

LEARNING AND CONTROL TECHNIQUES IN
PORTFOLIO OPTIMIZATION

by

Daniel S. Brown

Submitted to Brigham Young University in partial fulfillment
of graduation requirements for University Honors

Department of Computer Science

Brigham Young University

August 2011

Advisor: Sean Warnick

Honors Representative: Bryan Morse

Signature: _____ Signature: _____

ABSTRACT

LEARNING AND CONTROL TECHNIQUES IN PORTFOLIO OPTIMIZATION

Daniel S. Brown

Department of Computer Science

Bachelor of Science

This thesis provides an introduction to modern portfolio theory and focuses on two fundamental models: the Markowitz Mean-Variance model and the Mean Absolute Deviation model. The optimal portfolios generated by these models are compared and contrasted using S&P 500 historical stock returns from 2007 to 2010. The results show that the MAD and Markowitz models are very comparable in performance. However, the MAD model is computationally much faster making it more practical for most problems. Simulation results show significant return on investment for several of the portfolios studied. Using the standard MAD model with a 63 day rebalancing horizon, a return of 261.26% is achieved despite the recession. Several other unusual results are reported: it is shown that the Markowitz model while being slower, actually uses fewer stocks than the MAD model; also an unusual bump is found in the time complexity of the Markowitz model and it is shown that for very large

time horizons the time complexity of the Markowitz model is similar to the MAD model. Based on the overall results, the MAD model solved with the simplex method requires the least amount of computation time, is robust to overdiversification, and achieves the highest returns.

ACKNOWLEDGMENTS

I'd like to thank my advisor Dr. Sean Warnick for his continued support and suggestions throughout the entire research and writing process. I would also like to thank all of the undergraduate and graduate researchers in IDeA Labs for their support. A special thanks also is due to all those in my Honors 300 class and to Ryan for helping proofread many sections and for all their suggestions on how to make this a better flowing more readable thesis.

Contents

| | |
|--|------------|
| Table of Contents | v |
| List of Figures | vii |
| List of Tables | ix |
| 1 Introduction | 1 |
| 1.1 Portfolio Optimization | 1 |
| 1.2 Using Information to Make Optimal Decisions | 2 |
| 1.2.1 Decision and Control Problems | 3 |
| 1.2.2 Learning Problems | 3 |
| 1.3 Linear and Quadratic Programming | 3 |
| 2 Portfolio Optimization | 5 |
| 2.1 Investment Portfolios | 5 |
| 2.1.1 Simple Decision Problem with Perfect Information | 6 |
| 2.1.2 The Efficient Market Hypothesis | 6 |
| 2.1.3 Predicting the Future | 7 |
| 2.1.4 Markowitz and Risk Reward Trade-off | 8 |
| 2.1.5 The Mean Absolute Deviation Portfolio Model | 11 |
| 2.1.6 Other Methods Found in Literature Review | 13 |
| 3 A Comparison of the MAD and Markowitz Models | 14 |
| 3.1 Data | 15 |
| 3.2 Portfolio Optimization Models | 16 |
| 3.3 Computation | 17 |
| 3.4 Learning | 18 |
| 3.5 Rebalancing and Parameter Selection | 19 |
| 3.6 Results and Analysis | 21 |
| 3.6.1 Differences Between Portfolio Net Worths | 25 |
| 3.6.2 Simplex vs. Interior-Point Methods | 28 |
| 3.6.3 Starting Portfolio Trading at the Same Time | 30 |
| 3.7 Analysis of Portfolio Compositions | 33 |
| 3.7.1 Average Stocks in Portfolio | 33 |

| | | |
|----------|---|-----------|
| 3.7.2 | Similarity of Stock Choices | 35 |
| 3.8 | Analysis of Computational Complexity | 36 |
| 3.8.1 | Time Complexity Using Random Returns | 48 |
| 3.8.2 | Time Complexity Summary | 49 |
| 3.9 | Analysis of Different Test Horizons and Rebalancing Strategies | 49 |
| 3.9.1 | Results of Multiple Test Horizons | 49 |
| 3.9.2 | Changing the Learning Horizon | 50 |
| 3.9.3 | Results From Using Different Learning Horizons | 51 |
| 3.10 | Risk Adjusted Returns | 52 |
| 3.10.1 | Analysis of Portfolio Performance using the Sharpe Ratio | 53 |
| 3.10.2 | Analysis of Portfolio Performance using the M^2 Measure | 54 |
| 4 | Conclusion | 57 |
| 4.1 | Future Research | 59 |
| 4.2 | Epilogue | 59 |
| A | Tables and Lists | 61 |
| A.1 | Stocks from S&P 500 not included in study | 61 |
| A.2 | Exact Results from Experiment 1 | 63 |
| A.3 | Ticker Symbols of Stocks Used in MAD 63 Day High Rho Portfolio | 63 |
| B | Plots | 64 |
| B.1 | 63 Day Rebalancing Horizon Over Different Rho Values | 64 |
| C | Plots | 68 |
| C.1 | Separate plots for MAD-Simplex, MAD-IP, and Markowitz models showing trends | 68 |
| D | Plots | 70 |
| D.1 | Time complexity plots using random returns | 70 |

List of Figures

| | | |
|------|--|----|
| 2.1 | Optimal portfolios as a function of risk aversion μ | 10 |
| 2.2 | Trade-off between risk and reward in an optimal portfolio | 10 |
| 3.1 | Example of 252 day rebalancing strategy | 20 |
| 3.2 | Portfolio net worths with 252 day rebalancing | 26 |
| 3.3 | Portfolio net worths with 126 day rebalancing | 26 |
| 3.4 | Portfolio net worths with 63 day rebalancing | 27 |
| 3.5 | Portfolio net worths with 21 day rebalancing | 27 |
| 3.6 | MAD-IP and MAD-Simplex portfolio net worths with 21 day rebalancing | 29 |
| 3.7 | Time complexity as number of available stocks increases | 37 |
| 3.8 | Time complexity as Learning Horizon Increases using ρ_{High} | 38 |
| 3.9 | Portfolio compositions as time horizon changes | 40 |
| 3.10 | Number of iterations to compute optimal portfolio over changing horizon | 41 |
| 3.11 | Time Complexity using 100 stocks | 42 |
| 3.12 | Time Complexity using 200 stocks | 43 |
| 3.13 | Time Complexity using 300 stocks | 44 |
| 3.14 | Time Complexity using 400 stocks | 44 |
| 3.15 | Time Complexity using 485 stocks | 45 |
| 3.16 | Time Complexity using 21 days | 45 |
| 3.17 | Time Complexity using 63 days | 46 |
| 3.18 | Time Complexity using 126 days | 46 |
| 3.19 | Time Complexity using 252 days | 47 |
| 3.20 | Time Complexity using 504 days | 47 |
| 3.21 | Time Complexity using 1008 days | 48 |
| 3.22 | Final Portfolio Values for MAD and Markowitz portfolios using different rebalancing horizons for medium and high ρ values | 50 |
| B.1 | Portfolio Net Worths using ρ_{Low} | 65 |
| B.2 | Portfolio Net Worths using ρ_{Med} | 66 |
| B.3 | Portfolio Net Worths using ρ_{High} | 67 |
| C.1 | Time Complexity for MAD-Simplex over multiple horizons | 68 |
| C.2 | Time Complexity for MAD-IP over multiple horizons | 69 |
| C.3 | Time Complexity for Markowitz over multiple horizons | 69 |

| | | |
|-----|--|----|
| D.1 | Time Complexity using 100 artificial stocks with random normal returns | 70 |
| D.2 | Time Complexity using 200 artificial stocks with random normal returns | 71 |
| D.3 | Time Complexity using 300 artificial stocks with random normal returns | 71 |
| D.4 | Time Complexity using 400 artificial stocks with random normal returns | 72 |
| D.5 | Time Complexity using 485 artificial stocks with random normal returns | 72 |

List of Tables

| | | |
|------|--|----|
| 3.1 | Different Time Horizons for Learning | 18 |
| 3.2 | Portfolio Results for Multiple Time Horizons and Multiple ρ | 22 |
| 3.3 | Return on investment for Markowitz Portfolio over multiple time horizons and multiple ρ values. | 23 |
| 3.4 | Return on investment for MAD Portfolio over multiple time horizons and multiple ρ values. | 23 |
| 3.5 | Annualized return on investment for Markowitz Portfolio over multiple time horizons and multiple ρ values. | 23 |
| 3.6 | Annualized return on investment for MAD Portfolio over multiple time horizons and multiple ρ values. | 24 |
| 3.7 | Average Portfolio Differences (in \$ 1000) for Multiple Time Horizons and Multiple ρ | 28 |
| 3.8 | Percent of total days where the MAD portfolio outperformed the Markowitz portfolio | 28 |
| 3.9 | Return on investment for Markowitz Portfolio over multiple time horizons and multiple ρ values | 31 |
| 3.10 | Return on investment for MAD Portfolio over multiple time horizons and multiple ρ values | 31 |
| 3.11 | Average Portfolio Differences (in \$ 1000) for Multiple Time Horizons and Multiple ρ | 31 |
| 3.12 | Percent of total days where the MAD portfolio outperformed the Markowitz portfolio | 32 |
| 3.13 | Average Number of Stocks in portfolio for different horizons using ρ_{Low} | 33 |
| 3.14 | Average Number of Stocks in portfolio for different horizons using ρ_{Med} | 34 |
| 3.15 | Average Number of Stocks in portfolio for different horizons using ρ_{High} | 34 |
| 3.16 | Average percent of stocks in MAD-Simplex portfolio that were also included in the Markowitz portfolio for different horizons and ρ values | 36 |
| 3.17 | Average percent of stocks in Markowitz portfolio that were also included in the MAD-Simplex portfolio for different horizons and ρ values | 36 |
| 3.18 | Training multiples with higher final MAD portfolio value than a training multiple of 1.0 | 51 |
| 3.19 | Training Multiples with higher final Markowitz portfolio value than a training multiple of 1.0 | 52 |

| | | |
|------|---|----|
| 3.20 | Sharpe ratio of daily returns for Markowitz Portfolio over multiple time horizons and multiple ρ values. | 53 |
| 3.21 | Sharpe ratio of daily returns for MAD Portfolio over multiple time horizons and multiple ρ values. | 54 |
| 3.22 | M^2 measure of daily returns for the Markowitz Portfolio over multiple time horizons and multiple ρ values. | 55 |
| 3.23 | M^2 measure of daily returns for the MAD Portfolio over multiple time horizons and multiple ρ values. | 55 |
| 3.24 | M^2 measure of annualized returns for the Markowitz Portfolio over multiple time horizons and multiple ρ values. | 56 |
| 3.25 | M^2 measure of annualized returns for the MAD Portfolio over multiple time horizons and multiple ρ values. | 56 |
| A.1 | Portfolio Results (in \$100,000) for Low ρ | 62 |
| A.2 | Portfolio Results (in \$100,000) for Med ρ | 62 |
| A.3 | Portfolio Results (in \$100,000) for High ρ | 63 |

Chapter 1

Introduction

The recent chaos in the financial world has made increasingly clear the great value and importance of risk analysis when making financial decisions. Much research has been done in the field of computational finance in an attempt to accurately model financial markets and their inherent risks. Because we do not have perfect information about future events such as returns on derivatives and securities, financial analysts implement various learning algorithms in an attempt to forecast such behavior. These algorithms yield predictive models that analysts use to make optimal decisions. Researching ways to approach risk management and optimal decision-making through the use of learning algorithms and control techniques will be the underlying theme of this thesis. I will focus specifically on portfolio optimization.

1.1 Portfolio Optimization

Building optimal investment portfolios is a major focus in computational finance. Given a set of potential investments, a portfolio manager must decide on the number and proportion of investments to acquire in order to reach an objective. These objec-

tives are typically formulated as optimization problems where the portfolio manager desires to maximize returns given an admissible level of risk or to minimize risk given a required level of return. Portfolios often consist of stocks and bonds, but may include other investment vehicles such as contracts to buy or sell commodities, or actual physical resources such as power plants or retail stores. In 1990 Harry Markowitz received the Nobel Prize in Economics for his portfolio optimization model that addressed the tradeoff between portfolio risk and return [19].

Since Markowitz published his seminal paper in 1952, the field of portfolio theory has grown as researchers have criticized his model and proposed alternatives to his approach [21] [8] [14]. Researchers continue to question how to best estimate future returns and how to measure portfolio risk. In 1991 Konno and Yamazaki proposed an alternative to Markowitz's model that uses mean absolute deviation instead of variance as the measure of portfolio risk [14]. Their Mean Absolute Deviation (MAD) portfolio optimization model results in a more computationally tractable solution that can be found through linear programming techniques.

1.2 Using Information to Make Optimal Decisions

Finding the solution to a portfolio optimization problem often requires the use of algorithms that take available information and utilize it to make good decisions. These types of problems are referred to as decision and control problems. Portfolio optimization is typically referred to as an open-loop control problem because a portfolio manager can be thought of as a price taker whose decisions are influenced by the market, but whose actions have a negligible impact on the market as a whole.

1.2.1 Decision and Control Problems

A decision and control problem consists of three parts: (1) A clear objective, (2) A complete description of the available choices, and (3) An understanding of the consequences of available choices in terms of how they impact the objective. In order to solve decision and control problems, we often need information about the future to know how certain actions will affect our objective. When we do not have perfect information about future events and consequences, we must create a model we believe best predicts the future and use the outputs of this model in place of perfect information. This process of developing a model to make inferences about the future based on the past involves solving what is called a learning problem.

1.2.2 Learning Problems

A learning problem is the task of selecting out of a set of potential models the model that best fits the data of interest according to some quality metric. There are three fundamental pieces needed to solve a learning problem: (1) A parameterized family of mathematical models, (2) Data, and (3) A quality metric that scores a particular model given the available data. Scientists often develop learning algorithms that solve learning problems. These algorithms select a model that best fits the available data according to a desired metric.

1.3 Linear and Quadratic Programming

Throughout this thesis I assume the reader has a general familiarity with basic concepts in convex optimization such as objective functions, decision variables, constraints, etc. This thesis focuses particularly on Linear and Quadratic programming as methods for solving portfolio optimization problems. An excellent introduction to

these topics can be found in [28] and [4].

Chapter 2

Portfolio Optimization

The following chapter provides an in depth treatment of the portfolio optimization problem. I show two standard ways of formulating the problem: first, I derive Markowitz's Mean-Variance approach for portfolio optimization; and second, I derive the Mean Absolute Deviation (MAD) model. I compare and contrast these basic methods and provide a review of other adaptations that have been made since the Markowitz and MAD methods were first published. I also formulate the learning problem associated with the portfolio optimization problem.

2.1 Investment Portfolios

Consider an investor who wants to invest in the stock market and must select from a set of n possible investments. The investor faces the choice of which investments to pick and how much money to invest in each. Before formulating the portfolio optimization problem as a decision and control problem, I first introduce some notation and definitions that will be helpful throughout the rest of this chapter. Let R_j be a random variable representing the total return in the next time period for investment

j , $j = 1, \dots, n$. For example, if you invest \$1 today in stock j , it will be worth $\$R_j$ tomorrow. A portfolio is defined as a set of variables x_j , $j = 1, \dots, n$, that are nonnegative (no short selling) and sum to one. Each x_j represents the fraction of the portfolio invested in stock j . It follows that the return of this portfolio is

$$R = \sum_j x_j R_j.$$

2.1.1 Simple Decision Problem with Perfect Information

Our objective is to maximize the total return, R , of our portfolio. This can be formulated as the following optimization problem:

$$\begin{aligned} \max \quad & \sum_j x_j R_j \\ \text{subject to} \quad & \sum_j x_j = 1 \\ & x_j \geq 0 \quad j = 1, \dots, n. \end{aligned} \tag{2.1}$$

If we have perfect information about the returns R_j , the problem becomes trivial; the optimal portfolio is found by simply setting $x_j = 1$ where $j = \operatorname{argmax}_j R_j$. However, in the real world perfect information is hard, if not impossible, to obtain. If it is absolutely impossible to know anything about the future value of a stock's return, then we must resign ourselves to the fact that we will never find a portfolio that can be deemed optimal in any way. Whether or not financial time series can be accurately predicted is a topic of great debate in the financial community and revolves around what is referred to in finance and economics as the Efficient Market Hypothesis.

2.1.2 The Efficient Market Hypothesis

The Efficient Market Hypothesis (EMH) states that financial markets are efficient—all of the information related to a certain investment is contained in the current

investments price—and that because they are efficient, excess returns cannot be made using publicly available information. There are three versions of the hypothesis: weak, semi-strong, and strong. The weak efficient market hypothesis is that prices reflect all past publicly available data. The semi-strong efficient market hypothesis states that not only do prices reflect all public information, but that prices instantly change to reflect new information. This implies that neither technical analysis nor fundamental analysis will yield information about future returns since information such as paid dividends and company cash flows and price trends will already be incorporated into the current market price. The strong efficient market hypothesis asserts everything contained in the semi-strong hypothesis with the addition that prices instantly reflect insider information [2]. Several studies show evidence for and against the weak and semi-strong versions of the EMH, and there is much evidence against the strong EMH [10, 23]. Due to the recent financial crisis, many have begun questioning the validity of the EMH and it has come under scrutiny by the media [16] [5]. Throughout this thesis it is assumed that we can use historical data to gain some information regarding future stock returns.

2.1.3 Predicting the Future

As discussed previously, the solution to a learning problem involves choosing a specific model from a class of models that best describes the available data according to some metric. These solutions are often found through the use of a learning algorithm. In finance, we often assume that stock prices come from a Gaussian distribution characterized by a specific mean and variance. Once we have collected a time series of historical prices, our learning algorithm computes the sample mean and sample variance to determine a parameterized Gaussian distribution that best fits our data.

Given T samples of stock j the sample mean is defined as

$$\bar{R}_j = \frac{1}{T} \sum_{t=1}^T R_j(t), \quad (2.2)$$

where $R_j(t)$ is the historical return for asset j at time t . The sample covariance matrix C is defined as

$$C_{ij} = \frac{1}{T-1} \sum_{t=1}^T (R_i(t) - \bar{R}_i)(R_j(t) - \bar{R}_j), \quad i = 1, \dots, n \quad j = 1, \dots, n. \quad (2.3)$$

The question now is how to use these estimates in our decision and control problem. We could just substitute our sample mean in place of future returns in Equation 2.1; however, this would imply that we believe the solution to our learning problem is just as good as knowing the actual future returns. Because our estimate is not exact, we want to incorporate into our model a measure of how confident we are in the forecasts produced by our learning algorithm.

2.1.4 Markowitz and Risk Reward Trade-off

Harry Markowitz received the 1990 Nobel Prize in Economics for coming up with a solution to this problem and in the process showed mathematically why investors diversify their portfolios. Markowitz observed that investments with high expected returns in the long run can be very erratic in the short term. This led him to use the covariance of portfolio returns as a measure of risk. Because predictions of future returns are not perfectly accurate, Markowitz reformulated the decision and control problem to account for an investor's risk aversion. Rather than just substituting the solution to our learning problem in place of perfect information, we should weight our objective function to represent how close we believe our predictions match reality.

Markowitz's portfolio optimization problem can be formulated as follows:

$$\begin{aligned} \max \quad & \sum_{j=1}^n x_j \bar{R}_j - \mu \sum_{i=1}^n \sum_{j=1}^n x_i x_j C_{ij} \\ \text{subject to} \quad & \sum_{j=1}^n x_j = 1 \\ & x_j \geq 0 \quad j = 1, \dots, n. \end{aligned} \tag{2.4}$$

Using this formulation and setting μ equal to zero, we get the same result as just replacing future information with our predictions in Equation 2.1. Markowitz's insight was that because we don't believe our predictions to be perfectly accurate, we should discount them by including a term that quantifies the risk or unpredictability of our forecasts. The risk aversion parameter μ allows individual investors to adjust their portfolios according to their risk and reward preferences: when μ is small the portfolio will be chosen to focus more on maximizing reward; when μ is large more emphasis is put on minimizing the riskiness of the portfolio.

