FINA4380 Algorithmic Trading Strategy Report

Group 10

**Abstract**

We adopt a long-only trading strategy that aims to construct a portfolio with higher Sharpe ratio than that of S&P500 portfolio.

To start with, we perform principal component analysis on the return of all the S&P500 constituent stocks to find explanatory factors. Then, we forecast the values and covariance matrix of the factors for the next time period. At the same time, we also calculate the expected value and covariance matrix for the next period. Using these data, we calculate the expected return of each stock and the expected covariance between each pair of stocks for the next time period. With this information, we choose the tangency portfolio with the highest Sharpe Ratio for the next period. At the end of each period, we repeat the process and rebalance our portfolio.

To test our trading strategy, we perform back testing on the period 2019-10-31 to 2021-10-31 and compare our performance with that of S&P500, using Sharpe Ratio.

1. **The stock return model**

We model the return of stock *i* at time *t* as:



where *PCit* is the *i*th principal component score obtained calculated using the *i*th principal component, *βit* are state variables that follow random walk, and we assume that Cov(*εit, εjt*) = 0 for all *i* ≠ *j*.

1. **Principal Component Analysis**

In order to reduce the amount of computation and reduce the noise in the estimation of the covariance matrix of stock returns, we perform principal component analysis on the return of all the S&P500 constituent stocks to reduce the dimension of the dataset and find explanatory factors.

We aim to retain the most information possible while keeping the number of components selected low. After testing, it is determined that we would select the first *n* components so that the proportion of variance explained exceeds 0.8.

The following shows an example of the 8 principal components extracted after performing PCA on stock returns for the period September 2019-August 2021:

A screenshot of a computer

Description automatically generated with low confidence

1. **Forecast of Explanatory factors—ARMA Model**

We use ARMA model to perform 1-period forecast of the principal component scores and use DCC-GARCH model to forecast the covariance between principal component scores.

It is assumed that:

Text, letter

Description automatically generated

And DCC is used to model the covariance between different factors.

For example, by using the PC values from September 2019 to August 2021, we make a prediction of the PC values of September 2021:

Chart, line chart

Description automatically generated

1. **Betas—Kalman Filter**

We assume that ***β****i* vary over time following the model:

Text, letter

Description automatically generated

And for simplicity we assume that Cov(*βit, βjt*) = **0** for all *i* ≠ *j*.

We obtain ***Q*** and the forecast of ***β****it* using Kalman filter. Before we apply Kalman filter, the unknown parameters are estimated using expectation maximization algorithm.

1. **Expected Value of and Covariance Between Stock Returns**

With the above information, the expected return of stock *i* can be calculated using the formula



And the covariance of stock returns can be found using the formula below:

Text, letter

Description automatically generated

1. **Sharpe Ratio Maximization**

After obtaining the expected returns of all stocks and the covariance matrix of stock returns, we are able to calculate the expected return and volatility of all portfolios.

To find the tangency portfolio, we construct portfolios that minimizes volatility given a fixed return under the constraint that the weights allocated to all stocks are between 0 and 1 (i.e., no shorting allowed), and among these portfolios, we select the one with the highest Sharpe ratio.

1. **Our Strategy**

We form our basket of stocks using S&P500 constituent stocks, which are large-cap stocks from different industries. We believe that the large-cap nature of the stocks would imply lower volatility and the basket would also be diversified as the stocks are from different industries. We rebalance our portfolio each month in order to strike a balance between lowering transaction cost and capturing the change in stock return trend. We adopt a long-only strategy to avoid the infinite downside risk brought by short selling.

At the end of each month, we would feed the 5-year monthly return data of each stock in the basket into our algorithm to generate the portfolio with highest Sharpe ratio. We would then hold this portfolio until the end of the following month, at which we would repeat the process and rebalance our portfolio.

To prevent any potential issues caused by long periods of missing data, we would not trade a certain stock if the stock has not started trading in exchanges at the start of the 5-year period.

1. **Back testing**

To test our strategy, we perform back testing on the period 2019-10-31 to 2021-10-31 with an initial NAV of USD$1,000,000. The results of the back test are as follows:

|  |  |  |
| --- | --- | --- |
|  | Our Portfolio | SPY |
| Cumulative return | +89.73% | +46.11% |
| Annualized STD | 32.38% | 17.98% |
| Annualized Sharpe ratio | 1.3097 | 1.1388 |
| Max drawdown | 19.0288 | 6.5185 |

Chart, line chart

Description automatically generated

Based on the annualized Sharpe ratios we can conclude that the portfolio constructed using our algorithm can outperform S&P500.