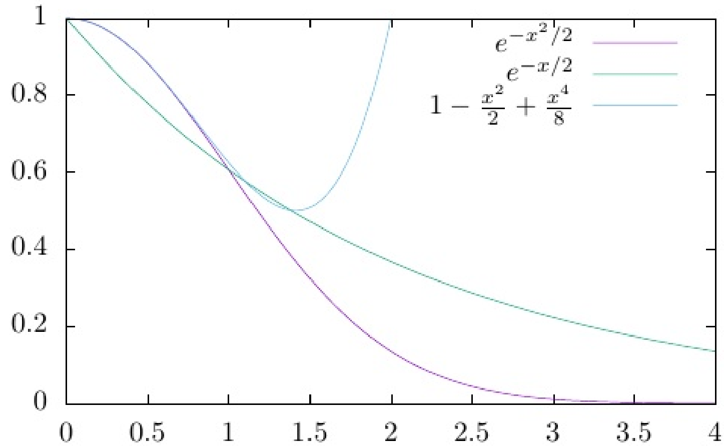

Figure 4.1: Unscaled standard normal p.d.f. and a tight upper bound for it

Hence, for $a \ge 1$, we can write

$$\Pr_{y \sim \mathcal{N}_{0,1}}[y > a] \le \frac{1}{\sqrt{2\pi}} \int_a^\infty e^{-x/2}\, dx \le \frac{1}{\sqrt{2\pi}} \left[-2e^{-x/2}\right]_a^\infty \le \frac{1}{\sqrt{2\pi}} 2e^{-a/2}.$$

On the other hand, for $0 \le a < 1$,

$$\begin{aligned}
\Pr_{y \sim \mathcal{N}_{0,1}}[y > a] &\le \frac{1}{\sqrt{2\pi}} \left(\int_a^1 \left(1 - \frac{x^2}{2} + \frac{x^4}{8}\right) dx + \int_1^\infty e^{-x/2} dx \right) \\
&= \frac{1}{\sqrt{2\pi}} \left(\left[x - \frac{x^3}{6} + \frac{x^5}{40} \right]_a^1 + 2e^{-1/2} \right) \\
&= \frac{1}{\sqrt{2\pi}} \left(\frac{103}{120} - a + \frac{a^3}{6} - \frac{a^5}{40} + 2e^{-1/2} \right).
\end{aligned}$$

We can substitute the bounds above to obtain

$$\begin{aligned}
&\int_0^\infty \Pr_{y \sim \mathcal{N}_{0,1}}[y > z/\sigma_t]\, dz \\
&= \int_0^{\sigma_t} \Pr_{y \sim \mathcal{N}_{0,1}}[y > z/\sigma_t]\, dz + \int_{\sigma_t}^\infty \Pr_{y \sim \mathcal{N}_{0,1}}[y > z/\sigma_t]\, dz \\
&\le \frac{1}{\sqrt{2\pi}} \int_0^{\sigma_t} \left(\frac{103}{120} - \frac{z}{\sigma_t} + \frac{z^3}{6\sigma_t^3} - \frac{z^5}{40\sigma_t^5} + 2e^{-1/2} \right) dz \\
&\quad + \frac{1}{\sqrt{2\pi}} \int_{\sigma_t}^\infty 2e^{-z/(2\sigma_t)} dz \\
&= \frac{1}{\sqrt{2\pi}} \left(\left[\frac{103z}{120} - \frac{z^2}{2\sigma_t} + \frac{z^4}{24\sigma_t^3} - \frac{z^6}{240\sigma_t^5} \right]_0^{\sigma_t} + 2\sigma_t e^{-1/2} + 4\sigma_t e^{-1/2} \right) \\
&= \frac{\sigma_t}{\sqrt{2\pi}} \left(\frac{103}{120} - \frac{1}{2} + \frac{1}{24} - \frac{1}{240} + 6e^{-1/2} \right) \\
&< 1.61 \cdot \sigma_t.
\end{aligned}$$

Hence, the expected regret can be bounded as

$$\mathbb{E}[R] < 6.5 \sum_{t=1}^T \sigma_t.$$

This is larger than the factor of 4 in the bound derived in Section 4.6, although we can improve the above factor of 6.5 further by including more terms of the Maclaurin series in $\hat{f}$. Analogous to the above, we can give an upper bound on the variance of regret; upper bounding the normal p.d.f. by $\hat{f}$, we will need to evaluate indefinite integral