Markowitz wasn't the first to tell people to diversify their portfolios—this was already a common practice. He was, however, one of the first to show mathematically why diversification is optimal. Figure 2.1 shows how changing the risk aversion parameter μ affects the composition of a portfolio obtained by solving Markowitz's optimization model (2.4). The x-axis is the risk aversion and the y-axis is the percent of the portfolio allocated to a particular stock. Each color represents a different stock and the proportion of the colored area represents the proportion of the portfolio that is invested in that particular stock. When the risk aversion is zero the portfolio consists of only one stock; however, as the risk aversion increases the portfolio becomes more diversified to account for the greater emphasis on risk rather than reward.

Figure 2.2 shows the trade-off between risk and reward when selecting an optimal portfolio. Each point along the blue line represents an optimal portfolio defined by its risk and reward. The entire set of these points is called the efficient frontier. As

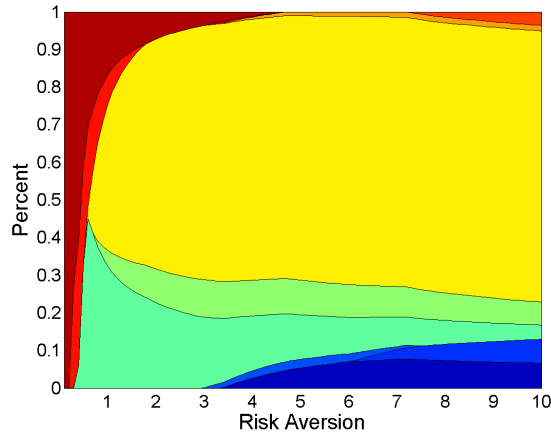


Figure 2.1 Optimal portfolios as a function of risk aversion μ

the reward of a portfolio on the efficient frontier increases so does its risk and vice versa. Each red dot represents a portfolio consisting entirely of one stock. As seen in Figure 2.2 all of the optimal portfolios are diversified except for the maximum reward portfolio (the upper right point on the blue curve) which corresponds to a portfolio consisting entirely of the stock with the highest expected reward and is found by setting μ to zero.

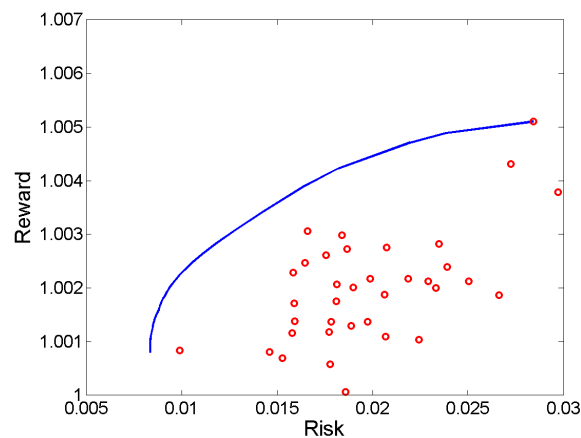


Figure 2.2 Trade-off between risk and reward in an optimal portfolio

2.1.5 The Mean Absolute Deviation Portfolio Model

Despite its prominence in portfolio theory, Markowitz's portfolio optimization model has come under criticism for several reasons. In 1991, Hiroshi Konno and Hiroaki Yamazaki pointed out three flaws to Markowitz's portfolio optimization model [14]:

1. *Computational Burden.* The Markowitz model has a quadratic cost function which can be very time intensive to solve depending on the number of stocks being considered. Konno and Yamazaki argued that this computational cost can be impractical for investors trading on a real-time basis.
2. *Investor's perception against risk and distribution of stock prices.* Using the variance as a measure of risk seems pessimistic since most investors are happy when a stock outperforms and yields a larger profit. Furthermore, studies of the Tokyo Stock Market showed that most returns are not normally distributed, so the mean and variance may not be an accurate representation of a stock's return distribution [13].
3. *Transaction/management cost and cut-off effect.* An optimal solution to the Markowitz model generally consists of many nonzero elements; however, often many of these elements represent only a fraction of a percent of the total portfolio. Solutions such as this are not realistic because of the significant amount of transaction costs associated with a very large portfolio of investments. Managing a very large portfolio can also be very time consuming. To try and avoid this problem an investor may choose to eliminate stocks from the optimal portfolio which have smaller weights; however, this distorts the optimal solution and may result in a portfolio with larger risk.

Because of the difficulties involved in using the Markowitz model directly, in [14], Konno and Yamazaki introduced the L_1 risk or mean absolute deviation from the

mean:

$$\mathbb{E} \left[\left| \left(\sum_{j=1}^n x_j (R_j - \mathbb{E}[R_j]) \right) \right| \right]. \quad (2.5)$$

Using this linear risk measure the portfolio optimization problem can be formulated as

$$\begin{aligned} \max \quad & \sum_{j=1}^n x_j \bar{R}_j - \mu \frac{1}{T} \sum_{t=1}^T y_t \\ \text{subject to} \quad & y_t - \sum_{j=1}^n x_j (R_j(t) - \bar{R}_j) \geq 0 \quad t = 1, \dots, T \\ & y_t + \sum_{j=1}^n x_j (R_j(t) - \bar{R}_j) \geq 0 \quad t = 1, \dots, T \\ & \sum_{j=1}^n x_j = 1 \\ & x_j \geq 0 \quad j = 1, \dots, n \end{aligned} \quad (2.6)$$

where $\bar{R}_j = \frac{1}{T} \sum_{t=1}^T R_j(t)$.

This formulation has many advantages. One of the most important is that, rather than computing an $n \times n$ covariance matrix and using quadratic programming techniques, this problem requires much less computational effort to formulate and can be solved quickly using linear programming techniques. This problem has $2T + 1$ constraints so by the Fundamental Theorem of Linear Programming there will be at most $2T + 1$ nonzero variables [28]. This means that an optimal portfolio generated by the Mean Absolute Deviation (MAD) model using the simplex method [6] will only contain $2T + 1$ stocks regardless of the value of n . This is a substantial improvement over the Markowitz model which has no limit on the number of stocks included in the optimal portfolio. This also means we can use T as a control variable to limit the number of assets in the portfolio.

2.1.6 Other Methods Found in Literature Review

Since the publications of Markowitz's Mean-Variance model and Konno and Yamazaki's MAD model there have been numerous papers published on the topic of portfolio optimization. In [9], Feinstein and Thapa show that a simple manipulation of the MAD model results in a linear program with only $T + 2$ constraints, further decreasing the computational cost of finding an optimal portfolio. In [25] and [24], Sharpe develops two linear programming approximations to Markowitz's mean-variance model. Stone also provides a linear programming approach to the portfolio selection problem and shows how it can be used to approximate a mean-variance-skewness portfolio model [27]. In [21], Peng adds transaction costs to the standard Markowitz model. He shows that the new model moderates the transaction volume and provides a significant improvement over the Markowitz model when dealing with transaction costs. In [8] Fazel formulates the portfolio optimization problem as a more general convex optimization problem and also develops a model that takes into account linear transaction costs.

Many other studies have focused on using different risk measures other than the variance or mean absolute deviation. A study of downside-risk measures is performed in [11]. Some specific alternate risk measures include Value at Risk [7], Tail or Conditional Value at Risk [15], and Shortfall [3]. In [1] Barmish develops a robust control paradigm for trading equities. Primbs [22] shows how the portfolio optimization problem can be formulated using stochastic receding horizon control.

Chapter 3

A Comparison of the MAD and Markowitz Models

In the previous chapter we discussed both the Markowitz model and the MAD model as potential solutions to the portfolio optimization problem. This chapter compares the performance of portfolios generated using the MAD and Markowitz models. The goal is to use historical data to determine whether one model produces portfolios that consistently perform better than the other.

In [14], Konno and Yamazaki state that the MAD model is comparable to the Markowitz model and show that if the returns are multivariate normal that the mean absolute deviation and the standard deviation measures of risk are essentially the same. There has been very little work recently comparing the performance of the MAD and Markowitz models. In 2003 Mansini *et al.* compared the Markowitz and MAD models and found that the MAD model performs slightly better than the Markowitz model [18]. Their study used only stocks in the Milan Stock Exchange and considers four sets of data from 1994-1998. In 2010 Hoe *et al.* performed an empirical comparison of different risk measures in portfolio optimization using only

54 stocks in the Kuala Lumpur Composite Index from January 2004 until December 2007 [12]. Hoe found that the portfolios generated by the MAD and Markowitz models use very similar stocks but have different portfolio weights. Most of the analysis in [12] focuses on risk measures other than mean absolute deviation and standard deviation, and little focus is given to comparing the MAD and Markowitz portfolios; however, Hoe's comparison does find that the Markowitz mean-variance portfolio has the lowest risk-adjusted performance overall.

There are two main problems with these papers: (1) They only calculate the optimal portfolios once and do not allow the portfolio to be periodically rebalanced, (2) They only use data before 2008, thereby avoiding data related to the recent market crash and ensuing recovery. This chapter will address these two issues and provide a more complete comparison of the MAD and Markowitz models using more recent data and allowing portfolios to be rebalanced throughout the study.

3.1 Data

In order to test Konno and Yamazaki's claim that the MAD and Markowitz models are comparable I collected historical price data from Yahoo Finance for all the stocks in the S&P 500 from January 2007 through December 2010. The S&P 500 was chosen to provide a representative sampling of commonly traded stocks from different sectors. The time period was chosen to obtain a large sample of historical data that includes the recent stock market crash and recession and the recent recovery. This data set allows the MAD and Markowitz portfolio optimization algorithms to be evaluated over a range of market conditions which are often avoided in financial studies. During the period of time from 2007–2010 there were several changes made to the S&P 500 listing as stocks were added and removed. Because of these discrepancies I chose to limit

this case study to companies which were included in the S&P 500 for the full four years of interest (2007–2010). After eliminating any stocks which were not included in the S&P 500 for the full length of the study, I was left with 485 stocks each with 1008 days of total returns. Throughout the rest of this thesis, this subset of the S&P 500 will be referred to for simplicity as the S&P 500, or stocks in the S&P 500, with the understanding that in actuality we are only considering 485 stocks. For a list of the stocks that were not included see Appendix A.

3.2 Portfolio Optimization Models

This case study seeks to provide an “apples to apples” comparison of the MAD and Markowitz models. To do this we must first formulate the problems in such a way that this comparison can be made. To achieve this goal I chose to formulate both models as minimization problems where the objective function is either the variance of the portfolio (Markowitz) or the average mean absolute deviation (MAD). The constraints are similar to those posed earlier in this thesis with the addition of a minimum desired return. The Markowitz model is formulated as follows:

$$\begin{aligned}
 & \text{minimize} && \sum_{i=1}^n \sum_{j=1}^n x_i x_j C_{ij} \\
 & \text{subject to} && \sum_{j=1}^n x_j \bar{R}_j \geq \rho M_0 \\
 & && \sum_{j=1}^n x_j = M_0 \\
 & && x_j \geq 0 \quad j = 1, \dots, n
 \end{aligned} \tag{3.1}$$

where $\bar{R}_j = \mathbb{E}[R_j]$, $C_{ij} = \mathbb{E}[(R_i - \bar{R}_i)(R_j - \bar{R}_j)]$, ρ is a parameter representing the minimal expected return required by an investor, and M_0 is the total amount of

money available for investing. It is important to note that the decision variables (x_j) now represent actual dollar amounts rather than a proportion of available funds.

The MAD model is formulated similarly:

$$\begin{aligned}
& \text{minimize} && \frac{1}{T} \sum_{t=1}^T y_t \\
& \text{subject to} && \sum_{j=1}^n x_j \bar{R}_j \geq \rho M_0 \\
& && y_t + \sum_{j=1}^n x_j (R_j(t) - \bar{R}_j) \geq 0 \\
& && y_t - \sum_{j=1}^n x_j (R_j(t) - \bar{R}_j) \geq 0 \\
& && \sum_{j=1}^n x_j = M_0 \\
& && x_j \geq 0 \quad j = 1, \dots, n \\
& && y_t \geq 0 \quad t = 1, \dots, T.
\end{aligned} \tag{3.2}$$

Both models have the same constraints: the average expected value of the portfolio must be above some benchmark ρ , the sum of the amount of money invested in each stock must equal the total money invested, M_0 , and decision variables must be positive (no short selling). Because the only difference between the two models is the measure of risk, the above formulations allow us to fix ρ and M_0 and accurately compare the performance of the portfolios generated by each model.

3.3 Computation

I ran all the experiments for this thesis using MATLAB. Solutions for the MAD and Markowitz portfolio optimization models were obtained using MATLAB's linear and quadratic programming routines. MATLAB allows the user to specify whether the

linear programming routine uses the simplex method or an interior-point method. Unless otherwise specified, the MAD model will be solved using the interior point method which is the default method for solving linear programs in MATLAB.

3.4 Learning

The purpose of this experiment is to compare the portfolio net worths of the MAD and Markowitz models. In order to gain insight into the relative performance of the MAD and Markowitz models, I performed several tests by varying the time horizon and the desired portfolio return ρ . Both the MAD and Markowitz portfolio models require an estimate of the expected return of each potential investment. In this experiment I used the arithmetic mean of past returns as a learning algorithm to estimate the true expected value. One of the choices that arises from using this type of estimator is how much data to use to solve the learning problem. For this study I chose to use the four different time horizons in Table 3.1. The data used to estimate the expected return is referred to as the ‘training data’ for the learning algorithm.

| Horizon | Days |
|---------------|------|
| Yearly | 252 |
| Semi-Annually | 126 |
| Quarterly | 63 |
| Monthly | 21 |

Table 3.1 Different Time Horizons for Learning

For example, when using the yearly time horizon of 252 days, the past 252 days of stock returns are used to calculate the average return for each stock (I do not use 365 because the stock market is not open on weekends or holidays). These average

returns are then used as inputs to the MAD and Markowitz models.

Another decision that must be made is how often to rebalance a portfolio. Because of changing market conditions, an optimal portfolio from a month ago may not be the optimal portfolio today. Deciding when and how to rebalance a portfolio as time progresses is an important problem in portfolio optimization.

3.5 Rebalancing and Parameter Selection

After learning on the past m days the optimal MAD and Markowitz portfolios are computed by solving equations (3.1) and (3.2). These portfolios are then held constant and tested on the next n days before the optimization is reevaluated. For this experiment, I decided to let $m = n$ so that the horizons for the training and test data are equal. Continuing with our previous example of using the yearly time horizon of 252 days, we would train on the first 252 days, then compute the optimal portfolio using the MAD and Markowitz models. These portfolios would then be tested on the next 252 days. We would then form a new training set consisting of the returns from days 253–504 (the second set of 252 days). After learning on this data we would compute a new portfolio and test it over the 3rd set of 252 days, and so on. In this manner we learn on n days, test on the next n days and then shift everything forward by n days and repeat until we reach the end of our data sample. Figure 3.1 shows a graphical example of a 252 day rebalancing strategy and illustrates how the training and test data sets change over time. This procedure is often called ‘back-testing’, and essentially takes the algorithm back in time so its performance can be tested on historical data.

The free parameters in (3.1) and (3.2) are M_0 and ρ , and must be chosen for each simulation. I chose to set $M_0 = \$100,000$ for every simulation, although any

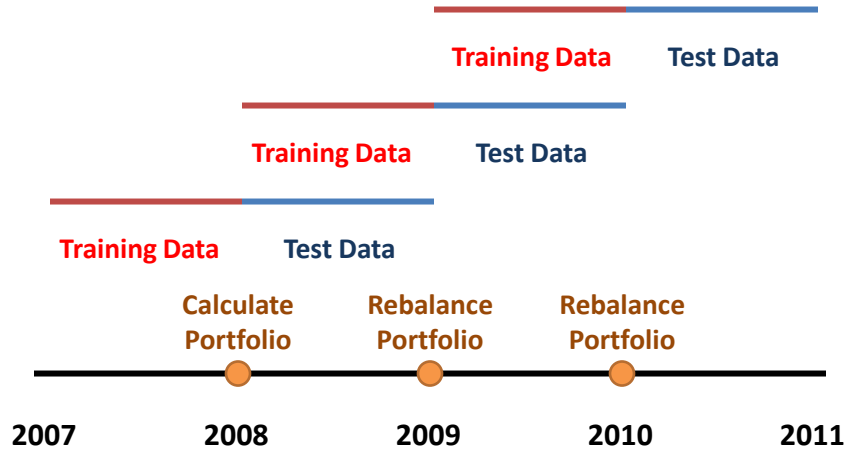


Figure 3.1 Example of 252 day rebalancing strategy

dollar amount would have sufficed. To decide on an appropriate value for ρ it is beneficial to remember that the ρ parameter determines how much expected return on investment we require from our portfolio. The higher we set ρ the more risk we are willing to accept in order to achieve our desired expected return; however, if we require our portfolio to have an expected return that is too high, our linear or quadratic program will be infeasible. For example, if we consider only two stocks, A and B, with expected total returns 1.02 and 1.07, respectively, it is impossible to find a portfolio which has an expected return greater than 1.07. Because of this fact, we see that for a given training set consisting of n time series $S_1 \dots S_n$, the upper and lower bounds on admissible values of ρ are as follows:

$$\begin{aligned}\rho_{\text{upper}} &= \max_i(\mathbb{E}[S_i]) \\ \rho_{\text{lower}} &= \min_i(\mathbb{E}[S_i]).\end{aligned}\tag{3.3}$$

Because the upper and lower bounds are functions of the training set, they will change over the course of the experiment. This experiment uses three different values of ρ for both the MAD and Markowitz models:

$$\begin{aligned}\rho_{\text{Low}} &= \rho_{\text{lower}} + (\rho_{\text{upper}} - \rho_{\text{lower}})/10 \\ \rho_{\text{Med}} &= \rho_{\text{lower}} + (\rho_{\text{upper}} - \rho_{\text{lower}})/2 \\ \rho_{\text{High}} &= \rho_{\text{lower}} + 9 \cdot (\rho_{\text{upper}} - \rho_{\text{lower}})/10\end{aligned}\tag{3.4}$$

These three values were chosen to provide a very conservative, moderate, and high target return portfolio to allow for a more thorough comparison of the MAD and Markowitz models over a range of possible investor preferences. The values of ρ_{Low} , ρ_{Med} , and ρ_{High} depend on the training data, but are consistent for both the MAD and Markowitz models. This is reasonable because it reflects how real investors update their views on how much return they expect from their portfolio based on recent data. Using the three values of ρ defined in 3.4, I compared the results from 12 different MAD portfolios (3 ρ values \times 4 update horizons) and 12 Markowitz portfolios.

3.6 Results and Analysis

Table 3.2 shows the results from comparing the MAD and Markowitz portfolios for different values of ρ and different rebalancing schedules. The MAD and Markowitz portfolios are compared based on three different performance categories: the average

value over the entire study, the average value of the portfolio at the end of each rebalancing horizon, and the final net worth of the portfolio. ‘MAD’ represents a test where the MAD portfolio performed better in all three categories. ‘Markowitz’ represents a test where the Markowitz portfolio performed better in all three categories. The ‘?’ represents a test which was inconclusive (see Appendix A for the exact values).

| Horizons | 252 | 126 | 63 | 21 |
|---------------|-----|-----------|-----------|-----------|
| ρ_{Low} | MAD | Markowitz | Markowitz | MAD |
| ρ_{Med} | ? | ? | Markowitz | MAD |
| ρ_{High} | ? | Markowitz | MAD | Markowitz |

Table 3.2 Portfolio Results for Multiple Time Horizons and Multiple ρ

As can be seen from Table 3.2, there are cases where the Markowitz portfolio performs better than the MAD portfolio and vice versa, but there is no clear winner. Tables 3.3 and 3.4 show the return on investment for the Markowitz and MAD portfolios. The 63 day rebalancing strategy did extraordinarily well when the required rate of return was equal to ρ_{High} . The entries highlighted in green show the highest return achieved for each portfolio and those highlighted in red show the lowest return achieved. The highest and lowest returns were obtained when the required minimum rate of return was set to ρ_{High} for both the Markowitz and the MAD portfolios. A higher value of ρ will force the portfolio to invest in stocks with higher average returns even if they are more volatile; however, despite the higher risk that comes from using ρ_{High} , the only portfolios with negative returns were the ones using a 252 rebalancing strategy—all the others had returns above 15%. Tables 3.5 and 3.6 show the annualized returns from tables 3.3 and 3.4 for the Markowitz and MAD portfolios.

