

# Optimal portfolios in higher dimensions - estimation and uncertainty

Erik Thorsén

September 20, 2021

## Acknowledgements

## List of papers



# Contents

<b>1</b>	<b>Introduction</b>	<b>7</b>
<b>2</b>	<b>Modern Portfolio Theory</b>	<b>9</b>
2.1	Relationship between assets and the (inverse) Covariance matrix . . . . .	10
<b>3</b>	<b>Statistical models and estimation</b>	<b>13</b>
3.1	Estimation - the name of the game . . . . .	14
3.2	Simulations, inverses and why stochastic representations are valuable . . . . .	14
<b>4</b>	<b>Higher dimensions or portfolios with many assets</b>	<b>15</b>
4.1	Random Matrix Theory and Stieltjes transforms . . . . .	15
<b>5</b>	<b>Summary of Papers</b>	<b>17</b>
<b>6</b>	<b>Future research</b>	<b>19</b>



# Chapter 1

## Introduction - making decisions and allocations

test test

Understanding variation





## Chapter 2

# Modern Portfolio Theory

Modern Portfolio Theory was introduced by Markowitz (1959). In his seminal work he argued that any portfolio which simply maximize its profit will result in a naive solution. Very much like the example in Chapter 1. The future is unknown and (one of) the best model(s) we have for it is stochastic Investing all your capital in the asset with the highest return is not sensible if you do not know the future. Such an investment will cause you to take an extreme amount of risk. He therefore argued that any well diversified portfolio should be preferred to any non diversified portfolio. Such a portfolio can be obtained through many different procedures but he proposed the use of the first two moments for the allocation problem. If an asset has high return on average, then that should be accounted for though not at the cost of large amounts of risk. If an asset implies a lot of risk, then we should try to invest small amounts in it.

We assume assets  $\mathbf{x}$  are random with mean  $E(\mathbf{x}) = \boldsymbol{\mu}$  and covariance matrix  $\text{Var}(\mathbf{x}) = \boldsymbol{\Sigma}$ . Although there is usually little restriction on  $\boldsymbol{\mu}$  there is usually very specific restrictions on the covariance matrix  $\boldsymbol{\Sigma}$ . Since the covariance matrix is a subject of its own we dedicate the next section to it and disregards these restrictions for now. We will merely say that it is well behaved. The restrictions on the mean will be commented on below. Using the two moments for the asset returns the portfolio distribution  $x = \mathbf{w}^\top \mathbf{x}$  has mean  $E(x) = \mathbf{w}^\top \boldsymbol{\mu}$  and variance  $\text{Var}(x) = \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}$ . Let  $\mu_0$  be the target return that the investor would like to achieve from their portfolio and  $\mathbf{1}$  column vector of ones with appropriate dimensions. Markowitz (1959) considered the following optimization problem

$$\begin{aligned} & \underset{\mathbf{w}}{\text{minimize}} && \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w} \\ & \text{subject to} && \mathbf{w}^\top \mathbf{1} = 1 \\ & && \mathbf{w}^\top \boldsymbol{\mu} \geq \mu_0 \\ & && w_i \geq 0, i = 1, 2, \dots, p \end{aligned} \tag{2.0.1}$$

This problem is a quadratic optimization problem with linear equality and inequality constraints. The objective is to minimize the variance of the portfolio. A natural question is to ask whether or not that implies diversification? As it turns out, minimizing the portfolio variance will encourage diversification. **Give example on using convex combinations on variances.**

The constraint  $\mathbf{w}^\top \mathbf{1} = 1$  essentially states that the investor must invest all available money. The weights are scaled according to the amount of cash invested. The disposition is very different whenever an inequality is used rather than equality. As Hult et al. (2012) states, if  $\mathbf{w}^\top \mathbf{1} \leq 1$ , then the investor could be throwing money away since there is a lot of opportunity left in the market when investing. The second constraint describes the investors expectations on the portfolio. As  $\mu_0$  grows, the return of the portfolio will grow. However, that has implications for the objective. Increasing  $\mu_0$  will change the amount of variance the portfolio can achieve. Depending on the value  $\mu_0$  we would be accepting more risk. The last constraint is rather simple though have quite large implications. It states that the weights can not be negative which means that we can only invest money we have. A negative value of  $w_i$  in the  $i$ th asset is called a short position. You borrow the asset from someone who owns it and then sell it. Later on, you can buy it back, hoping that the price is less so that you can make money on such a position. For certain types of investors

this constraint can be limiting and for others its a must. In this thesis, we exclude it altogether. That is, this thesis considers

