Portfolio Optimization

Long Zhao

02 September, 2021

Introduction

The principal of portfolio optimization goes beyond trading stocks to make money. It tries to allocate limited resources in an uncertain world efficiently. With appropriate data, the following problems all belong to this topic.

1. Given limited time, energy, and money, how should one spend their resources to maximize their overall happiness?
2. Given a limited marketing budget, how to choose the promotion vehicles to maximize advertising benefit?
3. Given a limited hiring budget, how to build a team to maximize the probability of success?
4. Given limited resources, how to balance R&D with daily operations?

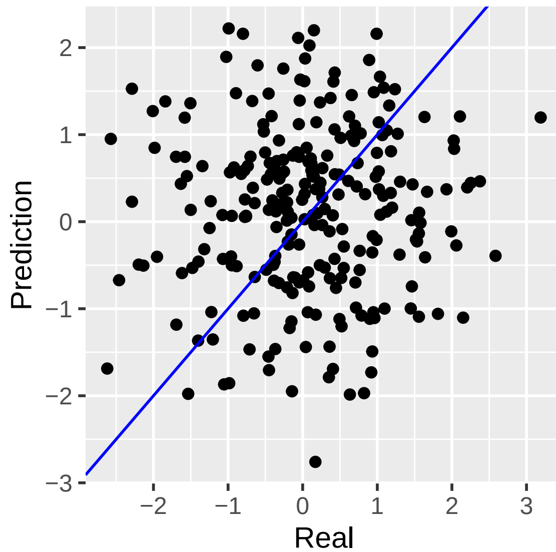
Due to data availability, in this project, we will focus on building a portfolio of risky assets. Through this project, we will understand overfitting of linear regression in terms of monetary values.

The rest of the project is structured as follows.

1. We will motivate the portfolio construction procedure using two assets.
2. By utilizing matrix form, we extend the idea towards multiple assets.
3. By a smart transformation, the optimization becomes linear regression.
4. We analyze the problem using a linear regression perspective and obtain the lasso and ridge portfolio.
5. Finally, we provide the specifics of this project.

Two Risky Assets

The mean-variance portfolio optimization proposed in Markowitz (1952) opened a new era of modern finance. It argues that one should make an explicit tradeoff between benefit and risk. Markowitz is awarded the Nobel price in economics in 1990 for this contribution. However, the expected returns of stocks, namely benefits, are infamously hard to estimate. Try to guess what Apple, Amazon, and Alphabet's stock price will be in the next month. It is commonly believed that a 2% **out-of-sample** R^2 will lead to a profitable trading strategy. Here is a visualization of 2% R^2



Noticing this issue, some start focusing only on minimizing the risk, and they hope the corresponding return will not be too disappointing. To illustrate the core idea, we will demonstrate how to build a portfolio of two risky assets within this framework. Denote the expected return and variance of the returns of these two assets as (μ_1, σ_1^2) and (μ_2, σ_2^2). We also denote the covariance between two assets returns as σ_{12}. Now, we could calculate the expected return and variance of a portfolio $\boldsymbol{w} = (w_1, w_2)^T$ (w_1 is the weight of the first asset) as

$$\begin{aligned} \text{Expected Return:} &\quad \mu_1 w_1 + \mu_2 w_2, \\ \text{Variance:} &\quad \sigma_1^2 w_1 + 2\sigma_{12} w_1 w_2 + \sigma_2^2 w_2. \end{aligned}$$

Mean-variance framework balances between the expected return and variance. Meanwhile, if one only minimizes variance, the optimization becomes

$$\begin{aligned} \min_{w_1, w_2} &\quad \sigma_1^2 w_1 + 2\sigma_{12} w_1 w_2 + \sigma_2^2 w_2 \\ \text{subject to} &\quad w_1 + w_2 = 1, \end{aligned}$$

where the constraint $w_1 + w_2 = 1$ means that all money should be invested in these two assets. For example, $w_1 = 30\%$ and $w_2 = 70\%$ means that 30% money in first asset while 70% in the second asset. This portfolio satisifes the constraint that 100% of money is invested.

Multiple Risky Assets

We could also write the above optimization in the matrix form. With

$$\begin{aligned} \Sigma &= \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \Rightarrow \boldsymbol{w}^T \Sigma \boldsymbol{w} = \sigma_1^2 w_1 + 2\sigma_{12} w_1 w_2 + \sigma_2^2 w_2, \\ \mathbf{1} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Rightarrow \boldsymbol{w}^T \mathbf{1} = w_1 + w_2, \end{aligned}$$

we have

$$\begin{aligned} \min_{\boldsymbol{w}} &\quad \boldsymbol{w}^T \Sigma \boldsymbol{w} \\ \text{subject to} &\quad \boldsymbol{w}^T \mathbf{1} = 1. \end{aligned}$$

