Portfolio Optimisation on 48 Industries

Mouktar ABDILLAHI

1 Introduction

Modern portfolio theory (MPT) provides investors with a portfolio construction framework that, under certain assumptions, achieves maximum returns for a given level of risk, through diversification. The Markowitz mean-variance portfolio theory states that in order to choose the optimal portfolio weights, an efficient tradeoff between returns modeled here as the mean and risk measured here through the variance-covariance matrix. Risk and returns are estimated empirically on a given timescale and the instability of these estimations could lead the investor to build a unefficient portfolio.

The following paper will focus on the global minimum variance to decide of the optimal portfolio (and not the maximum Sharpe ratio) and we will use penalty regularization on training data to build robust estimations of the assets' weights.

In this report, first, we will briefly explain the data as a whole and will present graphs and tables that describe the dataset. Then, we will tackle the methodology for calculating and penalizing global minimum variance portfolios. In the third part, we will present and discuss the results. Finally, the last section will focus on the 2008 Financial crisis and we are using values since 2001-2007 to train the dataset and the 2007-2008 financial crisis as testing dataset. We chose this interval to look at how a portfolio made before a financial crisis would work during the financial crisis.

2 Exploratory Data Analysis

The initial dataset consists of 48,198 daily returns for 48 industry asset groups. These assets include all listed in the 48_Industry_Portfolios_daily and the link to the dataset is available in the bibliography, it includes industries such as agriculture, healthcare, oil, guns, gold, etc.

The daily returns are from 1^{st} July 1926 to 31^{st} October 2017, however, I have noticed that every day is reported twice. The difference between the two is that in one case the industrial index is computed averaging over the stocks of that industry with equal weight, in the other case stocks are weighted by market capitalization. As we do not know which version of the day corresponds to which method of computing, we have decided to only keep the first appearance of every date and that left us with only 24,099 rows.

For the first analysis done on this dataset, we wanted to work on the whole timescale available, so we had to eliminate 8 industries out of the 48 because they contained a lot of missing values, these eight industries (soda, rubber, health, fabric products, guns, gold, personal service and paper) will be used along the 40 others in the second study of this paper that will focus on the 2008 financial crisis.

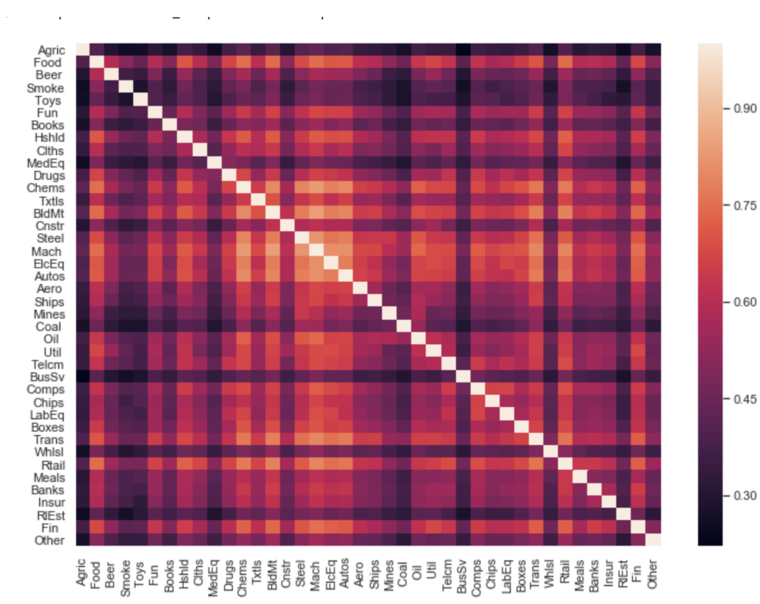


Figure 1: Correlation Heatmap

To minimize the portfolio risk, the portfolio built has to be well-diversified. Hence, you need to work upon the correlation of the different industries. Correlation can be used to identify the degree and type of relationships between assets and when choosing industries, you have to choose investments that are poorly correlated or negatively correlated. The Figure 1 is a heatmap that presents the correlation within the 40 industries for the whole timescale.

