Research On Finding Optimal Portfolios Based on Markowitz Mean-Variance Model -- Constructing Efficient Frontier and Portfolio with Highest Sharpe Ratio Using 4 Assets in Chinese Market

Xiaoyu Deng
Jinan University
xiaoyudeng6@outlook.com

Abstract. Markowitz mean-variance model is the foundation of modern investment science, and how to apply this model to construct optimal portfolios is always of interest. This paper uses the Markowitz mean-variance model and four typical assets in Chinese market to construct efficient frontier with highest expected return rate at each risk level and lowest risk at each expected return rate level, and the portfolios with the highest Sharpe ratio. There are three main results in this paper. Firstly, the efficient frontier has the U shape and V shape rotated 90° clockwise in the case of excluding and including risk-free asset respectively. Secondly, the theoretical efficient frontier obtained from mathematical formulas approximately coincide with those from the simulation experiments. Thirdly, the portfolio with the highest Sharpe ratio always occurs in the intersection point of the efficient frontier and its tangent line that passes through the risk-free asset. The results from this paper provide a convenient method of using mathematical formula for investor to construct optimal portfolios in real practice.

Keywords: Markowitz mean-variance model; efficient frontier; Sharpe ratio.

1. Introduction

Portfolio theory is the most important and valuable part in investment science. It answers the questions of how the investors can make the optimal investment decisions to maximize return based on known or estimated information about the assets in the market. In 1952, Harry Markowitz published the article portfolio selection in the Journal of Finance to provide strategies to choose efficient portfolios, which revolutionized modern investment theory [1]. In later years, he took deeper research in this field, improved and comprehensively expounded his mean-variance model in the article mean-variance analysis in portfolio choice and capital markets [2]. In 1961, a colleague of Markowitz, William Sharpe, proposed the single factor model in his article A simplified model for portfolio analysis [3]. This improvement made Markowitz’s model more applicable in real practice because of the significant reduction of computation. In 1981, another professor in economics, James Tobin, refined Markowitz’s model by considering the risk-free assets and cash when selecting portfolios in his article Liquidity preference as behaviour toward risk [4]. Tobin’s research not only made the analysis of investment more efficient, but also provided more solid foundation of investment theory.

The capital asset pricing is also an important part of investment research and decision making in real practice. In 1964, Willian Sharpe published the article Capital asset pricing: market equilibrium theory under risk, and explained his Capital Asset Pricing Model [5]. To be more specific, he used the market coefficient $\beta$ as a measurement of risk, and proposed that diversification of investment can only eliminate non-systematic risk, but not systematic risk. In 1976, Stephen A. Ross proposed another pricing model, named arbitrary pricing model, in his article Arbitrage pricing theory of capital asset [6]. This theory states that, the market is equilibrium if there is no opportunity of arbitrage.

How to use these powerful models to construct desirable portfolios and deal with investment problems is always of interest in modern society. In this paper, 3 typical risky assets and 1 risk-free asset in Chinese market are used to construct portfolios based on Markowitz mean-variance model.
One goal of this paper is to construct the efficient frontier. On one hand, by choosing the highest mean at each standard deviation level and the lowest standard deviation at each mean level in the simulation experiments in computer, the efficient frontier can be extracted from the whole feasible set. On the other hand, the mathematical expression of relationship between mean and standard deviation of return rate of the efficient portfolios can be obtained by solving the programming problem with estimated parameters mean, variance and covariance. It can be observed that the two efficient frontiers obtained from simulation experiments and theoretical derivation of the relational formula respectively have similar patterns. Another goal of this paper is to figure out the portfolio with the highest Sharpe ratio, since the portfolio with the highest ratio of expected return rate to risk is always the most preferable. The results of this paper shows that such portfolio always in the tangent point of the efficient frontier with the tangent line passing through the risk-free asset.

2. Method

2.1 Markowitz mean-variance model

The method used in the paper is Markowitz mean-variance model. This model is based a series of rigorous assumptions, and uses mean and variance to represent expected return rate and risk respectively. It establishes a quadratic programming problem to optimize the investment:

\[
\min \ x^T V x = \sum_i \sum_j x_i x_j \text{cov}(r_i, r_j)
\]

Subject to

\[
\begin{align*}
\text{subject to} \\
\sum_i x_i r_i = \mu \\
x^T 1 = \sum_i x_i = 1 \\
(x_i \geq 0)
\end{align*}
\]

Where \( V \) is the matrix of covariance, \( x \) is the vector of the relative amount invested in each single asset, \( R \) is the vector of return rate of each asset, and \( \mu \) is the expected return rate. It is important to note that the third constraint is only needed in the case that short selling is not allowed.

