

Portfolio Optimization via Markowitz Model Extensions

Jake Merry

June 9, 2025

1 Introduction

In 1952, Harry Markowitz published a paper in the *Journal of Financial* under the title “Portfolio Selection”. In this paper, Markowitz first presented his optimization model that would define what is known as modern portfolio theory, and for which he would later win the Nobel Memorial Prize in Economic Sciences. We will build off Markowitz’s ideas to optimize a simple stock portfolio, comparing different extensions as we go to ensure that we choose a good model for our data.

To achieve the Markowitz model, we start with an index set $\Omega = \{1, 2, \dots, n\}$ of financial assets which we call our universe. We associate with our universe a probability vector $\underline{\omega}$ such that for all $i \in \Omega$, ω_i is the proportion of our total wealth that we invest in asset i .

With respect to any universe, there exists a random vector $(R^{(1)}, R^{(2)}, \dots, R^{(n)})$ where each $R^{(i)}$ represents the return of an asset $i \in \Omega$. Then, for each $i \in \Omega$, we can define $\mu_i = \mathbb{E}(R^{(i)})$, the expected value of the return price of asset i . However, we are more interested in the total return of our entire portfolio, given by

$$R = \sum_{i \in \Omega} \omega_i R^{(i)}.$$

We can then find the total expected return of our portfolio as follows:

$$\mathbb{E}(R) = \mathbb{E} \left(\sum_{i \in \Omega} \omega_i R^{(i)} \right) = \sum_{i \in \Omega} \omega_i \mathbb{E}(R^{(i)}) = \sum_{i \in \Omega} \omega_i \mu_i = \underline{\mu}^T \underline{\omega}.$$

Similarly, we can solve for the total variance of the returns:

$$\text{Var}(R) = \text{Var} \left(\sum_{i \in \Omega} \omega_i R^{(i)} \right) = \sum_{i, j \in \Omega} \text{Cov}(\omega_i R^{(i)}, \omega_j R^{(j)}) = \sum_{i, j \in \Omega} \omega_i \omega_j \text{Cov}(R^{(i)}, R^{(j)}) = \underline{\omega}^T \underline{\underline{\Sigma}} \underline{\omega},$$

where $\underline{\underline{\Sigma}}$ is the covariance matrix of the returns vector.

One common form of the Markowitz model is the constant expected returns (CER) model. To derive this model, we start by choosing a return rate R which we will hold constant while we minimize the risk to our

investment. In other words, our goal is to find the weights $\underline{\omega}^*$ which minimizes the variance of the returns of our portfolio, while having a constant return rate of R .

The CER model is, hence, defined as follows:

$$\begin{aligned} & \text{minimize} \quad \underline{\omega}^T \underline{\Sigma} \underline{\omega} \\ & \text{subject to} \quad \underline{\mu}^T \underline{\omega} = R \\ & \quad \quad \quad \underline{1}^T \underline{\omega} = 1. \end{aligned}$$

This is the model that we will work with in this report.

2 Analysis

We begin by collecting stock data for the top 20 companies on the S&P 500 by weight. Data was collected monthly from nasdaq.com starting from May 1999. Then, for each asset in our portfolio, we have a collection of numbers $\{P_t^{(i)}\}$, representing the price of asset $i \in \Omega$ at a time t . We convert the prices to returns via the following formula:

$$R_t^{(i)} = \frac{P_t^{(i)} - P_{t-1}^{(i)}}{P_{t-1}^{(i)}}.$$

Taking the sample mean and covariances of each collection, we define

$$\hat{\mu}_i = \frac{1}{m} \sum_{t=1}^m R_t^{(i)} \quad \text{and} \quad \hat{\Sigma}_{ij} = \frac{1}{m-1} \sum_{t=1}^m (R_t^{(i)} - \hat{\mu}_i)(R_t^{(j)} - \hat{\mu}_j).$$

Hence, we have the sample mean vector $\hat{\mu} = (\hat{\mu}_1, \dots, \hat{\mu}_n)$ and sample covariance matrix

$$\hat{\Sigma} = \begin{bmatrix} \hat{\Sigma}_{11} & \dots & \hat{\Sigma}_{1n} \\ \vdots & \ddots & \vdots \\ \hat{\Sigma}_{m1} & \dots & \hat{\Sigma}_{mn} \end{bmatrix},$$

which we will use in place of μ and Σ in our analysis.