The lowest returns (highlighted in red) were when rebalancing yearly (252 days) and the highest returns were achieved when rebalancing every three months (63 days).

| Horizons | 252 | 126 | 63 | 21 |
|---------------|---------|--------|---------|---------|
| ρ_{Low} | 2.57% | 3.54% | 20.76% | 3.13% |
| ρ_{Med} | 1.37% | 25.23% | 50.60% | 38.32% |
| ρ_{High} | -33.04% | 19.01% | 192.88% | 133.23% |

Table 3.3 Return on investment for Markowitz Portfolio over multiple time horizons and multiple ρ values. Green is the best, red is the worst.

| Horizons | 252 | 126 | 63 | 21 |
|---------------|---------|--------|---------|---------|
| ρ_{Low} | 4.87% | 1.58% | 11.98% | 11.29% |
| ρ_{Med} | -0.11% | 28.35% | 40.13% | 52.91% |
| ρ_{High} | -32.23% | 15.43% | 261.26% | 124.22% |

Table 3.4 Return on investment for MAD Portfolio over multiple time horizons and multiple ρ values. Green is the best, red is the worst.

This shows that more frequent rebalancing tends to improve the return on investment. This makes sense intuitively because more frequent rebalancing allows a portfolio to adapt quickly to a changing market. It is also interesting to note that the 21 day rebalancing strategy was not as profitable as the 63 day rebalancing strategy. This seems to contradict the notion that more frequent updating is better; however, we must realize that with 21 day rebalancing we are only using the past 21 days to make

| Horizons | 252 | 126 | 63 | 21 |
|---------------|---------|-------|--------|--------|
| ρ_{Low} | 0.85% | 0.99% | 5.16% | 0.79% |
| ρ_{Med} | 0.45% | 6.64% | 11.54% | 8.64% |
| ρ_{High} | -12.51% | 5.10% | 33.18% | 24.14% |

Table 3.5 Annualized return on investment for Markowitz Portfolio over multiple time horizons and multiple ρ values. Green is the best, red is the worst.

| Horizons | 252 | 126 | 63 | 21 |
|---------------|---------|-------|--------|--------|
| ρ_{Low} | 1.60% | 0.45% | 3.06% | 2.77% |
| ρ_{Med} | -0.037% | 7.39% | 9.41% | 11.45% |
| ρ_{High} | -12.16% | 4.19% | 40.85% | 22.89% |

Table 3.6 Annualized return on investment for MAD Portfolio over multiple time horizons and multiple ρ values. Green is the best, red is the worst.

decisions versus the 63 days of historical data used for 63 day rebalancing. Later on in this thesis I investigate whether a portfolio with a better final return can be found by varying the learning and rebalancing horizons.

Right now is a good time to point out that all is not what it seems; this experiment has left out one important factor that faces real investors: namely, transaction costs. A conservative estimate of how the portfolio would have performed with transaction costs is to assume that every time the portfolio was rebalanced that 485 transactions were made (we will see later that the actual number of transactions is much lower than this). If the investor were charged \$10 per trade, rebalancing every 63 days, the total cost would be \$77600. This seems like a lot, but the final return on investment would still be 183.65% which is not too bad for a four year investment during the recession.

Figures 3.2-3.5 show how closely the MAD and Markowitz models track each other through the entire four year simulation. From these graphs we see that the portfolios using ρ_{Low} had their best performance when rebalancing every 252 days (Figure 3.2), but performed worse than the ρ_{Med} and ρ_{High} portfolios when using the 126, 63 and 21 day rebalancing strategies. Figure 3.3 shows that the portfolios using ρ_{Med} performed best when using the 126 day rebalancing strategy. The portfolios using ρ_{High} in Figure 3.3 performed very well at first but were hit hard by the market crash. Figure

3.4 shows that both the Markowitz and MAD models using a rebalancing strategy of 63 days and ρ_{High} were able to achieve phenomenal growth. The portfolios using ρ_{High} also performed very well when using the 21 day rebalancing strategy (see Figure 3.5). All four figures show that a portfolio requiring a higher expected return is more sensitive to market conditions and will lose more value in a recession, but will often make up for these losses with high gains as market conditions improve.

The performance of the Markowitz and MAD portfolios is extremely similar. Examining figures 3.2 - 3.5, it is apparent that both portfolios track each other closely. The only large discrepancy between the MAD and Markowitz portfolios was found when both portfolios were using a 63 day rebalancing schedule and ρ_{High} (Figure 3.4). When holding ρ constant and varying the rebalancing strategy, it is apparent that the 63 day rebalancing horizon performs the best for both the Markowitz and the MAD model with the one exception being the MAD model with a ρ_{Med} which achieves its best performance when using the 21 day rebalancing horizon (see Appendix B).

3.6.1 Differences Between Portfolio Net Worths

To determine how much difference there was between the net worths of the MAD and Markowitz portfolios I computed the average difference between the daily net worth values. The results are contained in Table 3.7. The values are reported in thousands of dollars where positive values represent cases where, on average, the MAD portfolio performed better than the Markowitz portfolio, and negative values are cases where the Markowitz portfolio performed better on average. As noted previously, the largest difference between the MAD and Markowitz portfolios is when ρ_{High} and 63 day rebalancing are used. The entries highlighted in green show the highest difference in favor of the MAD model and those in red represent the lowest difference. Negative values represent cases where the Markowitz model outperformed

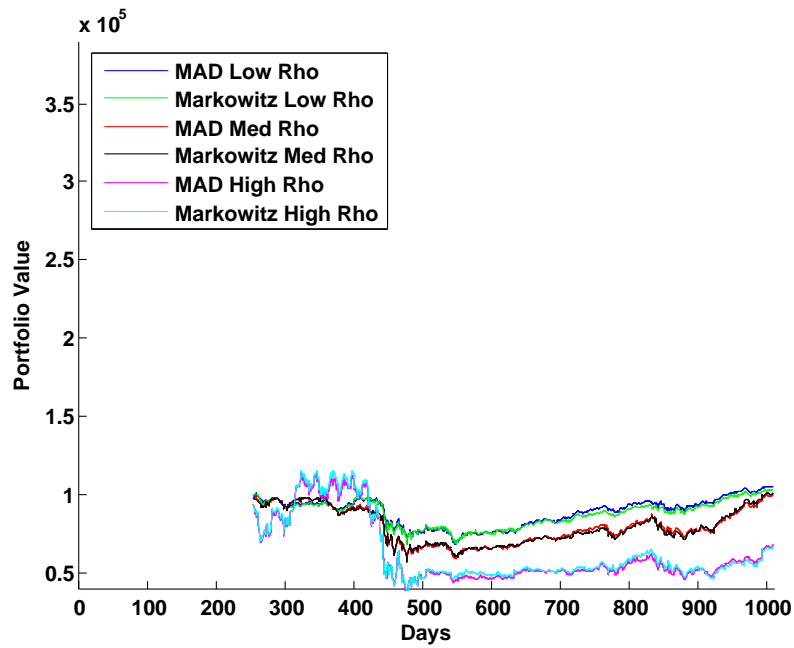


Figure 3.2 Portfolio net worths with 252 day rebalancing

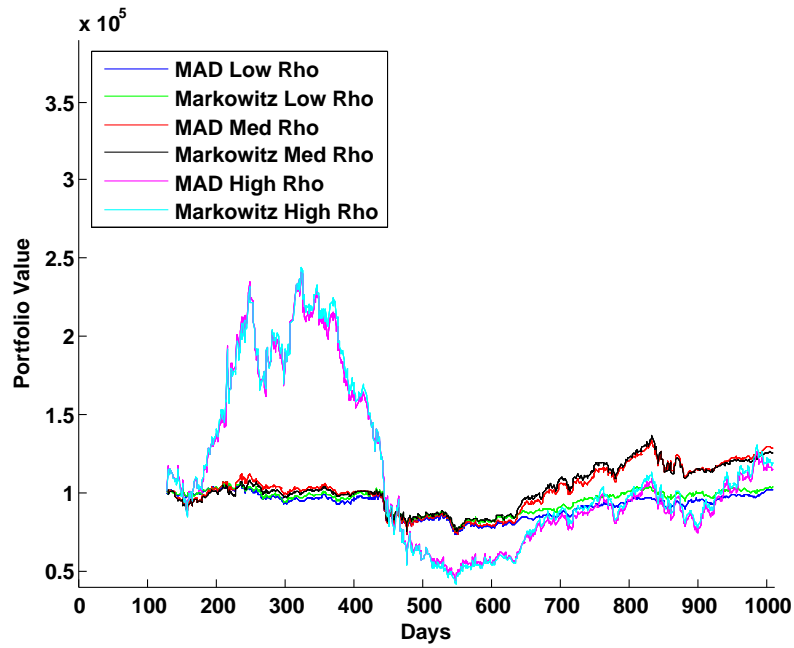


Figure 3.3 Portfolio net worths with 126 day rebalancing

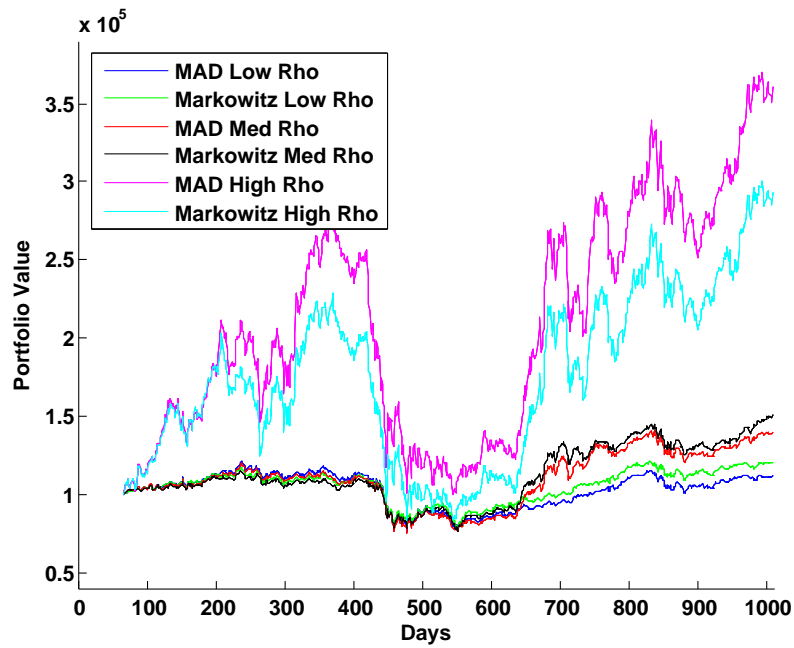


Figure 3.4 Portfolio net worths with 63 day rebalancing

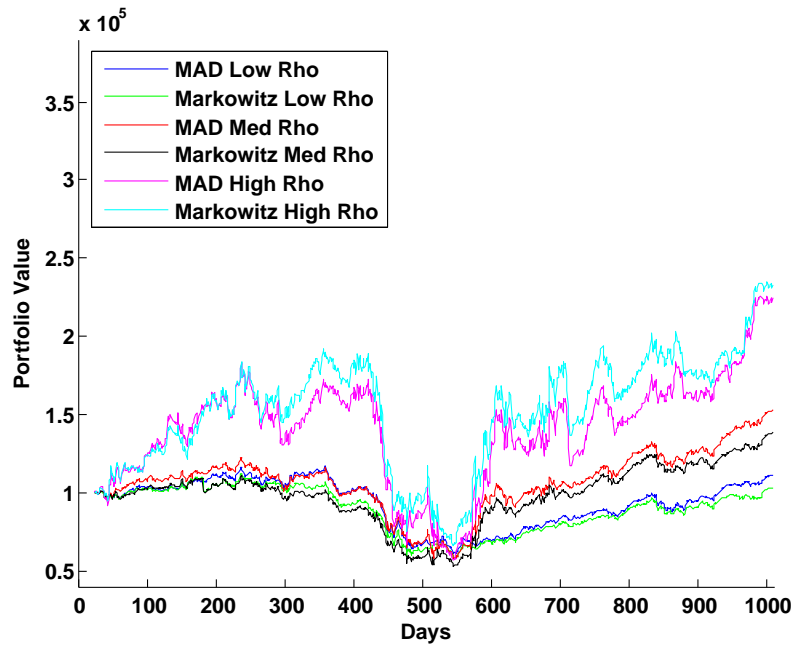


Figure 3.5 Portfolio net worths with 21 day rebalancing

| Horizons | 252 | 126 | 63 | 21 |
|---------------|--------|--------|--------|---------|
| ρ_{Low} | 1.047 | -3.025 | -2.603 | 3.942 |
| ρ_{Med} | 0.126 | -.475 | -1.793 | 8.145 |
| ρ_{High} | -1.050 | -1.682 | 34.141 | -11.533 |

Table 3.7 Average Portfolio Differences (in \$ 1000) for Multiple Time Horizons and Multiple ρ

| Horizons | 252 | 126 | 63 | 21 |
|---------------|--------|--------|--------|--------|
| ρ_{Low} | 78.44% | 2.27% | 33.86% | 96.35% |
| ρ_{Med} | 46.96% | 47.85% | 40.32% | 98.68% |
| ρ_{High} | 21.30% | 34.58% | 95.56% | 12.16% |

Table 3.8 Percent of total days where the MAD portfolio outperformed the Markowitz portfolio

the MAD model.

Table 3.8 compares the individual daily values of each portfolio to determine the percentage of days where the MAD portfolio performed better than the Markowitz portfolio. Examining this table, we see that there is a wide range of values; however taking the average of the averages to get a more overall view we find that MAD is better than Markowitz only 50.69 % of the time. This value is interesting because it shows that while there is a large variation between performance of the MAD and Markowitz algorithms in the 12 simulations, on average both portfolios perform very similarly.

3.6.2 Simplex vs. Interior-Point Methods

Earlier we mentioned that MATLAB provides two methods for solving linear programs, the simplex method, and an interior-point method (for a good treatment of

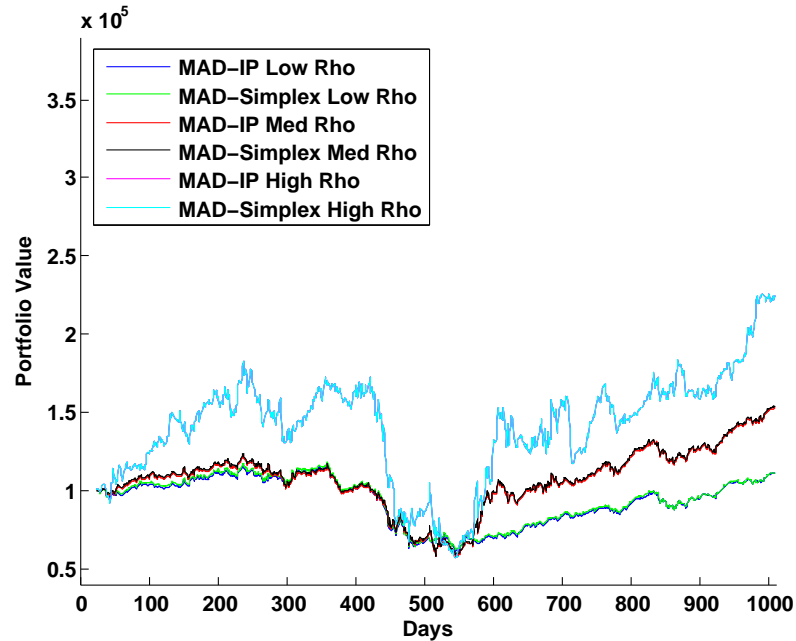


Figure 3.6 MAD-IP and MAD-Simplex portfolio net worths with 21 day rebalancing

these methods see [28]). Using the same parameters as before I computed the MAD portfolio twice, once using simplex and once using interior-points. These two methods will be referred to as MAD with Simplex and MAD with Interior Points, or simply MAD-Simplex and MAD-IP. Portfolios were generated for all four rebalancing horizons and all three required minimum returns. The resulting portfolios were identical except for the 21 day rebalancing portfolios which are shown in Figure 3.6.

In Figure 3.6 there is a slight difference between the portfolio net worths for low and medium ρ values, but the portfolios with high-rho values are identical. As can be seen in the figure, the MAD-Simplex portfolio is slightly better than the MAD-IP portfolio. Indeed, the MAD-Simplex portfolio outperforms the MAD-IP portfolio for 98% of the study, but the value of the MAD-Simplex portfolio is on average only 0.9% and 0.86% higher than the MAD-IP portfolio for the low and medium

ρ values respectively. These results demonstrate that for longer time horizons the MAD-Simplex and MAD-IP portfolios are identical; however for lower horizons the simplex method seems to perform slightly better.

3.6.3 Starting Portfolio Trading at the Same Time

All of the portfolios up until now have waited a certain number of days before starting to invest. For example, the 252 day rebalancing portfolio uses the past 252 days to make a prediction of the expected return of each stock. This means that the portfolio will wait one year and watch the market then start trading on the 253rd day. The 21 day rebalancing strategy will start trading on the 22nd day and will have rebalanced 11 times before the 252 rebalancing portfolio has made a single transaction. The advantage in using this methodology is that each portfolio is given the same amount of data and then the learning algorithm determines when to start investing.

Another approach is to use a single starting date for each portfolio. Each learning algorithm uses the past m days before this starting date to estimate expected returns. The advantage of this method is that each portfolio is held for the same amount of time and some portfolios do not start making/losing money before others; however, this also means each portfolio has access to different amounts of data which will also affect the results. To see if our conclusions still hold when starting each portfolio on the same date, I redid the previous experiment, starting every portfolio on day 253 and letting each portfolio have access to the necessary amount of previous data to start trading on the same date. This gives a study window of three years on which to test the MAD and Markowitz portfolios. The net worths of these new portfolios are recorded in tables 3.9 and 3.10.

These values are not as high as the four year study with different start dates, but it is interesting to note that the highest returns were obtained from the same

| Horizons | 252 | 126 | 63 | 21 |
|---------------|---------|---------|--------|--------|
| ρ_{Low} | 2.57% | 1.25% | 8.19% | -3.97% |
| ρ_{Med} | 1.37% | 18.99% | 37.22% | 30.22% |
| ρ_{High} | -33.04% | -46.26% | 70.96% | 38.83% |

Table 3.9 Return on investment for Markowitz Portfolio over multiple time horizons and multiple ρ values

| Horizons | 252 | 126 | 63 | 21 |
|---------------|---------|---------|--------|--------|
| ρ_{Low} | 4.87% | -.01% | -3.65% | .73% |
| ρ_{Med} | -0.11% | 18.35% | 21.89% | 31.57% |
| ρ_{High} | -32.23% | -48.14% | 79.86% | 36.53% |

Table 3.10 Return on investment for MAD Portfolio over multiple time horizons and multiple ρ values

parameters—63 day rebalancing and ρ_{High} . Computing the average portfolio differences for multiple horizons and multiple ρ values for portfolios starting on the same date produced the results in Table 3.11. The highest and lowest average portfolio differences are much smaller than those reported in Table 3.7. I also compared the individual daily values of each portfolio to determine the percentage of days where the MAD portfolio performed better than the Markowitz portfolio. Table 3.8 shows

| Horizons | 252 | 126 | 63 | 21 |
|---------------|--------|--------|---------|---------|
| ρ_{Low} | 1.047 | -2.789 | -6.7081 | 1.9134 |
| ρ_{Med} | 0.126 | -3.68 | -6.9441 | -0.6765 |
| ρ_{High} | -1.050 | -1.16 | 4.9258 | -7.2834 |

Table 3.11 Average Portfolio Differences (in \$ 1000) for Multiple Time Horizons and Multiple ρ

| Horizons | 252 | 126 | 63 | 21 |
|---------------|--------|--------|--------|--------|
| ρ_{Low} | 78.44% | .79% | 0.93% | 80.42% |
| ρ_{Med} | 46.96% | .93% | 0.26% | 42.33% |
| ρ_{High} | 21.30% | 26.98% | 96.43% | 1.72% |

Table 3.12 Percent of total days where the MAD portfolio outperformed the Markowitz portfolio

these percentages.