2.2 Theoretical derivation of formulas of efficient frontier

In this section, only the case of allowing short selling is taken into consideration.

2.2.1 Risk-free asset not included

In the case of not including risk-free asset, model (1) can be rewritten in a more concise way:

\[
\min \ x^T V x
\]

Subject to

\[
Px = d
\]

Where \( P = \begin{bmatrix} r_1 & r_2 & \cdots & r_n \end{bmatrix}, d = \begin{bmatrix} \mu \end{bmatrix} \).

Lagrange method can be used to solve this programming problem. In other words, finding solutions to the programming problem (2) is equivalent to find a solution \( [x, \lambda^*] \in \mathbb{R}^{n+2} \) such that it satisfies

\[
\begin{bmatrix} V & P^T \\ P & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda^* \end{bmatrix} = \begin{bmatrix} 0 \\ d \end{bmatrix}
\]

Where \( \lambda^* = [\lambda^*_1, \lambda^*_2] \) is the Lagrange multiplier vector.

After some operations of linear algebra, it can be found that the solution satisfying equation (3) is
\[
x = \frac{1}{D} (BV^{-1} - AV^{-1}R) + \frac{1}{D} (CV^{-1}R - AV^{-1}1) \mu
\]  
(4)

Where

\[
A = R^T V^{-1} 1
\]  
(5)

\[
B = R^T V^{-1} R
\]  
(6)

\[
C = 1^T V^{-1} 1
\]  
(7)

\[
D = \text{det} \begin{bmatrix} B & A \\ A & C \end{bmatrix}
\]  
(8)

Therefore, the relation between mean and standard deviation of return rate of the optimal portfolios can be obtained by substituting equation (4) into \( \sigma^2 = x^T V x \). That is,

\[
\sigma^2 = \frac{C}{D} (\mu - \frac{A}{C})^2 + \frac{1}{C}
\]  
(9)

Which is an equation of a hyperbola.

2.2.2 Risk-free asset included

In the case that risk-free asset is available in the investment, model (1) can be rewritten as

\[
\begin{align*}
\min & \quad x^T V x \\
\text{subject to} & \quad x^T R + (1 - \sum_{i=1}^n x_i) r_f = \mu
\end{align*}
\]  
(10)

Where \( V \) is the matrix of covariance between each two risky assets, \( x = [x_1, \ldots, x_n]^T \) is the vector of the relative amount invested in each single risky asset, (so \( 1 - \sum_{i=1}^n x_i \) is the relative amount invested in the risk-free asset), \( R = [r_1, \ldots, r_n]^T \) is the return rate of each risky asset, and \( r_f \) is the return rate of the risk-free asset.

Similarly, using the Lagrange method, model (10) can be transformed into the problem of finding a solution \([x, \lambda^*] \in R^{n+1}\) such that it satisfies

\[
\begin{bmatrix}
V \\
(R - r_f 1)^T \\
0
\end{bmatrix}
\begin{bmatrix}
x \\
\lambda^*
\end{bmatrix}
= \begin{bmatrix}
0 \\
\mu - r_f
\end{bmatrix}
\]  
(11)

Where \( \lambda^* \) is the Lagrange multiplier.

Similar operations of linear algebra can be performed to solve equation (11), and the solution is

\[
x = \left( \frac{\mu - r_f}{(R - r_f 1)^T V^{-1} (R - r_f 1)} \right) V^{-1} (R - r_f 1)
\]  
(12)

Substituting equation (12) into \( \sigma^2 = x^T V x \), the relationship between mean and standard deviation of the return rate of the optimal portfolios is

\[
\sigma = r_f \pm \sqrt{C(r_f - \frac{A}{C})^2 + \frac{D}{C} \mu}
\]  
(13)

