MMF1921 - Operations Research

Mean-Variance Portfolio Optimization under Different Factor Models

Chenglan Huang
Tianyi Long
Yuzhe Zhang

May 12, 2021

Contents

1 Introduction

This report discusses portfolio optimization under four different factor models, namely the OLS regression on all eight factors (OLS model), Fama-French three-factor model (FF model), Least absolute shrinkage and selection operator model (LASSO model), and Best subset selection model (BSS model). For each model respectively, the expected returns and covariance matrix are estimated, and are fed into the optimization as inputs. We then use Mean-Variance Optimization to determine the optimal weights for the portfolio.

Our investment universe consists of 20 assets and we will utilize the corresponding monthly adjusted closing prices from December-31-2005 to December-31-2016 to compute our observed asset monthly returns. Adjusted close price is a more accurate measure of stock performance, since it accounts for corporate actions such as stock split and dividend payments, which would normally cause discontinuity in the close prices.

Furthermore, we will demonstrate in detail about the methodology of obtaining the estimates under each of the four factor models, as well as implementation of the Mean-Variance Optimization strategy.

In Section 3, we present the In-sample and Out-of-sample analysis to evaluate and compare the resulting portfolios from these four models.

2 Methodology

2.1 Factor Models

Strategies to model security returns as the linear combination of factors are called factor investments. Factors could be technical, fundamental, macroeconomic or alternate to define a security's performance. By selecting appropriate factors, investors can decompose target assets risk and return into factor risk and return.

A generic multi-factor model with p factors and n assets for excess returns takes the form below:

$$r_i - r_f = \alpha_i + \sum_{k=1}^{p} \beta_{ik} f_k + \epsilon_i$$

where r_i is the return of asset i, r_f is the risk-free rate, α_i is the intercept from regression, f_k is the return of factor k, β_{ik} is the exposure term that dictates how much the the return on asset i moves with factor k, and ϵ is the residual of asset i measuring idiosyncratic risk.

The eight factors considered in this report are

Factor	Mkt_RF	SMB	HML	RMW	CMA	Mom	ST_Rev	LT_Rev
Description	Market	Size	Value	Profitability	Investment	Momentum	Short-term reversal	Long-term reversal

And tickers of 20 stocks are

F	CAT	DIS	MCD	KO	PEP	WMT	C	WFC	JPM
AAPL	IBM	PFE	JNJ	XOM	MRO	ED	T	VZ	NEM

2.1.1 OLS Model

The OLS model incorporates all eight aforementioned factors. Coefficients of regressors are chosen such that the sum of squared residuals are minimized. It is an unconstrained minimization problem,

$$\min_{\boldsymbol{B}_i} \|\boldsymbol{r}_i - \boldsymbol{X}\boldsymbol{B}_i\|_2^2$$

where $\|\cdot\|$ is the ℓ_2 norm operator,

$\boldsymbol{X} = [\boldsymbol{1}\boldsymbol{f}]$ is the data matrix,

$\boldsymbol{B}_i = \begin{bmatrix} \alpha_i \\ \boldsymbol{V}_i \end{bmatrix} = \begin{bmatrix} \alpha_i \\ \beta_{i1} \\ \vdots \\ \beta_{i8} \end{bmatrix}$ is the vector of regression coefficients for asset i.

Expanding the squared ℓ_2 norm, we get

$$(\boldsymbol{r}_i - \boldsymbol{X}\boldsymbol{B}_i)^T(\boldsymbol{r}_i - \boldsymbol{X}\boldsymbol{B}_i) = \boldsymbol{r}_i^T\boldsymbol{r}_i - 2\boldsymbol{r}_i^T\boldsymbol{X}\boldsymbol{B}_i + \boldsymbol{B}_i^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{B}_i$$

The original minimization is equivalent to

$$\min_{\boldsymbol{B}_i} \ -2\boldsymbol{r}_i^T\boldsymbol{X}\boldsymbol{B}_i + \boldsymbol{B}_i^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{B}_i$$

Let $f(\boldsymbol{B}_i) = -2\boldsymbol{r}_i^T\boldsymbol{X}\boldsymbol{B}_i + \boldsymbol{B}_i^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{B}_i$,
taking partial derivative w.r.t $\boldsymbol{B}_i$

$$\frac{\partial f(\boldsymbol{B}_i)}{\partial \boldsymbol{B}_i} = 2\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{B}_i - 2\boldsymbol{r}_i^T\boldsymbol{X}$$