Taking the average of all of these values the MAD portfolio outperformed the Markowitz portfolio only 33.12% of the time. This value is much lower than the previous experiment and implies that the Markowitz portfolios actually performed better than the MAD portfolios on average; however, the actual differences between the MAD and Markowitz portfolios is smaller, and there is little consistency among different parameter combinations.

In conclusion, I found that when using MATLAB's interior-point method to solve the MAD and Markowitz models there is no clear winner between the two. Sometimes MAD is better and sometimes Markowitz is better; however, we see that on average they both have similar performance. Studying the results of the individual simulations I found that the MAD and Markowitz portfolio values track each other and have similar net worths for every simulation. To get a more accurate picture of which model is best I present an analysis of the individual portfolio compositions of the MAD and Markowitz portfolios as well as an analysis of the differences between using the simplex and interior-point methods for computing the optimal MAD portfolio.

3.7 Analysis of Portfolio Compositions

We now turn our attention to the actual stocks that compose the portfolios we have been studying. Each portfolio can consist of anywhere from 1 to 485 different investments, each consisting of a certain percentage of the investor's wealth, allowing for an infinite number of possible compositions. Through an analysis of the portfolio compositions determined by the MAD and Markowitz models, I show that both portfolios use similar numbers of stocks, but that despite this similarity there is still variation between the types of stocks and the investment amounts chosen.

3.7.1 Average Stocks in Portfolio

In order to further investigate the similarities and differences between the MAD and Markowitz portfolios, I first examined the number of stocks included in each portfolio. Tables 3.13-3.15 show the average number of stocks in each portfolio using ρ_{Low} , ρ_{Med} , and ρ_{High} .

As can be seen from all three tables, the average number of stocks for the MAD-Simplex and MAD-IP are the same for every experiment except for the 21 day horizons using ρ_{Low} and ρ_{Med} . For these simulations the MAD-IP (highlighted in red) had several portfolios that used every single stock in the optimal portfolio thus increasing the average number. Another surprising result was the average number of stocks used

| Horizons | 252 | 126 | 63 | 21 |
|-------------|--------|--------|--------|--------|
| MAD-Simplex | 23.333 | 21.714 | 20.133 | 14.936 |
| MAD-IP | 23.333 | 21.714 | 20.133 | 34.681 |
| Markowitz | 21 | 20 | 20.333 | 14.766 |

Table 3.13 Average Number of Stocks in portfolio for different horizons using ρ_{Low}

| Horizons | 252 | 126 | 63 | 21 |
|-------------|-----|--------|--------|---------------|
| MAD-Simplex | 17 | 19.143 | 18.733 | 14.043 |
| MAD-IP | 17 | 19.143 | 18.733 | 33.766 |
| Markowitz | 15 | 16.857 | 17.933 | 13.894 |

Table 3.14 Average Number of Stocks in portfolio for different horizons using ρ_{Med}

| Horizons | 252 | 126 | 63 | 21 |
|-------------|-------|-------|--------------|--------------|
| MAD-Simplex | 5.333 | 4.286 | 5.133 | 5.383 |
| MAD-IP | 5.333 | 4.286 | 5.133 | 5.383 |
| Markowitz | 5 | 4.286 | 4.933 | 5.511 |

Table 3.15 Average Number of Stocks in portfolio for different horizons using ρ_{High}

in the Markowitz portfolios. Except for the portfolio that used the 21 day horizon and ρ_{High} (highlighted in blue), the Markowitz portfolios actually used fewer stocks than both of the MAD portfolios. This contradicts the third assertion made by Konno and Yamazaki in [14] that the Markowitz portfolio suffers from too many small portfolio values.

When I initially ran my simulations, the MATLAB quadratic programming optimization algorithm would reach the maximum number of iterations before converging and the resulting portfolio would consist of almost every stock; however, increasing the maximum iterations to 2000 eliminated the problem and allowed MATLAB's algorithm to converge to an optimal solution. The results in this thesis show that optimal solutions to the Markowitz model use slightly fewer rather than more stocks than solutions to the MAD model. The only instances of Markowitz portfolios which consist of small investments in almost every stock have been found when the opti-

mization algorithm terminated before finding an optimal solution. In contrast to the assertion made in [14], it appears that the interior point method used by MATLAB for solving linear programs occasionally falls into the trap of investing small amounts in every stock. The entries highlighted in red show that the MAD-IP algorithm can select too many stocks and come up with an unrealistic portfolio consisting of too many small investments.

One final observation is that the best portfolio considered so far (the MAD portfolio using 63 day rebalancing and ρ_{High}) used a very small number of stocks to achieve superior performance. The entries highlighted in green show that the MAD algorithm used on average 5.133 stocks in its portfolio, yet was able to achieve 261.26% returns as reported in Table 3.4. This shows that only a small number of stocks are needed to achieve high returns and that the 63 day rebalancing portfolio with high required returns was able to select out of 485 choices those 5 or 6 stocks which had exceptional performance. For a list of the stocks used in this portfolio see Appendix A.

3.7.2 Similarity of Stock Choices

We have seen that the average number of stocks used in the MAD and Markowitz portfolios are very similar. I next investigated whether the actual stocks used in each portfolio were the same or whether the portfolios were achieving similar performance with different stocks. I calculated the percent of stocks that were shared between the Markowitz portfolio and the MAD-Simplex portfolio. The results are shown in tables 3.16 and 3.17.

These values show that while the average number of stocks included in each portfolio is very similar, on average only 75% of the stocks in each portfolio match. This shows that while the MAD-Simplex and Markowitz portfolios exhibit similar returns, there is still variation between the two portfolios.

| | | | | |
|---------------|--------|--------|--------|--------|
| ρ_{Low} | 252 | 126 | 63 | 21 |
| ρ_{Med} | 68.98% | 62.72% | 66.43% | 72.98% |
| ρ_{High} | 74.10% | 65.48% | 65.29% | 71.96% |
| Markowitz | 91.67% | 77.86% | 88.87% | 79.83% |

Table 3.16 Average percent of stocks in MAD-Simplex portfolio that were also included in the Markowitz portfolio for different horizons and ρ values

| | | | | |
|---------------|--------|--------|--------|--------|
| ρ_{Low} | 252 | 126 | 63 | 21 |
| ρ_{Med} | 77.64% | 68.41% | 66.32% | 73.98% |
| ρ_{High} | 85.29% | 73.47% | 71.21% | 73.41% |
| Markowitz | 95.24% | 78.57% | 89.04% | 78.56% |

Table 3.17 Average percent of stocks in Markowitz portfolio that were also included in the MAD-Simplex portfolio for different horizons and ρ values

3.8 Analysis of Computational Complexity

I have shown that the resulting portfolios from the MAD and Markowitz models are similar and have comparable performance; however, the computational complexity of each algorithm must also be considered. Time complexity is an important topic in algorithm analysis because if one algorithm has higher performance than another algorithm, but takes twice as long to compute it may be better to use the algorithm with a lower performance in order to meet time constraints. For this experiment I compare and contrast the computational time required to compute each portfolio in MATLAB using the MAD and Markowitz optimization method. The Markowitz model is a quadratic program so its theoretical time complexity is much larger than the MAD model solved with an interior-point method. The simplex method has an exponential worst case time complexity, but typically performs well in actual practice [28]. I first compare how the time complexity of these three methods grows

as n , the number of stocks available for the portfolio, increases.

Using the timer feature in MATLAB I ran one optimization using the MAD-Simplex, MAD-IP, and Markowitz models for varying numbers of stocks. I started with 5 stocks and then continued to add 20 stocks until the maximum of 485 stocks was reached. Figure 3.7 shows the results of this experiment where the points for each individual experiment have been connected to approximate the time complexity curve. I set the required return to ρ_{High} and used the first 63 days of return values to calculate the each optimal portfolio.

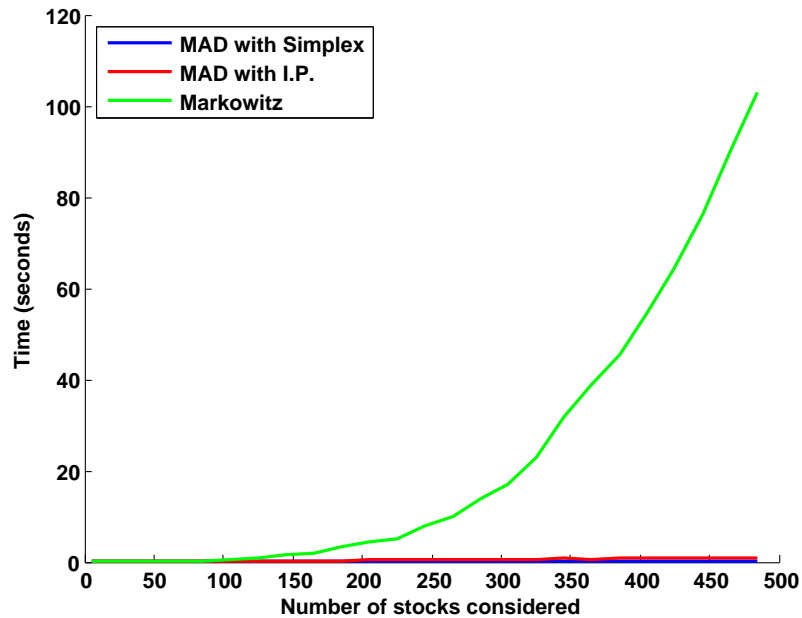


Figure 3.7 Time complexity as number of available stocks increases

Figure 3.7 shows the quadratic time complexity of the Markowitz model as the number of stocks considered in the portfolio grows. Both the MAD-Simplex and MAD-IP time complexity is linear, and much lower than the Markowitz time complexity. Also, even though the MAD-Simplex is an exponential time algorithm, in this case, it has an even lower time complexity than the MAP-IP algorithm. Tests

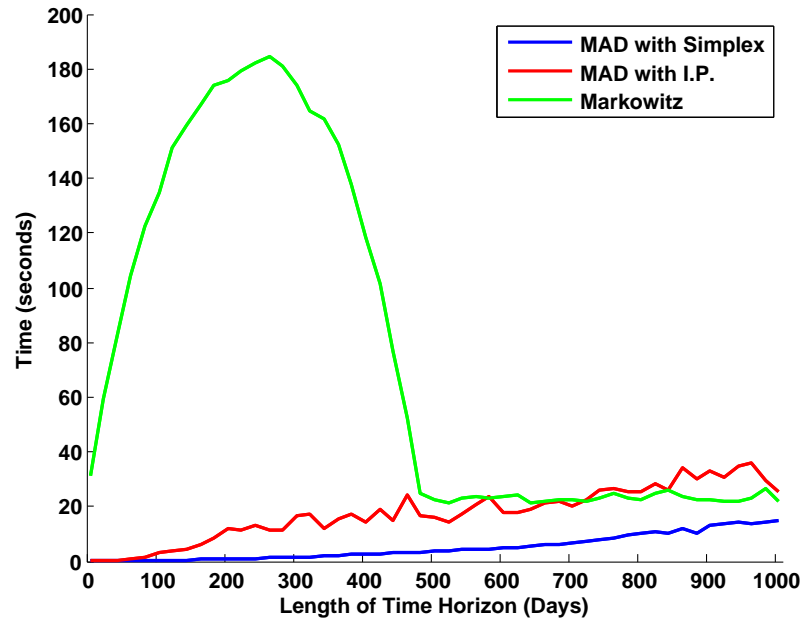


Figure 3.8 Time complexity as Learning Horizon Increases using ρ_{High}

were done using different values of ρ but showed no significant change from the results of Figure 3.7 so they are not included.

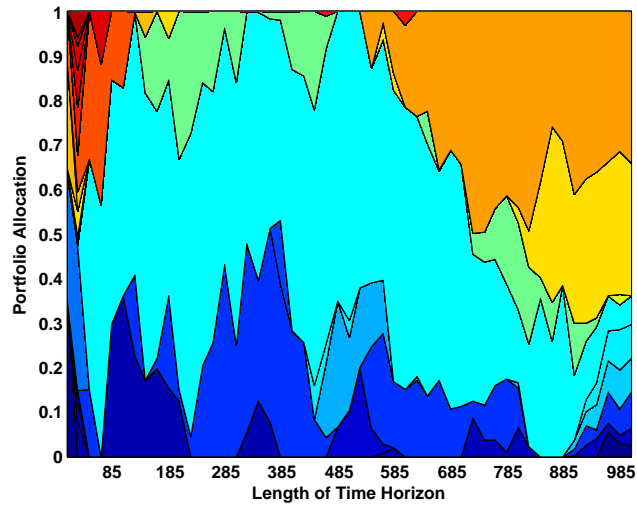
Figure 3.7 clearly shows that the Markowitz model has a quadratic computational complexity as the number of stocks considered for the portfolio increases. Another feature to consider is how the time complexity increases as the historical time horizon m increases. Using the same framework as above I kept the number of stocks fixed at 485 and varied m . Starting with $m = 5$ (the first 5 days of 2007) I continued to add 20 days until the S&P 500 data sample was exhausted. I repeated this simulation using ρ_{Low} , ρ_{Med} , and ρ_{High} as the required minimum return. The results of the portfolios using ρ_{High} are graphed in Figure 3.8. The portfolios using ρ_{Med} and ρ_{Low} were almost identical so they are not included.

Figure 3.8 shows that even though the MAD-Simplex and MAD-IP algorithms take less time to compute than the Markowitz model for smaller time horizons, for

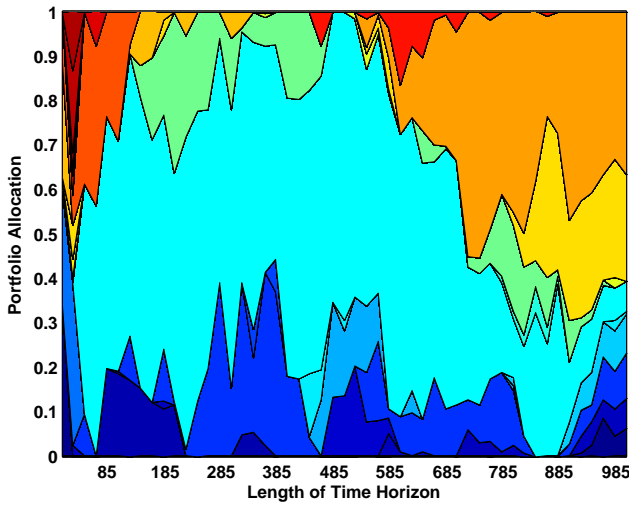
horizons longer than 500 days the three portfolios take roughly the same amount of time to find an optimal solution. The large bump in the time complexity for the Markowitz model seemed counterintuitive, and I thought at first that there had been a problem with MATLAB; however, using output from the MATLAB optimizations I verified that all the runs had converged successfully to an optimal solution.

To investigate this further I compared the compositions of the Markowitz portfolios with the compositions obtained by both of the MAD algorithms. All three algorithms showed very similar results in terms of portfolio composition. The MAD-Simplex and MAD-IP portfolios were identical except for the portfolio with a time horizon of 5 days. The Markowitz portfolio had very similar stocks but the actual amount invested in each stock was different than the two MAD portfolios. Because all three algorithms selected very similar stocks, this provides strong evidence that the portfolios being generated by the Markowitz model are optimal, and that the large bump in Figure 3.8 is not a result of the algorithm terminating early. The similarity in portfolio composition can be seen by comparing the graphs in Figure 3.9. The colored areas represent the proportion of the portfolio invested in different stocks. Comparing all three figures we see that despite minor differences, the overall compositions of the three portfolio are extremely similar. In fact the only difference between 3.9b and 3.9c is the initial portfolio composition with a time horizon of 5 days which is not noticeable in the figures.

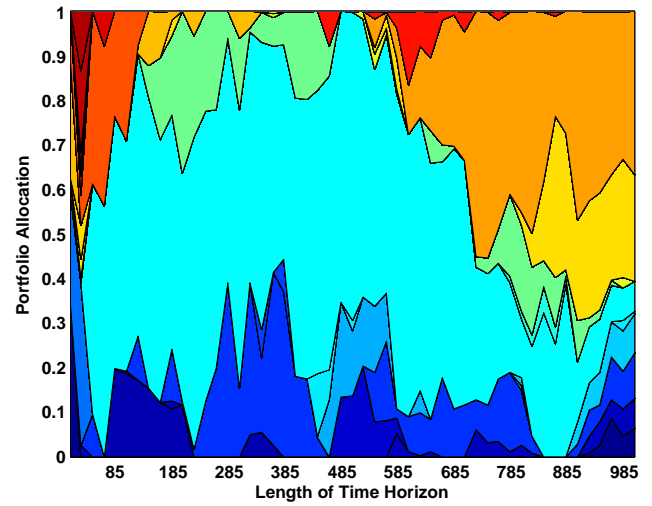
I also plotted the number of iterations needed by each algorithm to compute the optimal portfolio. Figure 3.10 shows that the number of iterations required to find the Markowitz portfolio and the MAD portfolio using interior points are relatively stable; however, the number of iterations to compute the MAD portfolio using simplex is steadily increasing as the horizon increases so this rules out the number of iterations as a possible cause of the bump in time complexity.



