Application of principal component analysis in portfolio management

Li Qichao[1]

[1] The Hong Kong University of Science and Technology
qlich@connect.ust.hk

1 Introduction

Principal component analysis (PCA) is the process of computing the principal components of a dataset and using them to represent it, so as to reduce the dimensional complexity while minimizing information loss[1]. Also, we know that in finance, a portfolio with high risk diversification is considered to be a good portfolio in stock market because of its high risk resistance capacity, especially after the global financial crisis in 2008[2]. In this report, we try to use PCA to find the main components of the stock market (Australian market) to represent uncorrelated risk sources, and then use these results to construct four different portfolios with strategies including “Stock investment with equal weight (EWS)”, “Risk Parity (RE)”, “Equal Wight Portfolio (EWP)” and “Diversified Risk Parity (DRP)” to do the comparison.

2 Theory

We investigate the constituents of the ASX200 index from September 2003 to November 2021 kept in it for the whole study period. After cleaning the data, we finally get only 66 unique stocks because of the frequently adding and deleting of the stocks[3]. We use the closing prices P to calculate the return r. The PCA is then performed on the correlation matrix of the return series since using a covariance matrix will lead to the small number of principal components dominated by variables with large variance, which has a bad effect on diversifying the market risk.

Table 1. The 10 components with eigenvalue larger than 1

component	1	2	3	4	5
eigenvalue	19.41182993	4.141725196	2.559374381	1.802498243	1.46310926
component	6	7	8	9	10
eigenvalue	1.384663399	1.277892387	1.244037328	1.186395071	1.039942665

After getting the eigenvalues and eigenvectors of the correlation matrix, we start to choose the number of components to retain. We first determine the desired cumulative variance to be 80% to get the number of main parts, which requires 33 stocks to be retained, a quite large number. Then we combine Kaiser's rule, a scree graph and log-eigenvalue (LEV)

diagram and finally leave 10 components to represent the major risk sources in Australian stock market. The first component represents the risk of the while market, which has the nearly average weight for each stocks. The Bi-plot (Fig. 1) show the relative weights of each stock in components 1 to 10.

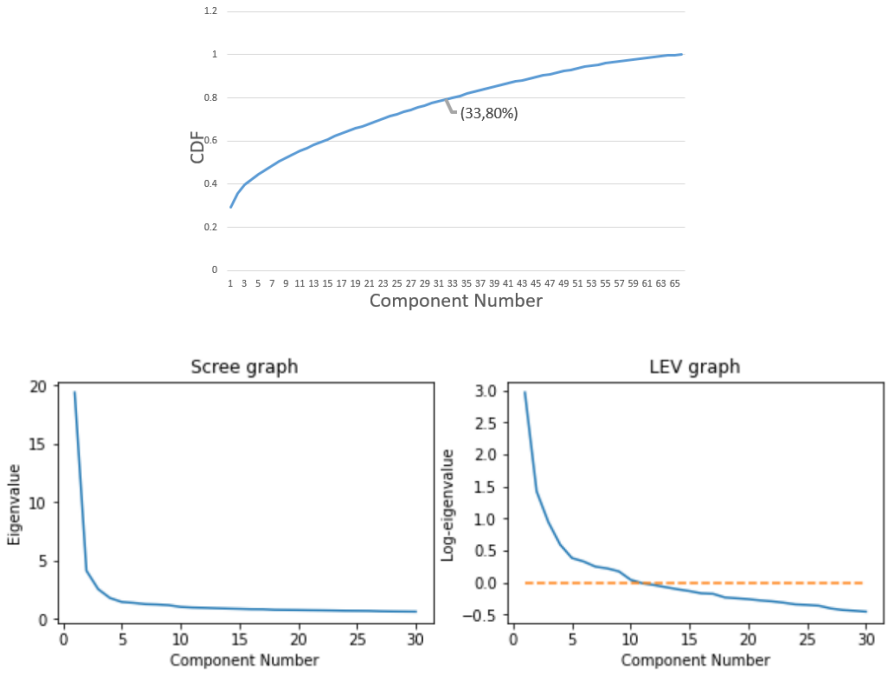

Fig. 2. (a) Percentage cumulative variance explained by principal components against component number (b) Scree graph (c) LEV diagram arising from PCA on a correlation matrix of 66 stocks

3 Principal Portfolios

In all of the following strategies, we consider the original wealth to be 1 and calculate the standard deviation, the minimum price, the final gain and Sharpe ratio for comparison based on the price of each portfolio at each time and its return. All these calculations are carried out on Excel.

3.1 Stock investment with equal weight

In this strategy, the capital allocated to each of the 10 stocks is the same, which is 1/66. Then volume of each stock can by directly calculated by the weight of this stock/initial price.

3.2 Risk Parity

In this strategy, the risk of stock i per unit of capital is estimated by its standard deviation σ_i, so the capital allocated to each stock is proportional to $1/\sigma_i$.

3.3 Equal Weight Portfolio

First 10 principal portfolios are selected and given equal weight to construct the portfolio.

3.4 Diversified Risk Parity

To diversify the risk, the capital allocated to each portfolio is proportional to $1/\sqrt{\lambda_a}$, considering that high eigenvalue represents high risk so as to invest less money.

4 Results

Table 2 shows the final gain, SD, and minimum of the four portfolios. We can find that the DRP portfolio shows the highest final gain as well as the highest SD, while RE has the lowest final gain with lowest SD. It seems that we need to take a higher risk in order to gain more. The Sharpe ratio of these portfolios are all less than 0, telling us these portfolios are underperforming the benchmark. The long-term holding of the stocks for nearly 19 years is likely to be the reason behind this situation.

