Master-Thesis

Axel Roth

2022-08-14

Contents

Preface

Blah blah blah

Chapter 1

Abstract

Things about this thesis. why and what question should be answered. and what are the answers. (zusammenfassung)

Chapter 2

Software information and usage

wie ich das buch schreibe, R markodwn bookdown und so und welche versionen ich nutze

2.1 R-Version and Packages

2.2 reproducibility

github und code im bookdown

2.3 R-functions

zb plotly_save

Chapter 3

Open Data Sources

yahoo finance

3.1   R-functions

Chapter 4

Mathematical Fundations

This chapter is a summary of mathematical calculations and conventions used in this Thesis. Its impotent to note that most mathematical formulas are expressed in matrix notation. Most of the time this will result in a direct translation into R-code. Additionally we state all assumptions needed for the modeled return structure, that can be easily used. It is important to say that the reality is too complex and can only partially be modeled. That is why we use fundamental and simple models that do not hold in practice. Nonetheless these models or improvements of it are widely used in finance and have proven to be useful. Additionally do these basic models produce more complex problems to solve by the PSO because for example factor models are often reduce the dimension of the dependency that are modeled.

4.1 Basic Operators

The following table is a compendium which will compare the used mathematical symbols to R-code and its meanings:

Latex/Displayed	R-Code	Meaning
×	%*%	Matrix or vector multiplication
A^T	t(A)	transpose of Matrix or vector A
·	*	Scalar or elementwise vector multiplication

4.2 Return calculation

Returns are fundamental for any portfolio optimization approach based on historical information. These returns reflect the percentage change from one point in time to another and are calculated from adjusted closing prices. Adjusted closing prices are reflecting dividends and are cleaned of by stock splits and rights offerings. These Returns enable the comparison of assets and are fundamental to analyze dependencies.

4.2.1 daily returns

We use exclusively simple returns in this thesis and the default timeframe for all rawdata is one workday. Suppose we have the adjusted closing price P of one asset on workday t_i and t_{i+1} so the simple returns can be calculated as follows:

$$R_{i+1} = \frac{P_{t_{i+1}}}{P_{t_i}} - 1$$

4.2.2 annualized returns

4.3 Markowitz Modern Portfolio Theory (MPT)

Harry Markowitz published his first groundbreaking work at 1952 and had a huge impact on the current financial point of view. Mainly by describing the impact of diversification and efficient portfolios. Efficient portfolios are those who have minimum risk for the given expected return or maximum expected return for a given risk target. Diversification can be described as "a portfolio has the same return but less variance as the sum of its parts". This holds if the assets are not perfectly correlated because bad and good movers can compensate each other which will result in a reduction of extreme situations. More detailed information can be found at (Maringer, 2005).

4.3.1 Assumptions of Markowitz Portfolio Theory

A good summarirization of Markowitz assumptions can be found in (Maringer, 2005):

1. perfect market without taxes or transaction costs.

2. short sales are disallowed.

3. assets are infinitely divisible.

4. Expected Returns, Variances and Covariances contain all information.

5. Investors are risk-adverse, they will only accept greater risk if they are compensated with a higher expected return.

It is not necessary to assume that the returns are normally distributed, but we will use it here, to simplify the problem. More information about the conditions for using other distributions can be found in (Maringer, 2005). It is clear that these assumption do not hold in practice.

4.4 Portfolio Math

This Section will provide proofs for the basic calculations needed for portfolio optimization like shown in (Zivot, 2021). Hopefully this will enable all readers to understand the stated formulas completely. We prefer a different notation of the returns than in most sources, because its the common data-format used in practice. The used notations are the random return vector R and the portfolio weights w for N assets:

$$R = \begin{bmatrix} R_1 & R_2 & \cdots & R_N \end{bmatrix}, \quad w = \begin{bmatrix} w_1 \\ w_2 \\ \cdots \\ w_N \end{bmatrix}$$

And we simplify each return to be normally distributed with $R_i = \mathcal{N}(\mu_i, \sigma_i^2)$ in this thesis. QUELLLLEEEEE It follows that linear combinations of normally distributed random variables are jointly normal distributed and can be fully described by its mean, variance and covariance.