(a) Markowitz Model



(b) MAD-IP Model



(c) MAD-Simplex Model

Figure 3.9 Portfolio compositions as time horizon changes

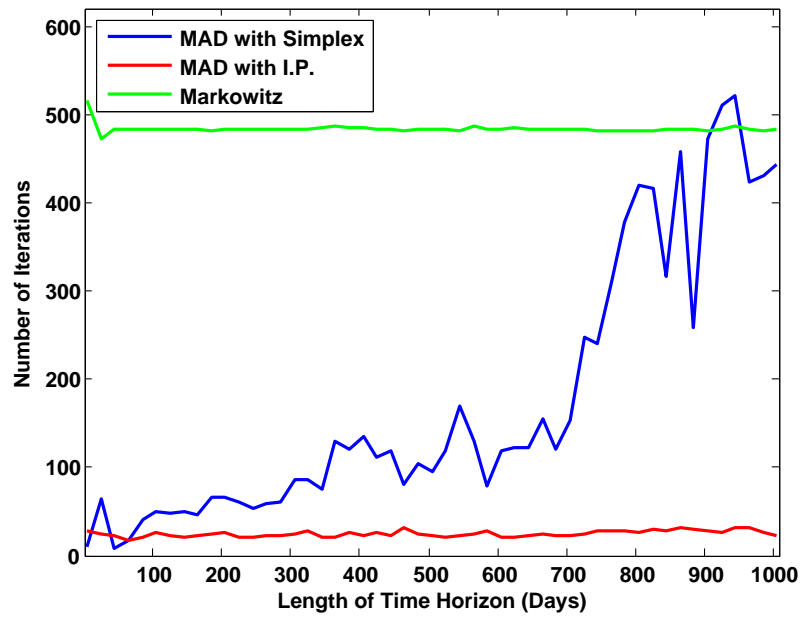


Figure 3.10 Number of iterations to compute optimal portfolio over changing horizon

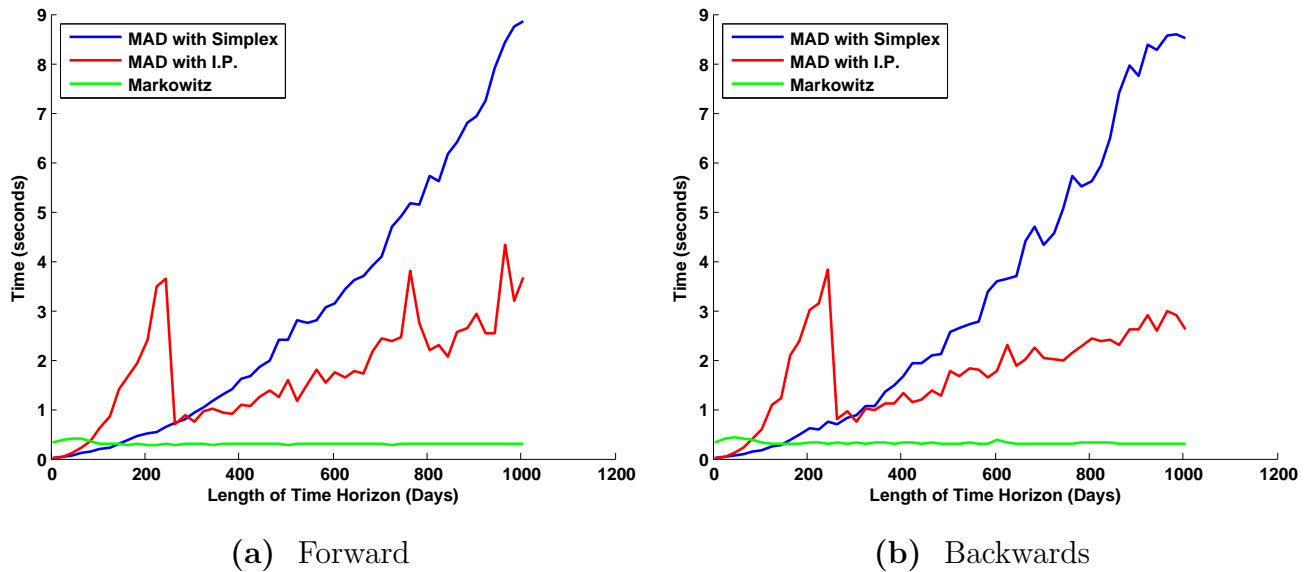


Figure 3.11 Time Complexity using 100 stocks

I also ran a series of experiments using different numbers of stocks rather than just using all 485. I first ran a series of experiments where I fixed the number of potential stocks and the progressively increased the time horizon starting with day 1 (Jan 1 2007) and adding consecutive days and stocks in increasing order. To check these results I also ran the same experiment backwards using the same values for number of stocks and number of days in the time horizon, but starting with day 1008 and stock number 485 and reversing the order that stocks and days were added. This provides a test to see if the observed trend is related to the specific data or if it is a more general trend that is true regardless of the specific data. For a fixed number of stocks less than 200 the computation time for the Markowitz portfolio was much smaller than both the MAD-IP and MAD-Simplex portfolios. Also the time required to generate the MAD-Simplex portfolio took longer than the time to generate the MAD-IP portfolio. Figure 3.11 shows the results when 100 stocks are used.

Above 200 stocks the bump in time complexity around the 200 day horizon mark

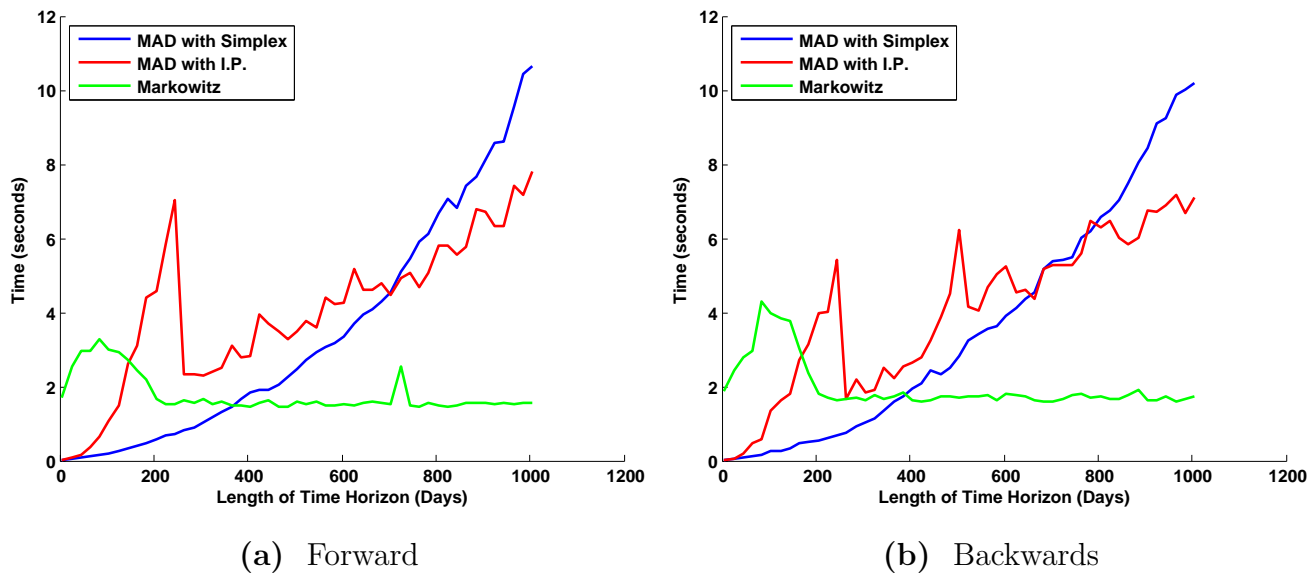


Figure 3.12 Time Complexity using 200 stocks

begins to grow and the time required to for the simplex algorithm drops below the interior point method. Even though there is a large bump in the time required for the Markowitz model, after the time horizon has reached 400 days, the time for the Markowitz model drops below the time required to compute the MAD portfolio using the interior-point method. Figures 3.12 and 3.19 show the results for 200, 300, 400 and 485 stocks. The results of running the experiments backwards are plotted alongside the original experiments and showed the same trend.

Figures 3.16 - 3.21 show the results of holding the time horizon constant and varying the number of stocks considered by the optimization algorithm. These experiments were run forwards and backwards and show the same pattern as Figures 3.11 - 3.14. Figures 3.16 - 3.20 show clearly the quadratic increase in time for the Markowitz model. Examining the range y-axis we also see the time required to compute the optimal portfolio is highest for the 252 day horizon (Figure 3.20).