Which is an equation of two straight lines.
3. Results

3.1 Measurement of single assets

In this paper, 4 assets in Chinese market are used to construct portfolios, including 3 risky assets: Kweichow Moutai (600519, abbreviated as Moutai), Fosun Pharma (600196, abbreviated as Fosun), ShunFeng Express (002352, abbreviated as SF) and 1 risk-free asset: government loan (100303, abbreviated as rf). Using the data of these 4 assets in the past 12 years (from year 2010 to 2021), the estimated mean and standard deviation of the return rate of each asset can be calculated and listed in Table 1.

|            | Moutai | Fosun | SF   | rf   |
|------------|--------|-------|------|------|
| mean       | 0.043  | 0.012 | 0.015| 0.001|
| standard deviation | 0.225  | 0.120 | 0.159| 0.010|

The results of the covariance and correlation between each two assets are summarized in Table 2 and Table 3 respectively. These results are the estimated indices of the corresponding assets.

|            | Moutai | Fosun | SF   | rf   |
|------------|--------|-------|------|------|
| Moutai     | 0.051  | 0.004 | 0.004| -0.0001|
| Fosun      | 0.004  | 0.014 | 0.007| -0.0001|
| SF         | 0.004  | 0.007 | 0.025| 0.0002|
| rf         | -0.0001| -0.0001| 0.0002| 0.0001|

|            | Moutai | Fosun | SF   | rf   |
|------------|--------|-------|------|------|
| Moutai     | 1.000  | 0.156 | 0.121| -0.051|
| Fosun      | 0.156  | 1.000 | 0.346| -0.082|
| SF         | 0.121  | 0.346 | 1.000| -0.133|
| rf         | -0.051 | -0.082| -0.133| 1.000|

3.2 Test of hypothesis

One hypothesis of the Markowitz mean-variance model is that, the return rate of each asset in a time period is normally distributed. The histograms in Figure 1 represent the density of the return rate of these 4 assets in the past 12 years, and the red lines are the normal distributions with their mean and standard deviation equal to those values of the distribution of the return rate of the corresponding asset. The results of Figure 1 show that although there are some bias, the distributions of each single asset basically conform to the normal distribution assumption.

![Fig. 1 Distribution of return rate of each single asset](image)
Another hypothesis of the model is, the return rate of each asset follows independently identical distribution. The auto-correlation of each asset are illustrated in Figure 2, and the blue dash lines represent the acceptable auto-correlation interval. The results of Figure 2 shows that the return rate of Fosun, SF and rf satisfy this assumption. The auto-correlation of lag 1 and 4 of Moutai exceed the range of the acceptable autocorrelation interval, so it does not satisfy this assumption. However, Moutai is a typical stock in Chinese market, so it should be considered as the special investment object and be included into the research sample.

Fig. 2 Auto-correlation of return rate of each single asset

3.3 Case 1: Portfolios constructed by 2 risky assets

In this section, the only available assets are Moutai and Fosun (2 risky assets).

3.3.1 Case 1.1: short selling not allowed

In the situation of not allowing short selling, the vector of invested weight of each single asset $x=(x_1, x_2)=(x_1, 1-x_1)$ has the condition $0 \leq x_1 \leq 1$. Figure 3 shows the relationship between mean and standard deviation of the feasible portfolio obtained from the simulation experiment in computer. It can be observed that the feasible set of constructed portfolios is a curve, and the two original assets are at the end points of the curve.

Fig. 3 mean-sd figure of case 1.1

3.3.2 Case 1.2: short selling allowed

Compared with the previous section, the condition $0 \leq x_1 \leq 1$ can be removed in this case. However, in the simulation experiment in computer, it is impossible to execute the simulated process without limitation on the range of $x_1$. For simplification, the situation of $-1 \leq x_1 \leq 2$ is simulated and the result is illustrated in Figure 4. It shows that the curve has a hyperbolic shape. Compared with case 1.1 (the grey part between Moutai and Fosun), the two original assets are no longer at the end points of the curve in this case.
3.3.3 Estimated formula of efficient frontier

Using the estimated indices of the Moutai and Fosun asset, the mean and variance of the return rate of feasible constructed portfolios are:

\[
 r_p = x_1 r_1 + (1 - x_1) r_2 = 0.042556 x_1 + 0.011972 (1 - x_1) \tag{14}
\]

\[
 \sigma_p^2 = x_1^2 \sigma_1^2 + (1 - x_1)^2 \sigma_2^2 + 2x_1(1 - x_1) \sigma_{12} = 0.0567x_1^2 - 0.0204x_1 + 0.0144 \tag{15}
\]

Which are parametric equations about \( x_1 \).