$$\int e^{-\sqrt{x}/k}\, dx,$$

for some constant k. Using integration by parts, we can show that

$$\int e^{-\sqrt{x}/k}\,dx = -2k\left(\sqrt{x}+k\right)e^{-\sqrt{x}/k} + c.$$

Similar to the expectation bound, using the above equations we get

$$\begin{aligned}
&\int_0^\infty \Pr_{y\sim\mathcal{N}_{0,1}}[y > \sqrt{z}/\sigma_t]\,dz \\
&= \int_0^{\sigma_t^2} \Pr_{y\sim\mathcal{N}_{0,1}}[y > \sqrt{z}/\sigma_t]\,dz + \int_{\sigma_t^2}^\infty \Pr_{y\sim\mathcal{N}_{0,1}}[y > \sqrt{z}/\sigma_t]\,dz \\
&\le \frac{1}{\sqrt{2\pi}}\int_0^{\sigma_t^2}\left(\frac{103}{120} - \frac{z^{1/2}}{\sigma_t} + \frac{z^{3/2}}{6\sigma_t^3} - \frac{z^{5/2}}{40\sigma_t^5} + 2e^{-1/2}\right)dz \\
&\quad + \frac{1}{\sqrt{2\pi}}\int_{\sigma_t^2}^\infty 2e^{-\sqrt{z}/(2\sigma_t)}dz \\
&= \frac{1}{\sqrt{2\pi}}\left(\left[\frac{103z}{120} - \frac{2z^{3/2}}{3\sigma_t} + \frac{z^{5/2}}{15\sigma_t^3} - \frac{z^{7/2}}{140\sigma_t^5}\right]_0^{\sigma_t^2} + 2\sigma_t^2 e^{-1/2}\right) \\
&\quad + \frac{1}{\sqrt{2\pi}}\left[-8\sigma_t(\sqrt{z}+2\sigma_t)e^{-\sqrt{z}/(2\sigma_t)}\right]_{\sigma_t^2}^\infty \\
&= \frac{\sigma_t^2}{\sqrt{2\pi}}\left(\frac{103}{120} - \frac{2}{3} + \frac{1}{15} - \frac{1}{140} + 2e^{-1/2} + 24e^{-1/2}\right) \\
&< 6.4\cdot\sigma_t^2\,.
\end{aligned}$$

Hence, we obtain that

$$\mathrm{Var}[R] < 52\sum_{t=1}^T \sigma_t^2\,.$$

Analogous to the analysis of expected regret, we can further improve the above factor of 52 by including more terms of the Maclaurin series in $\hat{f}$. For example, if we use the Taylor series expansion of $f(x)$ at $x = 1$ to upper bound $f(x)$,

$$\hat{f}(x) = \begin{cases} e^{-1/2}\cdot\left(\frac{39}{24} + \frac{x}{6} - \frac{5x^2}{4} + \frac{x^3}{2} - \frac{x^4}{24}\right) & \text{for } 0 \le x < 2, \\ e^{-x} & \text{for } x \ge 2, \end{cases}$$

will yield an expected regret bound with a constant of 2.1 and a variance of regret bound with a constant of 7.5, much tighter than the general bound derived in Section 4.6.

4.8 Portfolio Computation

The θ function can be viewed as a variant of the earth mover's distance, which, in turn, can be formulated as a transportation or flow problem and solved using linear programming (LP). Here, we present an LP for computing $\hat{b}$ (and, hence, for similarly computing b^*) by first computing θ. The input to the computation is the original allocation vector

$$w = (w_1, \ldots, w_m)$$

(corresponding to $K\hat{b}$, where K is the total wealth before rebalancing and $b \in \mathcal{B}$) and the target portfolio vector given as

$$q = (q_1, \ldots, q_m)$$

(with $\sum_i q_i = 1$). The variables of the LP are the wealth W resulting after the rebalancing and f_{ij}, for $i, j \in [m]$, that corresponds to wealth that needs to be transferred from Asset i to Asset j.

$$\begin{aligned}
&\max W \\
&\text{subject to} \\
&\sum_{j \in [m]} f_{ij} \leq w_i && \forall i = 1, \ldots, m && (4.6) \\
&f_{jj} + (1-c) \cdot \sum_{\substack{i \in [m] \\ i \neq j}} f_{ij} \geq W \cdot q_j && \forall j = 1, \ldots, m && (4.7) \\
&f_{ij} \geq 0 && \forall i, j = 1, \ldots, m && (4.8)
\end{aligned}$$

The constraints in (4.6) ensure that the wealth transferred out of each asset is bounded by the current wealth in that asset. The constraints in (4.7) ensure that the wealth that stays in each asset plus the wealth transferred into that asset, minus the incurred transaction costs, are sufficient to reach the target portfolio vector with a total wealth of W. Finally, the flow of wealth will always be positive by (4.8). Note that the sets of constraints in (4.6) and (4.7) will be satisfied tightly in an optimal solution. First of all, for any $i \in [m]$, total flow

$$\sum_{j\in[m]} f_{ij}$$

out of Asset i will be equal to w_i, because any increase in the total flow

$$\sum_{i,j} f_{ij}$$

can be distributed over the assets according to q, creating slack in each constraint in (4.7) and allowing a strictly larger value for W. Similarly, if the flow into any Asset j, given as

$$f_{jj} + (1-c)\cdot \sum_{i\in[m],i\neq j} f_{ij},$$

was strictly larger than $W \cdot q_j$, then this excess flow can be shifted to other assets to create slack in each constraint in (4.7), which, in turn, allows W to be increased. The fact that the constraints in (4.6) and (4.7) are tight for an optimal solution shows that all the wealth in the previous time step is used during rebalancing and the resulting portfolio distribution adheres to q. Finally, by the maximisation of W, we get that the optimal solution to the LP gives the value of θ, and also $\hat{b}$ (by summing up all of the flow in/out of each asset f_{ij}). In the case where there are multiple optimal solutions, we choose the one with the lowest

$$\sum_{j\in[m]} f_{ij},$$

for $i \in [m]$ sequentially; that is, we break ties by minimising the outflow from the smallest to the largest i.

Chapter 5

Derivatives Pricing

Another important problem in economics is the pricing of financial derivatives such as options, for which Black and Scholes [BS73] won the Nobel prize in 1997. Again, their approach relies on the assumption that the underlying price process follows a Geometric Brownian motion (GBM). In practice, this has often been observed to be inconsistent with the GBM assumption, for example, the existence of volatility smile in the options market [Hul06].