2.1 Vanilla Model

The first thing we notice about our derived model is that it does not include a nonnegativity constraint. Although this means that our optimal weight vector $\underline{\omega}^*$ may have negative components, this is not an accident. It may seem that, if this were to happen, that our model is telling us to take money from these assets and invest that money in other assets. But how is this possible? What a negative optimal weight really refers to is that this model allows for a technique called shorting. Shorting an asset is when an investor buys shares of that asset from a broker and sells them in an open market with the hope of later buying them back at a lower price and collecting on the difference.

Assuming that we continue with this model, computation is very simple. We define the Lagrangian of this problem as

$$\ell(\omega, \lambda) = \underline{\omega}^T \underline{\Sigma} \underline{\omega} + \lambda_1 (\underline{\mu}^T \underline{\omega} - R) + \lambda_2 (\underline{1}^T \underline{\omega} - 1).$$

Then, we can define the second-order sufficient conditions for our model as below.

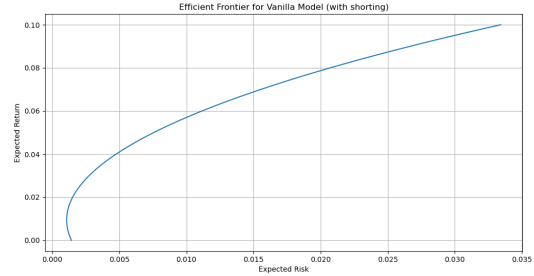
Suppose there exists some pair $(\underline{\omega}^*, \underline{\lambda}^*)$ such that the following conditions hold:

1. $\nabla_{\omega} \ell(\underline{\omega}^*, \underline{\lambda}^*) = 2\underline{\Sigma} \underline{\omega}^* + \lambda_1^* \underline{\mu} + \lambda_2^* \underline{1} = 0$,
2. $\underline{y}^T \nabla_{\omega\omega}^2 \ell(\underline{\omega}^*, \underline{\lambda}^*) \underline{y} = 2\underline{y}^T \underline{\Sigma} \underline{y} > 0$ for all $\underline{y} \in T(\underline{\omega}^*)$,
3. $\underline{\mu}^T \underline{\omega}^* = R$,
4. $\underline{1}^T \underline{\omega}^* = 1$.

Then ω^* is a strict local minimizer. Hence, we look for a vector $(\underline{\omega}, \lambda_1, \lambda_2)$ such that

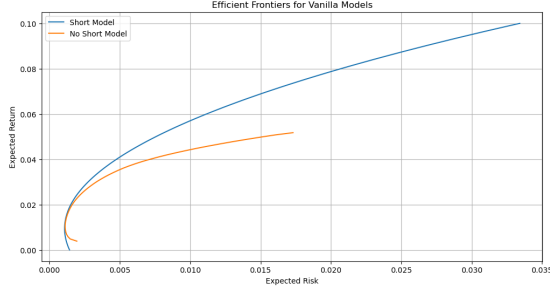
$$\begin{bmatrix} 2\underline{\Sigma} & \underline{\mu} & \underline{1} \\ \underline{\mu}^T & 0 & 0 \\ \underline{1}^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \underline{\omega} \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ R \\ 1 \end{bmatrix}.$$

We see that this problem can be solved using linear algebra alone, and although solving would involve the computation of a potentially large matrix inverse, this problem is very nice. Solving for all $R \in [0, 0.10]$, we plot the minimized risk against the return rate to find what is called the efficient frontier of our portfolio.



However, allowing shorting in our model can introduce more risk than the objective function accounts for. Therefore, it would be beneficial under our circumstances to require nonnegativity, i.e. add the constraint $\underline{\omega} \geq 0$ to our model. Although this problem can no longer be solved by linear algebra alone, we now obtain a quadratic program whose second-order sufficient conditions are the same the previous problem with the addition of the nonnegativity constraint.

In terms of feasibility, the original problem had optimal solutions for all R since the problem could be solved using a matrix computation. However, with the nonnegativity constraint, we find that it is impossible to achieve a feasible solution for any $R > \max_i \{\mu_i\}$ since we would require $\underline{\mu}^T \underline{\omega} > \max_i \{\mu_i\}$ with $\underline{1}^T \underline{\omega} = 1$ and $\underline{\omega} \geq 0$. Similarly, there is no feasible solution for $R < \min_i \{\mu_i\}$. Hence, we restrict our problem to $R \in [\min_i \{\mu_i\}, \max_i \{\mu_i\}]$. With this in mind, we plot the efficient frontier for the new, no shorting model and compare it with the original model.



Although the efficient frontier for the model without shorting falls entirely beneath that of the model which allows shorting, we will continue our analysis with the nonnegativity constraint due to the complicated nature of shorting.