$$\begin{aligned} & \underset{\mathbf{w}}{\text{minimize}} && \mathbf{w}^\top \Sigma \mathbf{w} \\ & \text{subject to} && \mathbf{w}^\top \mathbf{1} = 1 \\ & && \mathbf{w}^\top \boldsymbol{\mu} \geq \mu_0 \end{aligned} \quad (2.0.2)$$

which is what we refer to the mean-variance optimization problem. The solution to this problem is very often stated relative to another famous portfolio, namely the Global Minimum Variance (GMV) portfolio and its related quantities (see e.g. Bodnar and Gupta (2009); Bodnar et al. (2013); Bauder et al. (2020)). We will continue in the same manner. Let  $\Sigma^{-1}$  denote the inverse of  $\Sigma$  and

$$\mathbf{w}_{GMV} := \frac{\Sigma^{-1} \mathbf{1}}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}}, \quad R_{GMV} := E(\mathbf{w}_{GMV}^\top \mathbf{x}) = \frac{\mathbf{1}^\top \Sigma^{-1} \boldsymbol{\mu}}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}}, \quad V_{GMV} := \text{Var}(\mathbf{w}_{GMV}^\top \mathbf{x}) = \frac{1}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}}. \quad (2.0.3)$$

#### explain the quantities.

The solution to the mean-variance problem in (2.0.2) is equal to

$$\mathbf{w}_{MV} = \frac{\Sigma^{-1} \mathbf{1}}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}} + \frac{\mu_0 - R_{GMV}}{V_{GMV}} \mathbf{Q} \boldsymbol{\mu}, \quad \mathbf{Q} = \Sigma^{-1} - \frac{\Sigma^{-1} \mathbf{1} \mathbf{1}^\top \Sigma^{-1}}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}}. \quad (2.0.4)$$

.....

The moments of this portfolio is equal to

$$E(\mathbf{w}_{MV}^\top \mathbf{x}) = R_{GMV} + \frac{\mu_0 - R_{GMV}}{V_{GMV}} \boldsymbol{\mu}^\top \mathbf{Q} \boldsymbol{\mu}, \quad \text{Var}(\mathbf{w}_{MV}^\top \mathbf{x}) = V_{GMV} + \left( \frac{\mu_0 - R_{GMV}}{V_{GMV}} \right)^2 \boldsymbol{\mu}^\top \mathbf{Q} \boldsymbol{\mu}. \quad (2.0.5)$$

If we set  $\mu_0$  equal to  $\boldsymbol{\mu}^\top \Sigma^{-1} \mathbf{1} / \mathbf{1}^\top \Sigma^{-1} \mathbf{1}$  then the portfolio is equal to the GMV portfolio. Excluding the constraint  $\mathbf{w}^\top \boldsymbol{\mu} \geq \mu_0$  all together results in the same solution. From equation (2.0.5) we can see that all values  $\mu_0$  are rescaled according to the moments of the GMV portfolio. If the target return  $\mu_0$  is not  $\mu_0 > R_{GMV}$  then you are better off with the GMV portfolio in terms of return and risk. However, if you choose a value  $\mu_0 > R_{GMV}$  then the portfolio return will be more than  $R_{GMV}$ .

**Introduce merton and the efficient frontier and talk about  $\boldsymbol{\mu}^\top \mathbf{Q} \boldsymbol{\mu}$ .**

## 2.1 Relationship between assets and the (inverse) Covariance matrix

The covariance matrix  $\Sigma$  and the precision matrix  $\Sigma^{-1}$  are fundamental to mean-variance portfolios. In this section we discuss the restrictions we place on the covariance matrix and what the precision matrix actually represent.

For a vector  $\mathbf{x}$  with finite second moment, the covariance matrix is defined as  $\Sigma = E((\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^\top)$ . It contains the variances of each individual element of  $\mathbf{x}$  on the diagonal as well as the covariance between every pair of elements on the off-diagonal. That is, each diagonal element corresponds to the univariate case where the variance is equal to  $E((x_i - \mu_i)^2)$ . A distribution is usually called degenerate or singular if the variance is equal to zero, which holds true in the multivariate context as well. If any diagonal element is zero, then the covariance matrix is called singular. However, due to the fact that we involve covariances on the off-diagonal we are forced to work with a broader definition. Since we work with real matrices in this thesis, we limit the definition accordingly. From Harville (1997, ch 14.2) we say that a real symmetric  $p \times p$  matrix  $\mathbf{A}$  is called