The optimal coefficient vector is:

$$\boldsymbol{B}_i^* = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{r}_i$$

With residuals $\epsilon_i = \boldsymbol{r}_i - \boldsymbol{X}\boldsymbol{B}_i^*$, we can calculate the residual variances $\sigma_{\epsilon_i^2} = \frac{1}{T-p-1}\|\epsilon_i\|_2^2$
Putting it together for the 20 assets, we have

$$\boldsymbol{B}^* = \begin{bmatrix} \alpha^{*T} \\ \boldsymbol{V}^* \end{bmatrix}_{9\times 20} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{r}$$

The expected returns $\boldsymbol{\mu} = \alpha + \boldsymbol{V}^T\bar{\boldsymbol{f}}$ where $\bar{\boldsymbol{f}}$ is the vector of expected factor returns.

The covariance matrix $\boldsymbol{Q} = \boldsymbol{V}^T\boldsymbol{F}\boldsymbol{V} + \boldsymbol{D}$ where $\boldsymbol{D}$ is the diagonal matrix of residual variances, i.e. $\sigma_{\epsilon_i}^2$ on the diagonals and zeros on off-diagonals.

2.1.2 FF Model

The Fama-French three-factor model is a subset of OLS model, where we only use variables: Market, Size and Value. The FF model can be expressed as

$$r_i - r_f = \alpha_i + \beta_{im}(f_m - r_f) + \beta_{is}f_s + \beta_{iv}f_v + \epsilon_i$$

where r_i is the return of asset i, r_f is the risk-free rate, α_i is the intercept of regression, $f_m - r_f$ is the factor of excess market return with corresponding factor loading β_{im}, f_s is the factor of size with corresponding factor loading β_{is}, f_v is the factor of value with corresponding factor loading β_{iv}, ϵ_i is stochastic error term for asset i.

We have the same minimization objective as with OLS model, just with smaller dimensions as number of factors reduced from eight to three.

The optimal coefficient vector

$$\boldsymbol{B}^* = \begin{bmatrix} \alpha^{*T} \\ \boldsymbol{V}^* \end{bmatrix}_{4\times 20} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{r}$$

where the data matrix $\boldsymbol{X} = [\boldsymbol{1}\ \boldsymbol{f}] = [\boldsymbol{1}\ \boldsymbol{f1}\ \boldsymbol{f2}\ \boldsymbol{f3}]$.
The expected returns $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{Q}$ are obtained using the same approach as described in the previous section.

2.1.3 LASSO Model

The penalized form of LASSO on all eight factors, where the model is the following

$$\begin{aligned} \min_{\boldsymbol{B}_i} \quad & \|\boldsymbol{r}_i - \boldsymbol{X}\boldsymbol{B}_i\|_2^2 + \lambda \mathbf{1}^T \boldsymbol{y} \\ \text{s.t.} \quad & \boldsymbol{y} \geq \boldsymbol{B}_i \\ & \boldsymbol{y} \geq -\boldsymbol{B}_i \end{aligned}$$

where $\boldsymbol{y} \geq_{B\ i}$ and $\boldsymbol{y} \geq -_{B\,i}$ Align this Quadratic Programming for LASSO with the default MATLAB *quadprog*() function

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \frac{1}{2}\boldsymbol{x}^T H \boldsymbol{x} + \boldsymbol{f}^T \boldsymbol{x} \\ \text{s.t.} \quad & A\boldsymbol{x} \leq \boldsymbol{b} \\ & Aeq\boldsymbol{x} \leq \boldsymbol{beq} \\ & lowerbound \leq \boldsymbol{x} \leq upperbound \end{aligned}$$

To determine λ, we want to choose the values which result in a sparse factor model where two to five coefficients are non-zero. The LASSO determines whether we should include the intercept or not.

2.1.4 BSS Model

Best Subset Selection model impose an additional cardinality constraint on the portfolio by limiting the number of factors we can choose. This will require us to formulate a mixed-integer quadratic program (MIQP) by incorporating binary variables into mean variance optimization. It can be expressed as:

$$\begin{aligned} \min_{\boldsymbol{B}_i} \quad & \|\boldsymbol{r}_i - \boldsymbol{X}\boldsymbol{B}_i\|_2^2 \\ \text{s.t.} \quad & L\boldsymbol{y} \leq \boldsymbol{B}_i \leq U\boldsymbol{y} \\ & \|\boldsymbol{B}_i\|_0 \leq K \end{aligned}$$

where L and U are the lower and upper bound for regression coefficients and K is the cardinality constraint. We start with K = 4 and will test how other values of K affects the in-sample measures of fit and the out-of-sample portfolio performance. Note that K could be be ranged from 1 to 9 including the intercept α. We will let the BSS model determine whether we should include the intercept or not. To solve the above MIQP, we will leverage on the Gurobi optimization.

