Supervised learning task 2

Lin Zhao (16010906)

February 27, 2018

Introduction

Markovitz perfolio theory has always be a classic. However it is known that the optimal portfolio is not stable bringing issue on frequently rebalancing for investors. Following the lead of Fastrich, Paterlini, and Winker (Constructing optimal sparse portfolios using regularization methods, 2014), this study try to find global minimum variance portfolio under regularizations thus stablize the portfolio choice and reducing the rebalancing problem.

The regularizations used in this study are Lasso and Ridge. Resulting global minimum variance portfolios has lower variance than Markovitz portfolios and therefore confirmed the results from paper montioned above.

Data description and cleaning

The data used is French 48 industry portfolios daily data. After comparing variance of each industry and of the whole dataset, I choose to use equal weighted data since it has higher variance and hopefully the different among regularizations will be more obvious. After deleting any row that contains missing value, I have about 44 years data left which is a good amount.

In figure 1, I plotted standard diviation of all 48 industies and we can see that except for a few outliars, most industries standard diviation of returns lies in a interval between 0.75 to 1.5.

Figure 1: std of 48 industires show that except for a few, all industries has std with a steady interval

Next I plottd mean returns for each of 48 industies together with minimum and maxmum returns of each industry. The maxmum returns seem to has more fluctuation than the minimum.

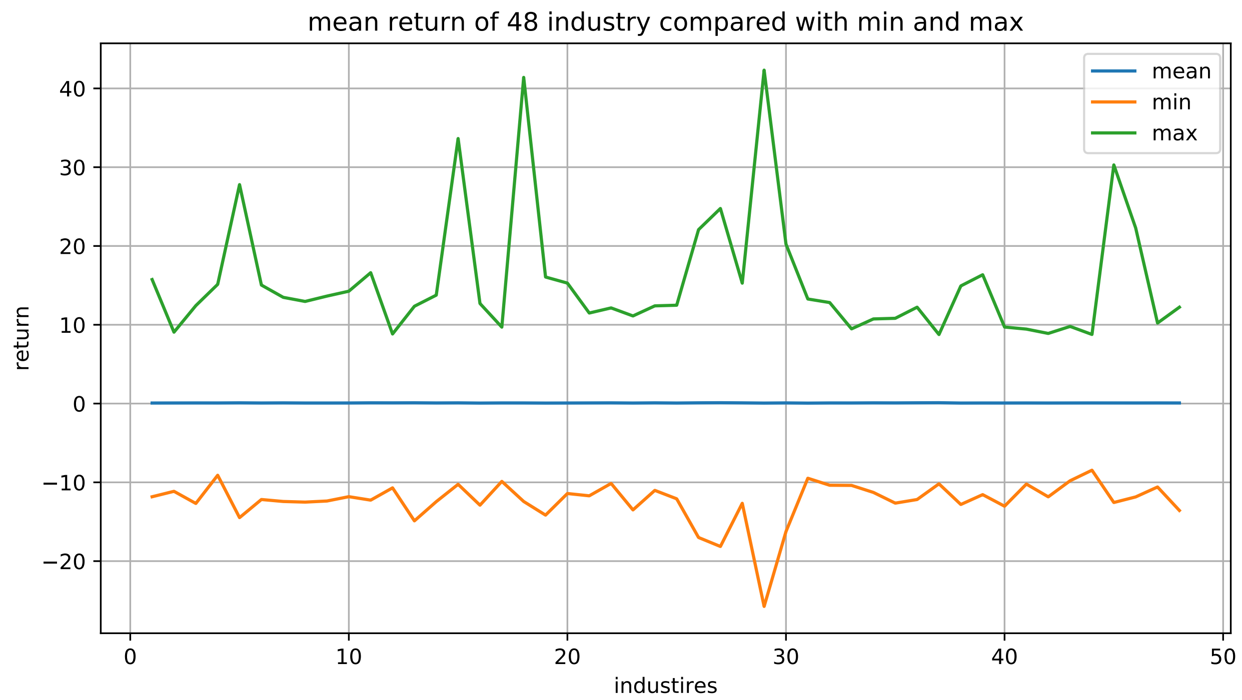

Figure 2: mean, minimum and maxmum returns of each industry

The industry that has highest variance is the coal and utililies has the lowest variance. Here is a plot to demonstrate that.