Combining equation (14) and (15), the theoretical formula of the relationship between mean and standard deviation of feasible portfolios can be represented as a hyperbolic function:

\[
 \sigma_p^2 = 60.6170 r_p^2 - 2.1184 r_p + 0.0311 \tag{16}
\]

In Figure 5, the black curve represents the feasible portfolios obtained from simulation experiment as in Figure 4, and the green curve illustrates equation (16). It can be observed that these two lines have similar patterns.

3.3.4 Portfolio with highest Sharpe ratio

Sharpe ratio is a measurement of the performance of the expected return rate of a portfolio together with its risk. And the formula is

\[
 \text{Sharpe ratio} = \frac{E(r_p) - r_f}{\sigma_p} \tag{17}
\]

Where \( E(r_p) \) is the yearly expected return rate of the portfolio, \( \sigma_p \) is the standard deviation of the return rate, and \( r_f \) is the yearly return rate of the risk-free asset.

The Sharpe ratio of a portfolio is the slope of the line joining that portfolio and the risk-free asset in the mean-sd figure, since the standard deviation of a risk-free asset is approximately 0. To extract
the portfolio with the highest Sharpe ratio in the feasible portfolios in case 1.2, searching method is used in the simulation experiment. It can be found the yellow point P in Figure 6 has the highest Sharpe ratio in the whole feasible set. Its mean and standard deviation are 0.0363 and 0.1965 respectively, and therefore, its Sharpe ratio is 0.18198.

![Fig. 6 Portfolio with highest Sharpe ratio (case 1.2)](image)

**3.4 Portfolios constructed by 3 risky assets**

In this section, Moutai, Fosun and SF (3 risky assets) are available investment assets.

**3.4.1 Case 2.1: short selling not allowed**

In the case when short selling is not allowed, the vector of weight invested in each asset is 
\[(x_1, x_2, x_3) = (x_1, x_2 - x_1 - x_2)\] where \(0 \leq x_1, x_2 \leq 1\). The yellow part in Figure 7 is the representation of the feasible portfolios constructed by 3 risky assets in this case.

![Fig. 7 mean-sd figure of case 2.1](image)

**3.4.2 Case 2.2: short selling allowed**

In this case, the condition \(0 \leq x_1, x_2 \leq 1\) in case 2.1 is not needed. Theoretically, \(x_1, x_2, x_3\) can be any real numbers that satisfies \(x_1 + x_2 + x_3 = 1\). For the purpose of analysis in simulation experiment, only the situation of \(-3 \leq x_1, x_2 \leq 4\) is simulated in the experiment, because the computer has no capability to deal with the infinite situation. The yellow part in figure 8 is the feasible set in case 2.1, and the grey part is the feasible set when short selling is allowed. It can be observed that the yellow part is only a small fraction of the grey part. Therefore, short selling can expand the selection range for investors significantly in theory.
Given a certain level of expected return rate, the investors always choose the one with the lowest risk. Such portfolios are represented in the black and red lines in Figure 9, which are known as the efficient frontier. If the standard deviation of the portfolio is fixed, the investors always choose the one with the highest expected return rate. Therefore, only the upper branch of the simulated efficient frontier - the red curve in Figure 10 - is of interest.

3.4.3 Estimated formula of efficient frontier

Using the estimated indices of these 3 assets in section 3.1 and the theoretically derived formulas of efficient frontier in section 2.2.1 when risk-free asset is not available in investment, the values of A, B, C, D can be calculated.