Note that

$$\|\boldsymbol{r}_i - \boldsymbol{X}\boldsymbol{B}_i\|_2^2 = \boldsymbol{r}_i^T \boldsymbol{r}_i - 2\boldsymbol{r}_i^T \boldsymbol{X}\boldsymbol{B}_i + \boldsymbol{B}_i^T \boldsymbol{X}^T \boldsymbol{X}\boldsymbol{B}_i$$

.

By ignoring the constant term and compare with Gurobi's default formulation

$$\min_{\boldsymbol{x}} \quad \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} + \boldsymbol{c}^T \boldsymbol{x}$$

We can find $\boldsymbol{x} = \boldsymbol{B}_i, Q = X^T X$ and $c = -2X^T r_i$.
By combing the auxiliary binary variables, our decision variable $\boldsymbol{x}$ will become

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{B}_i \\ \boldsymbol{y} \end{bmatrix}_{18\times 20}$$

where $\boldsymbol{y} \in \{0, 1\}$
Similarly, Q and c will also be resized by inserting $(p+1)$ by $(p+1)$ zero matrix.
Constraints are set with

$$\boldsymbol{A} = \begin{bmatrix} -\boldsymbol{I} & L\boldsymbol{I} \\ -\boldsymbol{I} & U\boldsymbol{I} \end{bmatrix} \boldsymbol{b} = \begin{bmatrix} \boldsymbol{0} \end{bmatrix}$$

and

$$\boldsymbol{Aeq} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{1} \end{bmatrix} \boldsymbol{beq} = \begin{bmatrix} \boldsymbol{K} \end{bmatrix}$$

2.2 Portfolio Optimization

After obtaining the estimates of expected returns $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{Q}$ for each factor model, we proceed to the Mean-Variance Optimization (MVO) problem. We aim to minimize the portfolio variance subject to exceeding the target expected return, i.e. average expected excess return of the market for the corresponding calibration period.

In matrix form, our optimization problem can be written as:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{x}^T \boldsymbol{Q}\ \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{\mu}^T \boldsymbol{x} \geq R \\ & \boldsymbol{1}^T \boldsymbol{x} = 1 \\ & (x_i \geq 0, \quad i = 1,\ ...,\ n) \end{aligned}$$

In addition to the constraint of achieving the target return R, we need the weight of each asset to sum up to one to form a complete portfolio. Moreover, we impose the lower bound of zero for the weights to disallow short sales.

To implement the optimization, we set up five time windows, each having a four-year calibration period and a one-year investment period. For each of the five investment periods starting Jan-2012, Jan-2013, ... , Jan-2016, we re-calibrate the expected returns $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{Q}$ using the four-year historical returns immediately precede and use as input variables for the corresponding MVO in order to calculate asset allocation $\boldsymbol{x}$ for the optimal portfolios. The calibration period rolls forward with the investment period. For example, our first investment period begins at 2012 and we use data from 2008 to 2011. Once this investment period is over, we re-calibrate new parameters using the most recent 4-year window returns.

Within each period, we construct four MVO portfolios for each factor model using the corresponding $\boldsymbol{\mu}$ and $\boldsymbol{Q}$.

3 Results

In this section, we will perform both in-sample and out-of-sample analysis to access the fitness of the factor models and financial performance of the associated portfolios respectively.

3.1 In-sample Analysis

The key metric we used for in-sample analysis is the adjusted R^2, $\bar{R}^2$. R^2 is a statistical measure of how well the regression predictions approximate the real data points that takes value from 0 to 1. The closer to 1, the better the regression fits. $\bar{R}^2$ is a less biased estimator of R^2 that has been adjusted for number of factors in the model.