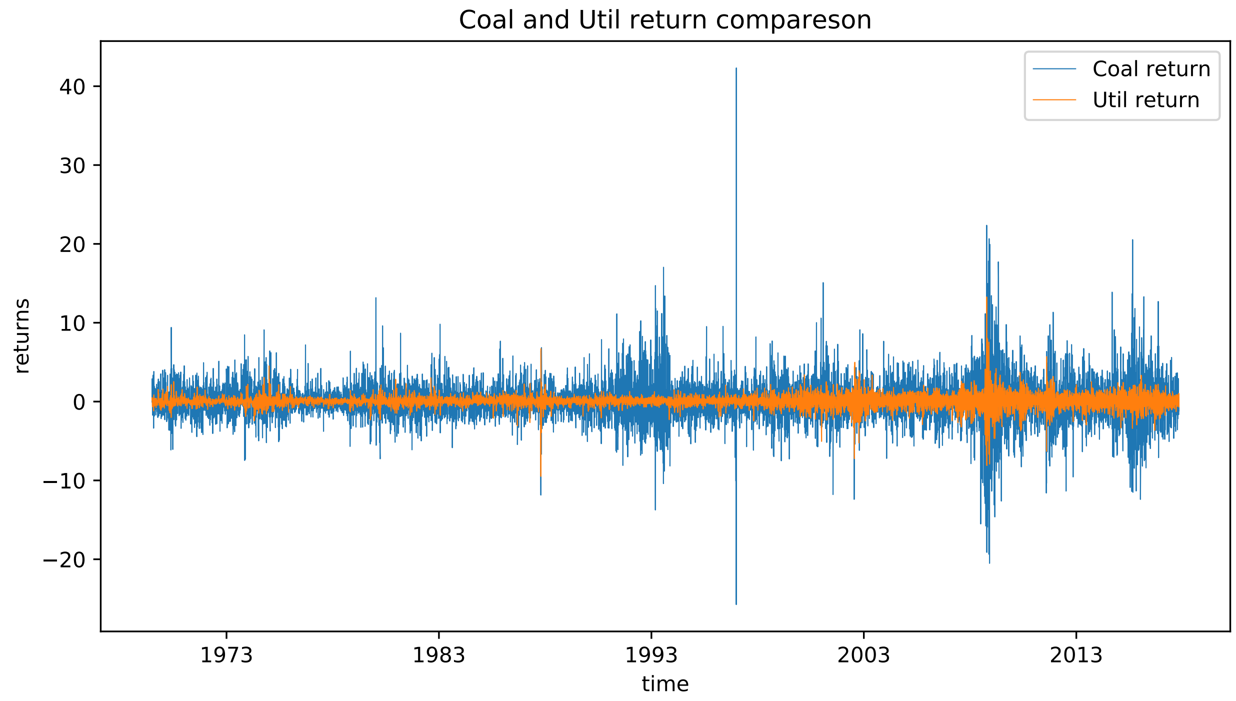

Figure 3: Variances of coal industry and utility industry.

Global minimus variance portfolio with different regulariztions

Three methods to select portfolios

In this study, I used three methods for selecting minimum variance portfolios. The frist one is Markovitz portfolio selection. This is a problem minimize variance of the portfolio under the constraint that sum of all the weights equals to 1. I also performed selection for long-only portfolios to use together with long-short portfolio as benchmarks. To find the long-only portfolios, the minimize problem need a extra constraint that all the weights greater or equal to 0. The second method I used is a Markovitz under Lasso regularization. This is a minimizing problem that minimizing sum of variance and a 1-norm penalty:

$$\mathbf{w}^{\mathbf{T}} \sum \mathbf{w} + \lambda \sum_{n=1}^{K} |w_n| \,,$$

where $\mathbf{w}$ is weight vector, $\mathbf{\Sigma}$ is the covariance matrix and K is the number of stocks in your investing universe. Parameter λ controls how much penalty to add. To find a portfolio, one need to solve this minimizing problem under the same constrain that sum of weights equals to one. The other method in this report is a Markovitz under Ridge regularization. In this minimizing problem, the target function is a sum of variance and 2-norm penalty:

$$\mathbf{w}^{\mathbf{T}} \sum \mathbf{w} + \lambda \sum_{n=1}^{K} |w_n|^2 \,,$$

and solving this under same budget constraint give a minimum variance portfolio under Ridge regularization.