\[ A = R^T V^{-1} = \begin{bmatrix} 0.0426 & 0.0120 & 0.0150 \end{bmatrix}^T \begin{bmatrix} 0.0507 & 0.0042 & 0.0066 & 0.0043 \end{bmatrix}^{-1} = 1.564020 \quad (18) \]

\[ B = R^T V^{-1} R = \begin{bmatrix} 0.0426 & 0.0120 & 0.0150 \end{bmatrix}^T \begin{bmatrix} 0.0507 & 0.0042 & 0.0066 & 0.0043 \end{bmatrix}^{-1} \begin{bmatrix} 0.0426 & 0.0120 & 0.0150 \end{bmatrix} = 0.043392 \quad (19) \]

\[ C = 1^T V^{-1} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T \begin{bmatrix} 0.0507 & 0.0042 & 0.0066 & 0.0043 \end{bmatrix}^{-1} = 91.180538 \quad (20) \]

\[ D = \det \left( \begin{bmatrix} C & A \\ A & C \end{bmatrix} \right) = 1.510358 \quad (21) \]

Substituting these values into equation (9), the formula of the estimated efficient frontier is
\[
\frac{\sigma^2}{1.91180538} - \left( \frac{\mu - 1.564020}{1.510358} \right)^2 = 1
\]  

(22)

Which can be transformed into

\[
\mu = \pm \sqrt{0.00018 \times (91.180538 \sigma^2 - 1)} + 0.0172
\]  

(23)

The green curve in Figure 10 is the representation of equation (23), which is the theoretically estimated efficient frontier. Comparing the simulated efficient frontier (the red and black curves) and the theoretically estimated one (the green curve), it can be observed that they approximately coincide with each other.

Fig. 10 Comparison between simulated efficient frontier and the theoretically estimated one

3.4.4 Portfolio with highest Sharpe ratio

Using formula (17) and the searching method, the portfolio with the highest Sharpe ratio can be extracted from the whole feasible set of constructed portfolios in case 2.2. The yellow portfolio P in figure 11 is the one with the highest Sharpe ratio. The blue line connects the rf asset and the portfolio P, thus its slope is the value of Sharpe ratio of the portfolio P. It can be found that the mean and standard deviation of the return rate of this portfolio P are 0.0367 and 0.1744 respectively, and thus, its Sharpe ratio is 0.2009.

Fig. 11 Portfolio with highest Sharpe ratio (case 2.2)

3.5 Portfolios constructed by 3 risky assets and 1 risk-free asset

In this section, Moutai, Fosun, SF (3 risky assets) and rf (risk-free asset) are all available in the investment.

3.5.1 Case 3: (short selling allowed)

In this section, all four assets are available in the investment, and only the case that allowing short selling is considered. Although \( x_1, x_2, x_3, x_4 \) can be any real number that satisfies \( x_1 + x_2 + x_3 + x_4 = 1 \), only the case of \( -1 \leq x_1, x_2, x_3 \leq 2 \) is simulated in the experiment, for the same reason as
stated in section 3.4.2. The feasible portfolios are illustrated in the grey part in Figure 12, and the feasible set in case 2.2 is also drawn in yellow for comparison. It can be observed that the grey set is an extension of the yellow part, since there is more asset available in the investment.

![Fig. 12 mean-sd figure of case 3](image)

The black and red lines in Figure 13 both have the lowest standard deviation at its corresponding mean level, and the red line have the greatest mean at its corresponding standard deviation level. It can be observed that the two branches starting from the rf asset have the shape of straight lines at first (extracted as blue lines). The reason why they bend when the standard deviation increases to a certain level is that, the grey part in figure 14 is not the whole feasible set of portfolios constructed by the four assets. Only the case of $-1 \leq x_1, x_2, x_3 \leq 2$ is taken into consideration in the simulation experiment.

![Fig. 13 Simulated efficient frontier in case 3](image)

3.5.2 Estimated formula of efficient frontier

Using the formula (13), and the value of A, C, D as calculated in (18), (20), (21), the theoretically estimated efficient frontier is

$$\sigma = 0.000534 \pm 0.196041\mu$$

(24)

The green lines in Figure 14 are the representation of formula (24). As shown in Figure 14, the theoretically estimated efficient frontier (green lines) and the simulated one (red and black lines) almost coincide with each other at the beginning of each branch. When the standard deviation increases to a certain level, they begin to diverge from each other, but this is because of the lack of some feasible points drawn in the figure.
3.5.3 Portfolio with highest Sharpe ratio

In this case, all the portfolios on the upper branch of efficient frontier have the greatest Sharpe ratio, which is equal to the slope of the upper efficient frontier. Using formula (17) and the estimated indices of the four assets, the slope of the upper efficient frontier can be calculated as 0.2009. Therefore, the highest Sharpe ratio within this feasible set is 0.2009.