Examining the MAD-Simplex, MAD-IP, and Markowitz portfolios individually,

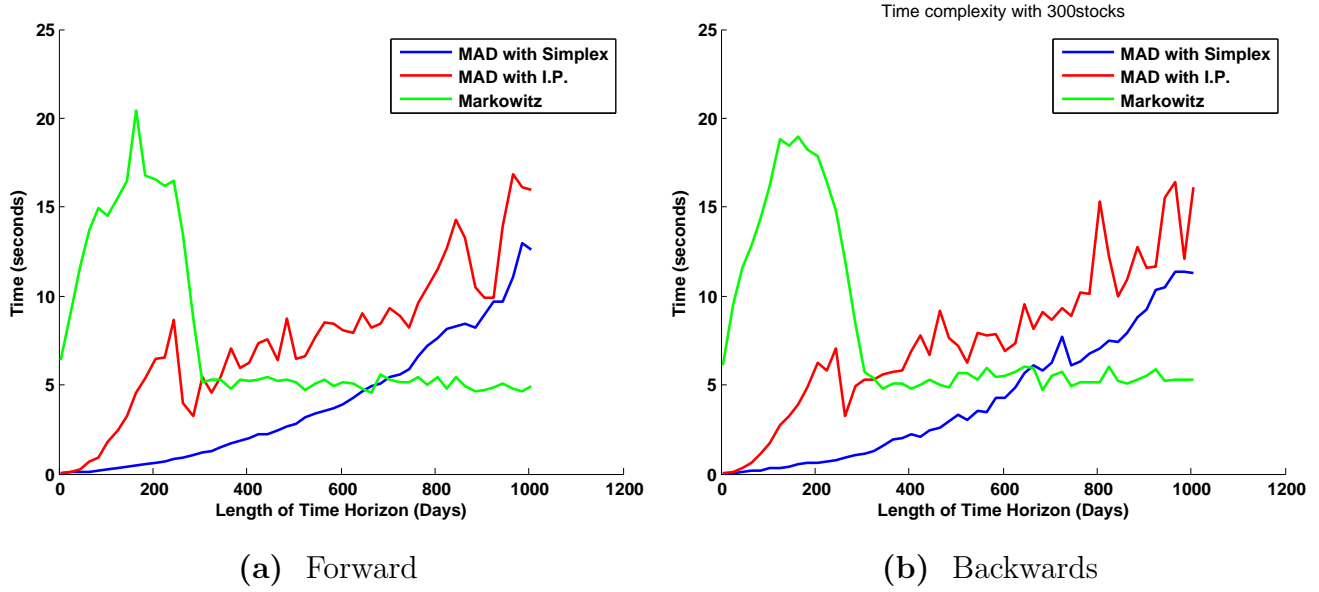


Figure 3.13 Time Complexity using 300 stocks

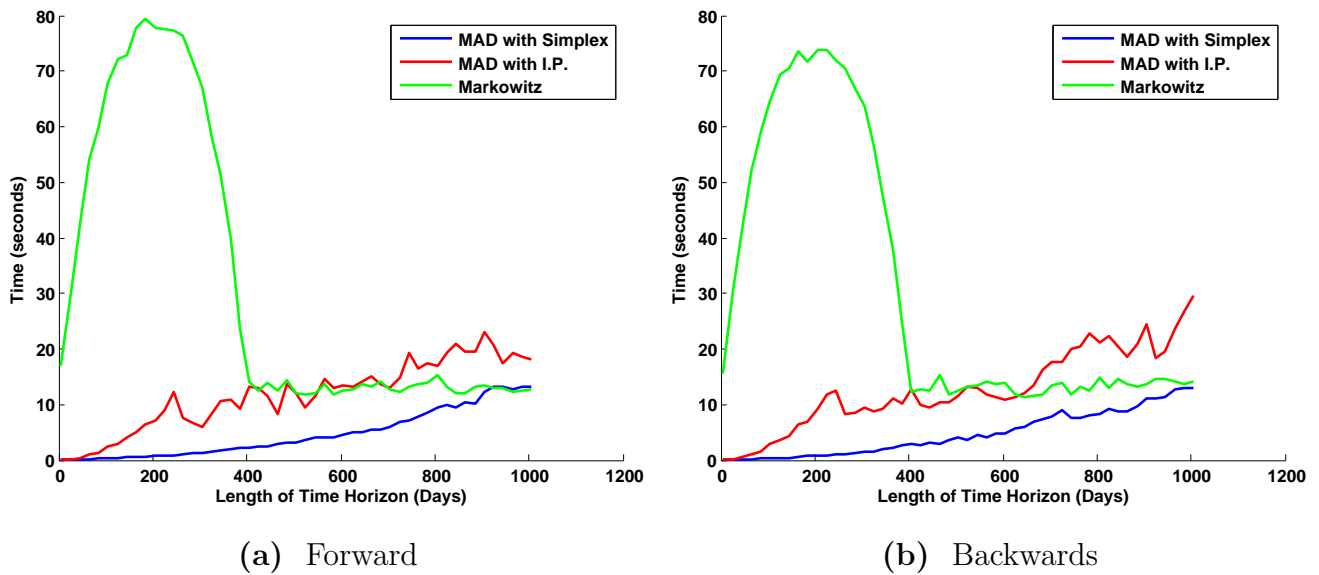


Figure 3.14 Time Complexity using 400 stocks

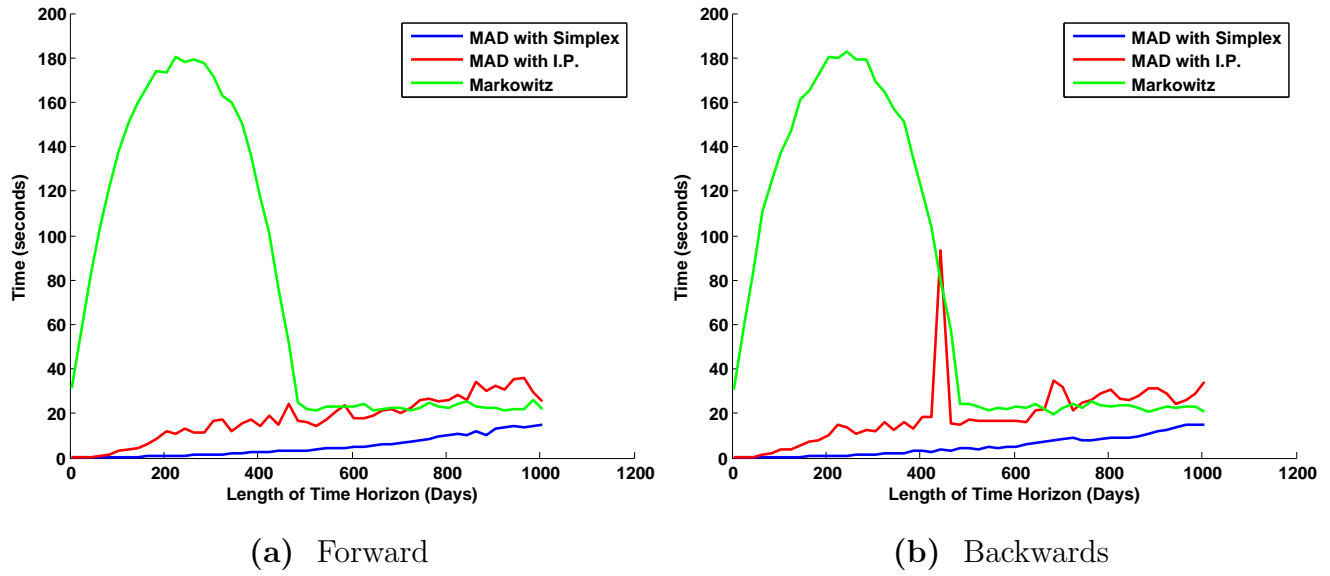


Figure 3.15 Time Complexity using 485 stocks

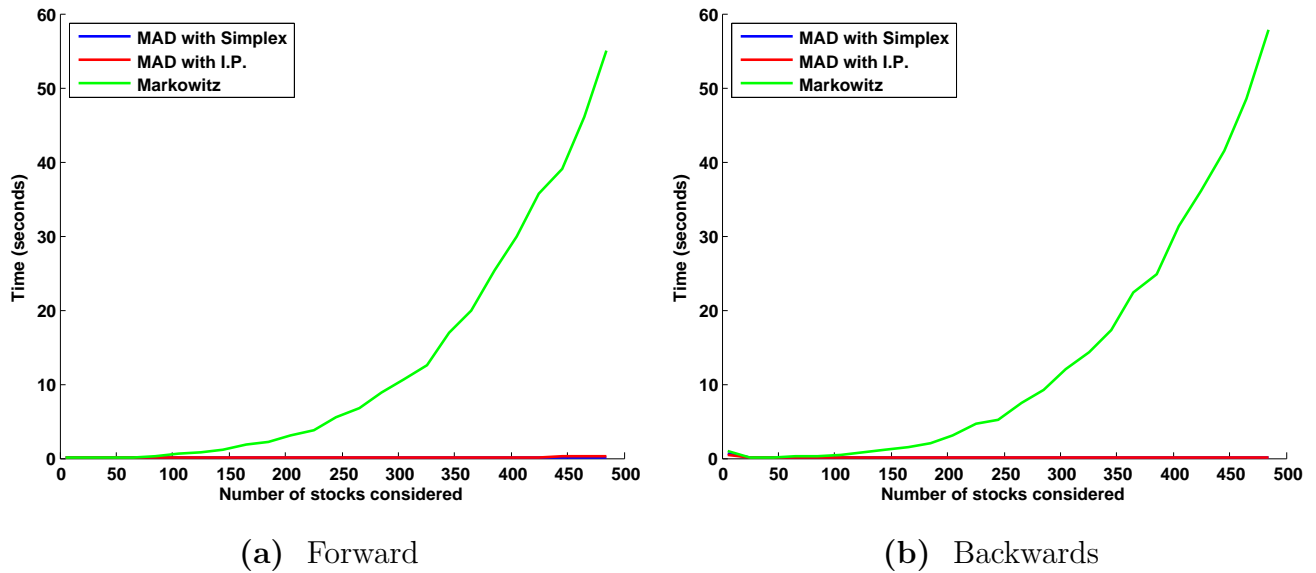


Figure 3.16 Time Complexity using 21 days

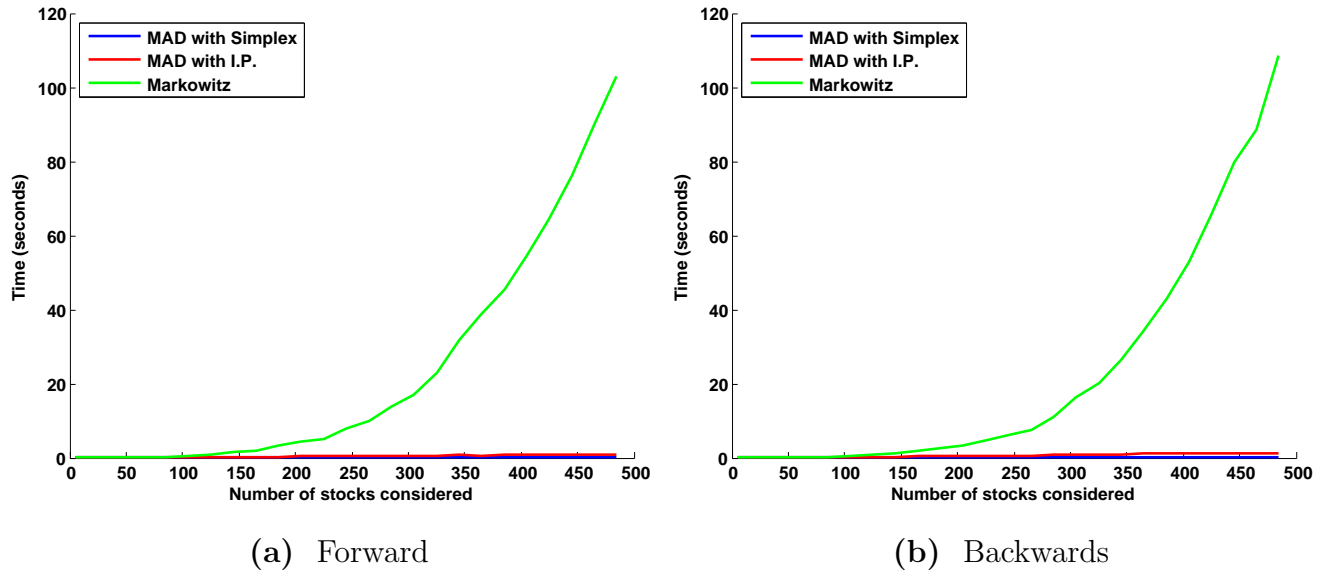


Figure 3.17 Time Complexity using 63 days

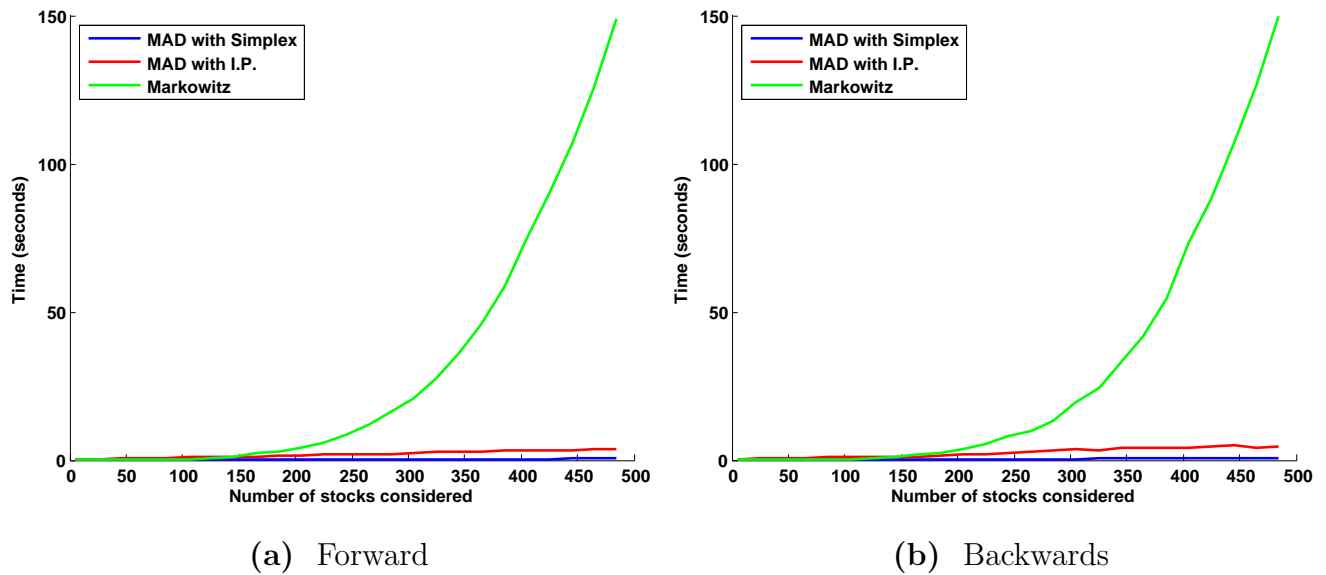


Figure 3.18 Time Complexity using 126 days

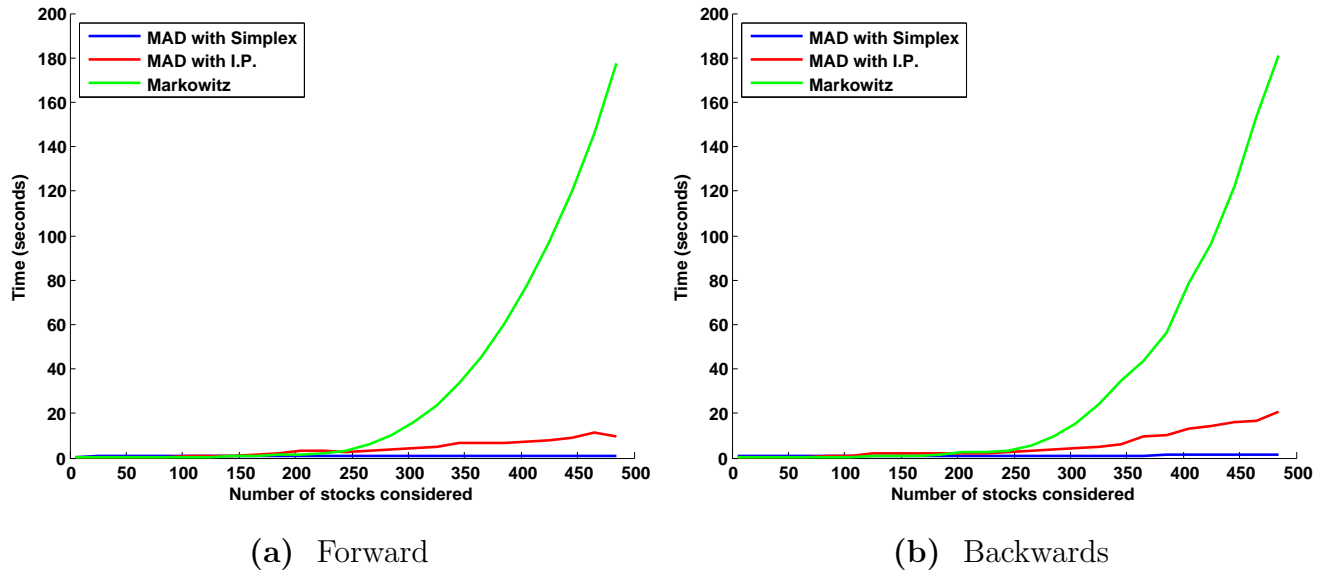


Figure 3.19 Time Complexity using 252 days

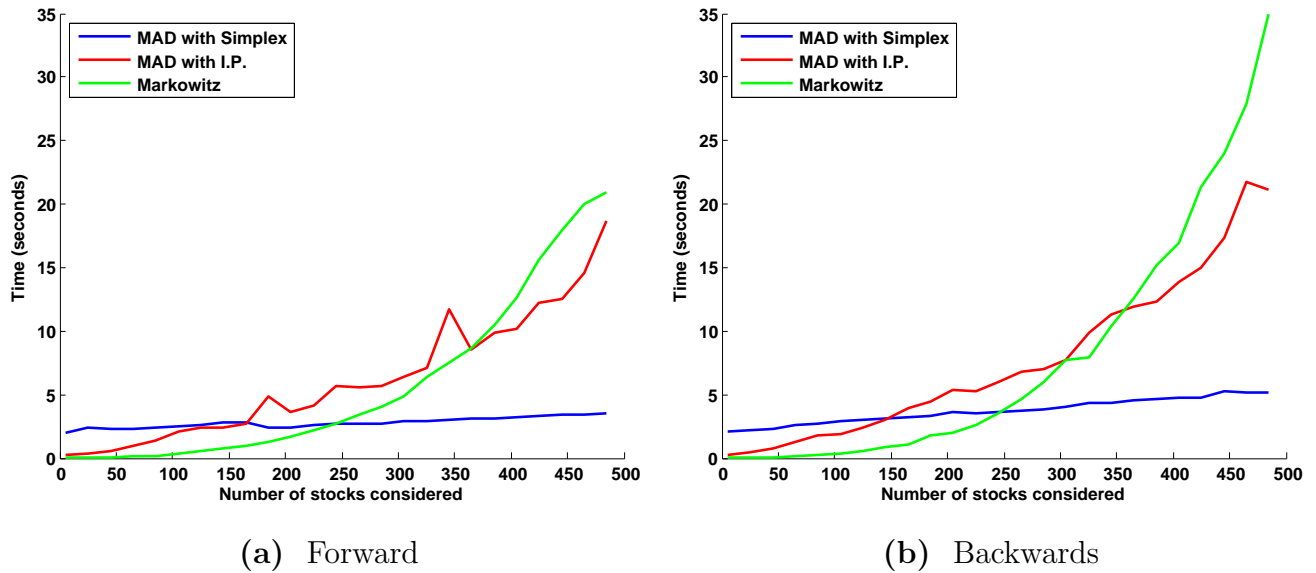


Figure 3.20 Time Complexity using 504 days

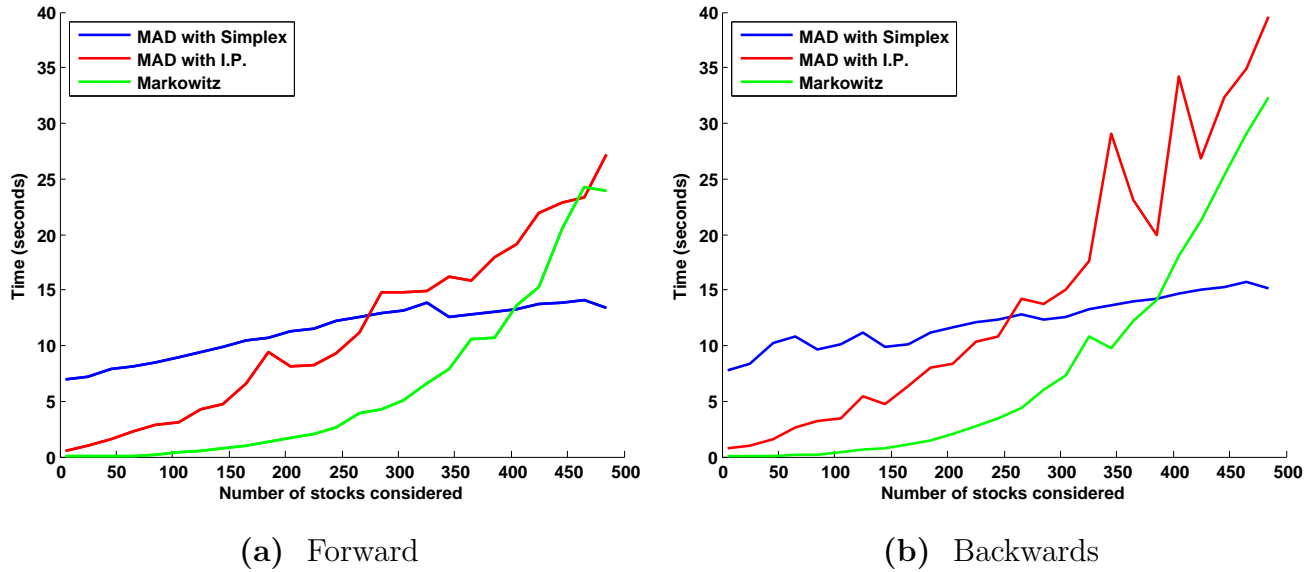


Figure 3.21 Time Complexity using 1008 days

there is a clear trend– the MAD-Simplex and MAD-IP portfolios take longer to compute as the time horizon increases, but the time to compute the Markowitz portfolio first increases and then decreases as the time horizon gets larger. For separate plots of each of these algorithms see Appendix C.

3.8.1 Time Complexity Using Random Returns

As one final test the unusual time complexity bump, I created a data set of random daily returns with 485 artificial stocks each with 1008 random returns sampled from a normal distribution with mean 1 and standard deviation 0.01. This data set was designed to be the same size as the S&P 500 data set to compare time complexity results. I tested 5 different numbers of stocks: 100, 200, 300, 400, 485. For each of these five experiments the number of stocks was held constant and the time horizon was slowly increased starting with 5 days and continuing to add 20 stocks until the entire data set was exhausted. The results closely match the time complexity results

for the S&P 500 data set and are included in Appendix D.

3.8.2 Time Complexity Summary

These results show that the large bump when using 485 stocks is not data specific, but is potentially related to the optimization algorithms used by MATLAB. Further investigation into this phenomenon would be instructional, but because understanding the time complexity of the Markowitz model is not the main goal of this thesis so I note this strange occurrence, but admit that I do not have any explanation for its cause.

3.9 Analysis of Different Test Horizons and Rebalancing Strategies

For this experiment I wanted to see if I could find a rebalancing strategy with higher returns than the MAD portfolio with 63 day rebalancing. I tested over the following rebalancing horizons: 5, 10, 15, 21, 35, 49, 63, 84, 105, 126. The training horizon for this experiment is equal to the testing horizon. Because portfolios using ρ_{Low} had the lowest average performance in previous experiments, I decided to test these horizons only using ρ_{Med} and ρ_{High} . I used the same four year data set of S&P 500 stock returns.

3.9.1 Results of Multiple Test Horizons

The final values of each portfolio are graphed in Figure 3.22. From this bar graph we see that the 63 day rebalancing horizon had the highest final return when using the parameter ρ_{High} . The 21 day and 15 day horizons also performed very well. The 15

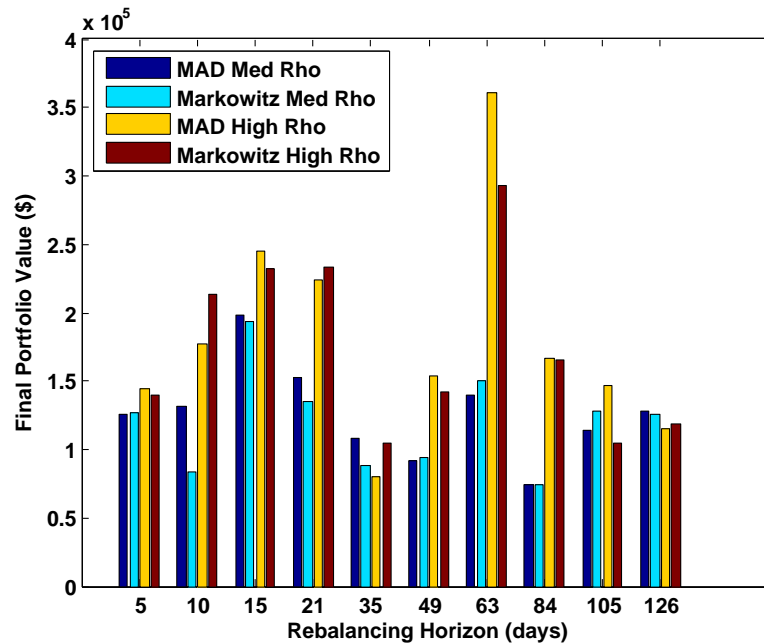


Figure 3.22 Final Portfolio Values for MAD and Markowitz portfolios using different rebalancing horizons for medium and high ρ values

day horizon had the highest final portfolio value when ρ_{Med} was used. It is interesting that the 35 and 49 day horizons performed so much worse than the 63, 21, and 15 day horizons. The reason for this is unclear, but from this analysis we see that the 63 day horizon had the best performance out of all other rebalancing horizons.

3.9.2 Changing the Learning Horizon

Next, I investigated whether changing the learning horizon had any impact on the performance of the different portfolios. The goal of this experiment is to test whether learning on more data improves portfolio performance. For this experiment I decided to use the original test horizons of 21 and 63 days but change the amount of data that is used to estimate the expected returns. The length of the training data was determined by multiplying the test horizon by a training multiple and rounding up to

the nearest whole number. I did experiments with the following training multiples: 0.5, 1, 1.5, 2, and 3. For example, when running the experiments for the 21 day test horizon I used training horizons of 11, 21, 32, 42, and 63 days. I also ran an experiment for each test horizon where the learning algorithm used all available historical data. For each run using all available data, I started with a training set equal to the test set and then continued to add days to the training set until the simulation was done. For example, when computing the optimal portfolio for the last 21 day rebalancing horizon, the test set consisted of returns from days 988 to 1008 and the training set consisted of days 1 to 987.

3.9.3 Results From Using Different Learning Horizons

Using the previously mentioned six different learning horizons and using low, med, and high ρ values, I ran each rebalancing strategy and recorded which strategies performed better than the original strategy with a training multiple of 1.0. The results can be seen in tables 3.18 and 3.19.

Examining both tables we see that six out of twelve portfolios achieved their highest final value using a training multiple of 1.0. One surprising result was that using a training multiple of 0.5 was better than a multiple of 1 in two cases, but in both cases using all the data was also better than a multiple of 1. These results seemed

| Horizons | 21 | 63 |
|---------------|-------------|----------|
| ρ_{Low} | 0.5, 3, all | 1.5, all |
| ρ_{Med} | - | 1.5 |
| ρ_{High} | all | - |

Table 3.18 Training multiples with higher final MAD portfolio value than a training multiple of 1.0

| Horizons | 21 | 63 |
|---------------|-------------|----|
| ρ_{Low} | 0.5, 3, all | - |
| ρ_{Med} | - | - |
| ρ_{High} | all | - |

Table 3.19 Training Multiples with higher final Markowitz portfolio value than a training multiple of 1.0

rather arbitrary and the main conclusion from this experiment was that a training multiple of 1 produced good results and that using all available data also has good performance. I had expected that on average portfolios using the 0.5 training multiple would have a lower performance than those using the 1 and 1.5 training multiples, and was surprised by the results. I also expected that more data will provide less accurate predictions of the expected value over time and would lead to lower portfolio performance which also was not true. Because there were no clear trends, it appears that the performance is highly dependent on the data and the other parameters, and that the learning horizon has less of an impact on portfolio returns.

3.10 Risk Adjusted Returns

I have shown that several of the portfolios in this study achieved phenomenal growth despite varying market conditions; however, one question still remains—did these portfolios actually beat the market? To answer this question it is necessary to not only consider the return of each portfolio, but to also consider each portfolio's risk. Having a risky portfolio is not necessarily a bad thing, as long as the returns gained from the portfolio provide adequate compensation for the risks taken. In order to evaluate the performance of the different portfolios studied in this thesis, I consider two measures of risk-adjusted performance: the Sharpe ratio, and the Modigliani

| Horizons | 252 | 126 | 63 | 21 |
|---------------|--------|--------|--------|--------|
| ρ_{Low} | 0.0080 | 0.0051 | 0.0190 | 0.0046 |
| ρ_{Med} | 0.0090 | 0.0216 | 0.0367 | 0.0246 |
| ρ_{High} | 0.0030 | 0.0217 | 0.0506 | 0.0412 |

Table 3.20 Sharpe ratio of daily returns for Markowitz Portfolio over multiple time horizons and multiple ρ values. Green is the best, red is the worst.

risk-adjusted measure [20, 26].

3.10.1 Analysis of Portfolio Performance using the Sharpe Ratio

Using daily returns for each portfolio I computed the Sharpe ratio for each portfolio.

The Sharpe ratio is defined as

$$\frac{\bar{r}_p - \bar{r}_f}{\sigma_p} \quad (3.5)$$

where \bar{r}_p is the average portfolio return, \bar{r}_f is the average return of the risk-free alternative, and σ_p is the standard deviation of portfolio returns or volatility [26].

The Sharpe ratio gives a measure of the excess returns achieved by a portfolio in relation to the amount of volatility in that portfolio where higher ratios indicate better risk-adjusted performance.

The results of calculating the Sharpe ratio for each portfolio are shown in tables 3.20 and 3.21. Yearly returns for 3 month T-Bills were converted into daily returns and used as a substitute for the risk-free rate. Comparing Sharpe ratio values we see that the best and worst Sharpe ratios correspond to the highest and lowest overall performance reported in tables 3.3 and 3.4.

The Sharpe ratio tells us the portfolios with the best risk-adjusted return; however, it tells us nothing of how the portfolios compared with the overall market. Using

| Horizons | 252 | 126 | 63 | 21 |
|---------------|--------|--------|---------|--------|
| ρ_{Low} | 0.0104 | 0.0033 | 0.0115 | 0.0106 |
| ρ_{Med} | 0.0077 | 0.0239 | 0.0303% | 0.0302 |
| ρ_{High} | 0.0028 | 0.0206 | 0.0577 | 0.0397 |

Table 3.21 Sharpe ratio of daily returns for MAD Portfolio over multiple time horizons and multiple ρ values. Green is the best, red is the worst.

daily returns for the S&P 500 index from 2007-2010, I found the Sharpe ratio to be -0.00092. Comparing this value with the Sharpe ratios for each portfolio we see that every portfolio had a higher risk-adjusted return than the market index.

3.10.2 Analysis of Portfolio Performance using the M^2 Measure

Because the Sharpe ratio can be difficult to interpret, another measure of performance, the Modigliani risk-adjusted measure or M^2 measure, is often used. The M^2 measure was developed by Leah and Franco Modigliani and gives a differential return relative to a benchmark index [20]. The M^2 measure is defined as

$$M^2 = \bar{r}_{P^*} - \bar{r}_M \quad (3.6)$$

where \bar{r}_{P^*} is the average return of a portfolio which has been adjusted to match the volatility of the benchmark index and \bar{r}_M is the average return of the benchmark index. The average return of the adjusted portfolio, \bar{r}_{P^*} , is calculated using the following formula:

$$\bar{r}_{P^*} = (\bar{r}_P - \bar{r}_f) \frac{\sigma_M}{\sigma_P} + \bar{r}_f \quad (3.7)$$

where σ_M is the volatility of the benchmark index.