4.4.1 expected returns

The expected returns of the vector of normal distributed random variables $R \in \mathbb{R}^N$ can be calculated as follows:

$$E[R] = \begin{bmatrix} E[R_1] & E[R_2] & \cdots & E[R_N] \end{bmatrix} = \begin{bmatrix} E[\mu_1] & E[\mu_2] & \cdots & E[\mu_N] \end{bmatrix} = \mu$$

expected portfolio return The linear combination of the expected returns μ with a weighting vector w (e.g. portfolio weights) can be calculated as follows:

$$\mu \times w = \begin{bmatrix} E[\mu_1] & E[\mu_2] & \cdots & E[\mu_N] \end{bmatrix} \times \begin{bmatrix} w_1 \\ w_2 \\ \cdots \\ w_N \end{bmatrix} = E[\mu_1] \cdot w_1 + E[\mu_2] \cdot w_2 + \cdots + E[\mu_N] \cdot w_N = \mu_P$$

4.4.2 portfolio returns

Let $R \in \mathbb{R}^{T \times N}$ denote a realized return Matrix of N assets and T days in the past. The portfolio return on each day can be calculated with the formula of the expected portfolio return. On all days T can this be done by:

$$R \times w = \begin{bmatrix} R_{1,1} & R_{1,2} & \cdots & R_{1,N} \\ R_{2,1} & R_{2,2} & \cdots & R_{2,N} \\ \cdots & & & \\ R_{T,1} & R_{T,2} & \cdots & R_{T,N} \end{bmatrix} \times \begin{bmatrix} w_1 \\ w_2 \\ \cdots \\ w_N \end{bmatrix} = \begin{bmatrix} R_1^P \\ R_2^P \\ \cdots \\ R_N^P \end{bmatrix} = R_P$$

Chapter 5

Activ vs Passiv Investing

The fundation of each Asset Management

passiv vs activ studie `https://www.scirp.org/journal/paperinformation.aspx?paperid=92983`

gut gut file:///C:/Users/Axel/Desktop/Master-Thesis-All/Ziel%20was%20beantwortet%20werden%20soll/Quellen%20nur%20wichtige/Rasmussen2003_Book_QuantitativePortfolioOptimisat.pdf

```r
source("global.R")
knitr::opts_chunk$set(cache = T, warning=F, message=F)
```

Chapter 6

Challenges of Passiv Investing

In this Chapter we will discuss two common challenges of Passiv-Investing and create simple examples to test the PSO. The first one is the mean-variance portfolio (MVP) from the modern portfolio theory of Markowitz which is simply said an optimal allocation of assets regarding risk and return. The second challenge is the index-tracking-problem which tries to construct a portfolio which has a minimal tracking error to a given benchmark.

6.1 Mean-variance portfolio (MVP)

Markowitz has shown that diversifying the risk on multiple assets will reduce the overall risk of the portfolio. This result was the beginning of the widely used modern portfolio theorie which uses mathematical models to archive portfolios with minimal variance for a given return target. All these optimal portfolios for a given return target are called efficient and create the efficient frontier.

6.1.1 MVP

We have N assets and its returns on T different days which creates a return matrix $R \in \mathbb{R}^{T \times N}$. Each element $R_{t,i}$ contains the return of the i-th asset on day t. The covariance matrix of the returns is $\sum \in \mathbb{R}^{N \times N}$ and the expected returns are $\mu \in \mathbb{R}^N$. The MVP with risk aversion parameter $\lambda \in [0, 1]$ like shown in (Maringer, 2005) can be formalized as follows:

$$\underset{w}{minimize} \quad \lambda \, w^T \sum w - (1 - \lambda) \, \mu^T w \tag{6.1}$$

The risk aversion parameter λ defines the trade-off between risk and return. With $\lambda = 1$, the minimization problem only contains the the variance term and

so on results in a minimum variance portfolio and $\lambda = 0$ transforms the problem to a minimization of the negative expected returns so on results in a maximum return portfolio. All possible $\lambda \in [0,1]$ represent the efficient frontier.