In 2006, DeMarzo et al. [DKM06, Man07] showed that the regret of a trading strategy naturally give rise to an upper bound for options price in the model-free sense. However, their bound directly depends on the volatility of the underlying asset, much like in the Black–Scholes framework. In particular, they showed

$$C(K,T) \leq \Theta(\sqrt{Q})\,,$$

where $C(K,T)$ denotes the price of a call option, without making any additional assumptions on the underlying price process, and Q is the quadratic variation (otherwise known as volatility) defined as

$$Q = \sum_{t=1}^{T} r_t^2\,,$$

where

$$r_t = x_t - 1$$

is the returns of the underlying asset between time $t-1$ and t. We will show

how to bound the options price (with high probability) against a quantity that doesn't directly depend on the volatility of the underlying asset. Instead we suppose that we have access to some predictions on the returns of the underlying asset, and follow the expected regret bound from Chapter 4 to obtain a model-free probabilistic bound on the option price,

$$\Pr\left[C(T) \leq e^{k\mu} - 1\right] \geq 1 - \frac{1}{k},$$

where $C(T)$ denotes the price of an at-the-money call option, and μ depends on the quality of these predictions (in particular, the tail probabilities of mis-estimation). We do this by constructing a portfolio that replicates the option payoff, and show that competitive regret is attainable against this benchmark, hence implying a bound on the options price as to preserve no-arbitrage.

This has the particular advantage that the option price is only limited by one's ability to predict the future, regardless of the underlying asset's future volatility, and does not limit the application of the results to particular stochastic price process (like the GBM assumption in Black–Scholes).

We will first prove a bound on the price of (at-the-money) call option, then extend this to a number of exotic derivatives. Throughout, we will assume zero interest rate. Note that the results can be easily extended to more general cases, for example, non-zero interest rate, arbitrary strike price, put options, other exotic derivatives, etc.

5.1 Arbitrage

Throughout the chapter we will assume *no arbitrage*. To define this precisely, given two securities (e.g. financial asset, portfolio model, etc), say X_1 and X_2, such that the payoff of X_1 *always* dominate X_2 on any future outcome (i.e., price paths). Then, we say there exists an *arbitrage* if the current value of X_2 is larger than the current value of X_1 (intuitively, meaning that an investor could buy X_1 and sell X_2 to guarantee a risk-free profit).

Conversely, there is *no arbitrage* if the current value of X_1 is larger than the current value of X_2. This will form the basis of constructing a bound on the options price based on the payoff of the corresponding trading algorithm.

5.2 Introduction to Options

A European call (or put) option is a contract that gives the buyer the right but not the obligation to buy (or sell) a certain units of the underlying asset at a pre-specified "strike" price K on the "expiration date" T. This essentially provides the buyer with insurance against the change in price of the asset. The buyer of the option pays for this right through the "premium" of the option: a fee for the seller to compensate for the risk of a potential loss. We denote the premium of a call option as $C(K,T)$. The payoff of a call option at expiry is given by

$$\max\{V_T - K, 0\},$$

where V_t is the value of the risky asset at time t. The question of interest is to determine the fair value (the premium) of the option at the initial time $t = 0$. Black and Scholes [BS73] showed how to calculate the exact current valuation of an option by replicating the option payoff using a dynamic trading strategy. They assumed the no-arbitrage condition and that the underlying asset price process follows a GBM. The fair option premium in their framework then depends on the volatility (standard deviation of the GBM) of the underying asset, among other things.

DeMarzo et al. [DKM06] took a different approach and made no assumption on the underlying price process, except a bound on the single-period return: that is,

$$|r_t| \leq M$$

for all $t \in [T]$. They showed how a regret bound can be translated into an upper bound for options premium. First define the optimal benchmark portfolio as the better of $(1,0)$ and $(0,1)$ across all time steps, where the first and second assets are the risky and risk-free assets, respectively. Note that this is an even more restrictive than the 2-asset CRP strategy (as previously discussed in Chapter 3).

We denote the wealth of this optimal benchmark at time t as S_t^*, and normalize S_t so that $S_0 = 1$. Throughout the rest of the chapter, we will assume

$$K = S_0$$

(at-the-money option), as the results can be easily generalized. We use $C(T)$ as shorthand for the fair value of an at-the-money option with time to expiry T. Without losing generality, we also normalize V_t so that $V_0 = 1$ (i.e., the underlying asset has initial price of 1), and thus the payoff of such at-the-money call option becomes $\max\{V_T - 1, 0\}$.

5.3 Black-Scholes

Perhaps the most well-known model for options pricing is Black–Scholes [BS73] where they showed that, given some assumptions, the fair (no-arbitrage) value of a call option is

$$C(K,T) = N(d_1)S_t - N(d_2)Ke^{-rT}$$

where

$$d_1 = \frac{1}{\sigma\sqrt{T}}\left[\log\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T\right],$$

$$d_2 = d_1 - \sigma\sqrt{T},$$

$N(\cdot)$ is the c.d.f. of a standard normal distribution,

T is the time to expiry,

r is the risk-free interest rate,

K is the strike price,

S_t is the price of the underlying asset at time t,

σ is the volatility of returns of the underlying asset.

