Review of Modeling Methods in Portfolio Analysis

Jason McCoy
Dennis Pfeiffer
Matthew Ward

Northwestern University School of Professional Studies
Master of Science in Data Science 460, Prof: Dr. Nathaniel Bastian

Abstract

This paper will be a review of the classical methods in portfolio optimization. Portfolio optimization allows us to develop models for risk and return for a given set of assets. This type of model is born of the financial industry. It is classically used to determine optimal investments in a stock, bond, treasury or other asset to maximize a rate of return, while minimizing risk. We will focus on a review of the modeling techniques and some of the potential issues and improvements seen in each model. This paper will focus on the application of these methods to a stock portfolio, but the primary goal is to explore the area of modeling and understand the development and use of this type of model.

Keywords: Portfolio, Optimization, Stocks, Markowitz, Sharpe, Investment

Introduction

Portfolio optimization is focused on the optimal utilization of assets in a fixed portfolio, while managing risk. Harry Markowitz, is largely considered to be the grandfather of this area of study. His 1952 essay in The Financial Journal set out the basic theory for this analysis (Markowitz, 1952). This analysis is based on the idea that a portfolio can be optimized through a review of the historical returns or losses and the covariance of those returns or losses to other potential stocks in the portfolio. This means that we are not only able to maximize the potential returns on a given set of assets, but that we can hedge our bets, by picking a set of assets that will not track together in the event of a market change.

The Analysis by Markowitz was groundbreaking and resulted in him receiving the Nobel Prize in 1990 (Investopedia, 2018). Markowitz's portfolio analysis methodology did however have some gaps and major assumptions. It assumed that risk was evenly and normally distributed. It did not account for taxes and fees, and did not allow for the investors risk preference or time horizon to be accounted for. These issues were taken on by other analysts.

William Shape's work addressed some of the issues with time horizon. By allowing a reference portfolio to be used Sharpe, was able to show the potential for an opportunity cost to the investment versus other potential investments (Sharpe, 1964). This method allows us to better isolate the risk associated with the portfolio itself rather than the overall market performance. Sharpe would also later provide work with the Quadratic Utility Function. This method allows the analyst to incorporate the investors' level of risk acceptance into the analysis (Sharpe, 2007).

Another issue with Markowitz model was the use of an evenly distributed normal distribution for risk. This issue was taken on by several researchers. In this paper, we will discuss the single sides risk strategies developed by Rom and Ferguson (1993). Rom and Ferguson, provided methods to look at single sides or down side risk. These methods may be better at estimating the real risk associated with a portfolio.

In this paper we will review these methods and explore their application on a set of randomly selected DOW stocks. Our goal is to develop a better understanding of these methods and the potential of new research in this area of study.

Literature Review

The volume of scholarly articles and industry publications on portfolio analysis is immense. The field of portfolio optimization is largely credited to Markowitz (1952) who laid out methods for portfolio optimization in his essay Portfolio Selection. Markowitz's work was largely focused on the optimization of a portfolio to minimize risk based on the use of covariance to hedge risk among various assets in a portfolio.

The majority of the work falling into this are, post Markowitz, are focused on the resolution of some of the gaps presented by the original model. Markowitz's model does not account for fees, taxes or for non-evenly distributed risk. Much of the work seeks to add the ability to include taxes and fees into the model. This is very desirable, given the significant effect these factors can have on the outcomes of the model selection. An example of this kind of work is Pogue's (1970) work for The Financial Journal, An Extension of the Markowitz Portfolio Selection Model to Include Variable Transactions' Costs, Short Sales, Leverage Policies and Taxes.

Another example of a missing element in Markowitz is the time element, or investment horizon. This is important as it relates to the liquidity need of the investor, but also in acknowledging that with and return profiles can change significantly given a time profile. It is also important as tying assets into a portfolio can carry some opportunity cost, for the lost opportunity to invest those assets into some other vehicle. Sharpe's (1964) work Capital Asset Prices, provided new models, which incorporated time and a comparison to stable low risk investments. Other work in this area has focused on better taking into account individual preferences for risk. These methods such as quadratic utility allow the result to be tailored more to the individual's level or risk acceptance. This dovetails well with Sharpe's other work as it allows investors with a shorter horizon, closer to retirement say, to alter their risk acceptance (Sharpe 2007).

More recent work in the area of portfolio optimization has focused on the problems with the normally distributed risk assumptions. Rom and Ferguson's (1993) article Post Modern Portfolio Optimization comes of age puts forth the idea of single sides risk calculations based on the downside risk only.

Current and future research in portfolio optimization is likely to focus on big data and machine learning techniques. The use of neural networks and sentiment analysis are looking to use the large amounts of data available to be more predictive of the potential risk profile of a given portfolio. Where Markowitz and previous researchers, depended soles on historical data, these methods will attempt to predict future behavior or at least potential behaviors. These methods have the potential to be transformative, but getting these predictions correct in a very complex world is likely to take some time. One example is the use of volatility indexes and sentiment analysis to set overall risk level. A higher risk would indicate the desirability of a lower risk profile overall for the optimal portfolio (B. Dumas, A. Kurshev and R. Uppal, 2009).