6.1.2 MVP example

We will analyze a small example to understand the meaning of the efficient frontier without going into detail how it was solved. First of all we are loading the daily returns of IBM, Google and Apple from the year 2020.

The cumulated daily returns are:
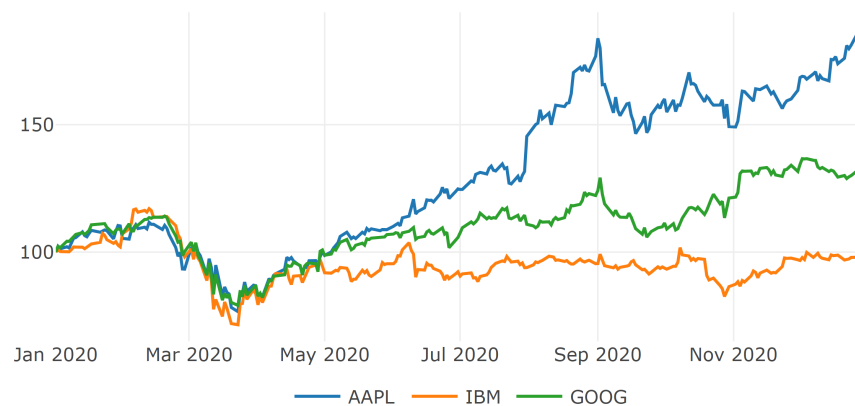
Now we can calculate the expected daily returns and the covariance matrix for the 3 assets:

```
mu <- as.vector((last(ret_to_cumret(returns))/100)^(1/nrow(returns))-1) %>%
  setNames(., colnames(returns))
mu
```

```
##            AAPL            IBM            GOOG
##   0.00237641721 -0.00004622149   0.00106873786
```

```
cov <- as.matrix(nearPD(cov(returns))$mat)
cov
```

```
##               AAPL           IBM          GOOG
## AAPL 0.0008635696 0.0004356282 0.0005337719
## IBM  0.0004356282 0.0006626219 0.0004086728
## GOOG 0.0005337719 0.0004086728 0.0005827306
```

We now have all the necessary data to solve the MVP (6.1) with $\lambda \in \{0.01, 0.02, ..., 0.99, 1\}$. We calculate all 100 portfolios by solving the quadratic minimization problem for each λ.