They proved this by replicating the options payoff as a dynamic hedging strategy, much like the approach in DeMarzo et al. [DKM06], although they assumed that the underlying asset price process follows a GBM. More details of the Black-Scholes model can be found in [Hul06].

5.4 Pricing from Trading Strategy

Now we provide a formal proof of the argument from [DKM06] that relates option price to the regret of a trading strategy, in the model-free sense.

Lemma 11 *Suppose $\hat{S}_T$ is the wealth obtained from a trading strategy on the risky asset with price V_t at time t, define $S^*_T = \max\{V_T, 1\}$, and the regret is bounded as*

$$\max_{x^T} \left(\log S^*_T - \log \hat{S}_T \right) \leq R_O \,.$$

Then, the at-the-money call option premium can be bounded as

$$C(T) \leq e^{R_O} - 1 \,.$$

Proof. Rearranging the regret term, we get

$$\hat{S}_T \geq e^{-R_O} S^*_T \,,$$

for all price sequences x^T, meaning that the trading algorithm incurs a loss of no more than e^{-R_O} fraction of the optimal strategy in the worst-case. Scaling up the trading algorithm and starting with $\hat{S}_0 = e^{R_O}$ wealth initially will ensure that it performs no worse than the optimal portfolio (that starts with $S^*_0 = 1$), that is, $\hat{S}_T \geq S^*_T$.

Let's suppose that the value of the call option (whose payoff at expiry is replicated by $S^*_T - 1$) exceeds $e^{R_O} - 1$, then the trader could sell the option at time $t = 0$ (for some premium $C(T) > e^{R_O} - 1$) and borrow e^{R_O} cash (at zero interest rate) to run the trading algorithm starting with $\hat{S}_0 = e^{R_O}$. The wealth at time T, after paying back the loan, is then

$$\begin{aligned}
&[C(T) - (S^*_T - 1)] + [\hat{S}_T - e^{R_O}] \\
&= [C(T) - (e^{R_O} - 1)] + [\hat{S}_T - S^*_T], \text{ from re-arranging} \\
&> 0, \text{ due to } C(T) > e^{R_O} - 1 \text{ and } \hat{S}_T \geq S^*_T
\end{aligned}$$

where

$C(T)$ is the premium received from selling the option,

$S^*_T - 1$ is the payoff of the option which the trader pays at expiry,

$\hat{S}_T$ is the wealth received from the trading algorithm,

e^{R_O} is paid back for the loan, interest free.

Therefore, the trader has managed to lock in a guaranteed gain no matter the price outcome. To preserve no arbitrage then implies that the price of the option at $t = 0$ cannot exceed $e^{R_O} - 1$. $\square$

5.5 Vanilla Options

We will now apply the results from Theorem 8 and Theorem 11 to derive an upper bound on the value of various options: firstly, for vanilla option, and then, for more exotic derivatives in the next section. The bounds will be probabilistic and will depend on the accuracy of the price predictions in the trading strategy.

Theorem 12 *Given predictions $\tilde{x}_t \sim \mathcal{D}_t(x_t)$, for some probability distribution $\mathcal{D}_t$ for each $t \in [T]$, and for every $0 \leq k \leq 1$,*

$$\Pr[C(T) \leq e^{k\mu} - 1] \geq 1 - \frac{1}{k}$$

where

$$\mu = 2 \sum_{t=1}^{T} \int_0^\infty \Pr_{\tilde{x}_t \sim \mathcal{D}_t(x_t)} [\tilde{x}_t \notin (e^{-z} x_t, e^z x_t)] \, dz \, .$$

Proof. Firstly, we bound the regret of the portfolio selection algorithm from Definition 5 (with wealth $\hat{S}_T$ against the optimal greedy portfolio with wealth S_T^G) against the log-wealth of the strategy that replicates an options payoff,

$$S_T^* = \max\{V_T, 1\},$$

except shifted by exactly one unit; as the payoff of an at-the-money call option is $S_T^* - 1 = \max\{V_T - 1, 0\}$. In particular,

$$\begin{aligned} R_O &:= \max_{x^T} \left(\log S_T^* - \log \hat{S}_T \right) \\ &= \max_{x^T} \left((\log S_T^* - \log S_T^G) + (\log S_T^G - \log \hat{S}_T) \right) \\ &\leq \max_{x^T} \left(\log S_T^G - \log \hat{S}_T \right) \\ &=: R_G \,, \end{aligned}$$

where the inequality arises from the fact that the optimal greedy portfolio (from Definition 5), with wealth S_T^G, will always perform no worse than S_T^*. In particular,

$$\log S_T^* = \log \left(\max\{\prod_{t=1}^{T} 1 + r_t, 1\} \right) = \max\{\sum_{t=1}^{T} \log(1 + r_t), 0\}$$

and

$$\log S_T^G = \sum_{t=1}^{T} |\log(1 + r_t)|,$$

therefore,

$$\log S_T^G \geq \log S_T^*$$

across all price paths. Using Theorem 8 gives an upper bound on the expected regret $\mathbb{E}[R_G]$ and, hence, also on $\mathbb{E}[R_O]$. Now, applying Markov's inequality,

$$\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a} \,,$$

with $X = R_G$ and $a = k\mathbb{E}[R_G]$, we obtain a probabilistic bound of R_O being within some multiple of the expected regret,

$$\Pr\left[R_O \leq k\mathbb{E}[R_G]\right] \geq \Pr\left[R_G \leq k\mathbb{E}[R_G]\right] \geq 1 - \frac{1}{k} \,.$$