Our work in this paper is a review of the major models used in the area of portfolio optimization. While this type of model can be applied to many problems we are focused on a classic financial example. Our scope is limited to a few stocks and does not include other potential investment options such as bonds or treasuries as outlined in Shape's methods. We do develop models based on the single sided risk methods pioneered by Rom and Ferguson. We do not develop models in this work which consider fees and taxes. We also do not look at any of the more cutting edge work using sentiment analysis or machine learning.

Methodology

Our Objective in this paper is to develop various models from a fixed data set. These models will cover a number of analysis, but focus on reducing potential risk in a portfolio. The purpose of these analysis is to explore the various classical methods for portfolio analysis. Portfolio optimization is considered a class of non-linear programming (NLP) as the optimal solution is a point on curve known as the efficient frontier. The efficient frontier represents the combination of securities that produce the maximum expected return for a given level of risk and is the basis modern portfolio theory.

For our models we will use index i to denote a particular stock, where $i = 1, 2 \ldots n$ corresponding to the stocks in our data set. Our decision variables are w_i, $i = 1, 2 \ldots n$ denoting the percentage of total money (or weights assigned in the portfolio) invested in stock i. Assuming a total of $1000 to invest, if a stock has a weight of 15%, it would mean that $150 is invested in that particular stock.

Let us make the following assumptions:

(i) We can trade any number of shares.

(ii) No short-selling is allowed.

These assumptions restrict the variables w_i to take on non-negative real values.

$$w_i \geq 0 \text{ for } i = 1,2 \ldots n$$

The return of a stock i is uncertain, and therefore, so is the return on our investment. Let us denote by r_i the random variable corresponding to the daily return for stock i. Then the return in stock i is $r_i w_i$, and the total return on our investment is:

$$\sum_{i=1}^{n} w_i r_i$$

Note that we have assumed that:

(iii) There are no transaction costs.

Now we need to quantify the notion of "risk" in our investment. Markowitz, in his Nobel prize winning work, showed that a rational investor's notion of minimizing risk can be closely approximated by minimizing the variance of the return of the investment portfolio. This variance of the portfolio is given by:

$$\sigma_p^2 = \sum_{j=1}^{n} \sum_{i=1}^{n} w_i w_j s_i s_j \rho_{ij}$$

Which can be re-written as:

$$\sigma_p^2 = \sum_{j=1}^{n} \sum_{i=1}^{n} w_i w_j \sigma_{ij}$$

where σ_{ij} is the correlation of the return of stock i with stock j, s_i is the standard deviation of returns of stock i, *and* σ_{ij} is the covariance of the return of stock i with stock j.

There will be some constraints in the optimization model for our problem:

1. Weights sum constraint, i.e. sum of weights must be 1.

$$\sum_{i=1}^{n} w_i = 1$$

2. Box constraint, i.e. i.e. weights assigned must be a minimum of w_{min}% and a maximum of w_{max}%. In our optimization model, we take w_{min} = 5 and w_{max} = 40.

We will also assume that the distribution of the return is normally distributed, although we do not know the exact distribution of the random return. We can obtain some statistical inference regarding it through analysis of historical data. In practice, the data collection and analysis process would involve very sophisticated statistical and time series models to obtain reliable estimates of the means and co-variances.

The data used in this analysis was collected from the NASDAQ website (NASDAQ, 2018). We selected ten DOW stocks and collected the daily returns over a five year period (December 2013 to November 2018). We calculated the daily returns of each of the ten DOW stocks based on the close price. We then calculated the covariance between each stock using software and Excel formula. The basic decision variable in this optimization is the percentage of the total portfolio allocated to each stock.

The first and most basic model we developed is a basic return on investment model. This model looks at the rate of return on the stocks over the five year period and selects weights assigned to stocks to maximize the expected portfolio return subject to weights sum and box constraints.

Maximize

$$\sum_{i=1}^{10} w_i r_i$$

Subject to

$$\sum_{i=1}^{10} w_i = 1$$

$$0.05 \leq \ w_i \leq 0.4$$

The second model we executed is a simple risk minimization problem. This analysis has the objective to minimize portfolio risk as measured by portfolio variance.

Minimize

$$\sum_{j=1}^{10} \sum_{i=1}^{10} w_i w_j s_i s_j \rho_{ij}$$

Subject to

$$\sum_{i=1}^{10} w_i = 1$$

$$0.05 \leq w_i \leq 0.4$$

As in standard Markowitz analysis, we have the return to be at least the desired return for the portfolio, but focused more on diversification by applying the box constraint. This analysis assumes that there are no taxes and fees and that the variance is evenly and normally distributed. This means that this type of analysis has some limitations. It does not account for the investors risk comfort, it simply defaults to minimum. This analysis also does not account for the time horizon of the investment. In our case, we used five years of data, but we may have expected to hold this portfolio for ten years. This model does not make any predictions about future performance, but should help to minimize risk by selecting stocks, which seem to offset each other in the event of a downturn.

