FINANCE WEB APP
PORTFOLIO CONSTRUCTION

Building the Markowitz efficient frontier

THIS web-app allows the user to build an efficient frontier given a series of historical prices. In particular the user uploads an Excel file containing historical stock prices. The procedure gives back the efficient frontier and randomly generated portfolios. The efficient frontier is computed in a range between the expected return of the minimum global variance portfolio and the highest expected return considering the stock prices used as input.

Background information

The Markowitz selection model delivers the portfolio with the lowest risk (measured by the variance), for a given level of expected return. The investor exploits the fact that the assets are not perfectly correlated, so that he can diversify the risk across a number of different securities.

The return model

We assume that log-returns of different stocks have a jointly normal distribution

$$\mathbf{r}\left(t, t+\Delta\right) \sim \mathcal{N}\left(\boldsymbol{\mu}_{\Delta}, \boldsymbol{\Sigma}_{\Delta}\right).$$

The mean vector $\boldsymbol{\mu}_{\Delta}$ is a vector with N components. The i-th component is denoted by μ_i and it represents the expected value of r_i

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_i \\ \vdots \\ \mu_N \end{bmatrix} = \begin{bmatrix} \mathbb{E}\left(r_1\right) \\ \vdots \\ \mathbb{E}\left(r_i\right) \\ \vdots \\ \mathbb{E}\left(r_N\right) \end{bmatrix}.$$

The covariance matrix $\boldsymbol{\Sigma}_{N\times N}$ is a squared and symmetric matrix; it is positive definite<[1] it is made of N variances (in the diagonal and denoted by σ_i^2) and $N\times\left(N-1\right)/2$ covariances (denoted by σ_{ij} with $\sigma_{ij}=\sigma_{ji}$).

$$\boldsymbol{\Sigma}_{N\times N} = \begin{bmatrix} \sigma_1^2 & & \cdots & & \sigma_{1,N} \\ & \sigma_2^2 & & & \\ \vdots & & \ddots & & \vdots \\ & & & \sigma_{N-1}^2 & \\ \sigma_{N,1} & & \cdots & & \sigma_N^2 \end{bmatrix}$$

where

$$\sigma_i^2 = \mathbb{V}\left(r_i\right) \text{ and } \sigma_{ij} = \mathbb{C}ov\left(r_i, r_j\right).$$

[1]This means that $\mathbf{x}'\boldsymbol{\Sigma}\mathbf{x} > 0\ \forall \mathbf{x} \in R^N, \mathbf{x} \neq \mathbf{0}$

The portfolio selection model

The portfolio selection problem consists in choosing the portfolio of weights $\mathbf{w}, \mathbf{w} \in R^N$ that solves the following quadratic problem

$$\min_{\mathbf{w}} \mathbf{w}'\mathbf{w}$$

subject to the following constraints

$$\mathbf{w}'\mathbf{1} = 1 \qquad \text{budget constraint} \tag{1}$$

$$\mathbf{w}'\boldsymbol{\mu} = m \qquad \text{return target} \tag{2}$$

The user can also decide if to impose the no-short sell constraint

$$\mathbf{w} \geq \mathbf{0}$$

and if to impose or not a maximum threshold $\mathbf{u}$ on the individual weights

$$\mathbf{w} \leq \mathbf{u}$$

where cleary the components u_i must satisfy $u_i \leq 1/N$.

Implementing the Markowitz model

The construction of the efficient frontier requires the knowledge of the vector of expected returns μ and of the covariance matrix Σ. They are never known. Therefore, they need to be estimated from a time-series of historical returns. Then the portfolio problem is solved, as if the estimates were the true parameters. Under the assumption that log-returns are jointly normally distributed through time, the maximum likelihood estimator of μ and Σ are the sample average and the sample var-cov matrix.

Let $\mathbf{R}_{\Delta}$ the $T\times N$ array containing past daily log-returns

$$\mathbf{R}_{T\times N} = \begin{bmatrix} r_1\left(0,\Delta\right) & & r_N\left(0,\Delta\right) \\ & r_j\left(i\Delta,\left(i+1\right)\Delta\right) & \\ r_1\left(T-\Delta,T\right) & & r_N\left(T-\Delta,T\right) \end{bmatrix}$$

where $r_j\left(i\Delta,(i+1)\Delta\right)$ is the log-return of stock j on day $\left(i\Delta,(i+1)\Delta\right)$. The array $\mathbf{R}$ has T rows and N columns.

