

# R Portfolio Optimization Project

Philipp Jung, Ilyana Pekova, Georg Velez

16.03.2019

## Contents

<b>1</b>	<b>Minimum-Variance-Portfolio</b>	<b>1</b>
1.1	Theory . . . . .	1
1.2	Solving the system of linear equations . . . . .	2
1.3	in-sample Sharpe ratio . . . . .	3
1.4	out-sample Sharpe ratio . . . . .	3
1.5	Implementation . . . . .	4
1.6	Datasets and data preparation . . . . .	4
1.6.1	Computing the in-sample Sharpe ratio . . . . .	5
1.7	Computing the out-sample Sharpe ratio . . . . .	5
1.8	Discussion . . . . .	6
1.8.1	Graphical Analysis . . . . .	6
<b>2</b>	<b>Bayes-Stein-Portfolio</b>	<b>7</b>
2.1	Theory . . . . .	7
2.2	Execution . . . . .	8
2.3	Discussion . . . . .	8
<b>3</b>	<b>Sample-based mean-variance-Portfolio</b>	<b>8</b>
3.1	Theory . . . . .	8
3.2	Execution . . . . .	8
3.3	Discussion . . . . .	9
	<b>References</b>	<b>9</b>

Introduction text to our project.

## 1 Minimum-Variance-Portfolio

An investment strategy based on a Minimum Variance Portfolio seeks to minimize the risk of an investment. There is thus no desired target-return or stock-forecast to be considered. The portfolio optimization process can be described as pure risk-minimization, where the goal is a determination of the weight-distribution yielding the lowest possible risk at any time. Graphically, one can think of a Minimum Variance Portfolio as the most left point of the mean-variance-frontier.

### 1.1 Theory

As written in the paper by DeMiguel et al.[DGR09], the weights were chosen according to the portfolio that minimizes the variance of return, e.g.

$$\min_{w_t} w_t^\perp \Sigma_t w_t \quad (1)$$

under the restriction that

$$1^\perp_N w_t \stackrel{!}{=} 1 \quad (2)$$

and

$$\Sigma_t w_t \stackrel{!}{=} 1, \quad (3)$$

where  $w_t \in \mathbb{R}^N$  is the weight-vector at time  $t \in \{1, 2, \dots, T\}$  with  $T \in \mathbb{N}$  period under observation and  $N \in \mathbb{N}$  number of assets considered.  $\Sigma_t \in \mathbb{N} \times \mathbb{N}$  names the covariance matrix of excess-returns  $R_\tau \in \mathbb{R}^{N \times \tau}$  at time  $t$  for a subset  $\tau$  of the period of Observation  $\tau \subseteq \{1, 2, \dots, T\}$ . In the equation above, 1 is thought of as the N-dimensional vector containing 1s.

We now want to investigate the restriction  $\min_{w_t}$  for one fixed time-period  $T$ , such that  $\tau = T$ . As a result, weights are not time-dependent anymore. Note that when constructing a Minimum Variance Portfolio, weights are in fact time-dependent. However, for the sake of legibility, we will for now treat the case where  $w_t = w$ .

Considering three risky assets and a weight vector  $w \in \mathbb{R}^3$ , such that

$$w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}, \quad (4)$$

the minimal variance restraint can then be interpreted as follows[Ziv13, p. 7]:

$$\begin{aligned} \min_{w_1, w_2, w_3} \sigma_{w_1, w_2, w_3}^2 = & w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + w_3^2 \sigma_3^2 + \\ & 2w_1 w_2 \sigma_{12} + 2w_1 w_3 \sigma_{13} + 2w_2 w_3 \sigma_{23} \end{aligned} \quad (5)$$