Tables 3.22 and 3.23 show the results of using the M^2 measure to compare the

| Horizons | 252 | 126 | 63 | 21 |
|---------------|---------|---------|---------|---------|
| ρ_{Low} | 0.0130% | 0.0176% | 0.0363% | 0.0115% |
| ρ_{Med} | 0.0150% | 0.0478% | 0.0678% | 0.0464% |
| ρ_{High} | 0.0034% | 0.0479% | 0.0924% | 0.0752% |

Table 3.22 M^2 measure of daily returns for the Markowitz Portfolio over multiple time horizons and multiple ρ values. Green is the best, red is the worst.

| Horizons | 252 | 126 | 63 | 21 |
|---------------|---------|---------|---------|---------|
| ρ_{Low} | 0.0176% | 0.0142% | 0.0230% | 0.0220% |
| ρ_{Med} | 0.0124% | 0.0519% | 0.0563% | 0.0561% |
| ρ_{High} | 0.0032% | 0.0459% | 0.1049% | 0.0726% |

Table 3.23 M^2 measure of daily returns for the MAD Portfolio over multiple time horizons and multiple ρ values. Green is the best, red is the worst.

Markowitz and MAD portfolios using the S&P 500 as the benchmark index. These values seem very small, but after annualizing the risk-adjusted returns we see in tables 3.24 and 3.25 that even the portfolios with the worst performance still had a slightly better risk-adjusted performance than the S&P 500 index. Furthermore, we see that the best performing portfolios had average excess returns of 26.20% and 30.25% above the average risk-adjusted S&P 500 index returns.

| Horizons | 252 | 126 | 63 | 21 |
|---------------|-------|--------|--------|--------|
| ρ_{Low} | 3.33% | 4.53% | 9.56% | 2.93% |
| ρ_{Med} | 3.85% | 12.81% | 18.63% | 12.40% |
| ρ_{High} | 0.86% | 12.81% | 26.20% | 20.86% |

Table 3.24 M^2 measure of annualized returns for the Markowitz Portfolio over multiple time horizons and multiple ρ values. Green is the best, red is the worst.

| Horizons | 252 | 126 | 63 | 21 |
|---------------|-------|--------|--------|--------|
| ρ_{Low} | 4.54% | 3.65% | 5.96% | 5.69% |
| ρ_{Med} | 3.16% | 13.98% | 15.25% | 15.19% |
| ρ_{High} | 0.80% | 12.26% | 30.25% | 20.07% |

Table 3.25 M^2 measure of annualized returns for the MAD Portfolio over multiple time horizons and multiple ρ values. Green is the best, red is the worst.

Chapter 4

Conclusion

This thesis has focused on the application of learning and control techniques for making decisions in portfolio optimization. In particular, I showed the importance of explicitly formulating a solution to both the learning problem and the decision and control problem when solving financial optimization problems. In Chapter 2, I formulated the portfolio optimization problem and compared the different risk measures found in the Markowitz and MAD optimization models.

In Chapter 3 I performed an extensive study of the MAD and Markowitz models using returns from the stocks in the S&P 500 from 2007 through 2010. This large data set provided an ideal test bed to run simulations and observe the performance of portfolios over changing market conditions. Rather than limit the study to a single optimized portfolio, I implemented various rebalancing strategies to obtain a more realistic comparison. Through the comparison of portfolios generated by the MAD and Markowitz optimization models I showed that both portfolios achieve similar net worths and consist of similar but slightly different portfolio compositions. I also showed that the basic MAD model using a 63 day rebalancing horizon and ρ_{High} achieved outstanding performance and had a final four-year return on investment of

261.26%. I also showed that this portfolio had the highest Sharpe ratio and that it had an average annualized risk-adjusted return that was 30.25% more than the average risk-adjusted return of the S&P 500 index. Using the M^2 performance measure, I showed that all 24 portfolios described in section 3.6 outperformed the market.

Through an analysis of the average number of stocks in the MAD and Markowitz portfolios I showed that contrary to Konno and Yamazaki's claim in [14], the Markowitz model implemented in MATLAB actually uses slightly fewer stocks on average than the MAD-IP and MAD-Simplex models. I also observed that the MAD-IP portfolio will occasionally over diversify and invest very small amount of money in almost every possible stock. I also investigated the time complexity of the MAD-Simplex, MAD-IP, and Markowitz models. I discovered an interesting bump in the time complexity of the Markowitz model. Using several simulations I confirmed that this is not related to the specific data set but appears to be related either to the optimization algorithms in MATLAB, or my method of implementation. Based on this analysis I showed that the MAD and Markowitz models achieve similar returns, but that the Markowitz model is much more computationally intensive. I also showed that in many cases MATLAB's simplex method requires less time to compute than the interior-point method, and that the simplex method is more stable in terms of the number of investments. Finally, I examined a wider range of learning horizons and rebalancing strategies, but was unable to find a portfolio that performed better than the MAD model with 63 day rebalancing and a high required return. Based on the overall results, the MAD model solved with the simplex method requires the least amount of computation time, is robust to overdiversification, and achieves the highest returns.

4.1 Future Research

As with any research paper, there comes a time when new ideas and research directions must be put on hold and saved for another day. There were several research areas that I initially investigated where the research and results did not progress far enough to be included in this thesis. In the future I would like to perform a more careful analysis of the Markowitz model's time complexity when implemented in MATLAB. It would be beneficial to also implement the algorithm using several other computational languages to compare the results to see if the bump in time complexity is a general phenomenon or if it is unique to my implementation in MATLAB. It would also be beneficial to carefully examine the covariance matrix of stock returns in the Markowitz model to check for near singularity which may be affecting the computation time.

The use of different learning algorithms to predict future stock prices was mentioned in this thesis but the analysis and case studies only used the arithmetic mean as an estimator. Future research would benefit from exploring the use of different learning algorithms and comparing the performance of the MAD and Markowitz models when using these predictors. Another area of potential future research would be to develop algorithms that pick their learning horizons and rebalancing horizons. One way of doing this would be to have the optimization adjust for transaction costs, thereby limiting the number of daily transactions but allowing more flexibility when trading.

4.2 Epilogue

You've probably been wondering how well these models would perform in the future. It is an interesting question with no perfect answer since we don't know what surprises the future will hold. However, if you happen to have \$100,000 lying around and

are feeling daring, here's what the MAD-Simplex model would pick using a 63 day rebalancing horizon and ρ_{High} : 'Biogen Idec Inc.', 'Cephalon, Inc.', 'Lorillard, Inc ', and 'National Semiconductor Corporation'. I'm not making any promises, but if you happen to make 200% returns over the next four years, remember me.

Appendix A

Tables and Lists

A.1 Stocks from S&P 500 not included in study

The total list of all 500 stocks in the S&P 500, as of March 2011, were obtained from the Stock Screener on FINVIZ.com. The following stocks were added to the S&P 500 between January 2007 and December 2010: V, TWC, TDC, SNI, QEP, PM, PEG, PCS, MMI, MJN, LO, DPS, DFS, COV, CFN. These stocks were not included in the research for this thesis to prevent missing data points and to only keep stocks which were in the S&P 500 for the full two years of the study and may be considered more ‘constant’ during this time period and representative of the market.

Table A.1 Portfolio Results (in \$100,000) for Low ρ

| | Average Value | Average Horizon Value | Final Networkh |
|----------|---------------|-----------------------|----------------|
| Mad 252 | 0.888 | 0.92 | 1.0487 |
| Mark 252 | 0.8775 | 0.9027 | 1.0257 |
| Mad 126 | 0.9212 | 0.9198 | 1.0158 |
| Mark 126 | 0.9515 | 0.9512 | 1.0354 |
| Mad 63 | 1.0298 | 1.0346 | 1.1198 |
| Mark 63 | 1.0558 | 1.0626 | 1.2076 |
| Mad 21 | 0.9345 | 0.9376 | 1.1129 |
| Mark 21 | 0.895 | 0.8965 | 1.0313 |

Table A.2 Portfolio Results (in \$100,000) for Med ρ

| | Average Value | Average Horizon Value | Final Networkh |
|----------|---------------|-----------------------|----------------|
| Mad 252 | 0.7957 | 0.9200 | 0.9989 |
| Mark 252 | 0.7944 | 0.9027 | 1.0137 |
| Mad 126 | 1.0232 | 0.9198 | 1.2835 |
| Mark 126 | 1.0279 | 0.9512 | 1.2523 |
| Mad 63 | 1.1047 | 1.0346 | 1.4013 |
| Mark 63 | 1.1226 | 1.0626 | 1.5060 |
| Mad 21 | 1.0695 | 0.9376 | 1.5291 |
| Mark 21 | 0.9880 | 0.8965 | 1.3832 |

Table A.3 Portfolio Results (in \$100,000) for High ρ

| | Average Value | Average Horizon Value | Final Networth |
|----------|---------------|-----------------------|----------------|
| Mad 252 | 0.6248 | 0.5757 | 0.6777 |
| Mark 252 | 0.6353 | 0.5743 | 0.6696 |
| Mad 126 | 1.1260 | 1.1748 | 1.1543 |
| Mark 126 | 1.1428 | 1.1948 | 1.1901 |
| Mad 63 | 2.0903 | 2.1858 | 3.6126 |
| Mark 63 | 1.7489 | 1.8224 | 2.9288 |
| Mad 21 | 1.4070 | 1.4141 | 2.2422 |
| Mark 21 | 1.5224 | 1.5301 | 2.3323 |

A.2 Exact Results from Experiment 1

A.3 Ticker Symbols of Stocks Used in MAD 63

Day High Rho Portfolio

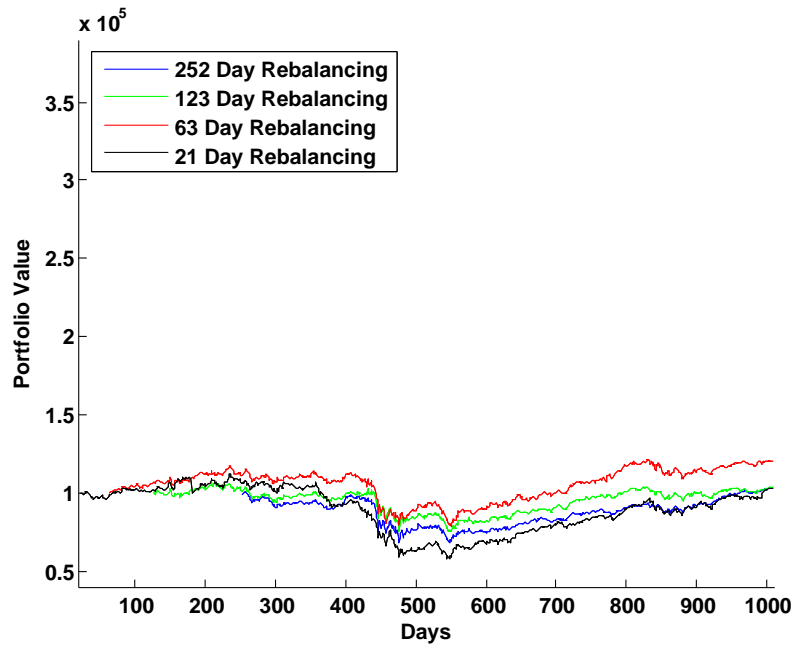
‘AAPL’ ‘ADM’ ‘AIG’ ‘AKAM’ ‘AKS’ ‘AMD’ ‘AMZN’ ‘APOL’ ‘BBT’ ‘BIG’ ‘BIIB’
‘CAH’ ‘CF’ ‘CHK’ ‘CLF’ ‘COG’ ‘CPWR’ ‘CTXS’ ‘EOG’ ‘EXPE’ ‘F’ ‘FDO’ ‘FFIV’
‘FHN’ ‘FIS’ ‘FSLR’ ‘FTI’ ‘GCI’ ‘GENZ’ ‘GNW’ ‘HBAN’ ‘HCBK’ ‘HIG’ ‘HP’ ‘HSY’
‘ISRG’ ‘JBL’ ‘JDSU’ ‘KMX’ ‘LEG’ ‘MA’ ‘MEE’ ‘NBR’ ‘NFLX’ ‘NVDA’ ‘PCLN’
‘PHM’ ‘PXD’ ‘R’ ‘RSH’ ‘S’ ‘SLM’ ‘SRCL’ ‘THC’ ‘TSN’ ‘TSO’ ‘WDC’ ‘WFMI’ ‘WHR’
‘WYNN’ ‘ZION’

Appendix B

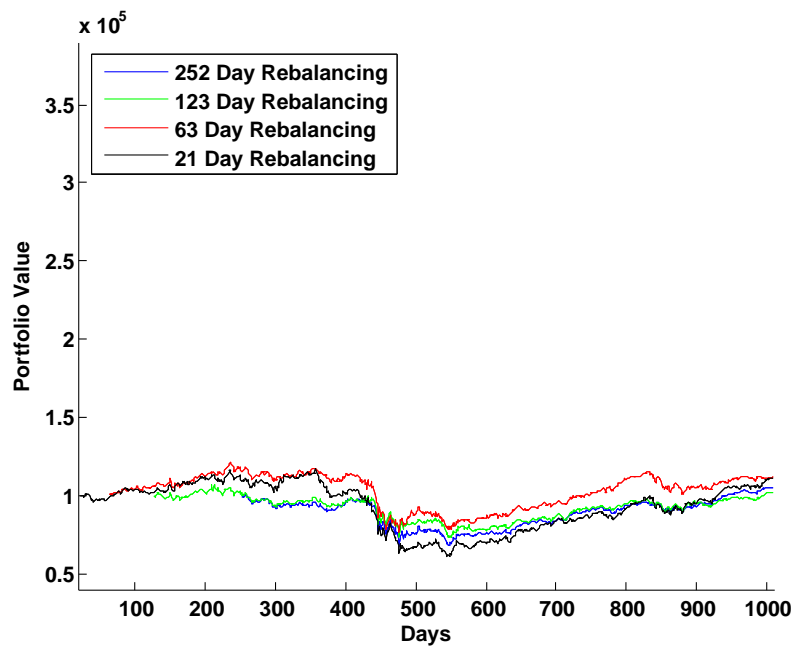
Plots

B.1 63 Day Rebalancing Horizon Over Different Rho Values

When holding ρ constant and varying the rebalancing strategy, it is apparent that the 63 day rebalancing horizon performs the best for both the Markowitz and the MAD model with the one exception being the MAD model with a ρ_{Med} which achieves its best performance when using the 21 day rebalancing horizon.