Mean vector and Sample mean

The mean vector is estimated using the sample mean vector:

$$\hat{\mu}_{\Delta}=\frac{1}{T}\mathbf{R}_{\Delta}^{\prime}\mathbf{1}_T \tag{3}$$

where $\mathbf{1}_T$ is the unit vector with T components

$$\mathbf{1}_{T\times 1}^{\prime}=\underbrace{\begin{bmatrix} 1 & 1 & \vdots & 1 & 1 \end{bmatrix}}_{T \text{ components}}.$$

Covariance matrix and the sample covariance matrix

The covariance matrix is estimated trough the sample covariance matrix. In vector notation it can be written as

$$\begin{aligned}\widehat{\Sigma}_{\Delta} &= \frac{1}{T}\left(\mathbf{R}_{\Delta}^{\prime}\left(\mathbf{I}_T-\frac{\mathbf{1}\mathbf{1}^{\prime}}{T}\right)\mathbf{R}_{\Delta}\right)\\ &=\frac{1}{T}\left(\mathbf{R}_{\Delta}^{\prime}\mathbf{R}_{\Delta}-T\hat{\mu}_{\Delta}\hat{\mu}_{\Delta}^{\prime}\right).\end{aligned} \tag{4}$$

where $\mathbf{I}_T$ is the $T\times T$ identity matrix.

Form Field

In this application the user has an input form (Figure 1).

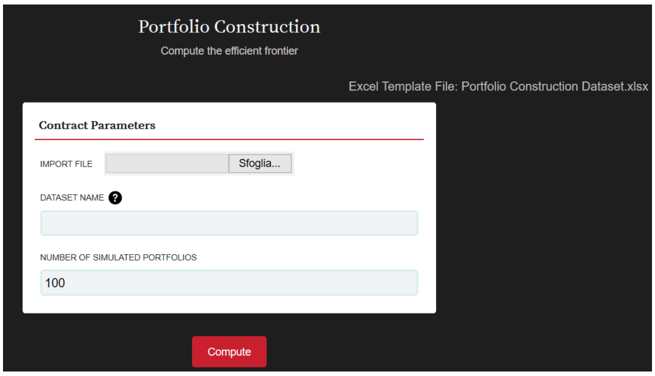
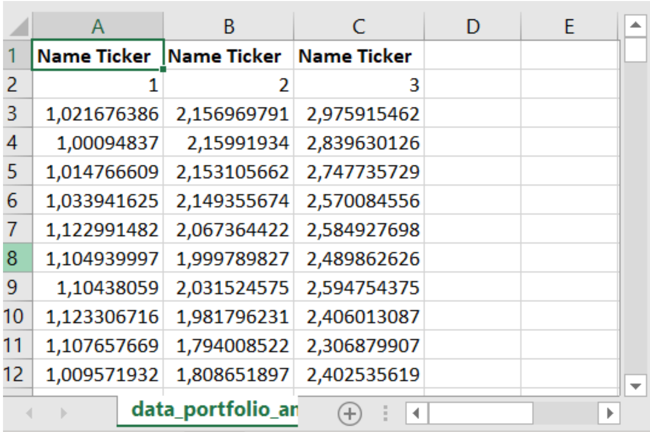


Figure 1: Input form

Input data characteristics

In the space next to the "Import File", the user can select an Excel file from a folder of his device and import it. The dataset contained in the file is used to estimate the mean vector and the covariance matrix and then to calculate the efficient frontier. The Excel file must comply with a specific template. An example is given Figure 2. The template file can be downloaded from the Portfolio Construction web page, by clicking on the link "Excel Template File". In particular, the first row must have the ticker name of each stock. Below each name, the Excel file contains the stock prices, starting from the oldest up to the most recent. The series must have the same number of observations. The sampling frequency is decided by the user. If prices are collected with weekly frequency, then the efficient frontier is built with a weekly horizon.



	A	B	C	D	E
1	Name Ticker	Name Ticker	Name Ticker		
2	1	2	3		
3	1,021676386	2,156969791	2,975915462		
4	1,00094837	2,15991934	2,839630126		
5	1,014766609	2,153105662	2,747735729		
6	1,033941625	2,149355674	2,570084556		
7	1,122991482	2,067364422	2,584927698		
8	1,104939997	1,999789827	2,489862626		
9	1,10438059	2,031524575	2,594754375		
10	1,123306716	1,981796231	2,406013087		
11	1,107657669	1,794008522	2,306879907		
12	1,009571932	1,808651897	2,402535619		

Figure 2: Portfolio analysis Excel template file

The template file contains an example of historical prices for different assets. The name of each asset must be reported at the head of each column. The columns must have the same length.

In the "Dataset name" field the user adds a description to the computation in order to store the calculations in the "Previous simulations" page.