In order to develop a model which takes into account the risk tolerance or appetite of the investor, we execute a third model based on maximizing Quadratic Utility. This model allows us to create different potential levels of risk. This is done by setting different values of risk aversion. Higher value of risk aversion mean that the investor is less risk loving. In this way we can model different potential portfolios as bounded by the acceptance of risk we have set. This type of analysis still contains, many of the drawbacks from the basic Markowitz analysis, but it does help us to understand the various potential preferences of the investor.

Maximize the quadratic utility function

$$U = E[R] - \frac{1}{2} A\sigma^2$$

Subject to

$$\sum_{i=1}^{10} w_i = 1$$

$$0.05 \leq w_i \leq 0.4$$

We took three different levels of risk aversion: 0.25, 10 and 100 which helped us to analyze how the optimal portfolio composition differs as the risk appetite of the investors change.

The fourth analysis we explored is the Sharpe Ratio. The Sharpe ratio is a ratio of the mean return of the portfolio to the standard deviation of the portfolio. This Ratio allows us to express the excess returns as compared to the risk free returns. We can then optimize the portfolio by maximizing the Sharpe ratio. This analysis has many of the same drawbacks as Markowitz analysis. It does help to provide some insight into the potential returns as compared to a safe investment, which should allow us to make more informed decisions on our portfolio. The advantage of using Sharpe ratio is that it takes into consideration both portfolio return and portfolio risk. It tries to maximize portfolio return and minimize portfolio variance or risk.

Maximize

$$\frac{\sum_{i=1}^{10} w_i r_i}{\sum_{j=1}^{10} \sum_{i=1}^{10} w_i w_j s_i s_j \rho_{ij}}$$

Subject to

$$\sum_{i=1}^{10} w_i = 1$$

$$0.05 \leq \; w_i \leq 0.4$$

The fifth and the final model we applied is an ETL (Expected Tail Losses) model. This can be thought of as a downside risk model. This model attempts to resolve the deficiency of evenly distributed risk in Markowitz. Markowitz model indeed revolutionized risk management at its time, however it has some drawbacks which we have already discussed. Two important drawbacks because it measures the risk in terms of variance of the portfolio are:

1. Variance is generally a useful risk measure when we have losses that are normally or symmetrically distributed. Since variance is measured in either direction, tail losses arising from skewed loss distributions are not taken into consideration.
2. Variance is also not a coherent risk measure because it is not monotone.

ETL model skews the risk towards the downside allowing the investor to better understand the potential losses. This type of analysis might be used to define the potential loss of a portfolio in a down year or to develop a very conservative model. The aim is to minimize ETL. In simple terms, ETL at 5% level would mean the expected return on the portfolio in the *worst* 5% of cases. It estimates the risk of an investment in a conservative way, focusing on the less profitable outcomes. We keep the same constraints, the weight sum and box constraints.

The developed models allow us to better understand the management of risk within our proposed 10 DOW stocks. They do not allow us to predict overall market performance and none of them allow us to compute taxes or fees. We have shown that we can build on the basic portfolio optimization models to explore the potential goals of our portfolio and to model the potential for a downside risk.