Combining this probability bound with Lemma 11, we obtain the desired probabilistic bound on the option premium. $\square$

Note that to obtain this bound for option price in practice does not require to actually implement the trading algorithm, hence does not require explicit knowledge of the predictions $\tilde{x}_t$. Rather, it just requires knowledge of how the predictions are distributed with respect to the price outcome, namely, $\mathcal{D}_t$. Recall from Section 4.3 that

$$\Pr_{\tilde{x}_t \sim \mathcal{D}_t(x_t)}[\tilde{x}_t \notin (e^{-z}x_t, e^z x_t)]$$

can be interpreted as the tail probabilities of $\mathcal{D}_t$, which are the mis-estimations of the prediction. We now simplify the statement of the theorem above in the case where the $\mathcal{D}_t$'s are parameterized.

Corollary 13 *Given predictions* $\tilde{x}_t \sim \ln \mathcal{D}_{\ln x_t, \sigma_t^2}$ *for each* $t \in [T]$*,*

$$C(T) \leq e^{O\left(\sum_{t=1}^T \sigma_t\right)} \approx O\left(\sum_{t=1}^T \sigma_t\right)$$

with high probability, where the last term holds for small $\sum_{t=1}^T \sigma_t$*.*

Proof. Recall that in Section 4.6, we showed that

$$\mathbb{E}[R_G] \leq 4 \sum_{t=1}^T \sigma_t.$$

Applying Theorem 12 with the fact that $\mu = \mathbb{E}[R_G]$ gives

$$C(T) \leq e^{4k \sum_{t=1}^T \sigma_t} - 1 = \left(\prod_{t=1}^T e^{4k\sigma_t}\right) - 1$$

with probability at least $1 - \frac{1}{k}$. Note also that for small enough $\sum_{t=1}^T \sigma_t$, the above bound can be approximated by the MacLaurin series as

$$C(T) \leq \left(\prod_{t=1}^T e^{4k\sigma_t}\right) - 1 \approx 4k \sum_{t=1}^T \sigma_t.$$

For large enough k, the above bound occurs with high probability. $\square$

First, let us assume that the underlying asset price evolves as a GBM and compare this to Black–Scholes. For zero interest rate, the value of at-the-money call option as implied by Black–Scholes can be simplified as

$$C(T) = S_0 \cdot \Pr_{x \sim \mathcal{N}_{0,1}} \left[x \in (-\sqrt{Q_\sigma/2}, \sqrt{Q_\sigma/2}) \right],$$

where

$$Q_\sigma = \sum_{t=1}^{T} \sigma_t^2$$

and σ_t is the standard deviation of x_t in this case. This can be derived easily from the general Black–Scholes formula from Section 5.3 by setting $S_t = K$ and $r = 0$. Note that, since probabilities are always between $[0, 1]$, we have $C(T) \leq S_0$ trivially. The quantity $\sqrt{Q_\sigma}$ can be thought of as the standard deviation of the returns of the underlying asset between $t = 0$ and $t = T$. Therefore, as one expects the underlying asset price to move more, the price of the option would be higher.

Secondly, the result from DeMarzo et al. [DKM06] showed that

$$C(T) \leq \Theta(\sqrt{Q})$$

using a regret bound from a portfolio selection algorithm which they call "Generic." Similarly to Black–Scholes, the price of the option also grows with the volatility of the underlying asset, although this does not necessarily guarantee an upper bound on S_0, although this can also be imposed trivially.

The main difference of both Black–Scholes and the bound from DeMarzo et al. [DKM06] with our result in Corollary 13 is that, our result has no dependency on the volatility but rather on the mis-estimations of the predictions $\tilde{x}_t$. Therefore, with good enough predictions, we can guarantee a (possibly tighter) bound on the fair options price in a model-free sense, regardless of how much the underlying asset price moves until time T. This gives a particular advantage when one has access to a learning model that is able to improve the predictive capability over time, or if the prediction accuracy $\mathcal{D}_t$ is independent of the volatility.

5.6 Exotic Derivatives

Now we will show that our results can be extended to give bounds on more complicated financial contracts: exotic derivatives. We consider n derivatives

$$X_1, \ldots, X_n$$

and their corresponding values at time t,

$$V_{1,t}, \ldots, V_{n,t}.$$

Without loss of generality, assume $V_{i,0} = 1$, for all $i \in [n]$ (by normalization), and assume that asset price is always non-negative: i.e., $V_{i,t} \geq 0$, for all $i \in [n]$ and $t \in [T]$. We will look at various exotic derivatives that have payoff at time T of the form $\max_i V_{i,T}$. We denote such a derivative as

$$\Psi(X_1, \ldots, X_n),$$

and its value at $t = 0$ as

$$\Phi(X_1, \ldots, X_n).$$

In the special case of a single asset (that is, $n = 1$), $\Phi(X)$ is simply the value of the derivative X at $t = 0$. Now we define a few of these derivatives.

EX An exchange option $\mathbf{EX}(X_1, X_2, T)$ allows the holder to exchange asset X_2 for asset X_1 at time T, making its payoff $\max\{V_{1,T} - V_{2,T}, 0\}$.

SH An (at-the-money) shout option $\mathbf{SH}(T)$ allows its holder to "shout" and lock in a minimum value for the payoff at one time $0 \leq t \leq T$ during the lifetime of the option. Its payoff at time T is, therefore, $\max\{V_{1,T} - 1, V_{1,t} - 1, 0\}$. If the holder does not shout, the payoff is $\max\{V_{1,T} - 1, 0\}$.