The last input is "Number of simulated Portfolios", i.e. the number of randomly generated portfolios that are generated in the standard-deviation-expected return plane, see Figure 3 are simulated by using randomly generated weights assuming no short selling. In practice, we simulate N independent uniforms in the range $[0,1]$, and then we normalize them so that their sum is equal to 1.

The output

Starting from the series of prices, the procedure computes the log-returns, the related returns vector and the variance-covariance matrix. The efficient frontier (Figure 3) and the portfolio composition (Figure 4) graphs are produced.

Efficient frontier

The dark blue points in Figure 3 are the combination of expected return and standard deviation of each efficient portfolio, starting from the expected return on the portfolio at global minimum variance[2] to the one having the largest expected return[3]. The light blue points represent the combinations of expected return and standard deviation of simulated portfolios. The orange points represent the expected return-standard deviation of the individual assets.

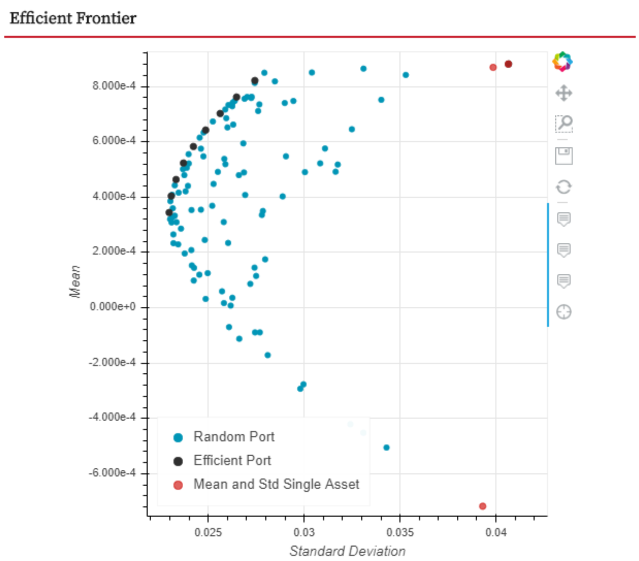


Figure 3: Efficient Frontier

Optimal weights

In the Figure 4 we show the efficient portfolio composition as the expected return increases. The weights are never greater than one (axis-Y) or negative because we do not allow for short selling. The legend contains the names of the individual stocks as given in the Excel file. The colours in the legend allow the user to identify the position of the individual stocks in the standard deviation expected return plane.

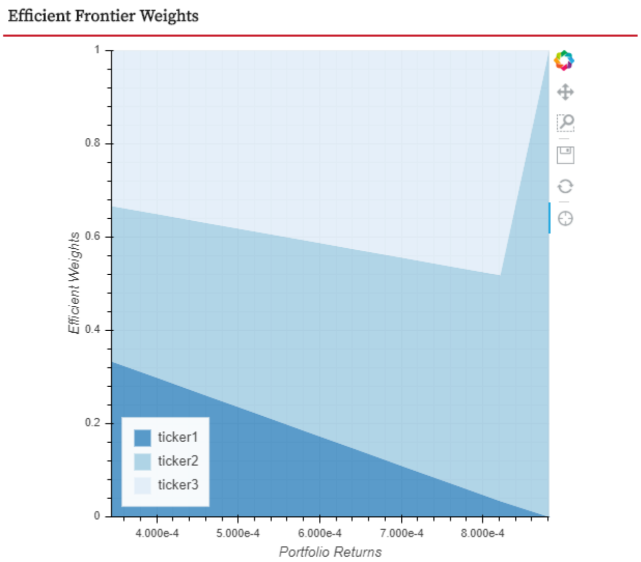


Figure 4: Efficient weights

Chart Tools

In each chart there are interactive tools positioned at the top right. Here are listed all of them starting from the first one.

- ❑ Bokeh Logo: hyperlink to access the bokeh site. Bokeh is the library used to create all interactive graph in the web-application.
- ❑ Pan Tool: the pan tool allows the user to pan the plot by left-dragging a mouse or dragging a finger across the plot region.
- ❑ Box Zoom: the box zoom tool allows the user to define a rectangular region to zoom the plot bounds too, by left-dragging a mouse, or dragging a finger across the plot area.
- ❑ Save: the save tool pops up a modal dialog that allows the user to save a PNG image of the plot.
- ❑ Reset: The reset tool turns off all the selected tools.
- ❑ Hoover Tool: the hover tool will generate a “tabular” tooltip where each row contains a label, and its associated value.

[2] This portfolio is found by solving the Markowitz problem without the expected return constraint

[3] This portfolio is fully invested in the stock having the largest expected return