Seven

Portfolio Optimization

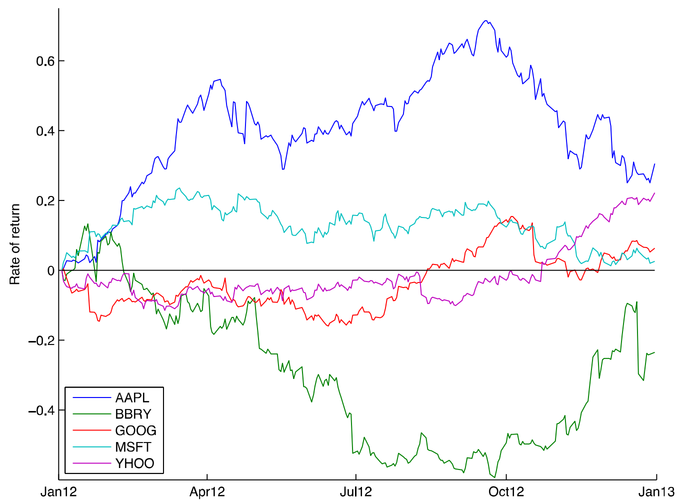

Figure 7.1: The historical rate of return of five technology stocks from the beginning of 2012.

1 Learning goals

In this case-study, we will learn how to compute portfolios of investments having either a maximum expected rate of return or a minimum overall volatility (i.e., risk) using real historical stock prices.

The exercises in this case-study will take us through an investigation of the relationship between the expected rate of return of a portfolio and its risk. We will start by downloading financial data from the Internet and using it to make a scatter plot of random portfolios. We will see that all portfolios lie in a region that is bounded by a curve called the *efficient frontier*. The portfolios on the efficient frontier are optimal in the sense that they maximize the expected rate of return for a fixed level of risk, and they minimize risk for a fixed level of expected rate of return. Finally, we will make use of dual multipliers to compute the *market portfolio* associated with a zero-risk investment.

This case-study will introduce students to the important class of *convex optimization problems*. We will learn how we can use the MATLAB-based convex optimization modeling system CVX to model and solve portfolio optimization problems. We expect students to have a good working knowledge of MATLAB, but we do not assume that students have studied portfolio theory or convex optimization before.

2 Background

Modern portfolio theory is based on the Markowitz model for determining a portfolio of stocks with a desired expected rate of return that has the smallest amount of variance. The main idea is that by *diversifying* (investing in a mixture of different stocks), one can guard against large amounts of variance in the rates of return of the individual stocks.

Suppose $p_1, \ldots, p_m$ are the historical prices of a stock over some period of time. We define the *rate of return* at time t, relative to the initial price p_1 by

$$r_t := (p_t - p_1)/p_1, \quad \text{for} \quad t = 1, \ldots, m. \tag{7.1}$$

The *expected rate of return* is the mean μ of the rates of return, and the *risk* is defined as the *standard deviation* σ of the rates of return:

$$\mu := \frac{1}{m}\sum_{t=1}^{m} r_t \quad \text{and} \quad \sigma := \sqrt{\frac{1}{m}\sum_{t=1}^{m}(r_t - \mu)^2}.$$

Given a collection of n stocks, let r_t^i be the rate of return of stock i at time t. Let r be the $n \times 1$ vector of the expected rates of return of the n stocks. In addition, let Σ be the $n \times n$ *covariance matrix* of the rates of return of the n stocks. Thus, r_i is the mean of the rates of return of stock i, Σ_{ii} is the variance of the rates of return of stock i, and Σ_{ij} is the covariance of the rates of return of stocks i and j:

$$r_i := \frac{1}{m}\sum_{t=1}^{m} r_t^i \quad \text{and} \quad \Sigma_{ij} := \frac{1}{m}\sum_{t=1}^{m}(r_t^i - r_i)(r_t^j - r_j).$$

We let x_i be the fraction of our investment money we put into stock i, for $i = 1, \ldots, n$. For the sake of this study, we assume there is no *short selling* (i.e., holding a stock in negative quantity). Thus, x is a vector of length n that has nonnegative entries that sum to one (i.e., $x \geq 0$ and $\sum_{i=1}^{n} x_i = 1$). The vector x represents our *portfolio* of investments. The expected rate of return and standard deviation of a portfolio x are then given by

$$\mu := r^T x \quad \text{and} \quad \sigma := \sqrt{x^T \Sigma x}.$$

Note that Σ is positive semidefinite, so there exists an $n \times n$ matrix R such that $\Sigma = R^T R$; if Σ is positive definite, then we can choose R to be the Cholesky factor of Σ. Then we can compute the standard deviation of portfolio x as

$$\sigma = \|Rx\|_2.$$

3 Problem description

We will consider various optimization problems in the following exercises. First, we will be interested in finding the portfolio x that has the minimum possible risk:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \sqrt{x^T \Sigma x} \\ \text{subject to} & \sum_{i=1}^{n} x_i = 1, \\ & x \geq 0. \end{array}$$

However, if we are seeking a minimum risk portfolio having expected return of at least μ, then we would want to solve:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \sqrt{x^T \Sigma x} \\ \text{subject to} & r^T x \geq \mu, \\ & \sum_{i=1}^{n} x_i = 1, \\ & x \geq 0. \end{array}$$

On the other hand, we may have a maximum risk that we are willing to tolerate. In this case, we want to find the portfolio x having maximum expected return having risk no more than σ:

$$\begin{array}{ll} \underset{x}{\text{maximize}} & r^T x \\ \text{subject to} & \sqrt{x^T \Sigma x} \leq \sigma, \\ & \sum_{i=1}^{n} x_i = 1, \\ & x \geq 0. \end{array}$$

Note that each of these problems are convex optimization problems.