To calculate the adjusted $\bar{R}^2$, we first need the two arguments - SS_{res} and SS_{tot} defined as:

$$SS_{res_i} = \sum_{t=1}^{T}(r_{it} - \boldsymbol{X}\boldsymbol{B}_i^*)^2$$

$$SS_{tot_i} = \sum_{t=1}^{T}(r_{it} - \bar{r}_i)^2$$

where $\bar{r}_i$ is the the average return on asset i over time T.
For in sample analysis, we will use 4-year historical returns as our calibration period. Given returns are monthly, T will be 48.

$$\bar{r}_i = \frac{1}{T}\sum_{i=1}^{T} r_{it}$$

$$R_i^2 = 1 - \frac{SS_{res_i}}{SS_{tot_i}}$$

$$\begin{aligned}\bar{R}_i^2 &= 1 - (1 - R^2)\frac{T-1}{T-p-1} \\ &= 1 - (\frac{SS_{res_i}}{SS_{tot_i}})\frac{T-1}{T-p-1}\end{aligned}$$

We calculated of $\bar{R}^2$ for each of the four factor models over the five investment periods. Within each period for each model, we have 20 $\bar{R}^2$ calculated for each of the 20 stocks respectively. To compare across different models, we took the average of the 20 $\bar{R}^2$s.
The averaged $\bar{R}^2$ calculated for the four factor models are shown below in Table 1.

Table 1: Average $\bar{R}^2$ for each factor models

	OLS	FF	LASSO	BSS
2012	0.480	0.436	0.433	0.397
2013	0.477	0.398	0.387	0.380
2014	0.436	0.347	0.313	0.392
2015	0.397	0.280	0.234	0.367
2016	0.439	0.341	0.284	0.377

Comparing the numbers in Table 1, we can see that the OLS model has the highest $\bar{R}^2$'s in all periods, possibly owing to inclusion of all eight factors. High $\bar{R}^2$'s indicate that the model is good at explaining the returns, and the resulting estimates, $\boldsymbol{\mu}$ and $\boldsymbol{Q}$, are therefore more reliable. BBS model has the lowest, but most stable $\bar{R}^2$'s, at approximately 0.4 throughout the investment period.

In order to investigate the impact of cardinality K on BSS model, we summarize a table for average R^2 as below. R^2 tends to increase with K at the beginning, but starts to decline when K goes beyond 5. Therefore, we conclude K = 5 is optimal.

Table 2: Average $\bar{R}^2$ for BSS models with different K

	K=3	K=4	K=5	K=6	K=7
2012	0.378	0.397	0.403	0.403	0.401
2013	0.362	0.380	0.401	0.400	0.398
2014	0.352	0.392	0.378	0.372	0.370
2015	0.325	0.367	0.352	0.345	0.341
2016	0.343	0.377	0.371	0.370	0.362

3.2 Out-of-sample Analysis

To quantify and compare portfolio performance associated with each factor model, we perform out-of-sample test based on portfolio returns and volatility. Out-of-sample test period will be exactly one year after the calibration period. ie. T = 12. We calculated the average return over the 5-year investment horizon, the standard deviation (volatility) of the returns, and the Sharpe Ratio for each portfolio; and the results are displayed in Table 3 below.

Table 3: Performance Metrics for Portfolios

	OLS	FF	LASSO	BSS
Average Return	0.71%	0.70%	0.71%	0.74%
Volatility	2.61%	2.62%	2.67%	2.73%
Sharpe Ratio	0.2401	0.2356	0.2360	0.2382

From the statistics in Table 2, we can see that the portfolio constructed using the BSS model provided the best average return. OLS and FF based portfolios yielded lower volatility and can be viewed as having lower risks. Also, the OLS model has the highest Sharpe Ratio, indicating high risk-adjusted excess returns.

To visualize the wealth evolution throughout the investment period, we have plotted portfolio values for each factor models against time in Figure 1 below.

