**The University of Hong Kong**

**Department of Statistics and**

**Actuarial Science**

**STAT6013 Financial Data Analysis**

**Group Project**

**Optimal Portfolio Based on Markowitz's Efficient Frontier and Single Index Market Model**

|  |  |
| --- | --- |
| Student’s Name | Student’s ID |
| **YU Shu hang** | **3035676456** |
| **WANG Yao** | **3035676389** |
|  |  |
|  |  |

**Contents**

1. Abstract

2. Motivation

3. Model

4. Data

5. Methodology

6. Analysis

7. Discussion

8. Conclusion

9. References

# 1. Abstract

In our project, we will construct two financial models with R language to find the optimal portfolios. One is Markowitz’s Efficient Frontier, it is the set of optimal portfolios that offer the highest expected return for a defined level of risk or the lowest risk for a given level of expected return. The other one is Single Index Market Model(SIMM), it is an asset pricing model, according to which the returns on a security can be represented as a linear relationship with any economic variable relevant to the security. We use 5 stocks and 5 indexes of different companies as our data source. At last we compare these 2 models.

# 2. Motivation

These two models are more and more important in making optimal investment. We can make predications and ensure that our investment has a high return with low risk. We choose 5 stocks and 5 indexes to find the optimal portfolio using the Markowitz’s Efficient Frontier and the Single Index Market Model.

# 3. Model

## Markowitz’s Efficient Frontier

In [modern portfolio theory](https://en.wikipedia.org/wiki/Modern_portfolio_theory), the efficient frontier (or portfolio frontier) is an investment [portfolio](https://en.wikipedia.org/wiki/Portfolio_(finance)) which occupies the 'efficient' parts of the [risk-return spectrum](https://en.wikipedia.org/wiki/Risk-return_spectrum). Formally, it is the set of portfolios which satisfy the condition that no other portfolio exists with a higher expected [return](https://en.wikipedia.org/wiki/Return_(finance)) but with the same [standard deviation](https://en.wikipedia.org/wiki/Standard_deviation) of return.

A combination of assets, i.e. a portfolio, is referred to as "efficient" if it has the best possible [expected](https://en.wikipedia.org/wiki/Expected_value) level of return for its level of [risk](https://en.wikipedia.org/wiki/Financial_risk) (which is represented by the standard deviation of the portfolio's return). Here, every possible combination of risky assets can be plotted in risk–expected return space, and the collection of all such possible portfolios defines a region in this space. In the absence of the opportunity to hold a [risk-free asset](https://en.wikipedia.org/wiki/Risk-free_interest_rate), this region is the opportunity set (the [feasible set](https://en.wikipedia.org/wiki/Feasible_set)). The positively sloped (upward-sloped) top boundary of this region is a portion of a [hyperbola](https://en.wikipedia.org/wiki/Hyperbola) and is called the "efficient frontier".

## (2) Single Index Market Model

The single index market model (SIMM) is a simple [asset pricing](https://en.wikipedia.org/wiki/Asset_pricing) model to measure both the risk and the return of a [stock](https://en.wikipedia.org/wiki/Stock). The model has been developed by [William Sharpe](https://en.wikipedia.org/wiki/William_F._Sharpe) in 1963 and is commonly used in the [finance](https://en.wikipedia.org/wiki/Finance) industry. Mathematically the SIMM is expressed as:

Where is the return on asset in period ,

is the return on index in period

and are constants while and are random variables

# 4. Data

In this project, we are interested in the five following industries: agroforestry, real estate, finance, real estate, IT and catering. In each industry, we pick one interested stock: GUO LIAN (300094), Vanke A (000002), Pin An Bank (000001), ZQ Game. COM (300052) and Quan Ju De (002186) in Shenzhen and Shanghai Securities Exchange. The daily trading data of the stocks in the past six years and the daily market indexes of the industries in the past six years are obtained via Tushare database. Table 1 and Table 2 show the preview of original data, respectively.

For simplicity, Index 1, Index 2, Index 3, Index 4 and Index 5, Stock 1, Stock 2, Stock 3, Stock 4 and Stock 5 may denote to the indexes of the five markets and the five stocks in the above order.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ts\_code | trade\_date | open | high | low | close | pre\_close | change | pct\_chg | vol | amount |
| 300094.SZ | 20131231 | 7.63 | 7.72 | 7.24 | 7.7 | 7.63 | 0.07 | 0.92 | 49512.95 | 37678.7492 |

Table 1. Preview of daily trading data of stock

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ts\_code | trade\_date | close | open | high | low | pre\_close | change | pct\_chg | vol | amount |
| 399231.SZ | 20131231 | 797.697 | 797.204 | 798.355 | 785.174 | 798.609 | -0.912 | -0.1142 | 977452.45 | 1217593.6 |

Table 2. Preview of daily index data of market

# 5. Methodology

# (1) Markowitz’s Efficient Frontier

Assume investors only consider to minimize the volatility of risky assets and there is no risky-free asset. In terms of stock return, represent the daily return for any stock where is close price in term , hence, can represent the annual return since there are annually 250 trading days in Chinese Stock Exchange.