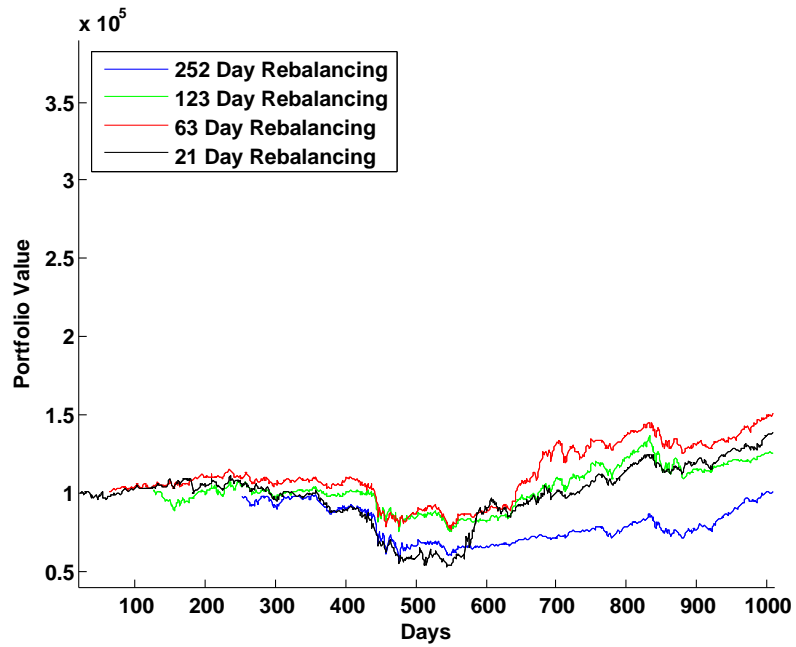


(a) Markowitz Model

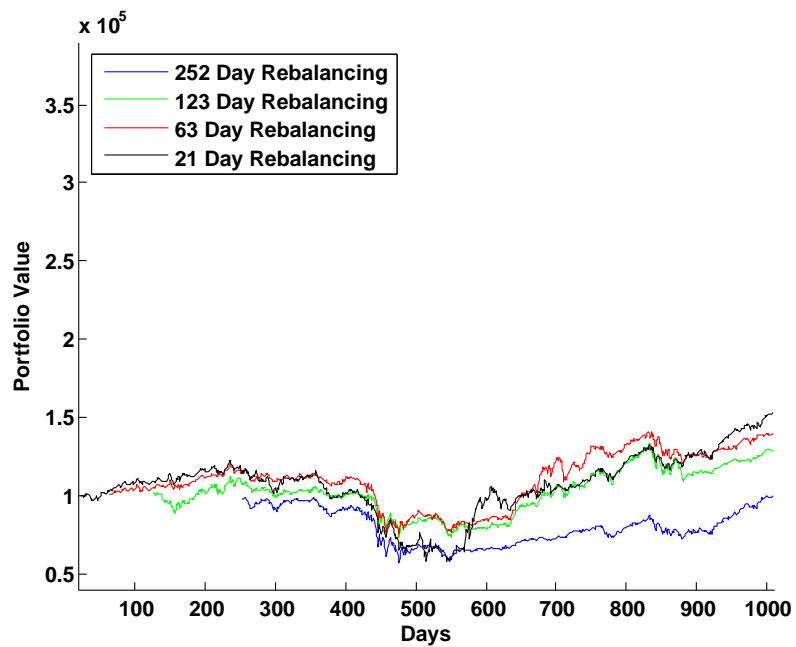


(b) MAD Model

Figure B.1 Portfolio Net Worths using ρ_{Low}

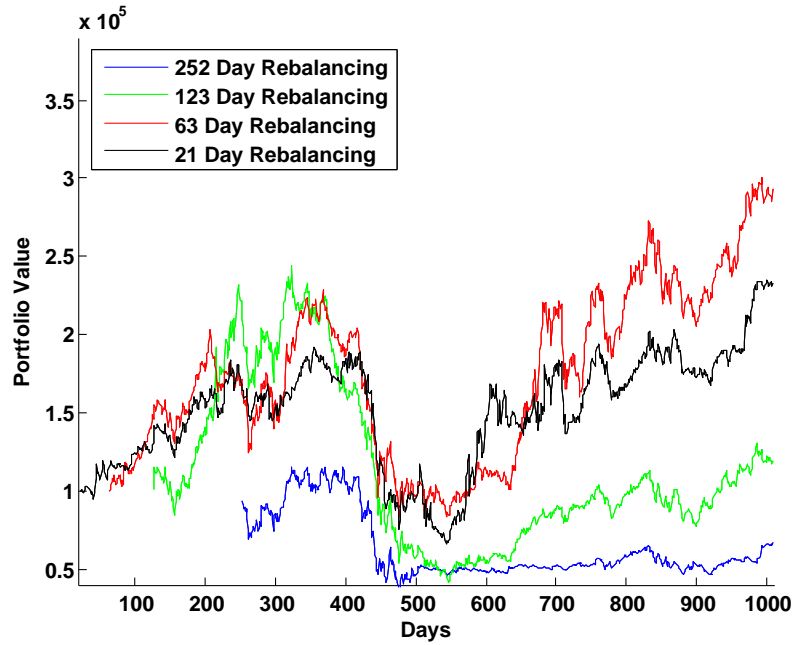


(a) Markowitz Model

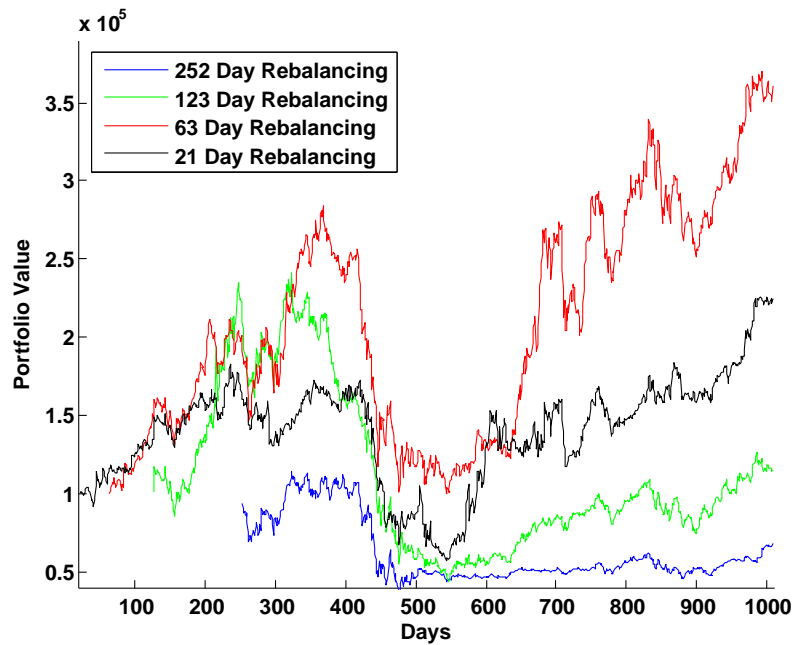


(b) MAD Model

Figure B.2 Portfolio Net Worths using ρ_{Med}



(a) Markowitz Model



(b) MAD Model

Figure B.3 Portfolio Net Worths using ρ_{High}

Appendix C

Plots

C.1 Separate plots for MAD-Simplex, MAD-IP, and Markowitz models showing trends

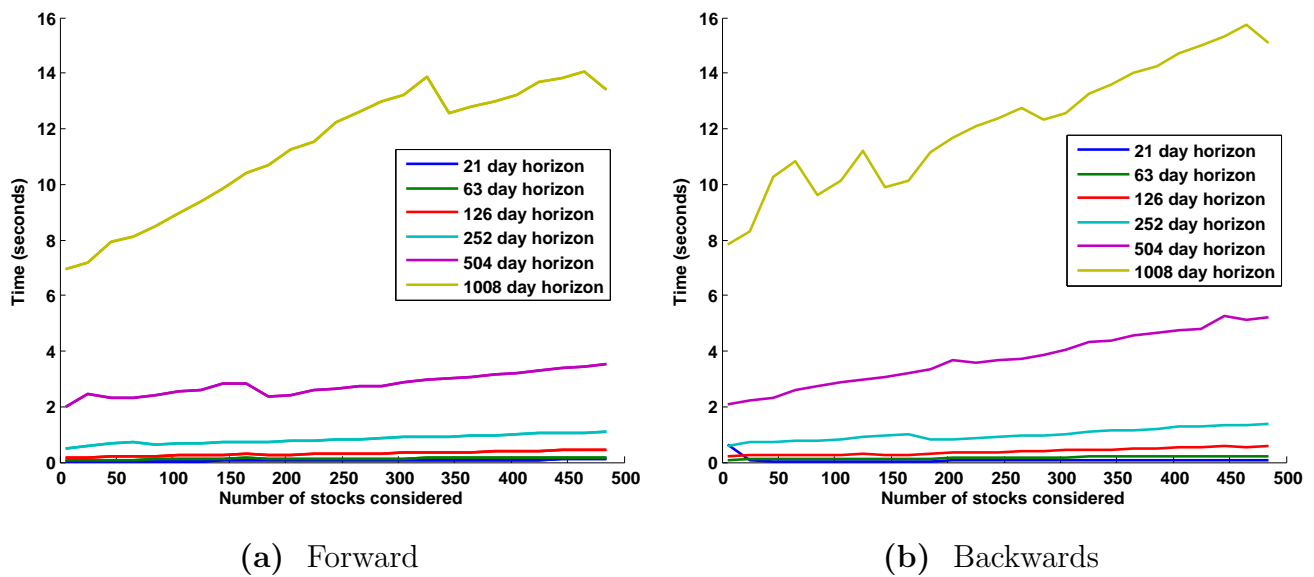


Figure C.1 Time Complexity for MAD-Simplex over multiple horizons

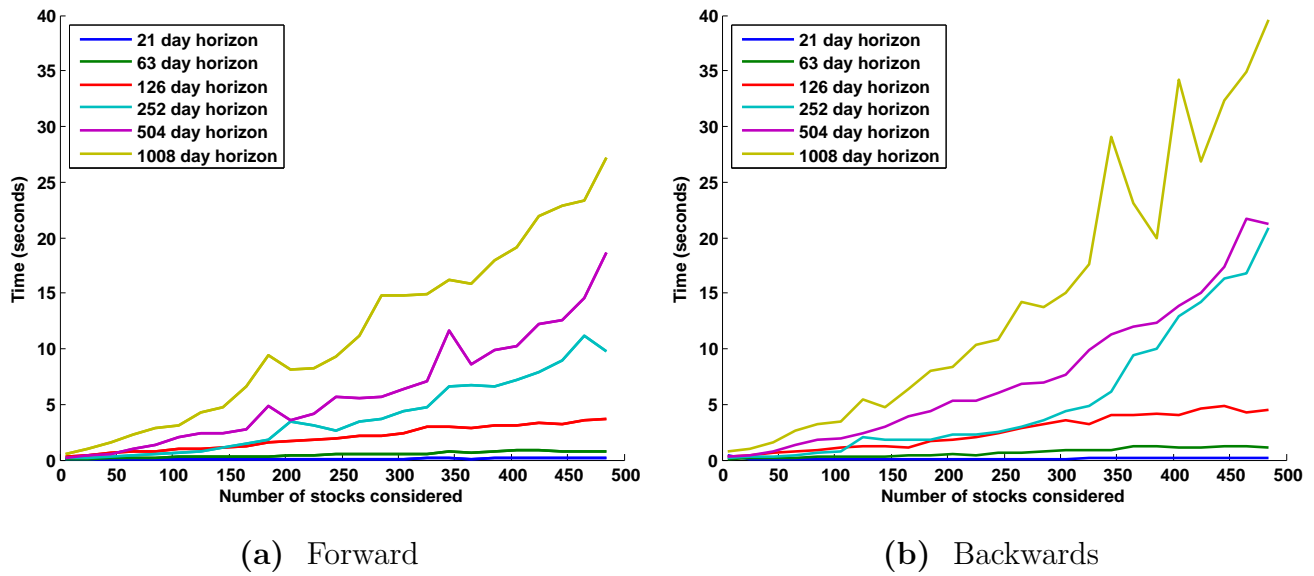


Figure C.2 Time Complexity for MAD-IP over multiple horizons

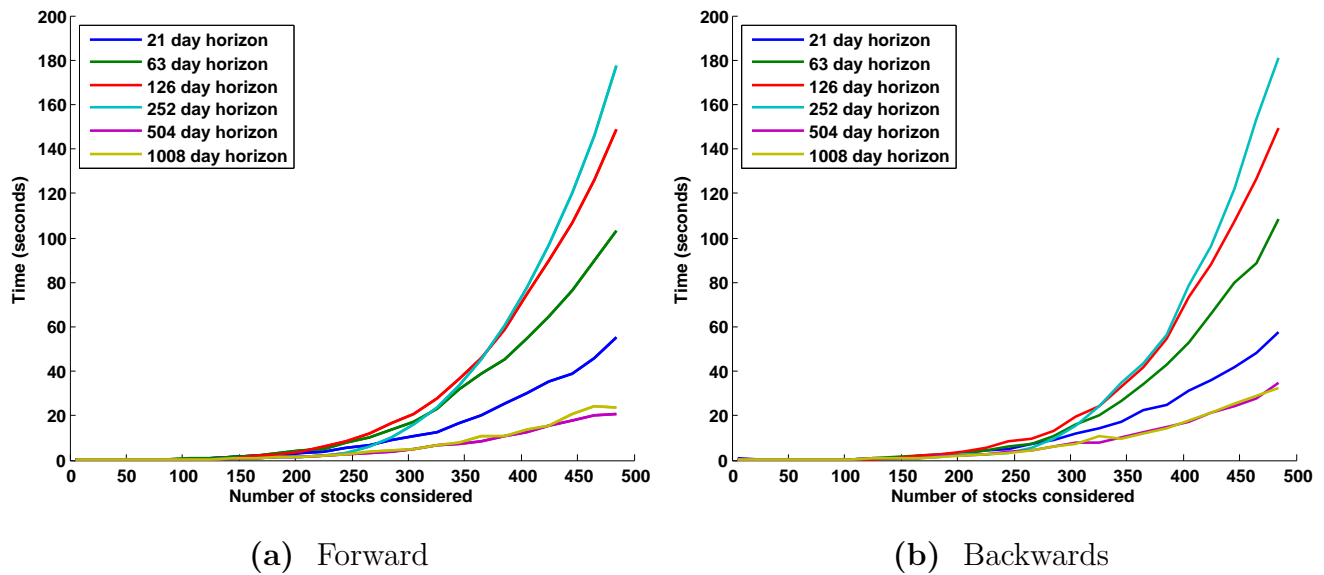


Figure C.3 Time Complexity for Markowitz over multiple horizons

Appendix D

Plots

D.1 Time complexity plots using random returns

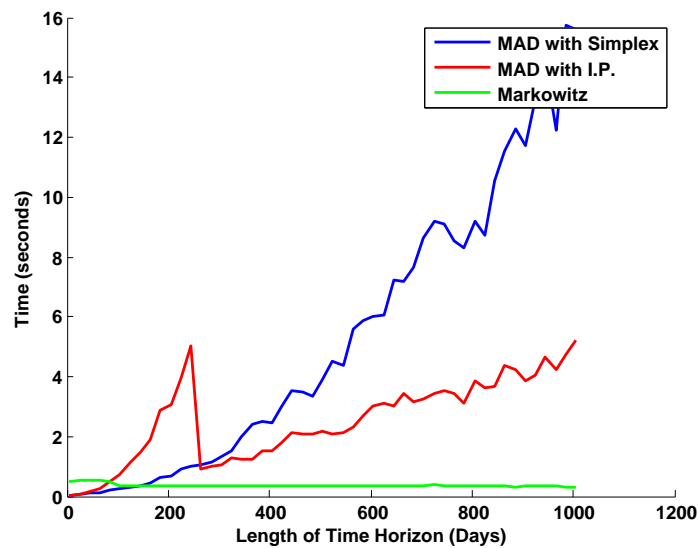


Figure D.1 Time Complexity using 100 artificial stocks with random normal returns

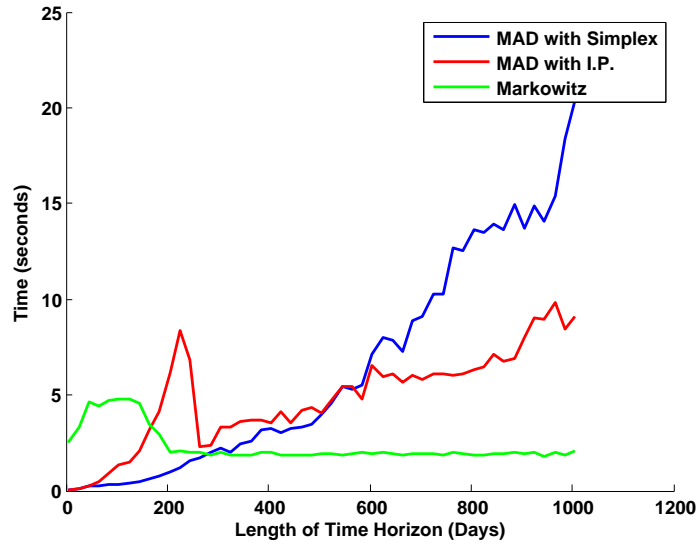


Figure D.2 Time Complexity using 200 artificial stocks with random normal returns

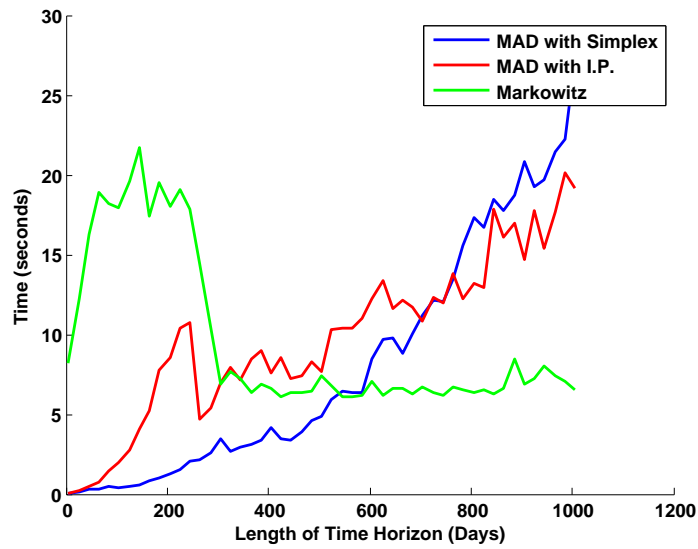


Figure D.3 Time Complexity using 300 artificial stocks with random normal returns

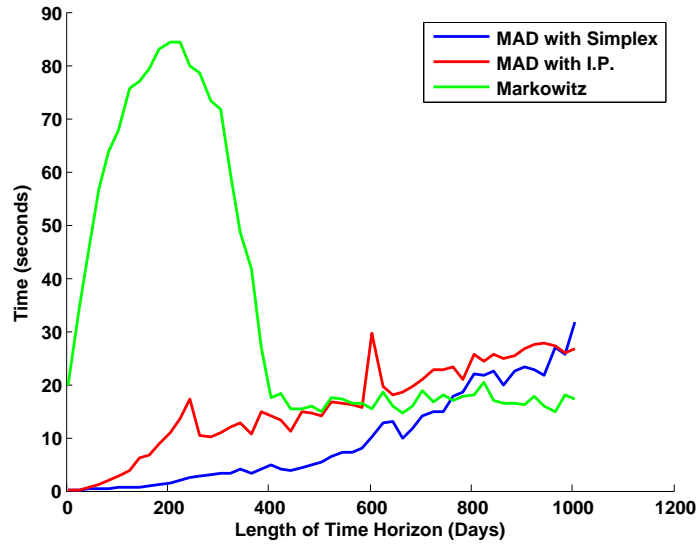


Figure D.4 Time Complexity using 400 artificial stocks with random normal returns

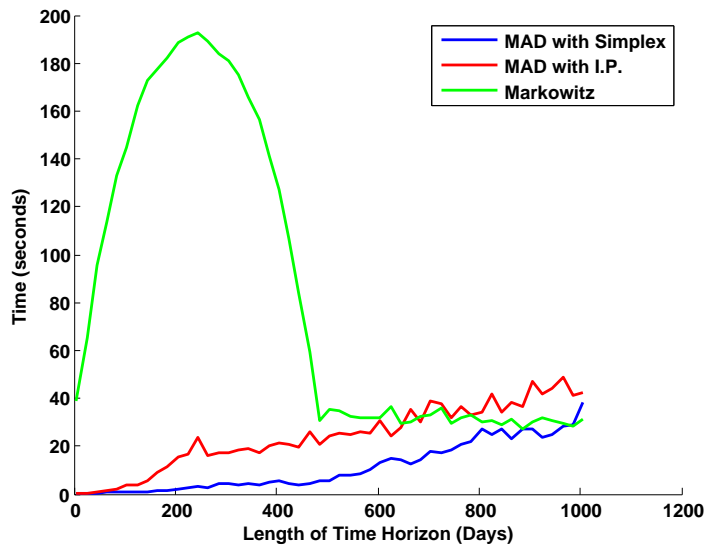


Figure D.5 Time Complexity using 485 artificial stocks with random normal returns

Bibliography

- [1] Barmish R., *On Trading Equities: A Robust Control Paradigm*. In Proc. of the 17th World Congress of The International Federation of Automatic Control, pages 16211626, July 2008.
- [2] Beechy M, Gruen D, Vickrey J., *The Efficient Market Hypothesis: A Survey*, Reserve Bank of Australia, 2000.
- [3] Bertsimas D., Lauprete G., Samarov A., *Shortfall as a risk measure: properties, optimization and applications*. Journal of Economic Dynamics and Control 28 (2004) 1353-1381.
- [4] Boyd, S., and Vandenberghe L., *Convex Optimization*, Cambridge Univ., 2005.
- [5] Cassidy J., *After the Blowup*, The New Yorker, 11 January 2010.
- [6] Danzig, G. B., *Linear Programming and Extensions*, Princeton University Press, Princeton, NJ, 1963.
- [7] Duffie D., and J. Pan. *An Overview of Value-at-Risk*. Journal of Derivatives. 4, 7-49, 1997.
- [8] Fazel, M. M. Logo, S. Boyd *Portfolio optimization with linear and fixed transaction costs*. Annals of Operations Research, special issue on Financial Optimization, 152(1):376-394, July 2007.

- [9] Feinstein C., and M. Thapa. *Notes: a reformulation of a mean-absolute deviation portfolio optimization model*. Management Sciences 39 (12) (1993) 1552-1553.
- [10] Francis Nicholson. *Price-Earnings Ratios in Relation to Investment Results*, Financial Analysts Journal. Jan/Feb 1968:105-109.
- [11] Harlow W. *Asset Allocation in a Downside-Risk framework*. Financial Analysts Journal. Vol. 47 No. 5 (Sept. - Oct., 1991) pp. 28-40
- [12] Hoe, L., Hafizah, J., Zaidi, I., *An empirical comparison of different risk measures in portfolio optimization*, Business and Economic Horizons, Vol 1, Issue 1, pp. 39-45, April 2010.
- [13] Kariya, T. *Distribution of Stock Prices in the Stock Market of Japan*. Tokyo Keizai Publishing Co., Tokyo, 1989.
- [14] Konno H., and H. Yamazaki *Mean-Absolute deviation portfolio optimization model and its applications to Tokyo stock market*. Management Science 37(5), 1991.
- [15] Krokmal, P., Palmquist, J., Uryasev, S., *Portfolio optimization with conditional Value-at-Risk criterion*, Journal of Risk 4 (2). Research Report 99-14, 2002.
- [16] Landau P., *Has 'guiding model' for global markets gone haywire?* Jerusalem Post, 11 June 2009.
- [17] Lauricella, T., and McKay, P. *Dow Takes a Harrowing 1,010.14-Point Trip*, Online Wall Street Journal, May 7, 2010.
- [18] Mansini, R., Ogryczak, W., Speranza, M. *LP solvable models for portfolio optimization: a classification and computational comparison*, IMA Journal of Management Mathematics (2003) 14.187-220

-
- [19] Markowitz, H. *Portfolio Selection*. Journal of Finance, vol. 7, pp. 77-91, 1952.
- [20] Modigliani F., Modigliani L., *Risk-Adjusted Performance*, Journal of Portfolio Management, pp. 45-54 Winter 1997.
- [21] Peng, H. *A New Optimal Portfolio Selection Strategy Based on a Quadratic Form Mean-Variance Model with Transaction Costs*. Optimal Control Applications and Methods, 2010.
- [22] Primbs, J.A., *Portfolio Optimization Applications of Stochastic Receding Horizon Control*, American Control Conference, 2007. ACC '07 , vol., no., pp.1811-1816, 9-13 July 2007.
- [23] Sanjoy B. *Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A test of the Efficient Markets Hypothesis*, Journal of Finance 32:663-682 (1977).
- [24] Sharpe, W. F., *A Linear Programming Approximation for the General Portfolio Analysis Problem*, The Journal of Financial and Quantitative Analysis, Vol. 6, No. 5, pp. 1263-1275, 1971.
- [25] Sharpe, W. F., *A Linear Programming Algorithm for Mutual Fund Portfolio Selection*, Management Science, Jan 1963.
- [26] Sharpe, W. F., *Mutual Fund Performance*, Journal of Business, January 1966, pp. 119-138.
- [27] Stone B. *A Linear Programming Formulation of the General Portfolio Selection Problem*, Journal of Financial and Quantitative Analysis, Sept 1973.
- [28] Vanderbei R. *Linear Programming: Foundations and Extensions*, Springer, New York, 2008.

-
- [29] Xiang S., Fu W., *Predicting the Stock Market using Multiple Models*, Proc. of International Conference on Control, Automation, Robotics and Vision, pp. 1-6 5-8 Dec 2006.