Here, the components' variance  $\sigma_i^2$  and standard deviation  $\sigma_{i,j}$  for  $i, j \in \{1, 2, 3\}$  are used to describe the relationship. The Lagrange-function of this problem can be written as

$$\begin{aligned} L(w_1, w_2, w_3, \lambda) = & w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + w_3^2 \sigma_3^2 + \\ & 2w_1 w_2 \sigma_{12} + 2w_1 w_3 \sigma_{13} + 2w_2 w_3 \sigma_{23} \\ & + \lambda(w_1 + w_2 + w_3 - 1) \end{aligned} \quad (6)$$

Now, the component-wise deviates can be found and be set to equal zero, yielding the first order conditions.

$$\begin{aligned} 0 &= \frac{\partial L}{\partial w_1} = 2w_1 \sigma_1^2 + 2w_2 \sigma_{12} + 2w_3 \sigma_{13} + \lambda \\ 0 &= \frac{\partial L}{\partial w_2} = 2w_2 \sigma_2^2 + 2w_1 \sigma_{12} + 2w_3 \sigma_{23} + \lambda \\ 0 &= \frac{\partial L}{\partial w_3} = 2w_3 \sigma_3^2 + 2w_1 \sigma_{13} + 2w_2 \sigma_{23} + \lambda \\ 0 &= \frac{\partial L}{\partial \lambda} = w_1 + w_2 + w_3 - 1 \end{aligned} \quad (7)$$

In this,  $\lambda$  is the Lagrange multiplier. Equation 7 can be expressed as a system of linear equations:

$$\begin{pmatrix} 2\sigma_1^2 & 2\sigma_{12} & 2\sigma_{13} & 1 \\ 2\sigma_{12} & 2\sigma_2^2 & 2\sigma_{23} & 1 \\ 2\sigma_{13} & 2\sigma_{23} & 2\sigma_3^2 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad (8)$$

Now we can clearly see that the equation has the form of

$$\begin{pmatrix} 2\Sigma & 1 \\ 1^\top & 0 \end{pmatrix} \begin{pmatrix} w \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (9)$$

where  $\Sigma$  is just the covariance-matrix of the matrix of excess-returns  $R_\tau$ .

## 1.2 Solving the system of linear equations

Above equation is of form

$$A \cdot x = b, \quad (10)$$

thus a system of linear differential equations. Since  $A = \Sigma$  and  $b = 1$  are known, we can solve this system to obtain  $x = w_t$  the weight-vector at time  $t$ . Note that the weight-vector  $\mathbf{x}$  may contain negative which are smaller than zero. These negative weights are interpreted as short sales<sup>1</sup>. While short sales are in reality strongly regulated and restricted, the Minimum Variance Portfolio created in this work will not consider any restrictions.

<sup>1</sup>[https://en.wikipedia.org/wiki/Short\\_\(finance\)](https://en.wikipedia.org/wiki/Short_(finance))

### 1.3 in-sample Sharpe ratio

The in-sample Sharpe ratio for one strategy  $k$  is a basic measure of performance. Only one weights-vector is computed for all results. According to DeMiguel et al. [DGR09, p. 1928] it can be computed by

$$\widehat{SR}_k^{IS} = \frac{\hat{\mu}_k^{IS\perp} \hat{w}_k}{(\hat{\mu}_k^{\perp\perp} \Sigma \hat{\mu}_k)^{1/2}}. \quad (11)$$

Here,  $\hat{\mu}_k^{IS\perp}$  is the mean-returns vector for strategy  $k$ , whereas  $\Sigma$  is the covariance-matrix of the excess-returns matrix  $R$ .  $\hat{w}_k$  is the mean weights-vector of strategy  $k$ .

### 1.4 out-sample Sharpe ratio

Computing the out-sample Sharpe ratio, the weights' time-dependency is addressed using a "rolling sample" approach [DGR09, p.1927]. These "rolling samples" are formed by defining continuous subsets  $t_i$  of the original time-series with fixed length of 120 consecutive time-steps,  $i \in \{1, 2, \dots, T - 120\}$  where  $T$  is the last time-step of the time-series. This time-interval is then shifted progressively through the entire length of the time-series.