Selection λ with cross validation

If λ equals to 0, both regularized method collaps to classic Markovitz. On the other hand, if λ is too large, which means the penalty term in the target function will dominate the variance term and thus impose to much bias. To select optimal λ, I performed a cross validation as did in the Fastrich paper. This cross validation is only meaningful under stationary assumption. First, I shuffle the whole training set and divide it into 10 equal folds. In each round, I take one of these 10 folds as test set and the rest 9 folds as training set. Then, in each round, for each value in a series of λ that I want to test, solve the minimizing problem under relative constraint and thus find a portfolio. Hold this portfolio in the test set and recieve a return series. Calculate the standard diviation of this return series since our aim is to find minimum variance portfolio. After 10 round, one will have 10 standard diviation for each λ, and taking mean of these gives the score of that λ. The optimal λ is the one with the lowest score. And a good searching interval should give a smile shape curve when plot all the score against λ. If the optimal λ is on the edge of searching interval, a change of interval is required to find a true minimum score. In the next to figure, I plotted scores against λ for a Lasso penalty and a Ridge penalty.

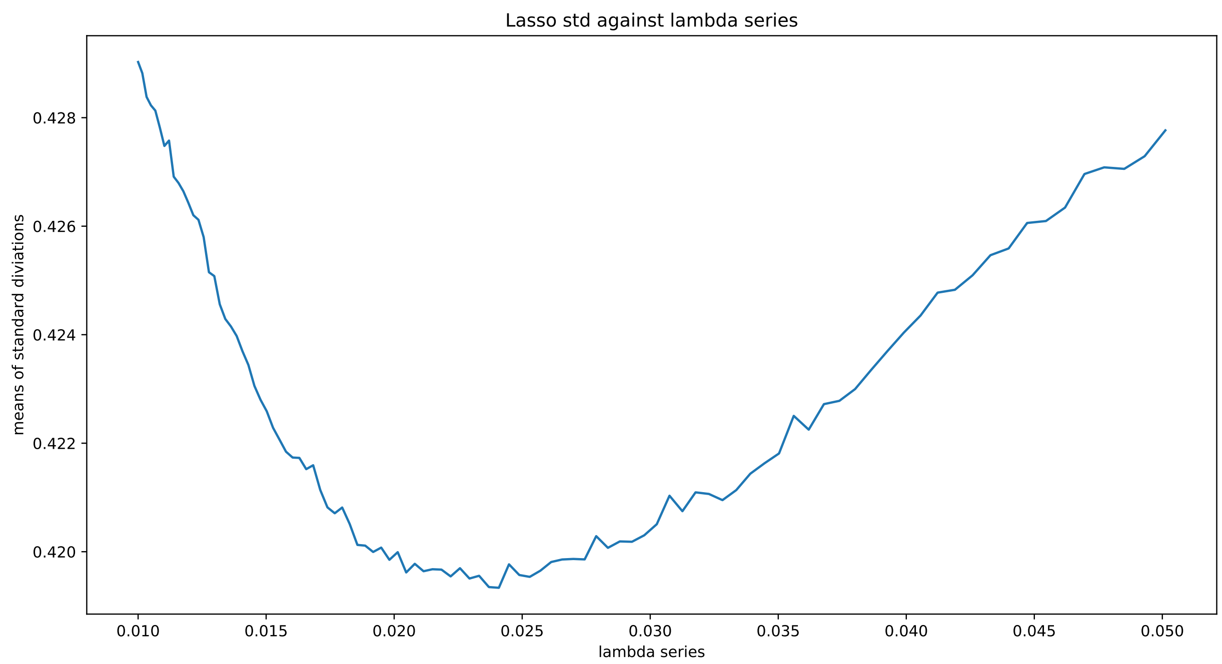

Figure 4: portfolio std with Lasso regularizaton. Training period 1 year. Searching interval 10^{-2} to $10^{-1.3}$ and searching steps 100.