AS An average strike call option $\mathbf{AS}(T)$ is a type of Asian option that allows its holder to get the difference between the final stock price and the average stock price, namely, a payoff of $\max\{V_{1,T} - \bar{V}_{1,T}, 0\}$ where $\bar{V}_{1,T}$ is the average of $\{V_{1,T}\}_{t \in [T]}$.

Notice that all of the above (and, in fact, most financial derivatives) have payoff functions that can be stated as a maximum of a number of different components. We will make this notion more precise in order to prove a bound on their value.

Definition 6 *Let* **A** *be a trading algorithm with initial value* $S_0 = 1$ *and let* $\beta_1, \ldots, \beta_n > 0$*.* **A** *is said to have* $(\beta_1, \ldots, \beta_n)$*-multiplicative regret w.r.t. derivatives* $X_1, \ldots, X_n$ *if, for every path path* x^T *and every* $i \in [n]$*,*

$$S_T \geq \beta_i V_{i,T}.$$

We have the following lemma from Gofer et al. [GM11a] that gives a bound on the value of these derivatives with multiplicative regret, as a function of the value of their constituent parts.

Lemma 14 ([GM11a]) *If there exists a trading algorithm with a* $(\beta_1, \ldots, \beta_n)$*-multiplicative regret w.r.t. derivatives* $X_1, \ldots, X_n$*, then*

$$\Phi(X_1, \ldots, X_n) \leq 1/\beta,$$

where $\beta = \min_{1 \leq i \leq n} \beta_i$*.*

Proof. By Definition 6 and $\beta = \min_{1 \leq i \leq n} \beta_i$, we have that

$$S_T \geq \beta \max_{1 \leq i \leq n} V_{i,T}.$$

Therefore, the payoff of the algorithm dominates β units of the corresponding derivative $\Psi(X_1, \ldots, X_n)$. By the arbitrage-free assumption,

$$\Phi(X_1, \ldots, X_n) \leq 1/\beta.$$

□

Note that Lemma 14 is an extension of Lemma 11 in the following sense. We consider again the scenario where X represents the (at-the-money) vanilla call option. Now S_T^* mimics the payoff of a call option, but only one that costs exactly one unit more in value (due to payoff function being exactly one unit more). The regret bound for R would tell us that $\beta = e^{-R}$, and, thus,

$$\Phi(X) \le 1/\beta = e^R.$$

But this is for the derivative with payoff exactly one unit more. Hence, shifting this down by one unit gives an upper bound of $e^R - 1$ as in Lemma 11.

As the payoff of most derivatives can be written as a maximum of a number of components, $\Psi(X_1, \ldots, X_n)$, we will show how to bound the price of such a derivative.

Theorem 15 *Given predictions $\tilde{x}_t \sim \mathcal{D}_t(x_t)$ for probability distributions $\mathcal{D}_t$ and the returns vector $x_t \in \mathbb{R}^n$ of the n derivatives, for each $t \in [T]$,*

$$\Phi(X_1, \ldots, X_n) \le e^{k\mu},$$

with probability at least $1 - \frac{1}{k}$, where

$$\mu = 2 \sum_{t=1}^{T} \int_0^\infty \Pr_{\tilde{x}_t \sim \mathcal{D}_t(x_t)} [\tilde{x}_t \notin (e^{-z} x_t, e^z x_t)]\, dz\,.$$

Proof. Theorem 8 implies that the trading algorithm from Definition 5 will guarantee a wealth of at least as large as μ fraction of the value of each asset, in expectation. Formally,

$$\mathbb{E}[\hat{S}_T] \ge e^{-\mu} V_{i,T}\,,$$

for each $i \in [n]$. This follows from $R_O \le R_G$ (as shown in the proof of Theorem 12). Hence, the expected regret of μ also holds against the benchmark portfolio model that has wealth defined as $\max_i V_{i,T}$. Applying Markov's inequality gives

$$\hat{S}_T \ge e^{-k\mu} V_{i,T}\,,$$

with probability at least $1 - \frac{1}{k}$. In other words, the trading algorithm has $(\beta, \ldots, \beta)$-multiplicative regret, where

$$\beta = e^{-k\mu},$$

with high probability (for large enough k). Applying Lemma 14 gives us

$$\Phi(X_1, \ldots, X_n) \leq 1/\beta = e^{k\mu}.$$

$\square$

Now we will apply the above results to give a probabilistic upper bound on the price of the three exotic derivatives introduced earlier.