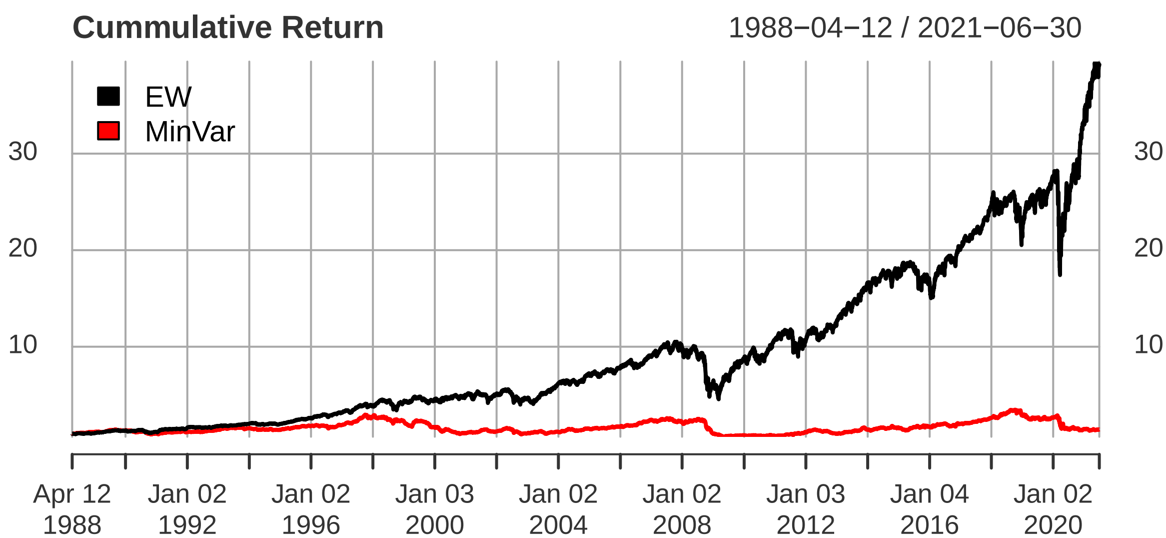
Formally speaking, Σ is the covariance matrix of the risky assets. This matrix formulation also applies to the cases where the total number of assets is more than 2. We call the solution as the **minimum-variance** portfolio.

Estimation of Σ

The theory leaves us an incomplete map: it does not tell us how to get Σ. It is tempting to replace Σ with the sample covariance matrix, $\hat{\Sigma}$. If so, we obtain the **estimated** minimum-variance, $\boldsymbol{w}_{MinVar}$ as the solution to

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^T \hat{\Sigma} \boldsymbol{w} \\ \text{subject to} \quad & \boldsymbol{w}^T \mathbf{1} = 1. \end{aligned}$$

Unfortunately, such a portfolio could have terrible performance. Take the **48 Industry daily** dataset from Prof. French's Data Library as an example.



The plot above is about the cumulative return of the minimum-variance (MinVar) and the equally-weighted (EW) portfolio. For the MinVar, we use 63 (approximately 3-month daily returns) observations to estimate the covariance matrix. Meanwhile, the equally-weighted portfolio naively invests 1/48 in each industry. That is to say, $\boldsymbol{w}_{EW} = \frac{1}{p}\mathbf{1}$, where p is the number of assets. The equally-weighted dominates the minimum-variance. In fact, DeMiguel, Garlappi, and Uppal (2009) show that none of 13 popular methods could consistently beat the equally-weighted. Next, we will use the perspective of linear regression to understand why.

Linear Regression Perspective.

Motivated by the good performance of the equally-weighted portfolio, we decide to focus on $\boldsymbol{w}_{EW} - \boldsymbol{w}$ instead of $\boldsymbol{w}$. To illustrate the procedure, let us go back to the two assets example. The constraint $w_1 + w_2 = 1$ could be satisfied by

$$w_1 = \frac{1}{2} - \beta_1 \quad \text{and} \quad w_2 = \frac{1}{2} + \beta_1.$$

The matrix form is

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} \beta_1 \Leftrightarrow \boldsymbol{w} = \boldsymbol{w}_{EW} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} \beta_1.$$

If there are three assets, we could have

$$\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} 1/3 - \beta_1 \\ 1/3 - \beta_2 \\ 1/3 + \beta_1 + \beta_2 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}.$$

If we define

$$N = \begin{pmatrix} I_{p-1} \\ -\mathbf{1}^T \end{pmatrix} \quad \text{and} \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_{p-1} \end{pmatrix}, (I_{p-1} \text{ is the } (p-1) \times (p-1) \text{ identity matrix})$$

then we have

$$\boldsymbol{w} = \boldsymbol{w}_{EW} - N\boldsymbol{\beta}, \tag{1}$$

which holds for any $p \geq 2$. Meanwhile, by the definition of covariance matrix, we have $\hat{\Sigma} \triangleq \frac{1}{n-1} R^T R$, where n is the number of observations and R is the **demeaned** return data[1]. Plug both into the optimization to have

$$\min_{\boldsymbol{\beta}} \|y - X\boldsymbol{\beta}\|_2^2, \text{ where } y = R\boldsymbol{w}_{EW} \text{ and } X = RN.$$

Amazingly, we convert the optimization into a **linear regression (OLS) without intercept** problem! Now, we could use all we know about linear regression to understand what happens to this portfolio. It is equivalently a linear regression with 47 covariates but only 63 observations. Clearly, it will overfit. Thus, one should use regularization to counter the overfitting issue.

