Markowitz Portfolio Optimization
Theory and Application

Yiming Xu[1,2] and Eric Adkins[1]

[1]Mathematics Department, Colorado State University
[2]Statistics Department, Colorado State University

April 20, 2015

Contents

1 Introduction

Suppose you have a data matrix comprised of several stock options over a set period of time. How do you choose the optimal collection of stocks such that you maximize your returns for a given level of risk? What Markowitz found was an elegant equation.

$$Min - \sum_{j} x_j(E(R_j)) + \mu \sum_{j,k} x_i x_j (E(\hat{R}_j \hat{R}_k)) \tag{1}$$

What we realized rather quickly is there does not exist a closed form solution to this problem. Instead we use the tried and tested linear approximation. By transforming this problem into Matrix multiplication, we are able to quickly and (with desired accuracy) approximate the optimal solution, using only linear algebra.

2 Problem Formulation

Given that Markowitz did the heavy lifting all that remains for us is to create a model that can take in data from Market analysis, transform that data into a quadratic problem, then optimize the approximation to the solution to a determined stopping point or (tolerance). To do so, we break this task into several sub steps.

1. Develop a solution technique for the optimization of quadratic programs
2. Cast the Markowitz portfolio problem as a quadratic program
3. Explore the expected returns for allowable risks.

2.1 Assumptions

- Only 11 investment opportunities from an assigned data matrix
- Allowable solutions are those determined to be within an allowable range of the true solution
- Customer's allowable risks can be captured by a single scalar parameter

2.2 Variables

All of the return data is held within a $M \times N$ matrix, $\mathcal{X}$, whose columns are different stocks and whose rows represent the change in the stocks value through time. The decision variable, x, will be the distribution of funds across the considered stocks. This will be a vector whose components are non-negative and sum to unity. The determination of this is the goal of this project. Additionally, there will be a risk parameter, μ, the purpose of which will be described later.

2.3 Interpretation as a quadratic program

The notion of risk and reward must be formalized quantitatively. A simple model for the reward is to consider the mean over time of each stock and the relative quantity of that stock bought. This will be achieved by introducing a vector whose components are the mean returns of each financial instrument.

$$c_i = \frac{1}{M} \sum_{j=1}^{M} \mathcal{X}_{ji}, \qquad 1 \leq i \leq N \tag{2}$$

Therefore the reward will be $\hat{R} = c^T x$. A reasonable characterization of risk is the variance of the expected reward.

$$\begin{aligned} Var(\hat{R}) \qquad &= E\left[\left(\hat{R} - E(\hat{R})\right)^2\right] \\ &= E\left[\left(\sum_j x_j (R_j - E(R_j))\right)^2\right] \end{aligned}$$

then by substituting $(R_j - E(R_j)$ by $\tilde{R}_j$ the variance may be simplified to a form where the decision variable may be extracted.

$$\begin{aligned} Var(\hat{R}) \quad &= E\left[\left(\sum_j x_j \tilde{R}_j\right)^2\right] \\ &= \sum_j \sum_k x_j x_k E(\tilde{R}_j \tilde{R}_k) \\ &= x^T Q x \end{aligned} \tag{3}$$

Where the expectation of the mean subtracted financial instruments has been wrapped up in the matrix Q. This problem may now be formulated as an optimization problem.

$$minimize \qquad \frac{1}{2} x^T Q x - c^T x \tag{4}$$

Finally the constraints of the problem must be applied. Since the decision variable is the relative distribution of funds its components must sum to one and be non-negative. The first may be enforced by the use of Lagrange multipliers and the second may be achieved by barrier methods. This amounts to including a monotonic term that grows very large for small values of the components. Here a logarithmic term will be used. Some optimal solutions will have components that are zero which can be handled by decreasing the effect of this term as the solution progresses. The inclusion of these terms results in the final quadratic program.

$$minimize \qquad \frac{1}{2}x^TQx - c^Tx + y^T(b - Ax) - \beta\sum_i ln(x_i) \tag{5}$$

Where y is the vector of Lagrange multipliers.

2.4 Iterative solution of the quadratic program

Taking partial derivatives of equation (4) with respect to the unknowns x and y yields a system of equations.

$$\begin{aligned} Qx - c - A^Ty - \beta\bar{X}^{-1}e = 0 \\ b - Ax = 0 \end{aligned} \tag{6}$$

Here, $\bar{X}$ is a diagonal matrix of the components of x and e is a column vector of ones. An inverted matrix will become difficult to deal with and as such a substitution may be made by introducing the variable z.

$$\begin{aligned} z &= \beta\bar{X}^{-1}e \\ \bar{X}z &= \beta e \\ \bar{X}Ze &= \beta e \end{aligned} \tag{7}$$

The non-linearity of this system does not lend it direct solution techniques, instead an initial guess will be made and the solution iteratively approached. The vectors x, y, z will be assigned some initial value and then replaced by $x + \Delta x, y + \Delta y, z + \Delta z$ in the system where $\Delta x, \Delta y, \Delta z$ are now the unknowns.