3 Methodology

In finance, people are mostly familiar with the concept of 'There ain't no such thinh as a free lunch', hence, higher risk is mostly associated with greater probability of higher return and vice-versa. Therefore, we have to understand the concept of risk-adjusted return that refines a portfolio's return by measuring the additionnal risk necessary when producing that return. Generally, the Sharpe-Ratio (or the Adjusted Sharpe-Ratio are used to expressed risk-adjusted return. However, in the case of this study, the focus is on finding the minimum risk portfolio.

How is the portfolio's volatility calculated? Let's take the case of a 2 assets portfolio, you will see in Figure 2, the matrix notation of the calculation done with ω_1 and ω_2, the weights for the two assets and the 2×2 matrix is the variance-covariance matrix.

[1] In Python, the sub-package scipy.optimize has a function called 'minimize' that I could use to choose the weights that would lead to the portfolio with the global minimum variance. Nevertheless, I have decided to use a more empirical method that would help us have a better visualization of what we were doing. In the first study, we have 40 industries and we have to decide what weight to allocate to each industry and in order to do so, we have randomly generated portfolios (25000 portfolios here) and for each one of them we have calculated the return and the risk associated to it.

$$\sigma_p^2 = \begin{bmatrix} w_1 & w_2 \end{bmatrix} \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_2^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} w_1\sigma_1^2 + w_2\sigma_{2,1} & w_1\sigma_{1,2} + w_2\sigma_2^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$
$$= w_1^2\sigma_1^2 + w_1w_2\sigma_{2,1} + w_1w_2\sigma_{1,2} + w_2^2\sigma_2^2$$
$$= w_1^2\sigma_1^2 + 2w_1w_2\sigma_{1,2} + w_2^2\sigma_2^2$$

Figure 2: Portfolio's volatility

Once the portfolio with the global minimum variance is defined, it has to be tested out-of-sample. Hence, we divided the dataset in a training dataset with 80% of the dates (from 1^{st} July 1926 to 4^{th} September 1998) and a testing dataset with the remaining 20% dates (from 5^{th} September 1998 to 31^{st} October 2017). Similarly to what one would do for cross-validation, the volatility of the optimal portfolio in the test set need to be low and the lowest the better.

[2] Once this is explained, the focus of this study is to try different regularization techniques and see if they improve our results (if the volatility is lower than the volatility of the optimal portfolio in the test set without any regularization technique). Then, additional constraints need to be added to the Markowitz portfolio theory and these constraintes are represented by a regularization parameter λ that goes along with a penalty function ρ (we have used LASSO, Ridge, Elastic Net, Lq and SCAD in this study and these functions are expressed in Figure 3). The optimal asset weights ω^* has to be adapted and is defined here by : $\omega^* = \text{argmin}\{\omega^T\overline{\Sigma}\omega + \lambda\sum_{i=1}^{n}\rho(\omega_i)\}$ with $\overline{\Sigma}$ being the variance-covariance matrix.

Penalty $\boldsymbol{\lambda\rho}(\omega_i)$	**Domain**
LASSO = $\boldsymbol{\lambda}\lvert\omega_i\rvert$	All
SCAD = $\boldsymbol{\lambda\lvert\omega_i\rvert}$	$\lvert\omega_i\rvert < \lambda$
$\frac{-\lvert\omega_i\rvert^2 + 2a\lambda\lvert\omega_i\rvert - \lambda^2}{2(a-1)}$	$\lambda < \lvert\omega_i\rvert < a\lambda$
$\frac{(a+1)\lambda^2}{2}$	$a\lambda < \lvert\omega_i\rvert$
$\mathbf{L_q}$ = $\boldsymbol{\lambda\lvert\omega_i\rvert^q}$, $\mathbf{0 < q < 1}$	All
Elastic Net = $\boldsymbol{\lambda(\frac{1-\alpha}{2}\lvert\omega_i\rvert^2 + \alpha\lvert\omega_i\rvert)}$	All
MC+ =	

Figure 3: Regularization Penalty Functions (Ridge is L_2)

To find the optimal value of lambda, we did a GridSearch and that means that for a given selection of lambda (from 0 to 1 with a jump of 0.01 between each value), we compute the volatility of the portfolio on the test set and choose the lambda that leads to minimum volatility. For certain regularization techniques such as SCAD, Lq and Elastic Net, the penalty function had another parameter and to optimize it too, we did the same process as for Ridge and Lasso but with a two parameters GridSearch.