Theorem 16 *Given predictions $\tilde{x}_t \sim \mathcal{D}_t(x_t)$ for probability distributions $\mathcal{D}_t$, for each $t \in [T]$,*

i. $\Phi(\mathbf{EX}(X_1, X_2, T)) \leq e^{k\mu_1} - 1$,

ii. $\Phi(\mathbf{SH}(T)) \leq e^{k\mu_2} - 1$,

iii. $\Phi(\mathbf{AS}(T)) \leq e^{k\mu_3} - 1$,

each with probability at least $1 - \frac{1}{k}$, where

$$\mu_i = 2 \sum_{t=1}^{T} \int_0^\infty \Pr_{\tilde{x}_t^i \sim \mathcal{D}_t(x_t^i)} [\tilde{x}_t^i \notin (e^{-z} x_t^i, e^z x_t^i)]\, dz,$$
$$x_t^1 = (r_t^1, r_t^2), x_t^2 = (r_t^1, r_t^1, 1), x_t^3 = (r_t^1, r_t^1),$$
$$r_t^i = \textit{return of asset } i \textit{ between time } t-1 \textit{ and } t.$$

Proof. We will prove the bound for the price of each of the exotic derivatives separately below.

i. Recall that for exchange option, X_1 and X_2 represent two risky assets. Then, by Theorem 15,

$$\Phi(\mathbf{EX}(X_1, X_2, T)) = \Phi(X_1, X_2) - \Phi(X_2) \leq e^{k\mu_1} - V_{2,0}\,.$$

ii. Let X_1 be the risky asset, and X_2 be the trading algorithm that buys the risky asset initially, and if the option holder shouts, sells it immediately.

Use X_3 to represent cash. Then,

$$\Phi(\mathbf{SH}(T)) = \Phi(X_1, X_2, X_3) - \Phi(X_2) \leq e^{k\mu_2} - V_{3,0} \, .$$

The prediction error after the option holder shouts can be bounded above by the prediction error of the underlying asset (since cash does not change in price).

iii. Let X_1 be the risky asset, and X_2 be the trading algorithm that buys the risky asset initially, and sells a fraction $\frac{1}{T}$ of the asset at each time $t \in [T]$ (keeping the remaining wealth as cash). This means that

$$\Phi(X_2) = \frac{1}{T} \sum_{t=1}^{T} S_t \, .$$

Note also that $\mathbf{AS}(T) = \mathbf{EX}(X_1, X_2, T)$. Then, by the above bound on exchange option, we have

$$\Phi(\mathbf{AS}(T)) = \Phi(X_1, X_2) - \Phi(X_2) \leq e^{k\mu_3} - V_{2,0} \, .$$

□

Chapter 6

Conclusion

In this thesis, we explored the application of techniques from algorithmic and statistical learning to address challenges in economics and financial mathematics. We provided a new framework for devising trading strategies and derivatives pricing based on financial predictions that, unlike most previous models in economics literature, do not make any assumption on the underlying process beyond the existence of such predictions. The performance of our models then hinge upon the quality of the predictions.

We derived a trading strategy (on a number of assets in a portfolio) in Chapter 3 that performs competitively against the best static trading strategy in hindsight, extending the previous results in literature to a more generalized setting. We then looked at a more general setting and characterized the performance (in probability) of a trading strategy against a (stronger) dynamic trading strategy in Chapter 4, and applied this strategy to price derivatives (in Chapter 5) in the model-free sense.

6.1 Further Research

There are a number of natural questions that arise from this work:

1. Most of the main results in this thesis were in the form of regret upper bounds. A possible improvement is to either show tightness of these bounds or provide tighter ones.

2. We derived upper bounds for derivative prices in Chapter 5, so it would be practically useful to also obtain the corresponding lower bounds.

3. We assumed the existence of the predictions without studying how these predictions can actually be obtained. It would be interesting to apply known techniques for generating such predictions (otherwise known as "alphas" in the financial industry), and observe how these algorithms perform in practice.

4. We assumed a discrete-time setting throughout, so it would be interesting to see how the results can be generalized to continuous time, perhaps similar to the techniques from Freund [Fre09] or Abernethy et al. [AFW12].

5. The transaction cost model here assumes a loss that is linear in the size of the transaction. Practitioners have often observed non-linear transaction costs in the financial market, so it may be worth extending the results to more realistic transaction cost models, or to account for market impact (this is the impact that the transaction induces on the market).

6. Lastly, this work is focused on maximizing the log-wealth of the trading strategies, but there is a large amount of work in online learning literature that look at more general (usually convex) loss functions. So it would be interesting to see how these results can be generalized for other utility functions, for example, to account for the risk of the strategy (like in the mean-variance framework of Markowitz [Mar52]).

Bibliography

[ABFW13] Jacob D. Abernethy, Peter L. Bartlett, Rafael M. Frongillo, and Andre Wibisono. How to hedge an option against an adversary: Black-scholes pricing is minimax optimal. In *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States.*, pages 2346–2354, 2013.

[AFW12] Jacob D. Abernethy, Rafael M. Frongillo, and Andre Wibisono. Minimax option pricing meets black-scholes in the limit. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 1029–1040, 2012.

[AH06] Amit Agarwal and Elad Hazan. Efficient algorithms for online game playing and universal portfolio management. *Electronic Colloquium on Computational Complexity*, 13(033), 2006.

[AHKS06] Amit Agarwal, Elad Hazan, Satyen Kale, and Robert E. Schapire. Algorithms for portfolio management based on the Newton method. In William W. Cohen and Andrew Moore, editors, *Machine Learning, Proceedings of the Twenty-Third International Conference (ICML 2006)*, volume 148 of *ACM International Conference Proceeding Series*, pages 9–16. ACM, 2006.

[BC12] Sébastien Bubeck and Nicolò Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends in Machine Learning*, 5(1):1–122, 2012.

[BK99] Avrim Blum and Adam Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35(3):193–205, 1999.

[BS73] Fischer Black and Myron S Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–54, May-June 1973.

[BS10] Andrew J. Bean and Andrew C. Singer. Universal switching and side information portfolios under transaction costs using factor graphs. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2010*, pages 1986–1989, 2010.