$$\begin{aligned} Q(x + \Delta x) - c - A^T(y + \Delta y) - (z + \Delta z) &= 0 \\ b - A(x + \Delta x) &= 0 \\ \bar{X}Ze + \Delta\bar{X}Ze + \bar{X}\Delta Ze + \Delta\bar{X}\Delta Ze &= \beta e \end{aligned} \tag{8}$$

Then by recognizing that the term $\Delta\bar{X}\Delta Ze$ should be small compared to all other terms, it may be dropped to linearize the system thus obtaining its final form.

$$\begin{aligned} Q\Delta x - A^T\Delta y - \Delta z &= -Qx + c + A^Ty + z \\ A\Delta x &= Ax - b \\ \Delta\bar{X}Ze + \bar{X}\Delta Ze &= \beta e - \bar{X}Ze \end{aligned} \tag{9}$$

This system may be written succinctly in block matrix form.

$$
\begin{pmatrix} A & 0 & 0 \\ -Q & A' & -I \\ Z & 0 & \bar{X} \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \begin{pmatrix} b - Ax \\ c + Qx - A'y + z \\ \beta * e - \bar{X} * Z * e \end{pmatrix}
$$

From this form progressively better solutions may be found by solving for $\Delta x, \Delta y, \Delta z$, updating x, y, z, and repeating until the prescribed tolerance is reached. This is done by summing the norms of the right hand side vectors and comparing to the tolerance.

$$
||b - Ax||_2 + ||c + Qx - A'y + z||_2 + ||\beta * e - \bar{X} * Z * e||_2 < TOL \tag{10}
$$

Finally, to ensure that the stepping of the variables forward does not result in negative values of x the values at each iteration will be updated with a term reminiscent of successive under relaxation i.e. $x_{new} = x + \epsilon^* \Delta x$ where the step size ϵ^* is calculated as such.

$$
\begin{aligned}
\epsilon &= \tfrac{1}{\min\limits_i \Delta\theta_i / \theta_i} \\
\epsilon^* &= min(\epsilon, 0.9)
\end{aligned} \tag{11}
$$

3 Model Development

In the application of this quadratic program framework to the Markowitz portfolio optimization the matrix Q may be pre-multiplied by some positive real scalar μ to vary its relative importance to the problem. Large values of μ then force the risk to play a greater role in the optimization thus returning a lower risk portfolio. The full effects of this risk parameter will be explored in later sections.

4 Results

μ	0.125	0.25	0.5	1	2	4	8	16	32	64	128	256	512	1024
goog	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.040	0.055	0.235	0.206	0.204	0.237	0.237
aapl	0.000	0.000	0.000	0.407	0.637	0.753	0.640	0.347	0.100	0.039	0.014	0.012	0.012	0.012
msft	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.012	0.016	0.026	0.028	0.027
gspc	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.422	0.303	0.348	0.394	0.393	0.402
djia	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.014	0.070	0.087	0.067	0.050	0.047
hon	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.021	0.051	0.073	0.062	0.047	0.032
ibm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.044	0.056	0.065	0.054	0.038	0.036
wmt	0.000	0.000	0.000	0.000	0.000	0.000	0.043	0.031	0.045	0.028	0.012	0.019	0.039	0.048
ual	0.000	0.000	0.000	0.000	0.000	0.000	0.141	0.471	0.199	0.109	0.111	0.084	0.064	0.065
ford	1.000	1.000	1.000	0.593	0.363	0.247	0.175	0.111	0.069	0.078	0.047	0.045	0.055	0.057
ba	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.015	0.019	0.021	0.032	0.037	0.037
Risk	0.00233	0.00233	0.00233	0.00080	0.00036	0.00024	0.00017	0.00009	0.00003	0.00002	0.00002	0.00002	0.00002	0.00002
Reward	1.0051	1.0051	1.0051	1.0040	1.0034	1.0031	1.0026	1.0018	1.0002	1.0000	0.9997	0.9996	0.9996	0.9996

The above table shows that increasing the risk also increases the reward. However, there is a huge caveat. What is demonstrated with Figure 1, is that for a small risk you are actually looking at taking on a loss. Where as the highest risk is only a return of .5 percent. This is not exceptionally exciting, a typical savings account can match those returns. A CD can easily return 1.3 percent.