Consider the Lagrange Multiplier Method and review Equality Constraint Problem:

Let

or

Subject to

We then have two steps to solve it:

Firstly, let , where is called Lagrange Multiplier;

Secondly, solve the following system of equations:

The above conditions are necessary and sufficient for the maximal (minimal) point if

is concave, i.e. ;

is linear, i.e. .

In this project, we are interested in four scenarios including global minimum variance portfolio (with short-selling and without short-selling) and portfolio with a specified expected return . With Lagrange Multiplier, the first three scenarios can be easily computed the optimal solution:

1. The portfolio with global minimum variance, if shot-selling is allowed:

Subject to

Introduce Lagrange Multiplier :

Solve it and get the final result:

Where , is the vector of portfolio weight, is the vector of return, is the covariance matrix and is the expected level of return.

1. The portfolio with global minimum variance, if shot-selling is not allowed:

Subject to for

For those assets to be short-sold, set For others, set Introduce Lagrange Multiplier and :

Guess , then set the column(s) of to be zeros and the diagonal element(s) to be -1. Solve it and get the final result:

1. The portfolio with a specified expected return , if shot-selling is allowed:

Subject to

Introduce Lagrange Multiplier and :

Solve it and get the final result:

Where

1. The portfolio with a specified expected return , if shot-selling is not allowed:

Subject to , for

Under this situation, it is too difficult to solve the problem with as many as constraints. Instead, use Sequential Least Squares Programming (SLSQP) optimizer, which is a sequential least squares programming algorithm which uses the Han-Powell quasi-Newton method with a BFGS update of the B-matrix and an L1-test function in the step–length algorithm, to come up with an optimal solution in non-linear way. It can be performed by the Module SciPy in Python.

## (2) Single Index Market Model

The portfolio formula:

Where is the return on asset in period ,

is the return on index in period

and are constants while and are random variables

Assumptions:

i) are random errors with and ;

ii) are independent of ;

iii) Since is random, we assume that has and

Properties:

Mean of asset return:

Variance of asset return:

Covariance between asset returns:

Under SIMM, the covariance matrix of the asset returns is

Therefore, the portfolio return in period is

Where , and

Hence, the portfolio return has

Mean:

Variance: where

# 6. Analysis

# (1) Markowitz’s Efficient Frontier

Initially, Table 1 shows the basic output of the mean, the variance and the covariance matrix of annual stocks return.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Mean | Standard Deviation | Covariance Matrix | | | | |
| Stock 1 | -11.62 | 38.60 | 1788.19 | 134.26 | 199.50 | 668.234 | 644.26 |
| Stock 2 | 3.99 | 27.10 |  | 881.60 | 83.76 | 595.91 | -117.66 |
| Stock 3 | 21.11 | 36.40 |  |  | 1590.09 | -59.03 | 561.40 |
| Stock 4 | -9.05 | 33.36 |  |  |  | 1335.73 | 176.46 |
| Stock 5 | -12.23 | 19.20 |  |  |  |  | 442.15 |

Table 1. Basic output of annual stocks return

In terms of annual return, two of them own positive mean returns. The largest mean is 21.11 corresponding to Stock 3 while its standard deviation is second largest. In terms of volatility, all stocks fluctuate dramatically and Stock 1 has the largest value 38.6 which is approximate to 36.4 of Stock 3. According to the covariance matrix, it can be demonstrated that most of them tend to be positively correlated, in particular only Stock 2 and Stock, Stock 3 and Stock 4 are negatively correlated.

Review four scenarios mentioned in Section 5(1):

Scenario 1: find the portfolio with global minimum variance, if shot-selling is allowed;

Scenario 2: find the portfolio with global minimum variance, if shot-selling is not allowed;

Scenario 3: find the portfolio with a specified expected return , if shot-selling is allowed;

Scenario 4: find the portfolio with a specified expected return , if shot-selling is not allowed.