[BS11] Andrew J. Bean and Andrew C. Singer. Factor graph switching portfolios under transaction costs. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2011*, pages 5748–5751. IEEE, 2011.

[BT16] Tugkan Batu and Pongphat Taptagaporn. Competitive portfolio selection using stochastic predictions. In *Algorithmic Learning Theory - 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016. Proceedings*, 2016.

[BT17] Tugkan Batu and Pongphat Taptagaporn. Derivatives pricing with predictive trading algorithms. 2017.

[CFH09] Kamalika Chaudhuri, Yoav Freund, and Daniel J. Hsu. A parameter-free hedging algorithm. In *Advances in Neural Information Processing Systems 22: 23rd Annual Conference on Neural Information Processing Systems 2009. Proceedings of a meeting held 7-10 December 2009, Vancouver, British Columbia, Canada.*, pages 297–305, 2009.

[CO96] Thomas M. Cover and Erik Ordentlich. Universal portfolios with side information. *IEEE Transactions on Information Theory*, 42(2):348–363, 1996.

[Cov91] Thomas M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.

[Cov96] Thomas M. Cover. Universal data compression and portfolio selection. In *37th Annual Symposium on Foundations of Computer Science, FOCS '96*, pages 534–538. IEEE Computer Society, 1996.

[CYL+12] Chao-Kai Chiang, Tianbao Yang, Chia-Jung Lee, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, and Shenghuo Zhu. Online optimization with gradual variations. In *COLT 2012 - The 25th Annual Conference on Learning Theory, June 25-27, 2012, Edinburgh, Scotland*, pages 6.1–6.20, 2012.

[DKM06] Peter M. DeMarzo, Ilan Kremer, and Yishay Mansour. Online trading algorithms and robust option pricing. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*, pages 477–486, 2006.

[Fre09] Yoav Freund. A method for hedging in continuous time. *CoRR*, abs/0904.3356, 2009.

[GM11a] Eyal Gofer and Yishay Mansour. Pricing exotic derivatives using regret minimization. In *Algorithmic Game Theory, 4th International Symposium, SAGT 2011, Amalfi, Italy, October 17-19, 2011. Proceedings*, pages 266–277, 2011.

[GM11b] Eyal Gofer and Yishay Mansour. Regret minimization algorithms for pricing lookback options. In *Algorithmic Learning Theory - 22nd International Conference, ALT 2011, Espoo, Finland, October 5-7, 2011. Proceedings*, pages 234–248, 2011.

[Gof14] Eyal Gofer. *Machine Learning Algorithms with Applications in Finance*. PhD thesis, Tel Aviv University, 2014.

[GW12] László Györfi and Harro Walk. Empirical portfolio selection strategies with proportional transaction costs. *IEEE Transactions on Information Theory*, 58(10):6320–6331, 2012.

[HAK07] Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.

[HK09] Elad Hazan and Satyen Kale. On stochastic and worst-case models for investing. In *Advances in Neural Information Processing*

Systems 22: 23rd Annual Conference on Neural Information Processing Systems 2009. Proceedings of a meeting held 7-10 December 2009, Vancouver, British Columbia, Canada., pages 709–717, 2009.

[Hul06] John C. Hull. *Options, futures, and other derivatives.* Pearson, Prentice Hall, Upper Saddle River (N.J.), 2006.

[KACS15] Yuri Kalnishkan, Dmitry Adamskiy, Alexey V. Chernov, and Tim Scarfe. Specialist experts for prediction with side information. In *IEEE International Conference on Data Mining Workshop, ICDMW 2015, Atlantic City, NJ, USA, November 14-17, 2015*, pages 1470–1477, 2015.

[KS91] Ioannis Karatzas and Steven Eugene Shreve. *Brownian motion and stochastic calculus.* Graduate texts in mathematics. Springer-Verlag, New York, Berlin, Heidelberg, 1991. Autres tirages corrigs : 1996, 1997, 1999, 2000, 2005.

[KV02] Adam Kalai and Santosh Vempala. Efficient algorithms for universal portfolios. *Journal of Machine Learning Research*, 3:423–440, 2002.

[KW99] Jyrki Kivinen and Manfred K. Warmuth. Averaging expert predictions. In Paul Fischer and Hans-Ulrich Simon, editors, *Computational Learning Theory, 4th European Conference, EuroCOLT '99*, volume 1572 of *Lecture Notes in Computer Science*, pages 153–167. Springer, 1999.

[Man07] Yishay Mansour. Learning, regret minimization and option pricing. In *Proceedings of the 11th Conference on Theoretical Aspects of Rationality and Knowledge (TARK-2007), Brussels, Belgium, June 25-27, 2007*, pages 2–3, 2007.

[Mar52] H.M. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):7791, 1952.

[Mer71] Robert C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373–413, December 1971.

[OC96] Erik Ordentlich and Thomas M. Cover. On-line portfolio selection. In *Proceedings of the Ninth Annual Conference on Computational Learning Theory*, COLT '96, pages 310–313, New York, NY, USA, 1996. ACM.

[RS13] Alexander Rakhlin and Karthik Sridharan. Online learning with predictable sequences. In *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12-14, 2013, Princeton University, NJ, USA*, pages 993–1019, 2013.

[SL05] Gilles Stoltz and Gábor Lugosi. Internal regret in on-line portfolio selection. *Machine Learning*, 59(1-2):125–159, 2005.