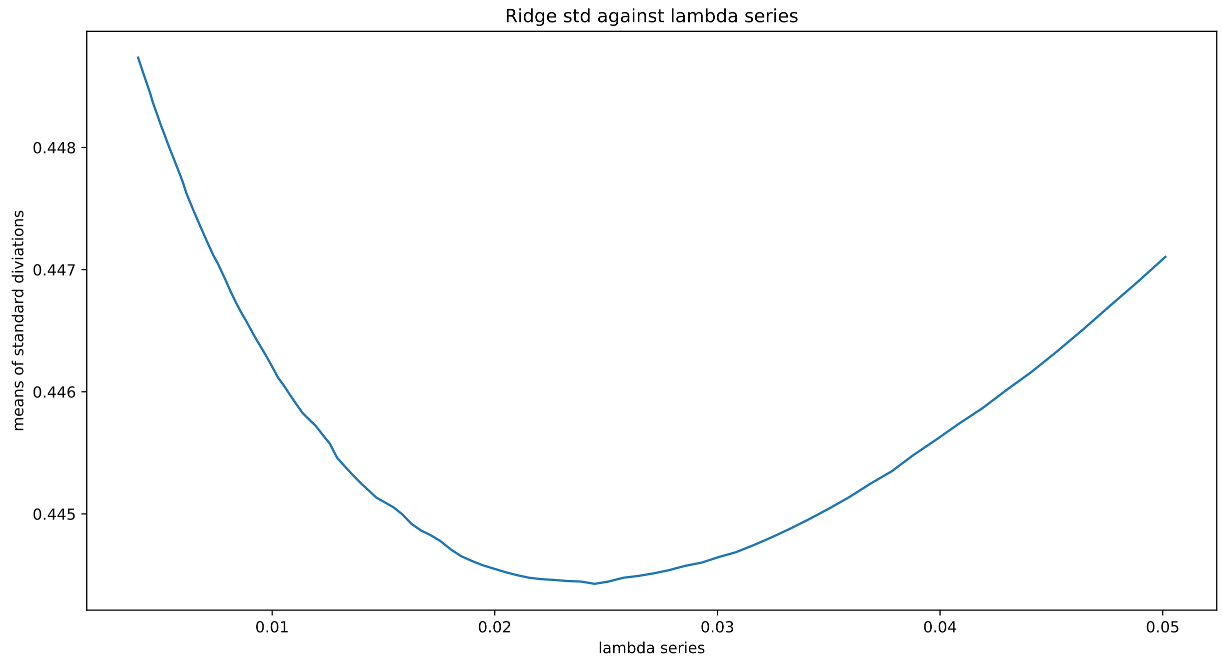

Figure 5: portfolio std with Ridge regularizaton. Training period 1 year. Searching interval 10^{-2} to $10^{-1.3}$ and searching steps 100.

Since there is randomness involved in the shuffling step, to find the reasonable searching interval, I run cross validaton for both regularizatons multiple times to make sure all the optimal values are with in the interval. When using on the whold data set, if the optimal solution is on the edge, the program will through a warning, shift both boundaries of interval to the coresponding direction, and perform another search.

Out-of-sample test

Out-of-sample tests are performed on the data with vary training windows and holding periods. In the plot below is return series trained on a one year window and hold for 3 months.

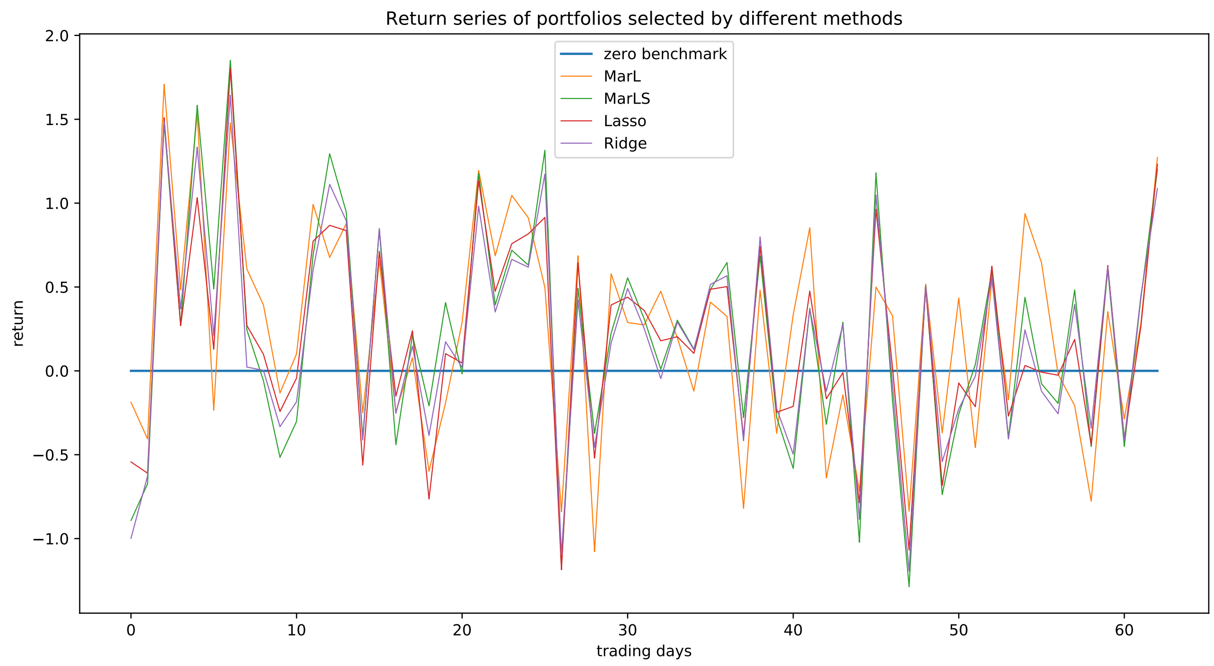

Figure 6: return series trained on 250 days and hold for 63 days. Cross validation interval, Lasso: $10^{-2.1}$ to $10^{-1.2}$, Ridge: $10^{-2.3}$ to $10^{-1.3}$. Searching steps: 100 for both.