Remark: $\boldsymbol{\beta}$ does not have an intuitive explanation. It is the result of the mathematical derivation.

Regularization

We applied two popular regularization methods, lasso and ridge, to address the overfitting issue:

$$\text{Lasso: } \min_{\boldsymbol{\beta}} \|y - X\boldsymbol{\beta}\|_2^2 + \lambda_1 \|\boldsymbol{\beta}\|_1$$
$$\text{Ridge: } \min_{\boldsymbol{\beta}} \|y - X\boldsymbol{\beta}\|_2^2 + \lambda_2 \|\boldsymbol{\beta}\|_2^2.$$

Define the corresponding solutions with **cross-validated** λ_1 and λ_2 as $\boldsymbol{\beta}_{Lasso}$ and $\boldsymbol{\beta}_{Ridge}$, respectively. Then we could use Eq. (1) to obtain $\boldsymbol{w}_{Lasso} = w_{EW} - N\boldsymbol{\beta}_{Lasso}$ and $\boldsymbol{w}_{Ridge} = w_{EW} - N\boldsymbol{\beta}_{Ridge}$.

The Specifics (Todolist for you)

0. Download the **48 Industry daily** dataset from Prof. French's Data Library.
    - The file contains **two** datasets. The one at the top is the value-weighted part. **You should keep this value-weighted part and delete the other part**.
    - Once again. You should only have **one** dataset from July 1st, 1926 **instead of two**.
1. Choose 6 industries (from 48) and build portfolios **for 2021-01-04**.
    - You will use the CV plots for both Lasso and Ridge with good explanations to convince me about overfitting.
2. Repeat the above procedure for 24 as well as 48 industries.

[1] Demeaned means that the sample mean of each asset return is 0.

- Do you have any explanation for the patterns?

3. Which of the following are the wrong approach to pursue better performance regarding variance? Please also explain your choices.
   - Use 42 observations instead of 63 as training.
   - Use 84 observations instead of 63 as training.
   - Use polynomial regression to create more features.
   - Use the sum of the absolute deviations as loss function because it is more robust.
   - Use PCR instead of Lasso and Ridge.
4. Something that you find interesting or important. [Choose one, no need to do multiple.]
   - Shall be somehow related to prediction or portfolio optimization.
   - You could recreate the cumulative return plot for EW, MinVar, Lasso, Ridge.
   - At least, you could implement the right approaches from Question 3.

Report Requirements

- Question 1-3 account for only 70% of grade.
  - Playing safe will not be enough!
- Pages: 6 pages (excluding cover & reasonable font) without penalty
  - You do not need to report all things in detail. Just the ones that fit your storyline.
- The target of this project is to quantify overfitting of linear regression in terms of monetary values. Based on this objective, I have the following requirements for you.
  - It should be an analysis about overfitting using portfolio optimization as a showcase. That is to say, you need to **organize your results to demonstrate the phenomenon of overfitting and how to mitigate the problem**. Just listing results might incur some huge penalty.
  - **You need to why some approaches are wrong in the specifics.**
- It should clearly state which are the chosen 6 or 24 industries.
  - I recommend a random sampling. Two groups with identical choice of industries will be scrutinized.
- Upload a .zip file consisting of **your report (a pdf file)** and **code**.
  - Please include your group number in your report.

Details about Dataset:

1. You should use the **value-weighted** version. It will combine all stocks within one industry by its market value to represent this industry.
2. Missing data are indicated by -99.99 and -999. Luckily, in this project, missing values will not make a difference.

Some Numbers to Help You Debug

1. If you use the whole dataset starting "1926-07-01," then "2020-12-31" is the 24896th row.
2. If you choose the first 6 industries[2], the MinVar portfolio for them on "2020-12-31" is -0.029, 0.542, 0.261, 0.067, 0.081, 0.077.
3. This portfolio's return on 2021-01-04 is -1.886.

[2] Namely, "Agric", "Food", "Soda", "Beer", "Smoke", and "Toys".

Some Functions to Help You Code

1. Demean data.
    - scale() in Python
    - scale() in R
2. Lasso and Ridge with cross-validation.
    - You could do it manually as in the DataCamp course. The following might be easier.
    - cv.glmnet() in R
    - In Python, you might need to use sklearn.linear_model.Lasso, sklearn.linear_model.Ridge, sklearn.model_selection.cross_val_score.
    - The idea is to find the relationship between the mean squared error and -log(alpha) using alpha values of np.logspace(-8,-8,100) for both Lasso and Ridge.

Reference.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal. 2009. "Optimal Versus Naive Diversification: How Inefficient Is the 1/n Portfolio Strategy?" *The Review of Financial Studies* 22 (5): 1915–53.

Markowitz, Harry. 1952. "Portfolio Selection." *The Journal of Finance* 7 (1): 77–91.