	Date	A_returns	B_returns
	01-2001	0.002	-0.001
t_1	02-2001	0.005	-0.005
	03-2001	-0.001	-0.008
t_3	04-2001	0.005	0.008
	05-2001	0.006	0.010
t_5	06-2001	0.007	0.001
	07-2001	0.000	0.004

Figure 1: Illustration of the rolling-window approach for a time-series containing seven time-steps filled with mock-data. Five subsets of length 3 divide the time-series.

This approach is schematically described in figure 1.

Finally, for each subset  $t_i$ , the weights are computed by solving the system of linear equations as in 10. Out-of-sample excess returns are then calculated using the actual excess returns of the  $i + 1$ 'th time-step in an attempt to validate if a prediction in the past had been successful.

Lastly, the out-of-sample returns' arithmetic mean  $\hat{\mu}$  is divided by its standard deviation  $\hat{\sigma}$  to derive the out-of-sample Sharpe ratio like so:

$$\hat{SR} = \frac{\hat{\mu}}{\hat{\sigma}} \quad (12)$$

## 1.5 Implementation

In this section we want to display and discuss the implementation of the approach described in the previous chapter in the R programming language. We will compute several performance measures of the Minimal Variance Portfolio we're about to derive from the data.

## 1.6 Datasets and data preparation

The first dataset upon which this analysis is based is called "Ten sector portfolios of the S&P 500 and the US equity market portfolio". It has been created by Roberto Wessels and was obtained from Mendeley Data<sup>2</sup>. Note that for this dataset, it is necessary to subtract the Treasury Bill-rates from each entry to obtain adjusted returns stripped of the risk-free rate. We call this dataset **TSP**.

The other dataset concerned is provided by Kenneth French and downloaded from his website<sup>3</sup>. The dataset is called "SMB and HML portfolios and the US equity market portfolio" and contains the factors "Small Minus Big" (SMB) and "High Minus Low" (HMB) from the Fama-French three-factor model<sup>4</sup> as well as the whole S&P500 portfolio. We call this dataset short **SHS**. This dataset is already cleared from the risk-free rate, thus a Treasury Bill-rate adjustment is not necessary.

In listing 1, it is showcased how data is imported and loaded into a dataframe. In line 1, the .csv-file containing the data is loaded via the `read.csv` function, where a separator as well as the header-parameter are specified.

```
1 return_matrix = read.csv('.. /ten-roberto-wessels.csv', header=TRUE, sep=',')
2 sub_return_matrix = subset(return_matrix, select = -c(13))
3 sub_return_matrix = apply(sub_return_matrix[, -1], 2, '-', return_matrix[, 13])
4 return_matrix = cbind(return_matrix$Date, sub_return_matrix)
5 colnames(return_matrix)[1] = 'Date'
```

Listing 1: Loading the data and subtracting the risk-free rate from the columns of returns yields a dataframe containing the matrix of excess-returns.

In case of dataset TSP, it is necessary to subtract the risk-free rate from the returns. To do this, a subset of the `return_matrix` is created in line 2, excluding the monthly risk-free rate which is stored in column 13 of the `return_matrix`. In line 3, the risk-free rate is subtracted from every column the `return_matrix`' subset, except for the column containing the date. Line 4 shows the reconstruction of the `return_matrix`, now containing risk-free

<sup>2</sup>Data can be downloaded here. It's also attached to this report.

<sup>3</sup>The dataset from Kenneth French's website can be downloaded here and is also attached to this report.

<sup>4</sup>For an introduction to the Fama-French three-factor model, see Wikipedia.

returns. After the dates' column-name has been reset in line 5, the `return_matrix` now contains only risk-free returns.