Table 2. Final gain, SD and minimum of the four portfolios.

Strategies	EWS	RE	EWP	DRP
Final gain	52.90	33.71	218.97	413.67
SD	0.86	0.80	0.93	0.94
Minimum	0.92	0.92	0.86	0.91
Sharpe ratio	-0.10612	-0.13139	-0.08092	-0.07207

The index of Australian stock market keeps steady in long term as shown in Fig.3, which means that it is a mature market. In addition, there are two main periods that have a sharp decrease, corresponding to 2008 financial crisis and the COVID-19 epidemic. Fig. 4 shows the result of each portfolio at different time. The fact that our four portfolios all show a good return may result from the 66 stocks we first choose. Since the index keeps including these stocks, it is possible that most of these 66 stocks are good stocks, giving us good return no matter what strategy we choose.

Fig. 3. The performance trajectory of the ASX200 index

We also find the trend of EWP and DRP still follows the performance trajectory of the ASX200 index, decreasing sharply as the index does.

However, what we expect is it to be steady instead of fluctuating so much. This may because the ASX200 index is no longer a weighted index of these 66 stocks we first choose since we have delete too many stocks from the market.

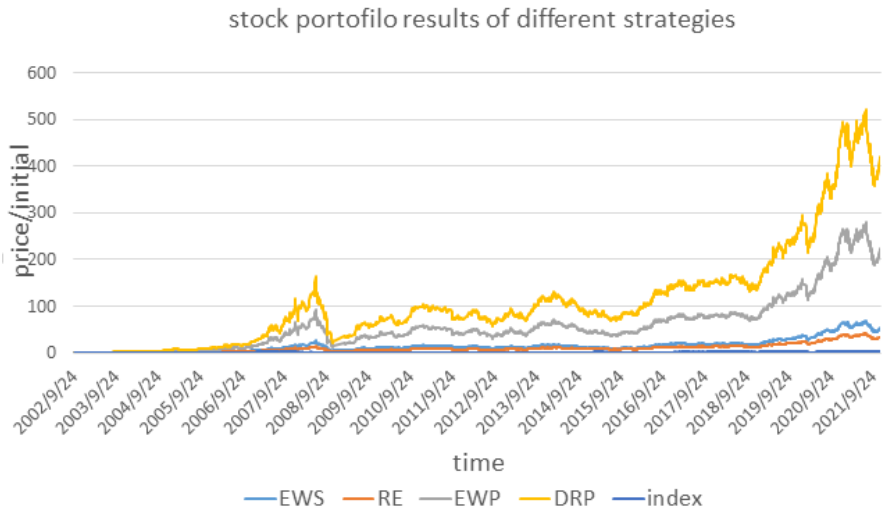

Fig. 4. The price of each portfolio at different time

5 Conclusion

10 components can represent these 66 selected stocks according to the Kaiser's rule. Because of the long time period, the Sharpe ratio is less than 0 for each portfolio, which means that the risk we take is higher than

the value we receive. So, for sound investors, it is not suggested to hold a stock for too long time.

The ASX200 index can no longer represent the market consisting of these 66 stocks due to the lack of too many stocks. So, to really find the effect of PCA in risk diversification. It is more reasonable to reduce the lasting time so as to consider enough stocks. Or you can just create a new index by yourself to reflect the 66-stocks market and compare your portfolio return with the new index to test the effect. These ideas can be considered for further study in future.

6 Appendix

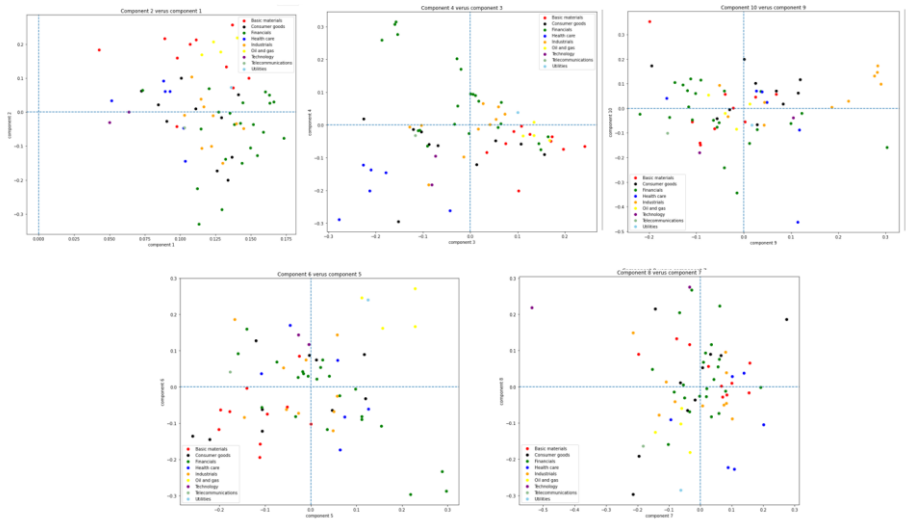

Fig. 1. Bi-plots of relative weights of each stock in components 1 to 10.

References

1. Jolliffe, I. T. (2002). Principal Component Analysis. Springer Series in Statistics. New York: Springer-Verlag. doi:10.1007/b98835. ISBN 978-0-387-95442-4.
2. Lowenfeld, H. (1909). Investment, an exact science. London: The Financial review of reviews.
3. Yang L. An Application of Principal Component Analysis to Stock Portfolio Management: A Thesis Submitted in Partial Fulfilment of the Requirement for the Degree of Master of Commerce in Finance, Department of Economics and Finance, University of Canterbury[M]. University of Canterbury,2015.