4. Discussion

In the simulation experiment, the portfolio set is constructed in a discrete and finite way. In other words, the invested weight for each asset except the last one is set to be a finite arithmetic sequence, and the weight of the last asset is dependent of the previous weights, since the sum of the weight of all assets has to be equal to 1. As the number of available assets increase, if the dimension of the weight sequence of each asset is fixed, the size of the whole feasible set of constructed portfolios would grow significantly, and this would cause a problem if it is too tremendous for the computer to manipulate. Therefore, it is necessary to decrease the dimension of the weight sequence for each asset as the number of available assets increases. As a result, one disadvantage of this simulation experiment is that, its performance in solving problems when the number of available assets is greater than 5 in ordinary computer is poor, since the step of the arithmetic sequence is set to be large in order to reduce the dimension of each weight sequence, which would make it difficult to observe the pattern of the feasible set.

The second disadvantage of the simulation experiment is that the case of short selling cannot be simulated perfectly. When short selling is allowed, the invested weight of each asset can be any real number as long as the sum of weights of all assets is 1 in theory. However, a limitation of the range is required in the simulation experiment, and this would lead to a deformation of the feasible set as in Figure 14.

The third disadvantage is that the calculation is not accurate. The efficient frontier and the portfolio with the highest Sharpe ratio are extracted by searching the whole constructed finite portfolio set. Rounding is used in these steps and therefore the solutions are the approximate ones.

Another method used in this paper is to find the theoretically estimated efficient frontier using the estimated indices obtained from the sample. This method provides a more convenient method to find the efficient frontier as the estimated mean, variance and covariance can be obtained from the sample of previous data. Although this method still has a size limitation since it requires the calculation of matrix inversion, this can be ignored because it satisfies most of the needs in real practice. And this method also gives an expression of the desirable vector of invested weight of each asset in the efficient frontier.

Comparing the simulated efficient frontier and the theoretically estimated one (Figure 6, 11, 15), it can be observed that they have similar patterns, and especially in Figure 11 and 15, the two lines almost coincide with each other. Therefore, it concludes that these two methods are valid to some extent. They both give a geometric intuition for people to find the efficient frontier and the portfolio with the highest Sharpe ratio.
5. Conclusion

This paper uses the Markowitz mean-variance model and four typical assets in Chinese markets to construct optimal portfolios. The simulation experiment and theoretical mathematical formulas are used to construct the efficient frontier and portfolio with the highest Sharpe ratio. There are some conclusions from the results of this paper. Firstly, from the results of simulation experiments, it can be observed that the efficient frontier has a U shape that rotates 90° clockwise when risk-free asset is not available, and a V shape that rotates 90° clockwise when risk-free asset is available. To be more specific, the upper branch of the efficient frontier is of interest since people always aim at the portfolios that has the highest expected return rate at each risk level, and the lowest risk at each expected return rate level. Secondly, the efficient frontier obtained from the simulation experiment and the theoretical formula coincide with each other, which means that both methods used in this paper are valid to some extent. Thirdly, the portfolio with the highest Sharpe ratio occurs in the tangent point of efficient frontier with the tangent line passing through the risk-free asset when risk-free asset is not available in the investment. When risk-free asset is available, all portfolio in the upper efficient frontier has the highest Sharpe ratio.

Therefore, the theoretical formula provides people a convenient and reliable method to construct optimal portfolios. In other words, people can use formula (9) or (13) to figure out the efficient frontier, and the corresponding vector of invested weight of each asset can be calculated by formula (4) or (12). Finding the portfolio with the highest Sharpe ratio can be converted into a problem of finding tangent point, which can be easily done in mathematics.

References

[1] H. Markowitz: Portfolio selection.
[2] H. Markowitz: Mean-variance analysis in portfolio choice and capital markets.
[3] J. Tobin: Liquidity preference as behaviour toward risk.
[4] W. Sharpe: A simplified model for portfolio analysis.
[5] W. Sharpe: Capital asset pricing-market equilibrium theory under risk.
[6] S. A. Ross: Arbitrage pricing theory of capital asset.
[7] S. Hubbert: Essential mathematics for market risk management.
[8] Z. Cao, Q. Han: Management of investment portfolio.