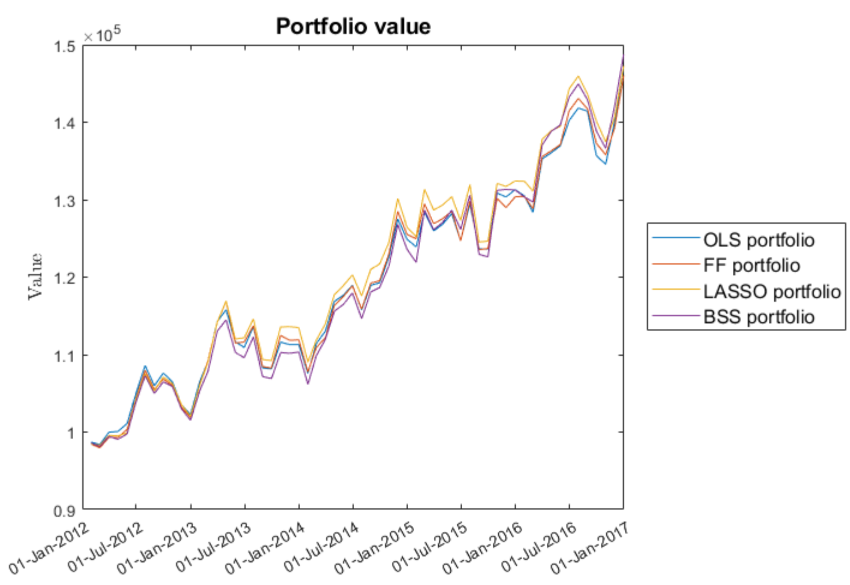

Figure 1: Portfolio Value Evolution

The evolution paths of portfolio value for the four factor models are highly similar. They share the same overall trend and ups and downs throughout the investment period. From year 2013 to 2016, the LASSO portfolio has noticeably higher total value than the others, and the BSS portfolio appears to be most volatile. The differences elsewhere are subtle.

We have also plotted portfolio weights against time to illustrate how the percentage of holdings for each stock evolves over time.

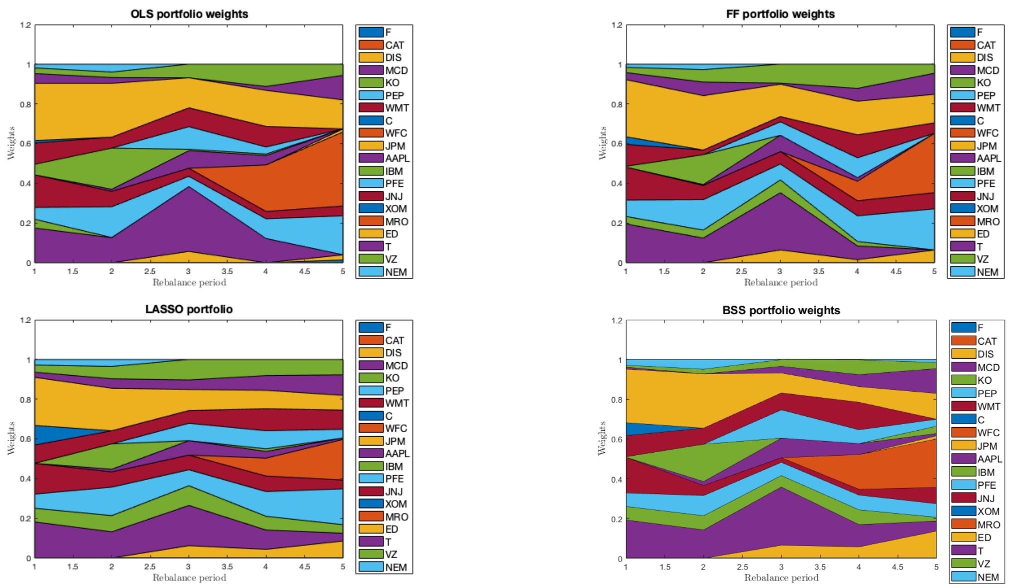

Figure 2: Composition of Portfolios

From the weight allocation plots (Figure 2), we observed some universal patterns that apply to portfolios based on all four factor models. Noticeably, the weights on stock MCD climbs up starting from rebalance period 2, spike at period 3, and declined afterwards, forming a mountain-like shape. Also, starting from the beginning of period 3, stock WFC weighs significantly more in the each portfolio, especially in the OLS portfolio.

Despite the similarities in the evolutions, the magnitude of weight allocation differs hugely for the four portfolios. The OLS portfolio appears to be the most unevenly allocated, as stock ED and MCD comprise of a large part of the portfolio. For the other portfolios, especially the LASSO portfolio, weights on each stock are more evenly distributed.

4 Conclusion

In terms of model fitness, the OLS model outperformed other models in explaining the asset returns. In terms of financial performance, the BSS portfolio yielded the best average returns and the OLS gave the highest risk-adjusted returns. However, the LASSO model has its unique advantage of diversification.

There are potential improvements that might be worthy of further research. For instance, we applied nominal MVO for our optimal portfolio construction which relies on the estimated parameters and ignores estimation error. The discrepancy between estimated expected returns and the true values may lead to a large deviation from optimal. The potential amendment to this is robust MVO which takes into consideration uncertainty, and it is beyond the scope of this report.