The variance of portfolios are 4.5392, 3.9004, 3.6182 and 3.3498 in percentage respectly for Markovitz long-short, Markovitz long-only, Lasso, and Ridge. As expected, portfolio found under regularizations have lower variance during the holding period. This is due to the barrier set by penalty term blocks the noise in the data, thus in the variance covariance matrix, to enter the results freely. A small surprise here is that with more constraint, the long-only portfoio has lower variance than the long-short portfolio. This is not consistant across all the back tests I performed, but it indicates the instability of Markovitz solution.

Next, I performed a series of out-of-sampel tests with moving windows. I tested different training window to see whether having longer training window affects any methods. I also test different holding periods to see whether some methods perform better when the selected portfolios were rebalanced more often. Since Fastrich paper only used 5 to 6 years of data, and considering runing time, I dicided to run these tests on the last 5 years from the whole data set.

Out-of-sample test result

Training window (days)	Holding window (days)	window number	Marlovitz long-short	Markovitz long-only	Lasso	Ridge
500	63	11	3.5831	4.3738	3.4346	3.4377
250	63	15	3.4866	4.3458	3.2998	3.2627
125	63	17	4.1919	4.3288	3.4325	3.4712
250	125	8	3.9520	4.5060	3.6145	3.7762
250	63	15	3.4866	4.3459	3.2840	3.2682
250	21	47	3.3769	4.1968	3.1406	3.1526

Table 1: Out-of-sample test with moving windows. Data use are the last 1250 observations of the data set

From the whole table, we can see that the result is robust to the change of training table and holding period. In all the tests cross various training period and holding length, the moving window out-of-sample tests denmonstrate that regularized portfolio constently out perform the Markovitz portfolio. Among two bench marks, long-only protfolio has higher variance because it cannot take adventage of shorting thus has less ability to diverge the risk. Between Lasso and Ridge, there is no clear result about which one provide portfolios with lower variance. The first three row of the table seems to indicate that shorter training period make the regularized portfolio out perform even more. This is not enough to evidence to conclude, but one potential reason could be that in a short training period, the true information in the covariance matrix is more mingled with noise and thus the Markovitz method has more difficulty to select real optimal portfolio. And because of the penalty term, regularized methods has a more restrict standard for only the significant information to pass through, thus allow them to out perform the Markovitz even more. For the changing of holding period, there is no clear result about whether a certain length of holding benifit a particular method.

At last, I performed a out-of-sample test on the whole 44 years data. Resulting variance are 2.1290, 3.5056, 2.0711, and 2.1074 for Markovitz long-short, Markovitz long-only, Lasso, and Ridge. As usual, portfolios under regulirizations out performed both benchmarks. The order of variances for the four portfolios and magnitude of differences match the results of Fastrich, Paterlini and Winker paper.

Both Fastrich paper and Brodie paper mentioned sparseness as another feature as portfolio chosen by regularized Markovitz methods, however, set threshold at 0, neither set of portfolios selected under regularizations show this feature. I did not expore more from here but it might related to the solver used. Also, a position with weight less than 10^{-2} is practically non-active, so the sparse feature can also rely on the choice of threshold.

Conclusion and potential further questions to answer

Observed global minimum variance portfolios found under different regularizations, I confirmed that regularized protfolios out perform both Markovitz benchmarks in the out-of-sample tests thus found optimal portfolios has lower variance in various the holdidng periods. Markovitz solutions are not stable and contains noice from the data so regularization provides practical benefit to investors. Another benefit of regularizations proposed by both Fastrich paper and Brodie paper is it gives sparse portfolios thus reduce the transition cost. This is not observed here and can be a interesting further topic.

References

[1] B.Fastrich, S.Paterlini & P.Winker (2014) *Constructing optimal sparse portfolios using regularization methods.*

[2] J.Brodie et al. (2009) *Sparse and stable Markovitz portfolios.*