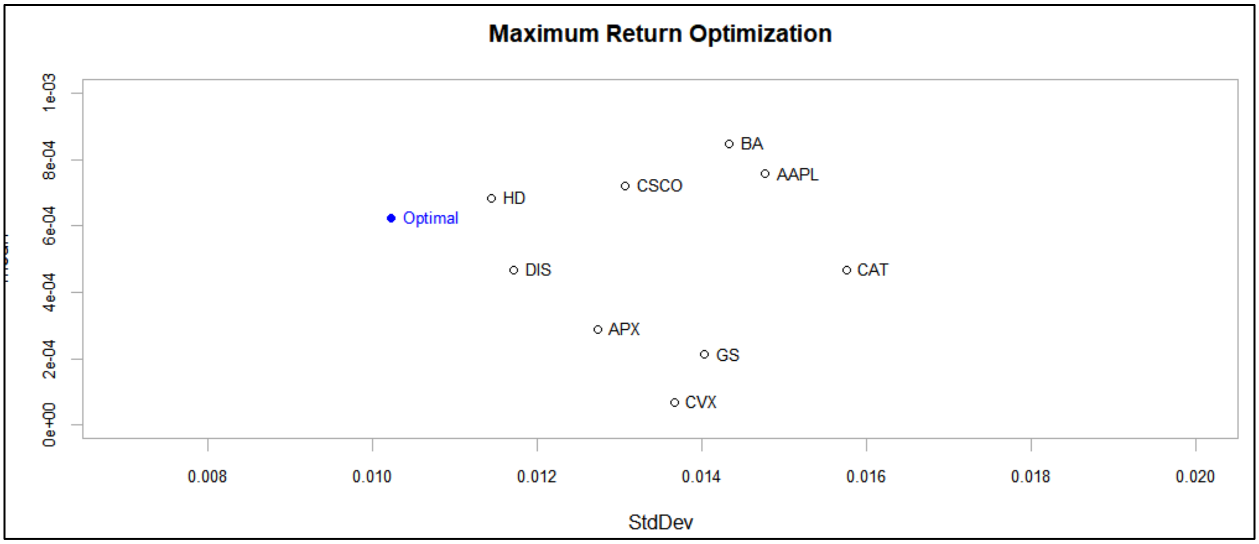
Computational Experiment and Results

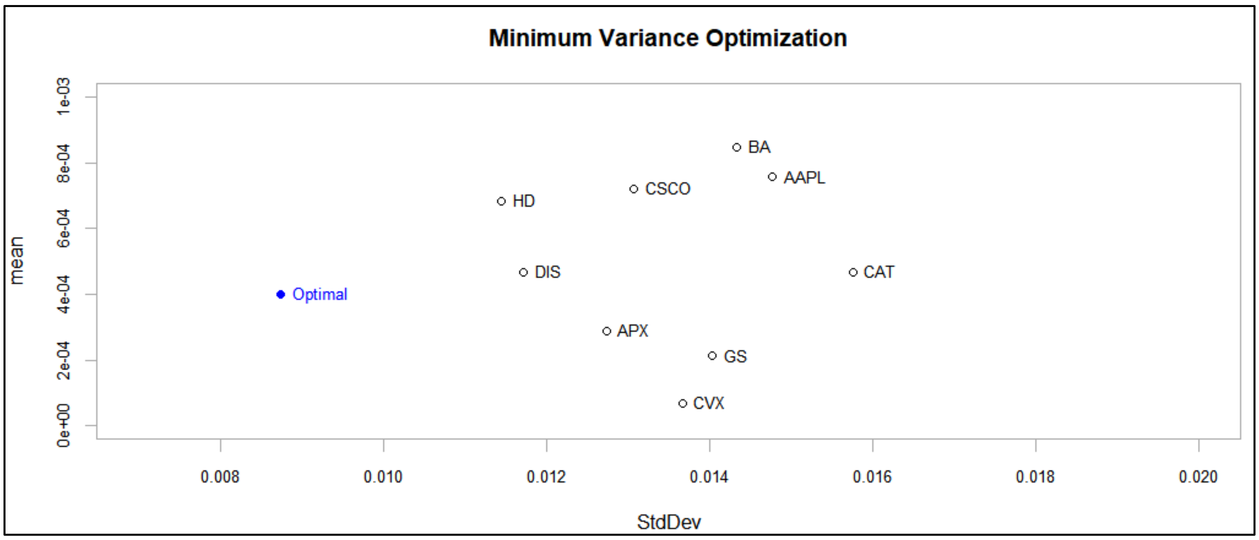
In this analysis we have developed five models (described in the Methodology section) to give is better insights into portfolio analysis. We executed the models primarily in R, however excel solver solutions were also used as checks. These models differ, likely due to the basic assumptions of the models. For the purpose of this analysis we will refer to the models as developed in R. Table 1 shows the weights assigned to different stocks in the optimized portfolio under each of the five models that we executed. The relevant optimized values are listed in the last three rows of Table 1.

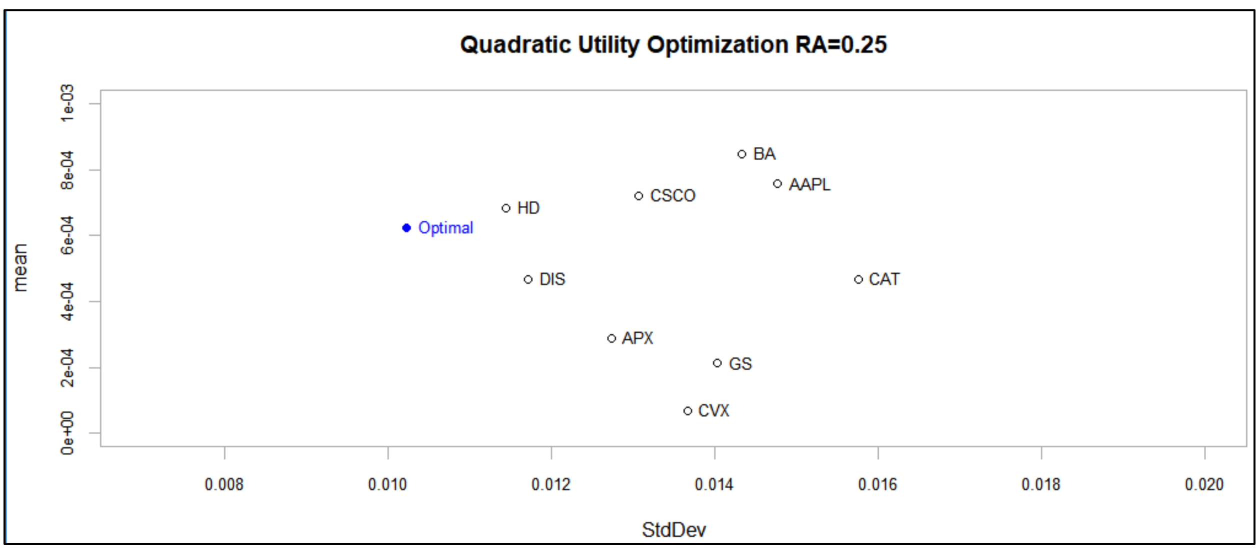
	Model 1	Model 2	Model 3a	Model 3b	Model 3c	Model 4	Model 5
AAPL	20%	6.89%	20%	11.86%	8.14%	13.80%	5%
APX	5%	10.66%	5%	5%	10.44%	5%	5.55%
BA	40%	5%	40%	10.93%	5%	20.76%	5%
CAT	5%	5%	5%	5%	5%	5%	5%
CSCO	5%	5%	5%	10.62%	5%	10.72%	5%
CVX	5%	9.05%	5%	5%	8.22%	5%	13.10%
DIS	5%	18.88%	5%	12.68%	19.52%	5%	15.70%
GS	5%	5%	5%	5%	5%	5%	5%
HD	5%	20.79%	5%	29.92%	22.89%	24.73%	20.25%
IBM	5%	13.73%	5%	5%	10.79%	5%	20.40%
Portfolio Return	0.000625		0.000625	0.000557	0.000433	0.000591	
Portfolio Stdev		0.008735	0.01022	0.008987	0.008745	0.009279	
ETL							0.02175