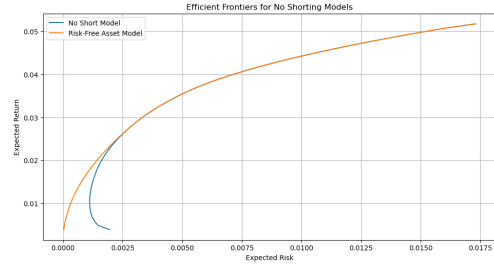
2.2 Risk-Free Asset Model

We now introduce our first extension of the Markowitz model. We construct a risk-free asset model by adding to our previous model an extra asset whose return variance is (close to) zero. This asset typically represents cash, so we will denote it as ω_{cash} . We update the constraints of our model to

$$\underline{\mu}^T \underline{\omega} + r_f \omega_{cash} = R, \quad \underline{1}^T \underline{\omega} + \omega_{cash} = 1, \quad \underline{\omega}, \omega_{cash} \geq 0 \quad (r_f \text{ small})$$

and solve for the efficient frontier.

Since the form of the problem has not changed by adding the extra asset, we solve the program in the same way as before. Then, we can plot the efficient frontier for the risk-free asset model and compare it to the no shorting model from before.



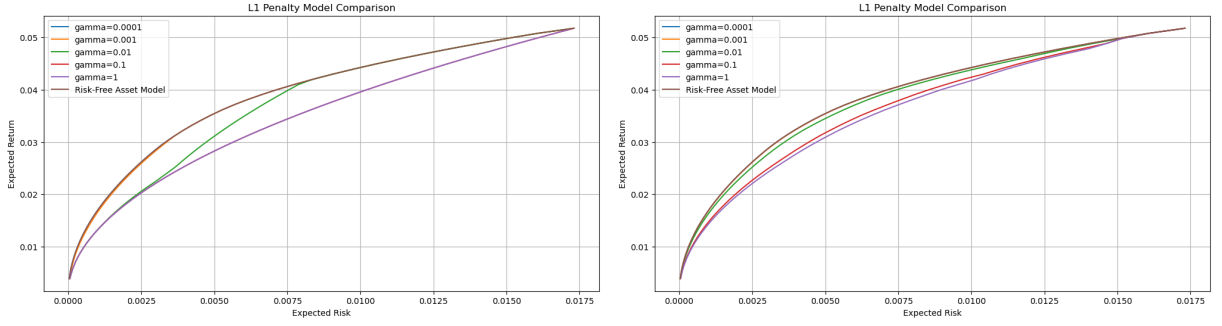
Our model has the expected result of extending the previous efficient frontier to one in which a portfolio exists such that the return and variance of the investment are both (close to) zero.

2.3 Comparison of Penalties

For the final part of our project, we will compare the constructed model to those with different added penalty terms to see if any of them benefit our model. Specifically, we will compare the ℓ_1 norm for sparsity and the square of the ℓ_2 norm for data sensitivity. The models, respectively, have objective functions

$$\underline{\omega}^T \underline{\Sigma} \underline{\omega} + \gamma \|\underline{\omega}\|_1 \quad \text{and} \quad \underline{\omega}^T \underline{\Sigma} \underline{\omega} + \gamma \|\underline{\omega}\|_2^2,$$

where $\gamma \geq 0$. For our analysis, we choose a small set of potential γ values and plot all efficient frontiers against that of the standard risk-free asset model. Here are the results:



We see that neither penalty decreases the variance of a portfolio at a fixed return rate, but this does not mean that these models are entirely invalid. Both penalties have models that, at very low values of γ are nearly identical to the original model, but would add other, potentially beneficial characteristics to our model.

3 Conclusion

After testing multiple different extensions, we found the best model for our data to be that of the risk-free asset model with no added penalty terms. This model behaved as expected, extending the efficient frontier of the no shorting model to the origin, but when we attempt to add an ℓ_1 or ℓ_2 norm penalty term to the model, the efficient frontier moved in a negative direction. This would return portfolios with the same expected return rate at a higher variance level.

References

- [1] Portfolio Optimization: Notes by Justin Wyss-Gallifent
- [2] Chong, E. K.P., Lu, W., & Zak, S. H. (2024). *An Introduction to Optimization with Applications to Machine Learning*. John Wiley & Sons, Inc.
- [3] Wright, S. J., & Recht, B. (2022). *Optimization for Data Analysis*. Cambridge University Press.
- [4] An Introduction to Portfolio Theory: Notes by Paul Atzberger
- [5] Nasdaq, www.nasdaq.com/. Accessed 28 May 2025.