4 Results

We have decided to allow short-selling so the weights in the following table can be negative. In Table 1, the weights for each allocation are represented and the value of lambda is the one that minimizes the volatility of the test set.

Asset	**No penalty**	**LASSO**	**Ridge**	**Elastic Net**	**Lq**	**SCAD**
Lambda	xxxxxx	0.33	0.74	0.72	0.24	0.93
Agric	3.91	8.81	4.05	7.13	2.64	6.47
Food	11.05	13.43	7.58	10.52	8.58	5.81
Beer	8.49	4.84	-3.4	-1.13	13.39	7.11
Smoke	3.58	1.31	8.02	6.12	11.08	2.57
Toys	-3.81	-0.92	-7.07	2.59	7.45	-9.56
Fun	-7.73	-10.07	7.17	-7.39	-8.21	0.
Books	8.65	-7.38	13.66	0.01	0.17	7.8
Hshld	-0.92	-4.93	-0.4	9.79	9.05	17.77
Clths	9.52	9.9	11.35	8.02	12.08	6.95
MedEq	11.08	5.34	8.54	7.87	13.85	5.56
Drugs	-1.95	6.19	-0.39	0.6	2.54	-4.87
Chems	-6.23	12.64	12.12	2.62	4.66	11.48
Txtls	9.31	4.68	11.38	7.95	11.26	8.48
BldMt	2.13	12.11	11.1	5.33	3.26	7.88
Cnstr	-2.22	-3.46	3.2	-10.35	-5.89	-10.13
Steel	7.48	-0.86	-10.96	-8.03	-13.74	-7.06
Mach	-0.52	-3.11	4.13	8.98	-1.14	-5.35
ElcEq	-4.17	-7.69	-5.67	-4.14	-11.76	-6.53
Autos	1.09	-9.68	-11.05	-8.22	2.09	-17.97
Aero	-3.71	-4.44	-8.35	-9.06	-3.36	-7.44
Ships	4.89	13.7	-12.68	0.68	5.74	7.28
Mines	-0.8	11.06	9.18	6.34	8.67	17.21
Coal	2.65	-3.19	5.02	2.08	-7.6	-11.42
Oil	10.7	-3.44	-4.64	10.70	7.69	6.15
Util	8.04	10.85	-1.67	4.65	12.75	16.39
Telcm	9.12	13.11	6.4	5.16	5.73	12.88
BusSv	3.85	-0.04	-0.71	-6.98	-9.21	16.31
Comps	-3.87	-3.88	-10.47	2.26	-3.43	-5.76
Chips	-2.85	-4.82	1.85	-3.12	-14.	-10.38
LabEq	10.96	3.81	9.54	4.72	11.92	-11.46
Boxes	-7.46	14.42	7.72	-1.81	-7.66	4.7
Trans	3.28	-13.9	5.5	10.11	-11.55	-13.39
Whlsl	2.6	7.37	-10.16	2.04	5.82	-0.56
Rtail	-0.52	8.02	9.28	9.60	10.91	5.98
Meals	-3.64	2.42	9.86	-1.28	11.85	8.93
Banks	-3.18	2.88	5.53	5.34	14.18	11.1
Insur	6.76	11.02	3.44	1.50	7.66	16.23
RlEst	-2.26	-5.94	-7.51	0.13	0.35	-9.52
Fin	5.42	5.74	13.61	9.38	-3.08	8.67
Other	11.27	4.14	5.9	9.28	-4.75	11.66

Table 1: Global Minimum Variance Portfolio weights allocation - Comparison of regularization techniques

For Lq, the best fit was with q = 0.6, for SCAD we have a = 0.8 and for Elastic Net, we have $\alpha = 0.6$.