4 Exercises

Software prerequisites. You will need to install CVX, a MATLAB-based modeling system for convex optimization.

4.1 (*In class*).

1. Download financial data (`csv` files) from Yahoo! Finance for the following twenty stocks:

   - Technology: AAPL, BBRY, GOOG, MSFT, YHOO
   - Services: AMZN, COST, EBAY, TGT, WMT
   - Financial: BMO, BNS, HBC, RY, TD
   - Energy: BP, CVX, IMO, TOT, XOM

   Store the `csv` files in a directory called '`data`'.

2. Download and complete the function `load_stocks.m`:

```
[X, dates, names] = load_stocks(dirname, startdate,
    enddate)
```

   This function must read the *adjusted closing prices* of all stocks in the given directory between the start date and end date, and compute the rates of return as in equation (7.1).

   Use the following start and end dates:

```
startdate = '2012-01-03';  enddate = '2012-12-31';
```

   Plot your results using `disp_stocks.m`:

```
disp_stocks(X, dates, names)
```

   You may compare your output against the solution `load_stocks_soln.p`.

3. Create a function `meancov.m` that returns the $n \times 1$ vector `r` of means and the $n \times n$ covariance matrix `Sig` of the rates of returns of n stocks given by `X`:

```
[r, Sig] = meancov(X)
```

4. Download and complete the function `portfolio_scatter.m`:

```
h = portfolio_scatter(r, Sig, num)
```

   This function must generate random portfolios and make a scatter plot of their expected rates of return and standard deviation. Each random portfolio is generated by randomly allocating a fraction of the overall investment among a small set of 5 randomly chosen stocks. Make a scatter plot with `num = 1000` points. This function returns a handle `h` to the figure. You may compare your output against the solution `portfolio_scatter_soln.p`.

4.2 (*Homework*).

1. Use CVX to compute the portfolio with minimum risk. What is the expected rate of return and standard deviation of this portfolio? Plot the rate of return of this portfolio over the entire time period. What is the portfolio with maximum possible expected rate of return? Create a function `return_range.m` that returns `num` linearly spaced rates of return between the rate of return of the portfolio with minimum risk and the maximum possible rate of return:

   ```
   rrange = return_range(r, Sig, num)
   ```

   You may compare your output against the solution `return_range_soln.p`.

2. Given a desired expected rate of return, we can see from the scatter plot that there are many portfolios that we can choose that have this expected rate of return. However, each of these portfolios have a different level of risk, or standard deviation. Among these, the most *efficient* portfolio is the one giving us the least amount of risk.

   Each expected rate of return determines a different efficient portfolio. Plotting the expected rate of return and standard deviation of each of the efficient portfolios will give us a curve called the *efficient frontier*.

   Download and complete the function `efficient_frontier.m`:

   ```
   [Y, rates, sigs] = efficient_frontier(r, Sig, num)
   ```

   This function will compute `num` efficient portfolios with linearly spaced rates of return (obtained from `return_range.m`). These portfolios will be stored in the `n`$\times$`num` matrix `Y`, and their corresponding expected rates of return and standard deviation in vectors `rates` and `sigs`. Plot `sigs` and `rates` on the scatter plot, with `num = 12`:

   ```
   h = portfolio_scatter(r, Sig, 1000);
   [Y, rates, sigs] = efficient_frontier(r, Sig, num);
   figure(h); hold on; plot(sigs, rates, 'ro-');
   ylim([0 0.5]); xlim([0 max(sigs)]);
   ```

   Display your results using `disp_portfolios.m`:

   ```
   h = disp_portfolios(Y, rates, sigs, names)
   ```

   You may compare your output against the solution `efficient_frontier_soln.p`.

4.3 (*In class*). Add a risk-free investment called 'RF' to the collection of stocks with a 3% rate of return. Use your `efficient_frontier.m` code from Exercise 2 to determine the new efficient frontier and plot it on the same plot with the original efficient frontier. You will notice that the new efficient frontier has two pieces: (1) a linear piece, and (2) a nonlinear piece that coincides with the original efficient frontier. What does the linear piece represent? The portfolio where these two pieces join is called the *market portfolio*. Download and complete the function `market_portfolio.m` that computes the market portfolio corresponding to a risk-free rate of return `f`:

```
    x = market_portfolio(f, r, Sig)
```

You may compare your output against the solution `market_portfolio_soln.p`. Plot the line that is tangent to the original efficient frontier at the market portfolio. What does the top half of this tangent line represent?