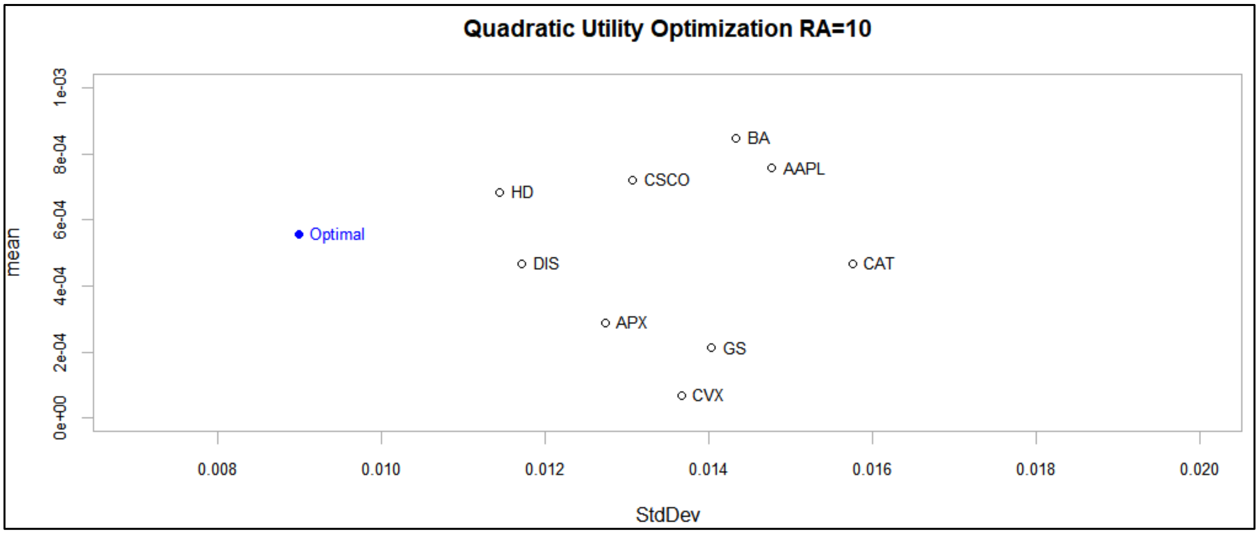
Table 1: Portfolio Optimization Results. Model 1: Maximize expected portfolio return. Model 2: Minimize portfolio variance. Model 3a: Maximize quadratic utility with risk aversion 0.25. Model 3b: Maximize quadratic utility with risk aversion 10. Model 3c: Maximize quadratic utility with risk aversion 100. Model 4: Maximize Sharpe Ratio. Model 5: Minimize ETL.

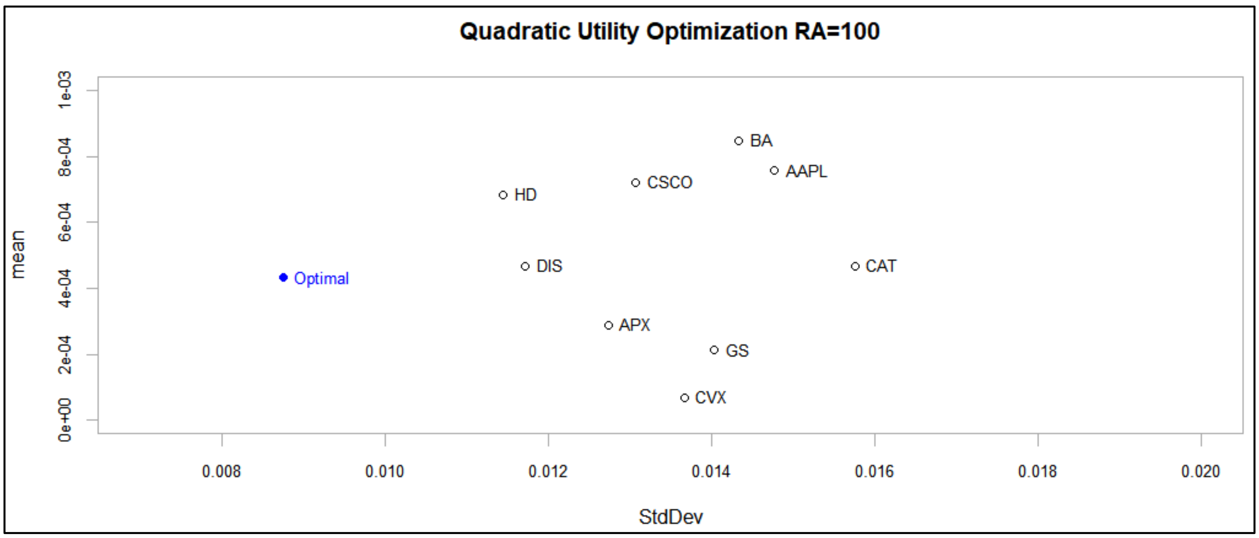
The following diagrams shows the optimized portfolio for each of the models executed. The diagrams show the position of the optimized portfolio along with the position of all the 10 DOW stocks that were used in the analysis.