As we talked about in the Methodology section, the risk-adjusted return is the focus of most investors as they want to make the largest return with the tiniest risk. In the following table, we will be comparing the various classical and regularization global minimum variance portfolios to the optimal portfolio trained on the test set.

Optimization method	**Out-of-sample volatility**	**Sharpe Ratio**
Optimal Portfolio - Test set	66.80	5.645
No penalty	112.72	2.759
Lasso	106.06	3.288
Ridge	123.52	2.875
Elastic Net	107.26	3.023
Lq	95.62	3.568
SCAD	103.22	3.401

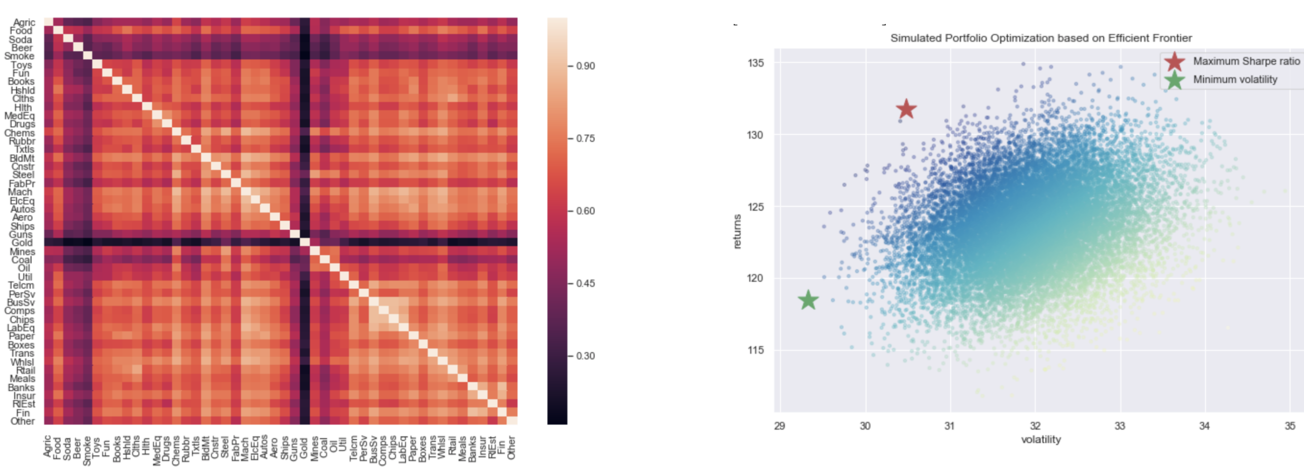
Table 2: Comparison of volatility and sharpe ratios of the various calculated global minimum variance portfolios

Generally, a Sharpe ratio between 1 and 3 is very good for an investment. However, it has to be compared to the benchmark performance of the optimal portfolios, as we did here. The benchmark performance is very good here, the out-of-sample volatility is very low compared to the others and the sharpe ratio is excellent. Overall, during the time-period, it appears that all the regularization techniques improve the Sharpe Ratio when compared to the 'No penalty' benchmark. However, the Ridge technique has the highest out-of-sample volatility so we should avoid this technique. The Lq method stablish the best portfolio allocation for the given data both in terms of Sharpe Ratio and Variance.

5 Focus on the 2008 Financial Crisis

For the second part of the study, we will be studying the financial crisis of 2007-2008. We are using values since 2001-2007 to train the dataset and the 2007-2008 financial crisis as testing dataset. The idea here is to study an extreme case. [3] Indeed, the global financial crisis of 2007-2008 is considered by many economists to have been the most serious financial crisis since the Great Depression of the 1930s. Hence, if an investor was working with the same dataset in early 2007 and wanted to build a very diversified portfolio. He has no idea that the following two years are going to be a nightmare and we want to see if any regularization technique could maybe improve the weights allocation to reduce his loss and the risk he will be facing.

For this study, we will be using the 48 industries available in the dataset and as for the previous study, the correlation between the industries for this particular time period has a great importance. So, in the following figure you will be finding the correlation matrix in a heatmap form and an example of the random portfolio generalization we had for this study (it is a generic example without any regularization method).