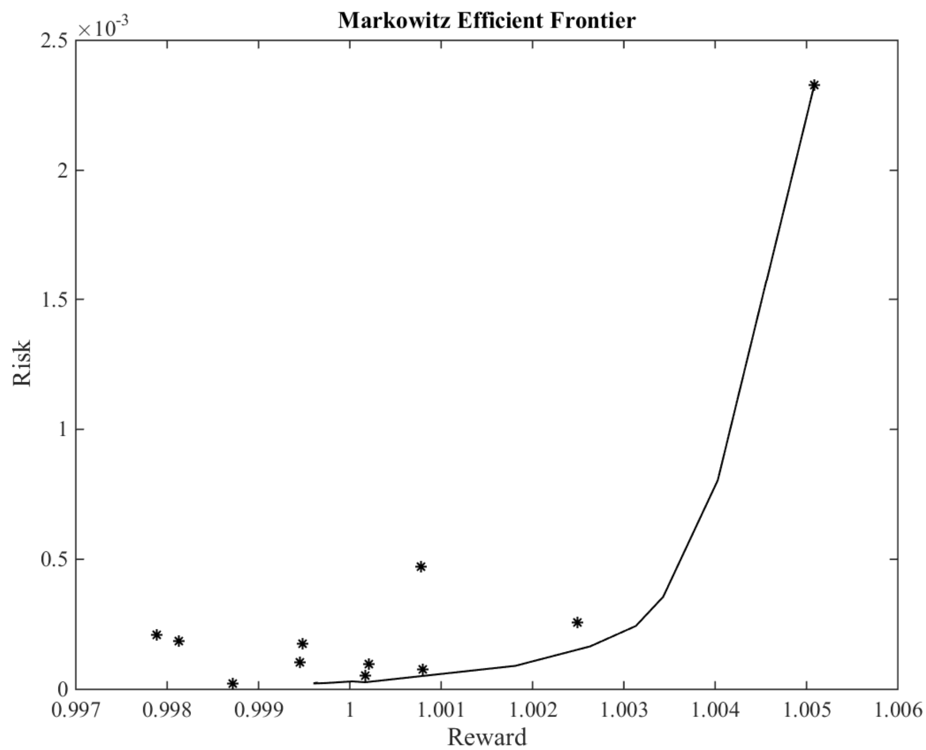

Figure 1: The solid line indicates the notable efficient frontier from standard Markowitz theory. Only optimal portfolios will lie on this border. Asterisks denote the expected returns from investing in only single financial instruments.

Figure 1 illustrates the famous efficient frontier provided by Markowitz portfolio theory. This curve represents the highest yielding portfolios for each allowable risk. The individual financial instruments (shown as asterisks) all lie on the lower yielding side of the curve as investing in only a single stock has a lower return than a diverse portfolio. The only exception to this is at the highest risk Markowitz theory shows that only the Ford stock should be purchased.

5 Robustness

The easiest way to test our model is to try other quadratic functions, and note the amount of steps to converge to the correct solution. Checking the vector $\vec{x}$ will show the two zero's for the quadratic function.

6 Model Evaluation

6.1 Added Value

There is more to investing than looking at the average return of an instrument. The variance of the return is very important to consider, even though it is not typically thought of. Using this model, the return variance is taken into consideration. This model adds an extra important variable to the decision making process when investing in the stock market.

6.2 Model Strengths

- Well written to easily use
- Easy to adjust tolerance
- Easy to keep track of amount of runs

6.3 Model Weaknesses

- Data used does not necessarily represent whole market data

7 Extensions of the Model

The most important extensions to the model are predictive methods. Three such methods have been implemented showing interesting deviations from the standard model. First, a naive method is simply to ignore some of the old data under the notion that it will have less bearing on the present behavior. In the sample calculated the first quarter of the data has been ignored. A second more interesting approach is to try extrapolating the return data and using that augmented matrix to perform the Markowitz analysis on. The third method is simply a combination of the first two. For a fixed risk of $\mu = 1$ the various compare as shown in the table below.

Case	Standard	Truncated	Extrapolated	Combined
goog	0.000	0.000	0.000	0.000
aapl	0.407	0.535	0.157	0.209
msft	0.000	0.000	0.000	0.000
gspc	0.000	0.000	0.000	0.000
djia	0.000	0.000	0.000	0.000
hon	0.000	0.000	0.000	0.000
ibm	0.000	0.000	0.000	0.000
wmt	0.000	0.000	0.000	0.000
ual	0.000	0.000	0.000	0.000
ford	0.593	0.465	0.843	0.791
ba	0.000	0.000	0.000	0.000

As the risk parameter is raised the different methods all start to homogenize and yield very similar answers. These methods were tested by withholding some of the most recent data and then evaluating the return that would have been obtained. For a risk of $\mu = 6$ the combined method yielded the best return at 1.0033 with a portfolio as shown.

μ	1
goog	0.0000
aapl	0.2507
msft	0.0000
gspc	0.0000
djia	0.0000
hon	0.0000
ibm	0.0000
wmt	0.1501
ual	0.1876
ford	0.4116
ba	0.0000

These extensions require much more testing and validation.

7.1 Extrapolation Code

While truncation is a simple task, the details of extrapolation are worthy of mention. The code simply augments the return data matrix $\mathcal{X}$.

```
load('Portfolio.mat');
[M, N] = size(X);
L = 20;
leastSquares = zeros(2,N);
Alsqr = ones(L,2);
for k=1:L
    Alsqr(k,1) = k;
end
for k=1:N
    blsqr = X(end-L+1:end,k);
    leastSquares(:,k) = lsqr(Alsqr, blsqr);
end
for k=1:L/2
    for i=1:N
        X(M+k,i) = k*leastSquares(1,i)+leastSquares(2,i);
    end
end
```

8 Conclusions

Through this investigation of optimal portfolio selections we have found that it is not possible to yield great returns with this method. Our returns, while mostly positive, were less than 1% for all cases. We believe this is caused in part by the data used to test our model. In further studies it would be advantageous to choose stocks that represent all sectors of the market.