Here, is set to be 10.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Mean | Variance | Weight of Stock 1 | Weight of Stock 2 | Weight of Stock 3 | Weight of Stock 4 | Weight of Stock 5 |
| Scenario 1 | -25.60 | 13.70 | -0.55 | 0.56 | -0.64 | -0.25 | 1.88 |
| Scenario 2 | -8.51 | 267.24 | 0 | 0.23 | 0 | 0 | 0.77 |
| Scenario 3 | 10.00 | 569.00 | 0.03 | 0.54 | 0.41 | -0.05 | 0.07 |
| Scenario 4 | 10.00 | 570.56 | 0.02 | 0.50 | 0.42 | 0 | 0.05 |

Table 3. Output of four scenarios

Table 3 indicates the optimal portfolios based on four given scenarios including their corresponding expected return and variance. Accordingly, if short-selling is allowed, the global minimum variance can be obtained with the value of 13.7 while its relative expected return is negative. If short-selling is not allowed, Stock 1, Stock 3 and Stock 4 are no longer considered in the optimal portfolio. The expected return is still negative and the variance dramatically increases to 267.24.

In order to make positive return, investors should bear more risks as the relevant variance becomes approximate to 570 which is two times higher than no required expected return. If short-selling is additionally allowed, then only Stock 4 is shortly sold as the weight is -0.05. If no short-selling, Stock 4 is not in the portfolio, but the result is not explicitly distinct compared to the previous situation.

图片包含 屏幕截图

描述已自动生成

Figure 1. Efficient frontier based on two conditions

Concentrating on making positive return only, Figure 1 shows the efficient frontier based on the conditions of short-selling and no short-selling. Given the expected return between 0 and 50, the variances of them both exponentially grow up. However, the difference is that if short-selling is allowed, the maximum return can be obtained although the corresponding variance would be greatly large, whereas if short-selling is not allowed, the expected return cannot go beyond about 20, meanwhile, the relative variance is larger when investors expect to make return approaching to 20. Table 4 involves more details about the portfolio, based on four extreme kinds of results, which are recognized in Figure 1, including maximum expected return and minimum variance (with short-selling and without short-selling).

Considering positive expected return between 0 and 50 only,

Extreme scenario 1: portfolio with the minimum variance, if short-selling;

Extreme scenario 2: portfolio with the maximum expected return, if short-selling;

Extreme scenario 3: portfolio with the minimum variance, if no short-selling;

Extreme scenario 4: portfolio with the maximum expected return, if no short-selling.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Mean | Variance | Weight of Stock 1 | Weight of Stock 2 | Weight of Stock 3 | Weight of Stock 4 | Weight of Stock 5 |
| Extreme scenario 1 | 0 | 293.16 | -0.13 | 0.54 | 0.12 | -0.10 | 0.58 |
| Extreme scenario 2 | 49.00 | 2539.49 | 0.66 | 0.51 | 1.56 | 0.17 | -1.90 |
| Extreme scenario 3 | 0 | 320.07 | 0 | 0.44 | 0.15 | 0 | 0.41 |
| Extreme scenario 4 | 21.11 | 1590.09 | 0 | 0 | 1 | 0 | 0 |

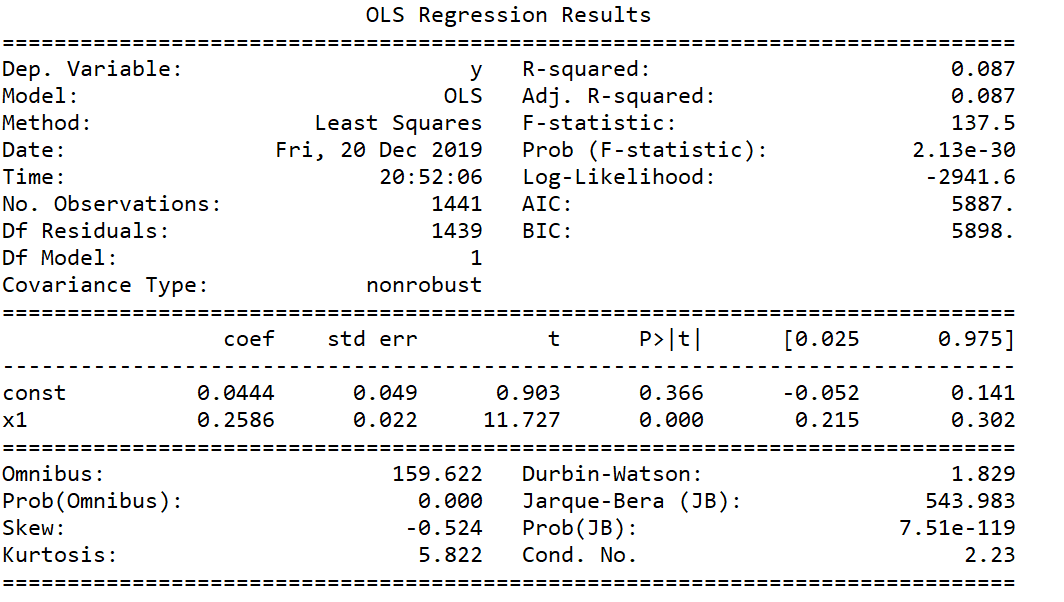
Table 4. Output of four extreme scenarios