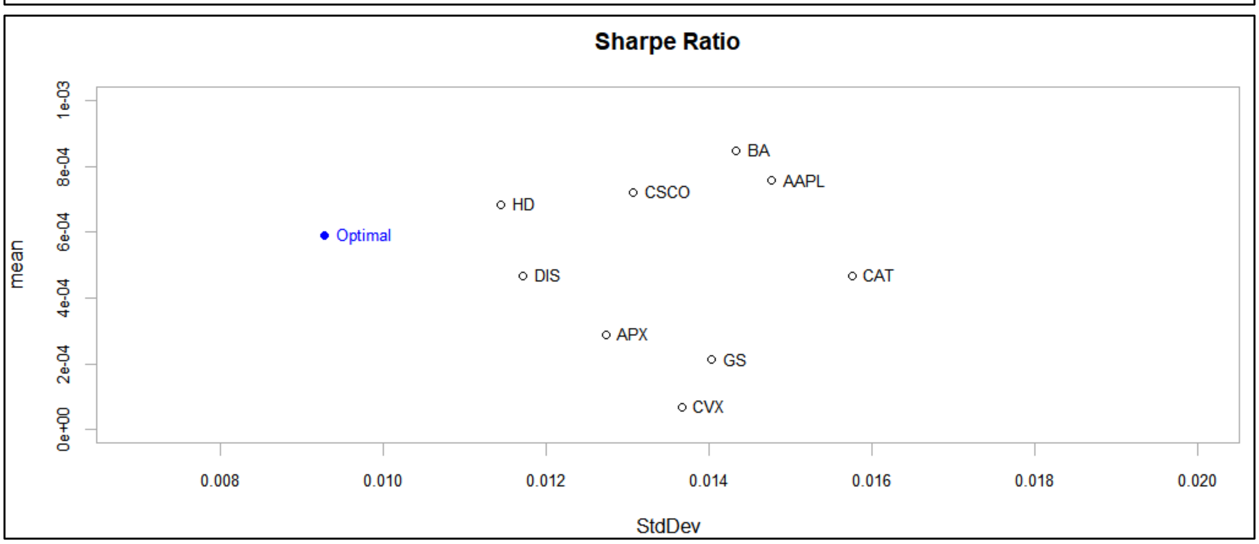

Maximum Return Optimization
1e-03
8e-04
6e-04
4e-04
2e-04
0e+00
BA
AAPL
CSCO
HD
Optimal
DIS
CAT
APX
GS
CVX
0.008
0.010
0.012
0.014
0.016
0.018
0.020
StdDev