Following that, we convert the daily returns and standard deviation to yearly returns and standard deviation before plotting the efficient frontier.
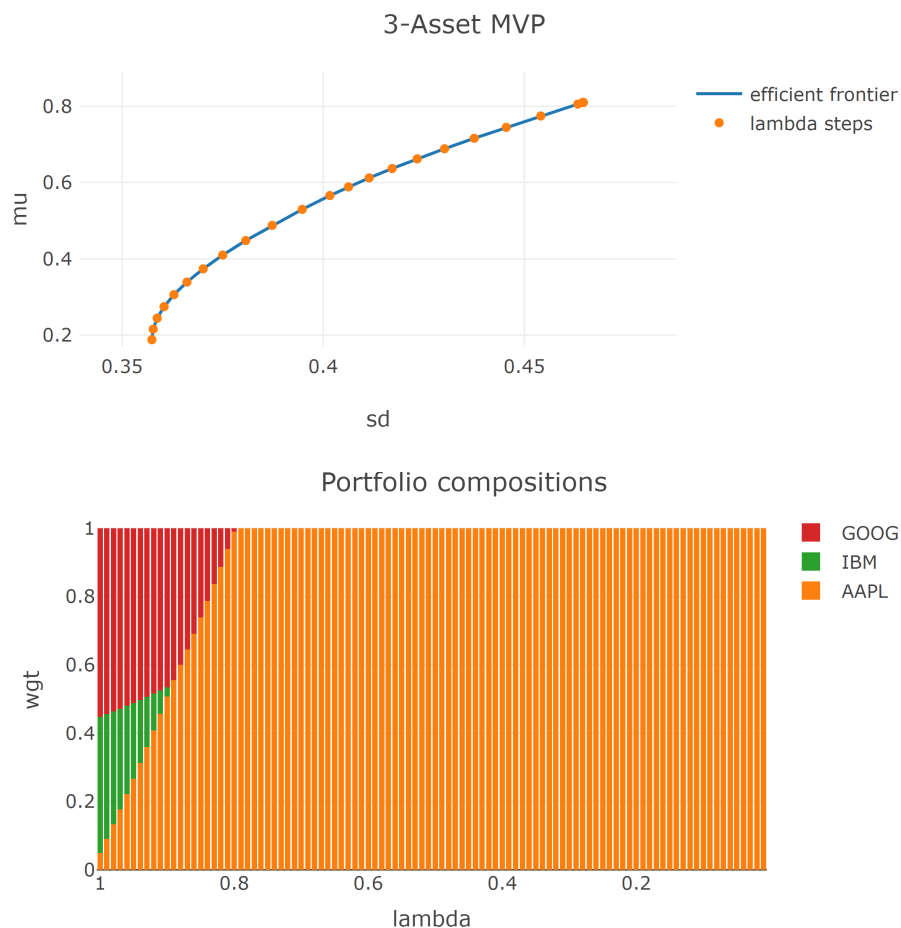
6.2   Index-tracking portfolio (ITP)

Indices are asset baskets that are used to track the performance of a specific asset group. The well-known Standard and Poor's 500 index (short: S&P 500), for example, tracks the top 500 stocks in the United States. All indices are not investible and only serve to visualize the performance of these asset groups without incurring transaction costs. Asset managers use such indices as benchmarks to compare the performance of their funds. Each fund has its own benchmark, which contains roughly the same assets that the manager could purchase. If the fund underperforms its benchmark, it may be an indication that the fund manager made a poor decision. That is why all fund managers strive to outperform their benchmarks through carefully chosen investments.

The past has proven that this is rearly achived with activ managemnt after costs (Desmond Pace and Grima, 2016). This is the reason why passiv managed funds with the goal to track there benchmarks are becoming more frequent. This is why passively managed funds with the purpose of tracking their benchmarks are becoming more common. This can be accomplished through either full or sparse replication. In most circumstances, a full replication that achieves the exact performance we seek is not achievable because not all assets in an index are investable. And, if so, it would be unwise because benchmarks with numerous indexes can contain over ten thousand separate assets, resulting in a massive amount of transaction costs. A sparse replication of the performance is the most prevalent approach. To do so, the portfolio manager must define his benchmark, which should overlap with his fund's investing universe. Following that, he will reduce this universe using investor principles such as liquidity and availability. Now he can begin to optimize a portfolio, taking into account the investor constraints, in order to match the benchmark performance. Typically, this is accomplished by lowering the variance between the ITP's daily returns and the benchmark:

$$minimize \ \ Var(r_p - r_{bm})$$

To obtain the portfolio weights w, we must first substitute r_p as shown below:

$$r_p = R * w$$

The Variance is then solved up until a quadratic problem dependent on the portfolio weights w is represented:

$$Var(r_p - r_{bm}) = Var(R*w - r_{bm}) = Var(R*w) + Var(r_{bm}) - 2 \cdot Cov(R*w, r_{bm})$$

We must now solve each of the three terms, beginning with the easiest.