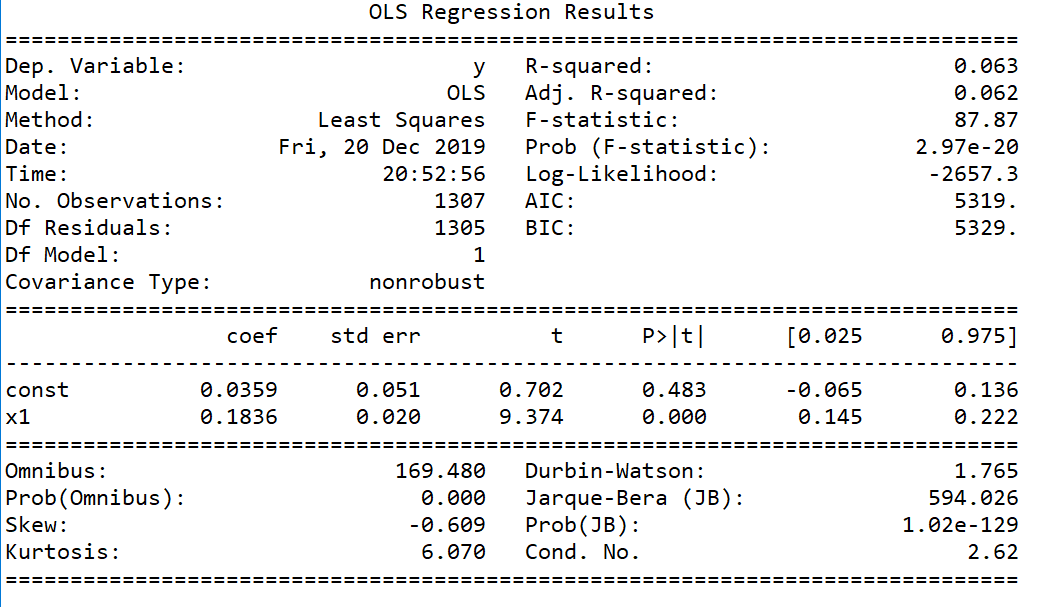
# (2) Single Index Market Model

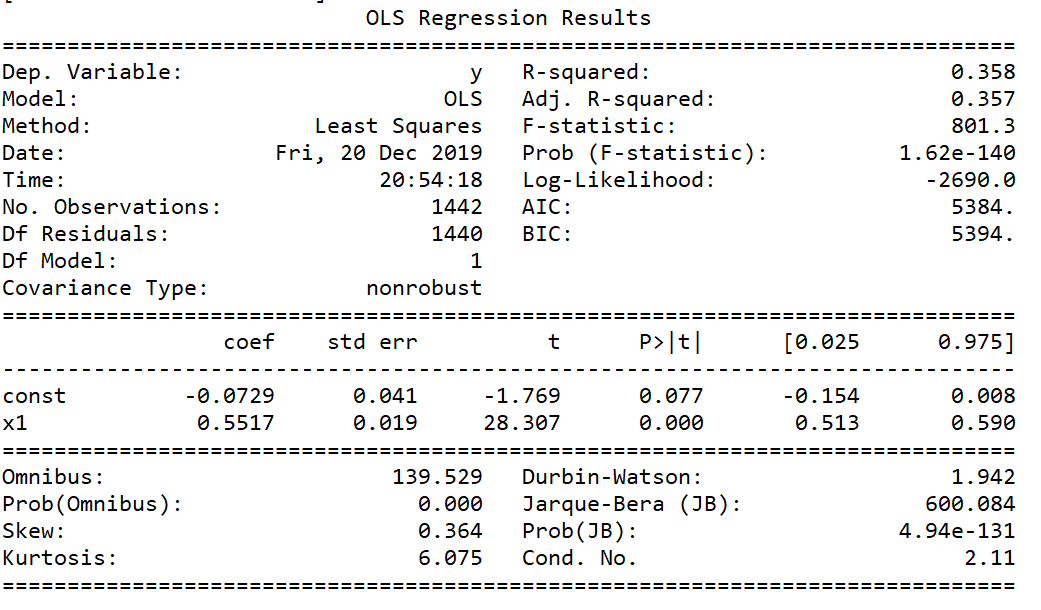
We try to use SIMM find the tangency portfolio. We assume that the 5 top stocks we chose from different industry have a relationship with the total market index. Therefore, the 5 individual stock can be replaced by the market index.