### 1.6.1 Computing the in-sample Sharpe ratio

A very basic way of measuring portfolio performance is provided by the in-sample Sharpe ratio. By setting  $\tau = \{1, 2, \dots, T\}$ , the entire time-series is considered when computing the covariance matrix of the matrix of excess-returns  $\Sigma_t$ , effectively removing the weight-vectors' time-dependency.

Following the procedure leading to the system of linear equations in equation 10, we first compute a covariance matrix of the entire risk-free return-matrix. Line 1 of listing 2 does just this, excluding the column containing the time-series dates. A vector of dimension equal to the amount of assets listed in the `return_matrix` is constructed, where every component of the vector equals one.

```

1 cov_return_matrix = cov(return_matrix[, -1])
2 b = vector(length=nA) + 1
3 weights = solve(cov_benchmark_matrix, b)
4 weights = weights/abs(sum(weights))
5 mean_returns = data.matrix(apply(benchmark_matrix[, -1], 2, mean))
6 weights = data.matrix(weights)
7 Sharpe_ratio_IS = t(mean_returns) %*% weights / sqrt(t(weights) %*% cov_benchmark_matrix %*% weights)
8 print(Sharpe_ratio_IS)
```

Listing 2: This example shows how the in-sample Sharpe ratio for a matrix of expected returns, computed in the R programming language.

Line 3 of listing 2 uses the `solve()` function to solve the linear equation by inverting the covariance matrix. The resulting `weights`-vector is normalized in line 4. A vector containing the mean-values of `return_matrix` is created in line 5 and the `weights`-vector is converted to a `data.matrix` in line 6. This is done to facilitate the vector- and matrix-multiplication performed in line 7 to calculate the in-sample Sharpe ratio according to equation 11, which is printed in line 7.

### 1.7 Computing the out-sample Sharpe ratio

For computing the out-sample Sharpe ratio as described in section 1.4, the approach for deriving the in-sample Sharpe ratio needs to be altered to incorporate time-dependency.

Considering listing 3, in lines 1 and 2, the number of assets as well as the length of the time-series are assigned respectively. `M` contains the length of the rolling-sample, the iterator-variable `t` is initialized in line 4 and adjusted to `M`.

```

1 nA = dim(return_matrix[, -1])[2] # number of assets
```

```

2 T = dim(return_matrix)[1] # no of months
3 M = 120 # size of rolling-sample
4 t = M + 1 # adjust t_0 to size of rolling-sample
5 while (t <= T - 1) {
6   tStart = t - M
7   tEnd = t
8   return_matrix_future = return_matrix[c(tEnd+1),]
9   cov_return_matrix_t = cov(data.matrix(return_matrix[c(tStart:tEnd),-1]))
10  b_t = vector(length=nA) + 1
11  x_t = solve(cov_return_matrix_t, b_t)
12  x_t = x_t/sum(x_t)
13
14  if(t == M+1) {
15    history_of_weights = data.frame(t(x_t))
16    mu = data.frame(return_matrix[,1][t], x_t %*% return_matrix_future[-1])
17  }
18  else {
19    history_of_weights = rbind(history_of_weights, x_t)
20    mu = rbind(mu, data.frame(return_matrix[,1][t], x_t %*% return_matrix_
    future[-1]))
21  }
22  t = t+1
23 }

```

Listing 3: Computing the out-sample Sharpe ratio from the matrix of expected returns in R.

In line 5, the sample gets rolling: By setting a `while`-loop, we iterate over the whole time-series. The last date included in the rolling-sample is  $T-1$ , since it is the last value for which the expected returns of the next month are known. From line 6 to 12, the weights for time  $t$   $x_t$  are computed. They are stored in dataframe `mu` in line 15 and 19 respectively. Meanwhile, the weights are saved in a dataframe in line 15 and 19.

## 1.8 Discussion

To validate the implementations of the minimum-variance portfolio, in this section metrics derived from the data will be compared with results from DeMiguel et al.[DGR09]. Additionally, results will be graphically displayed and compared.