$$Var(r_{bm}) = \sigma_{bm}^2 = constant$$

The variance of the portfolio can be solved by looking at the Portfolio math Using Matrix Algebra section in (Zivot, 2021):

$$Var(R * w) = w^T * Cov(R) * w$$

And the last term can be solved in general as (https://bookdown.org/compfin ezbook/introcompfinr/Multivariate-Probability-Distributions-Using-Matrix-Algebra.html 3.6.5):

$$Cov(A*a, b) = Cov(b, A*a) = E[(b-\mu_b)(A*a-\mu_A*a)] = E[(b-\mu_b)(A-\mu_A)*a] = E[(b-\mu_b)(A-\mu_A)]*a = Cov(A,$$

```r
A = matrix(c(1,4,2,4,6,3,8,4,4,10), ncol=2)
a = c(0.2, 0.8)
b = c(4,4,5,5,7)

cov(A %*% a, b)
```

```
##      [,1]
## [1,] 2.15
```

```r
t(a) %*% cov(A, b)
```

```
##      [,1]
## [1,] 2.15
```

```r
t(cov(A, b)) %*% a # das hier wird gebraucht
```

```
##      [,1]
## [1,] 2.15
```

This results in the final formula of the ITP:

$$
\begin{aligned}
Var(r_p - r_{bm}) &= Var(R \times w - r_{bm}) \\
&= Var(R \times w) - 2 \cdot Cov(R \times w, r_{bm}) + Var(r_{bm}) \\
&= w^T \times Cov(R) \times w - 2 \cdot Cov(r_{bm}, R)^T \times w + \sigma_{bm}^2
\end{aligned}
\tag{6.2}
$$

The minimization problem of the ITP in the general stricture which all optimizers need is:

$$min(\frac{1}{2} \cdot b^T \times D \times b - d^T \times b)$$

Because it is a minimization we can ignore constant terms and stretching coefficients and still find the same minimum. This results in the general substitution of the ITP as follows:
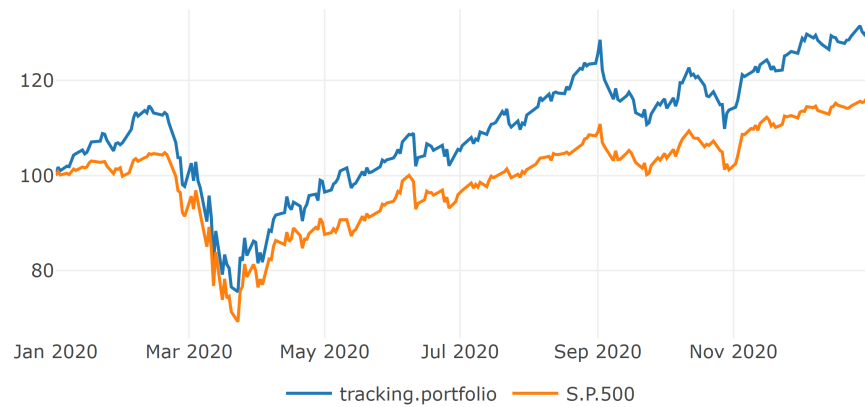
$$D = Cov(R)$$

and

$$d = Cov(r_{bm}, R)$$

Now we need to add some basic constraints like in the MVP, to sum up the weights to 1 and being long only.

6.2.1 Example ITP

We will show the results of tracking the S&P 500 with a tracking portfolio that can only invest in IBM, Apple and Google without going into details:

```
##      AAPL      IBM      GOOG
## 0.2677928 0.4041880 0.3280192
```

Chapter 7

Analytical_Solver

example to solve problems with analytical solvers

Chapter 8

Simple_Particle_Swarm_Optimization

first pso examples and explainations

Bibliography

Desmond Pace, J. H. and Grima, S. (2016). Active versus passive investing: An empirical study on the us and european mutual funds and etfs.

Maringer, D. (2005). *Portfolio Management with Heuristic Optimization.*

Zivot, E. (2021). *Introduction to Computational Finance and Financial Econometrics with R.*