

# SDGB 7844 HW 5: Portfolio Optimization

Instructor: Prof. Nagaraja

Due: 12/7 in class

Submit two files: (a) .Rmd R Markdown file with answers and code and (b) Word document of knitted R Markdown file. Email both files to [fordhamRcomputing@gmail.com](mailto:fordhamRcomputing@gmail.com) by the start of class with the subject line “HW5-[Full Name]-[Class Time]” and include HW 5 and your name in the file names (time of class is either 1:15 PM or 5:45 PM).

Please email your solutions only once! Complete your work individually and comment your code for full credit. For an example of how to format your homework see the files related to the Lecture 1 Exercises on Blackboard.

Given the uncertainty about the future performance of financial markets, investors typically diversify their portfolios to improve the quality of their returns. In this assignment, you will be constructing a portfolio composed of two ETFs (exchange traded funds) that tracks US equity and fixed income markets. The equity ETF tracks the widely followed S&P 500 while the fixed income ETF tracks long-term US Treasury bonds. You will use optimization techniques to determine what fraction of your money should be allotted to each asset (i.e., portfolio weights).

The Sharpe ratio is a widely used metric to gauge the quality of portfolio returns. It was developed by William F. Sharpe (a Nobel prize winner) and provides one way to construct an optimal portfolio. This computation allows investors to determine expected returns while taking into account a measure of how risky that investment is. Heuristically, it is computed by taking the expected return of an asset/portfolio, subtracting the risk free rate (i.e., how much you would make if you simply kept your money in the bank and let it collect interest) and dividing by the standard deviation of the portfolio returns. The higher the Sharpe ratio value, the better the investment is considered to be. This formula assumes that returns are normally distributed (not always the best assumption to make with financial data).

The data for the two assets were downloaded from Google Finance (ticker symbols: SPY, TLT) and the federal funds interest rate, representing the risk free rate, was taken from the U.S. Federal Reserve Bank website. We will be looking at weekly returns, as opposed to daily returns, as they are less correlated with each other; monthly returns would be even less correlated, however, we would need a much longer time period of data to have enough observations to do a robust analysis. Note that if we used a different time period to compute our optimal rates, or we used daily instead of weekly returns, we may get very different results.

1. Upload the data in “asset\_data.txt” into R and call the data frame `data.x`. The columns are defined as follows: `date`, `close.spy` is the closing price for the S&P500 ETF, `close.tlt` is the closing price for the long term treasury bond ETF, and `fed.rate` is the federal funds interest rate in percent form. Next, use the command:

```
data.x$date <- as.Date(data.x$date, format="%Y-%m-%d")
```

to turn the first column from a character string to a date<sup>1</sup>. Extract only the observations where the federal funds rate is available (so you are left with weekly data); this is the data you will use for the rest of the analysis. What is the start date and end date of this reduced data set? Graph the federal funds interest rate as a time series (i.e., date on  $x$ -axis; set `type="l"` in plot for connected points). Describe what you see in the plot and relate it briefly to the recent financial crisis.

2. Now we will split the data into training and test sets. The training data will be used to compute our portfolio weights and our test set will be used to evaluate our portfolio. Make two separate data frames; (a) the training set should contain all observations before 2014 and (b) the test set should contain all observations in 2014. How many observations are in each subset?
3. The federal funds interest rate is in percent form so convert it to decimal (i.e., fractional) form. Then, for the S&P 500 and long term treasury bonds ETF assets, compute the returns using the following formula:

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}}$$

where  $r_t$  is the return at time  $t$ ,  $p_t$  is the asset price at time  $t$ , and  $p_{t-1}$  is the asset price at time  $t - 1$  (i.e., the previous period). Add both sets of returns to your training set data frame. These returns are also called total returns. For both assets, construct time series plots of the returns (separate plots). Make the plots have identical  $y$ -axis limits. Add a dotted, horizontal line at  $y = 0$  on both plots as well. Compare the two returns series. What do you see?

---

<sup>1</sup>Just like the data types numeric, logical, etc., R has one for date/time data.

4. The Sharpe ratio calculation assumes that returns are normally distributed<sup>2</sup>. Construct two normal quantile plots, one for training set returns of each asset. Is this assumption satisfied? Justify your answer.
5. Compute the correlation between the S&P500 and long term treasury bond returns in the training set and interpret it<sup>3</sup>. Now, we will compute a rolling-window correlation as follows: compute the correlation between the two asset returns only using the first 24 weeks of data (i.e., weeks 2 to 25), next compute the correlation between the two asset returns for data from week 3 through 26, then week 4 through 27, and so forth.