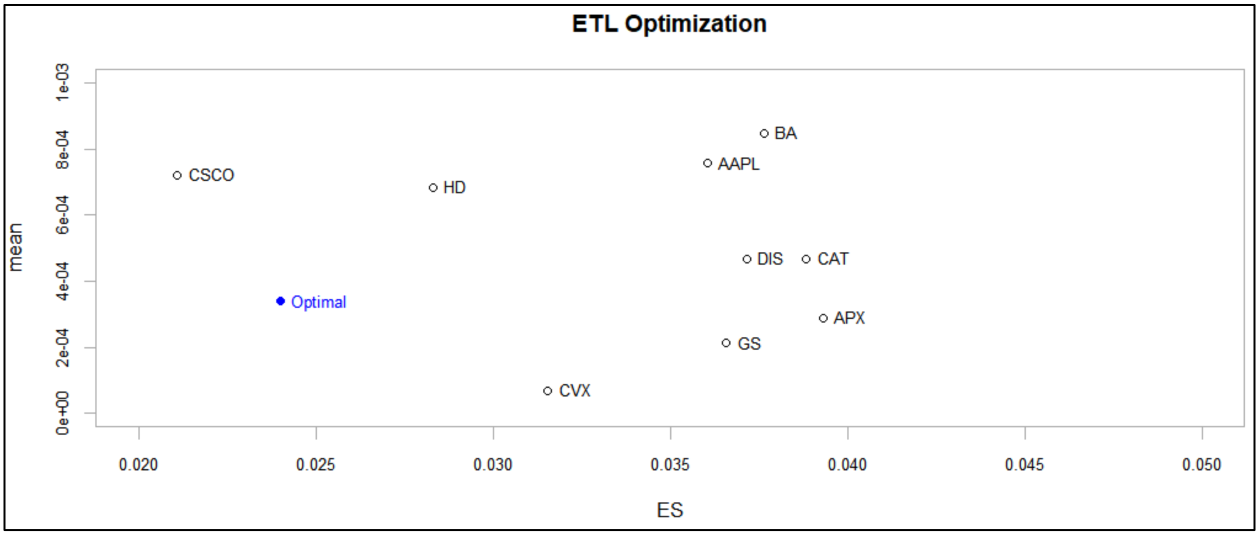

Minimum Variance Optimization
mean
1e-03
8e-04
6e-04
4e-04
2e-04
0e+00
BA
AAPL
CSCO
HD
DIS
CAT
Optimal
APX
GS
CVX
0.008
0.010
0.012
0.014
0.016
0.018
0.020
StdDev


Quadratic Utility Optimization RA=0.25
mean
1e-03
8e-04
6e-04
4e-04
2e-04
0e+00
BA
AAPL
CSCO
HD
Optimal
DIS
CAT
APX
GS
CVX
0.008
0.010
0.012
0.014
0.016
0.018
0.020
StdDev


Quadratic Utility Optimization RA=10
mean
StdDev
Optimal
HD
CSCO
BA
AAPL
DIS
CAT
APX
GS
CVX
0e+00
2e-04
4e-04
6e-04
8e-04
1e-03
0.008
0.010
0.012
0.014
0.016
0.018
0.020


Quadratic Utility Optimization RA=100
mean
StdDev
Optimal
HD
CSCO
BA
AAPL
DIS
CAT
APX
GS
CVX
0e+00
2e-04
4e-04
6e-04
8e-04
1e-03
0.008
0.010
0.012
0.014
0.016
0.018
0.020


Sharpe Ratio
mean
StdDev
Optimal
HD
CSCO
BA
AAPL
DIS
CAT
APX
GS
CVX
0e+00
2e-04
4e-04
6e-04
8e-04
1e-03
0.008
0.010
0.012
0.014
0.016
0.018
0.020



Discussion and Conclusions

Our analysis highlights the many options available to optimize a portfolio. If we look at the optimized results of Models 1, 2, 3a, 3b, 3c and 4, we can easily conclude that no model is preferred over the other. A model would be strictly preferred over another model if it gives higher portfolio return as well as lower portfolio variance (risk) than other model. We find that if any model gives higher portfolio return than the other, then it also has higher risk or portfolio variance. The choice of the model is subjective. It depends on the risk appetite of the investor. In Model 3, we observe that as risk aversion increases, i.e. investor becomes less risk loving, the expected return of the optimized portfolio decreases along with the a decrease in portfolio variance or risk. This aligns with our expectations. Models 1, 2, 3a, 3b, 3c and 4 have their own drawbacks because of which we introduced Model 5. We conclude that Model 5 where ETL is minimized is a conservative approach to portfolio optimization. It has the least portfolio return compared to all other models discussed.

There is no single correct model in this case. The goals of the investment and the preferences of the investor must be considered when setting up portfolio analysis. These models however, provide a solid framework for the development of various models to look at when considering portfolio investment and diversification.

This analysis, provided us with a much better understanding of the methods and underlying concepts of portfolio analysis. Our data set was far too small (only 10 DOW stocks) to provide a good basis for a real portfolio, but it was sufficient to allow us to exercise and understand the tools. In a more realistic analysis we would have included more stocks, a wider variety of stocks and other investment preferences and constraints.

The limited population and relative stability of the selected stocks meant that we had an overall very low variance on investment. This might have been more interesting if we had selected a mix of stable stocks like the DOW and potentially more dynamic or volatile stocks, such as those from the S&P.

The amount of time we used for this analysis was likely also an issue. We used a five year period, which was good, but a ten year period might have captured more volatility over an economic cycle. This time horizon, might be more important, for a longer term investment strategy. We likely should have better defined out initial investment goals to aid in defining the correct strategy. A better problem statement might have been;

Develop a five year investment strategy which utilizes DOW stocks to develop a portfolio to achieve an average ROI of .1% while achieving minimum risk.

This kind of more specific and bounded problem statement would have aided in developing the correct model for this type of optimization.

For future work we would be very interested to explore some of the work being developed in this area using machine learning techniques, sentiment analysis and evolutionary algorithms. As discussed in the literature review, emerging methods in this area are largely focused in these areas. These types of analysis have the potential to transform this type of analysis, much as Markowitz did in the 1950s.

In conclusion, I think that we can say that we were able to develop several of the classic portfolio optimization models, using real world data. Our exploration of these models allowed us to develop multiple risk based portfolios, which could be used to make better decisions about which investments should be selected from the given pool.