### 1.8.1 Graphical Analysis

As a first graphical display of results, a portfolio appreciation graph is constructed by cumulatively summing the returns and scaling them by a factor 100. Is is similar to having invested 100\$ in a specific portfolio at the beginning of the time-series considered. In figure 2, the portfolio appreciation graph of the minimum variance portfolio, which has



been created in previous chapter 1.7, as well as the portfolio appreciation graph of the naive investing strategy are displayed. When applying a naive investing strategy as in [DGR09], the weight-vector  $w_{naive}$  is of dimension  $N$  and every component of  $w_{naive}$  is of value  $1/N$ , where  $N$  is the number of assets considered.

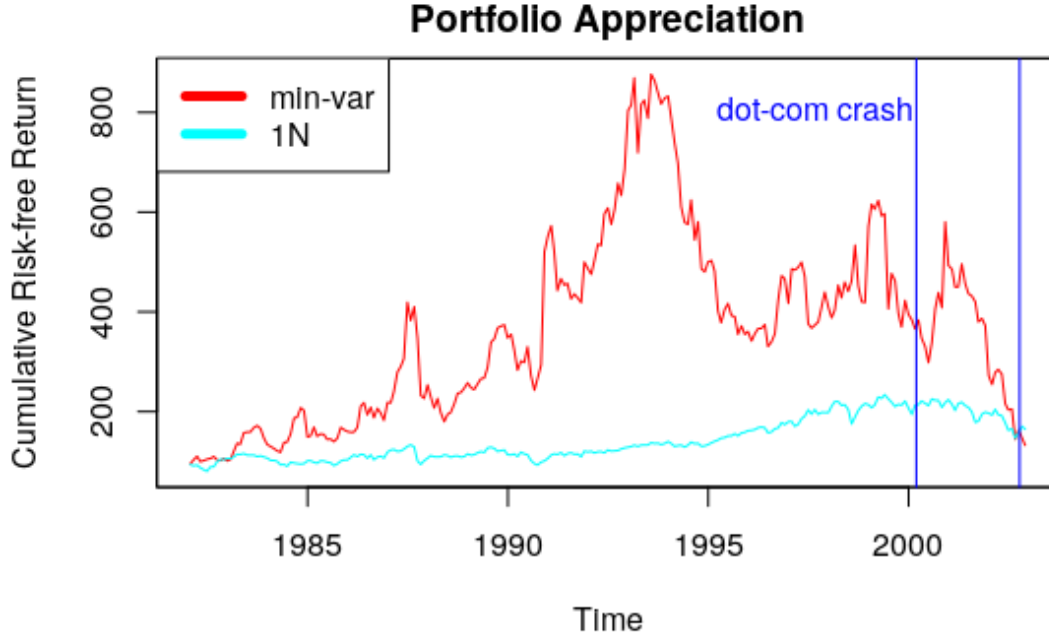


Figure 2: Illustration of the rolling-window approach for a time-series containing seven time-steps filled with mock-data. Five subsets of length 3 divide the time-series.

## 2 Bayes-Stein-Portfolio

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

### 2.1 Theory

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat,

sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

## 2.2 Execution

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

## 2.3 Discussion

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

# 3 Sample-based mean-variance-Portfolio

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

## 3.1 Theory

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

## 3.2 Execution

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata

sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

### 3.3 Discussion

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

## References

- [DGR09] Victor DeMiguel, Lorenzo Garlappi, and Uppal Raman. “Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?” In: *The Review of Financial Studies* 22 (5 2009), pp. 1915–1953. DOI: <http://dx.doi.org/hhm075>. URL: <https://hal.archives-ouvertes.fr/hal-01850926/document>.
- [Ziv13] Eric Zivot. *Portfolio Theory with MatrixAlgebra*. 2013. URL: <https://faculty.washington.edu/ezivot/econ424/portfolioTheoryMatrix.pdf>.