Once you compute the rolling-window correlations, make a time series plot of the rolling-window correlation with each point plotted on the *last* day of the window. Add a horizontal, dotted, gray line at 0 to your plot. Is the correlation or rolling-window correlation a better way to describe the relationship between these two assets? Justify your answer.

6. Compute the Sharpe ratios<sup>4</sup> for each asset on the training set as follows:

Step 0. Let  $r_t$  be the return and  $y_t$  be the federal funds interest rate for week  $t = 1, \dots, T$ .

Step 1. Compute the excess returns,  $e_t$ , for each week in the data set<sup>5</sup>:

$$e_t = r_t - \frac{y_{t-1}}{52}$$

Excess returns are returns that you earn in excess to the risk free rate.

Step 2. Convert the excess returns into an excess returns index,  $g_t$ :

$$\begin{aligned} g_1 &= 100 \\ g_t &= g_{t-1} \times (1 + e_t) \end{aligned}$$

Step 3. Compute the number of years of data,  $n$ , by taking the number of weeks for which you have *returns* (i.e., number of observations in your training set minus 1) and dividing by 52 (since there are 52 weeks in a year); therefore the number of years of data can be a fractional amount.

Step 4. Compute the compounded annual growth rate, CAGR:

$$CAGR = \left( \frac{g_T}{g_1} \right)^{1/n} - 1$$

---

<sup>2</sup>We will continue with our analysis regardless of whether this assumption is satisfied.

<sup>3</sup>If you construct a scatterplot of these two particular assets, you will get a linear relationship.

<sup>4</sup>The Sharpe ratio can be positive or negative.

<sup>5</sup>We divide the federal funds interest rate by 52, the number of weeks in a year, to turn it into a weekly rate from an annual rate.

Step 5. Compute the annualized volatility,  $\nu$ :

$$\nu = \sqrt{52}SD[e_t]$$

where  $SD[e_t]$  is the standard deviation of the excess returns.

Step 6. Compute the Sharpe Ratio,  $SR$ , which is the ratio of the compounded annual growth rate and the annualized volatility:

$$SR = \frac{CAGR}{\nu}$$

Which asset is a better investment? Justify your answer.

7. Write a function which takes the following inputs: (a) a vector of portfolio weights (call this argument  $\mathbf{x}$ ; weights are between 0 and 1), (b) a vector of returns for asset 1, (c) a vector of returns for asset 2, and (d) a vector of the corresponding weekly federal funds interest rates. The function will then do the following: for each weight value in your vector  $\mathbf{x}$ , you will compute the Sharpe ratio for the corresponding portfolio. To obtain the returns for the portfolio, use the following equation<sup>6</sup>:

$$r_{t,portfolio} = xr_{t,S\&P500} + (1 - x)r_{t,treasury}$$

That is,  $x$  proportion of the funds will be invested in the S&P500 ETF and  $(1 - x)$  proportion of the funds will be invested into the treasury bond ETF. After you compute the returns for the portfolio, apply the steps in question 6 to get the Sharpe ratio for that portfolio. Your function should output a vector of Sharpe ratios, one for each portfolio weight in  $\mathbf{x}$ .

Use the `curve()` function to plot the Sharpe ratio ( $y$ -axis) for weights between 0 and 1 ( $x$ -axis) using the training set data as part of your function input. Do you see a portfolio weight that produces the maximum Sharpe ratio?

8. Using the training set, use `optimize()` to determine the optimum weight for each asset using the function you wrote in question 7; how much of the funds should be allocated to each asset? What is the Sharpe ratio of the overall portfolio? According to your analysis, is it better to invest in S&P500 only, long term treasury bonds only, or your combined portfolio? Justify your answer.
9. For the remainder of this assignment, we will be evaluating our portfolio using the test set data. We will be comparing three strategies: investing only in the S&P500, investing only in long term treasury bonds, and investing in the combined portfolio (computed in question 8).

---

<sup>6</sup>Note to people with a finance background: these computations correspond to a setting where you are rebalancing the portfolio daily.

In your test set, convert the federal funds interest rate from percent to decimal form and compute the returns series for each of the three assets (see question 3 for more details). Next, compute the excess returns index for each asset in the test set (as outlined in question 6). Plot the excess returns index all on the same time series plot and include a legend so you can differentiate between the three series. Add a dotted, horizontal, gray line at  $y = 100$ . Describe what you see.

10. The excess returns index can be interpreted as follows: if you invested in \$100 in at time  $t = 1$ , the index value at time  $T$  represents how much you have earned in addition to (i.e., in excess of) the risk free interest rate. If you invested \$100 in each asset (portfolio, all in long term treasury bonds, or all in S&P500) in the first week of January, 2014 , how much would you have at the end of the test set period for each asset in addition to the risk-free interest rate? Did your portfolio perform well in the test set? Justify your answer.