References

Rom, B. & Ferguson, K. (1993) Post-Modern Portfolio Theory Comes of Age, The Journal of Investing Winter 1993, 2 (4) 27-33; DOI: https://doi.org/10.3905/joi.2.4.27

Dumas, B. , Kurshev, A. & Uppal, R. (2009) Equilibrium Portfolio Strategies in the Presence of Sentiment Risk and Excess Volatility, The Journal Finance, Vol 64, No 2, pp. 579-629 Retrieved from https://doi.org/10.1111/j.1540-6261.2009.01444.x

Investopedia (2018) Harry Markowitz, Investopedia/ Investing/ Financial Analysis, Feb7, 2018, Retrieved from https://www.investopedia.com/terms/h/harrymarkowitz.asp

Markowitz, H. (1952) Portfolio Selection, The Journal of Finance, Vol. 7, No. 1. (Mar., 1952), pp. 77-91. Retrieved from http://links.jstor.org/sici?sici=0022-1082%28195203%297%3A1%3C77%3APS%3E2.0.CO%3B2-1

Nasdaq (2018) Historical Prices, Nasdaq.com, retrieved from https://www.nasdaq.com/symbol/dwdp/historical

Pogue, G. A. (1970) An Extension of the Markowitz Portfolio Selection Model to Include Variable Transactions' Costs, Short Sales, Leverage Policies and Taxes, The Financial Journal, Vol 25, No 5 , pp.1005-1007, Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1970.tb00865.x

Sharpe, W. F. (1964) Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, The Journal of Finance, Vol. 19, No. 3. (Sep., 1964), pp. 425-442. Retrieved from https://www.jstor.org/stable/2977928?seq=1#page_scan_tab_contents

Sharpe, W. F. (2007).Expected Utility Asset Allocation, Stanford, Standford.edu September, 2006, Revised June 2007, Retrieved from https://web.stanford.edu/~wfsharpe/art/euaarevised.pdf

Appendix

R Code

```
install.packages("PortfolioAnalytics")
install.packages("DEoptim")
install.packages("ROI")
install.packages("ROI.plugin.glpk")
install.packages("ROI.plugin.quadprog")
library("PortfolioAnalytics")
library("DEoptim")
library("ROI")
require(ROI.plugin.glpk)
require(ROI.plugin.quadprog)

returns.portfolio = read.csv("Data for R.csv",header=TRUE,sep=",")
rownames(returns.portfolio)=returns.portfolio[,1]
returns.portfolio = returns.portfolio[,c(2:11)]

#Mean and Covariance
meanReturns = colMeans(returns.portfolio)
covMat = cov(returns.portfolio)

#Initial Portfolio Object
R = returns.portfolio
funds = colnames(returns.portfolio)
init = portfolio.spec(assets = funds)
init = add.constraint(portfolio=init, type="leverage", min_sum=1, max_sum=1)
init = add.constraint(portfolio=init, type="box", min=0.05, max=0.4)

#Maximize Mean return with ROI
maxret <- add.objective(portfolio=init, type="return", name="mean")
opt_maxret <- optimize.portfolio(R=R, portfolio=maxret,optimize_method="ROI",trace=TRUE)
print(opt_maxret)
plot(opt_maxret, risk.col="StdDev", return.col="mean",
    main="Maximum Return Optimization", chart.assets=TRUE,
    xlim=c(0.007, 0.02), ylim=c(0,0.001))

#Minimize Variance with ROI
minvar <- add.objective(portfolio=init, type="risk", name="var")
opt_minvar <- optimize.portfolio(R=R, portfolio=minvar,
                           optimize_method="ROI", trace=TRUE)
```

```
print(opt_minvar)

plot(opt_minvar, risk.col="StdDev", return.col="mean",
    main="Minimum Variance Optimization", chart.assets=TRUE,
    xlim=c(0.007, 0.02), ylim=c(0,0.001))

#Maximize quadratic utility with ROI
qu <- add.objective(portfolio=init, type="return", name="mean")
qu <- add.objective(portfolio=qu, type="risk", name="var", risk_aversion=0.25) #risk aversion
measure can be changed for different risk appetite
opt_qu <- optimize.portfolio(R=R, portfolio=qu,
                             optimize_method="ROI",
                             trace=TRUE)
print(opt_qu)
plot(opt_qu, risk.col="StdDev", return.col="mean",
    main="Quadratic Utility Optimization", chart.assets=TRUE,
    xlim=c(0.007, 0.02), ylim=c(0, 0.001))

#Maximize Sharpe Ratio
qu <- add.objective(portfolio=init, type="return", name="mean")
qu <- add.objective(portfolio=qu, type="risk", name="var")
sharpe <- optimize.portfolio(R=R, portfolio=qu,
                             optimize_method="ROI",
                             trace=TRUE, maxSR=TRUE)
print(sharpe)
plot(sharpe, risk.col="StdDev", return.col="mean",
    main="Sharpe Ratio", chart.assets=TRUE,
    xlim=c(0.007, 0.02), ylim=c(0, 0.001))

#Minimize expected tail loss with ROI
etl <- add.objective(portfolio=init, type="risk", name="ETL")
opt_etl <- optimize.portfolio(R=R, portfolio=etl,
                              optimize_method="ROI",
                              trace=TRUE)
print(opt_etl)
plot(opt_etl, risk.col="ES", return.col="mean",
    main="ETL Optimization", chart.assets=TRUE,
    xlim=c(0.02, 0.05), ylim=c(0,0.001))
```