(a) Correlation Heatmap

(b) Random Portfolio Generation

Figure 4: Few plots

For the random portfolio generated in Figure 4, the lightest the blue and the lowest is the Sharpe Ratio.

In the following table, we will be comparing the various classical and regularization global minimum variance portfolios to the optimal portfolio trained on the test set.

Optimization method	Volatility (1)	Volatility (2)	Sharpe Ratio
No Penalty	25.47	169.85	-0.1003
Optimal Portfolio - Test set	xxxxx	120.82	0.0801
Lasso ($\lambda = 0.48$)	26.85	183.49	-0.187
Ridge ($\lambda = 0.81$)	25.84	156.82	-0.232
Elastic Net ($\lambda =$, $\alpha = 0.5$)	26.1	162.12	-0.176
Lq ($\lambda = 0.54$, $q = 0.8$)	26.85	138.12	-0.028
SCAD ($\lambda = 0.09$, $a = 0.2$)	25.36	168.54	-0.249

Table 3: Comparison of volatility and sharpe ratios of the various calculated global minimum variance portfolios

In Table 3, Volatility(1) is the volatility computed on the training dataset and Volatility(2) is the volatility computed on the testing dataset. You will notice that there is no value for Volatility(1) for 'Optimal Portfolio - Test set' because that row is the optimal portfolio for the financial crisis period (hence the testing dataset, as we do not work backwards and apply model of the future to the past). The Sharpe Ratio is computed for the testing dataset each time.

Overall, we can say that the financial crisis is very harsh as even the optimal portfolio of the test set has a very low Sharpe Ratio (slightly higher than 0 so a very weak ratio). All the other Sharpe Ratio are negative and that shows that this period of our recent history was one of the worst for any investor in any industry. During the time-period, it appears that for the Sharpe ratio, only the Lq method improves it when compared to the 'No Penalty' benchmark. The Lasso method is the worst one here as its score both in term of Sharpe Ratio and out-of-sample volatility is not good and even worse than without any regularization technique. For SCAD, Ridge and Elastic Net, we can appreciate the improvement of the out-of-sample volatility but the Sharpe Ratio is not good at all. To sum up, the Lq method establishes the best portfolio allocation for the given data both in terms of Sharpe Ratio and Variance.

6 Conclusion

In conclusion, we have seen that the modern portfolio theory of Markowitz can get a lot of benefit using the regularization techniques such as Lasso, Ridge or Elastic Net. These regularization techniques are widely used in Supervised Machine Learning when overfitting happens. [4] Overfitting happens when model learns signal as well as noise in the training data and wouldnt perform well on new data on which model wasnt trained on and the same problem can be faced in portfolio optimization. We have seen in the first study that these penalty functions can really improve the returns of an investment when tested in real situations and as any rational investor is risk averse, using these techniques can really reduce the risk he will be facing with his portfolio. For the second study, we have seen that these techniques are really helpful in extreme cases such as the 2007-2008 Financial crisis as they also increase the money you will be saving when the economical conditions are very bad. However, it is not a miracle recipe as the way the best lambda (and also the other parameters) is quite empirical.

References

[1] Ricky Kim, *Efficient Frontier Portfolio Optimisation in Python*, available on Medium at `https://towardsdatascience.com/efficient-frontier-portfolio-optimisation-in-python-e7844051e7f`

[2] Alexandra Malindi Colyer from Oxford University, *Markowitz Mean-Variance Portfolio Optimization Using Regularization Techniques*, available on Academia.edu at `https://www.academia.edu/36207052/Markowitz_Mean-Variance_Portfolio_Optimization_Using_Regularization_Techniques?auto=download`

[3] *Financial crisis of 2007-2008*, available on Wikipedia at
`https://en.wikipedia.org/wiki/Financial_crisis_of_2007-2008`

[4] Anuja Nagpal, *Over-fitting and Regularization*, available on Medium at
`https://towardsdatascience.com/over-fitting-and-regularization-64d16100f45c`