Finding a good index that has a strong relationship with all the stocks is important to make this model precisely. We collected some popular market index, including SHANGZHENG 50 ETF (sz50), HUSHEN 300 (hs300) and industries index. We finally choose the industry index, which is better and make sense to explain from economics perspective.

Fit the individual stock with market index by OLS methods. Some regression results are as below:







Explanation for the parameters:

The parameter results from OLD regression can be used as the indication for individual stock. The return and variance of stock can be calculated by the parameters and the performance of the index.

Explanation for low R^2:

Because the individual stock can only reflect part of the situation of the whole market. It’s a normal thing that one stock cannot reflect and fit the whole market. As we analysed from the scatter plot before, the industry market index already reflects a good relationship between stocks and index. And in the real world, it’s hard to find one good stock that can fit the whole market.

The return and non-system risk of individual stock as below:

|  |  |  |
| --- | --- | --- |
| Industry | Return | Non-system Risk T^2 |
| Arg | 9.321997008227378 | 3667.4679607152057 |
| Finance | 11.52972260748959 | 5003.546394875989 |
| House | 20.548157492354733 | 4464.636811000912 |
| IT | 19.643019169329076 | 3543.208381397129 |
| Rst | 18.04383922383923 | 3521.996856438093 |

According to the formula we can calculate the tangency portfolio weights. Because the stocks we chose are nearly top 10 stock from each industry. A good result should be longing all the 5 stocks. Because different industry can disperse the risk. Our result is as our expectation. And the result is below:

|  |  |
| --- | --- |
| Industry | Weight |
| Arg | 0.1328871010988875 |
| Finance | 0.08839297144526322 |
| House | 0.19823701786648268 |
| IT | 0.3008501108661973 |
| Rst | 0.2796327987231694 |

# 7. Discussion

According to the analysis of Markowitz’s Efficient Frontier, it is easy to find out these five stocks tend to have negative annual returns and high volatilities which may leads to a poor result to portfolios. Even the global minimum variance reaches 13.7, the expected return of the optimized portfolio is unexpectedly negative with short-selling. If there is no short-selling, the number of risky assets with positive weights will be declined so much, and obviously, the risk of investment will be increased due to it.

In this case, a positive return becomes more difficult to make because investors should bear extremely high volatility. To make positive return up to 50, the portfolio is not practical. On the one hand, short-selling of a stock in hand cannot be larger than 100% (100% means entire sale out), while the weight of Stock 5 in Extreme scenario 2 is -190%. On the other hand, the maximum expected return can be reach is only 21 without short-selling, even if investors just consider Stock 3, which has the most outstanding mean of annual return, and ignore volatility.

Above all, the portfolio of these five stocks can only make positive expected annual return when 293.16, as the minimum value of variance, can be accepted by investors. Otherwise, investors lose their money. To solve this problem, it is so important to reconsider the selection of stocks in the risky bundle.

According to the analysis of single index model, it’s very important to analyse the portfolio firstly. The assumption of SIMM is more limited. Using the plot to verify whether a linear regression exists is the key step to determine whether we can use this model.

In this case, we compare many indexes with our stocks. Although we use the best indexes we found. The tau-square is still very large. The large tau-square will make the portfolio’s variance very large. Therefore, the non-system risk, which is very large and cannot disperse, will make the model imprecisely.

In this case, under the strong assumption, we can still find the tangency portfolio. The portfolio return is about 17%. And the variance from the market part is 38477.961210278045. Because we have over 1400 data and the tau-square from the regression is large, the variance of the portfolio is higher. The finance stock has the worst performance. The stock from IT industry is the best choice for the investor the same as the Markowitz’s method.

# 8. Conclusion

We use Markowitz’s Efficient Frontier and Single Index Market Model to get the optimal portfolio. By comparing these two expected returns, we can find the advantages and disadvantages between two methods, thus making further predications more exactly.

# 9. References

[1] scipy.optimize.minimize

<https://docs.scipy.org/doc/scipy-0.18.1/reference/generated/scipy.optimize.minimize.html#scipy.optimize.minimize>