- positive definite if  $\mathbf{z}^\top \mathbf{A} \mathbf{z} > 0$
- positive semi-definite if  $\mathbf{z}^\top \mathbf{A} \mathbf{z} \geq 0$

for all nonzero vectors  $\mathbf{z} \in \mathbb{R}^p$ . In the multivariate case we need to assert that a quadratic form is (strictly) positive in comparison to the univariate setting where we can observe it through the variance. Positive- or semi-positive definite can be quite cumbersome to work with. We need to assert that the conditions

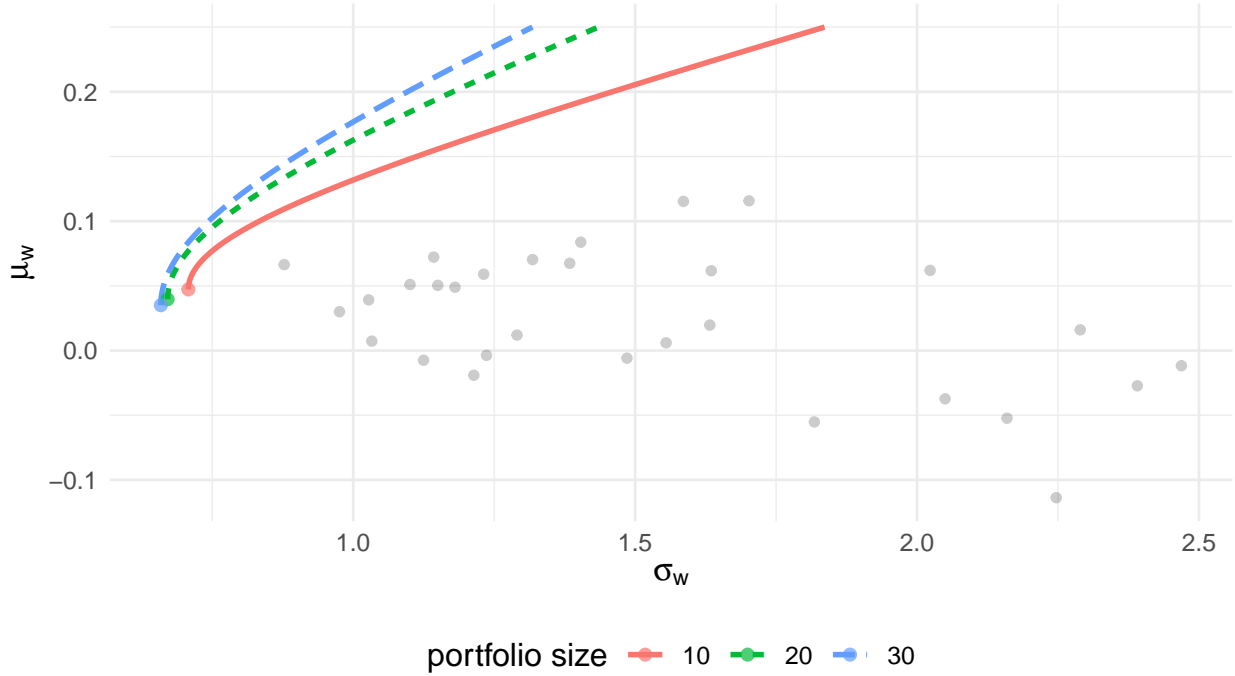


Figure 2.1: Three different efficient frontiers for different portfolio sizes. The stocks are randomly selected from the S&P500. The individual means and std. are displayed as points in the background.

holds for all vectors  $\mathbf{z}$ . One necessary condition for a matrix to be positive definite can be derived using the eigenvalues of a matrix and its eigenvalue decomposition. As described in Harville (1997, ch. 21), an eigenvalue (or characteristic root)  $\lambda$  is the solution to

**Definition 2.1.1.** Let  $\mathbf{A}$  be a  $p \times p$  matrix. The characteristic roots (with multiplicity) are given by the solutions to

$$|\mathbf{A} - \lambda \mathbf{I}| = 0$$

where  $|\cdot|$  is the determinant of a matrix.

Let  $\lambda_i, i = 1, 2, \dots, p$ , denote the *ordered* eigenvalues of the matrix  $\mathbf{A}$  such that  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ . Given an eigenvalue, the eigenvectors  $\mathbf{u}_i$  are defined by  $\mathbf{A}\mathbf{u}_i = \lambda_i\mathbf{u}_i, i = 1, 2, \dots, p$ . Let  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$  and  $\mathbf{U} = (\mathbf{u}_1^\top, \mathbf{u}_2^\top, \dots, \mathbf{u}_p^\top)^\top$ . Furthermore, we say that  $\mathbf{B}$  is an inverse of  $\mathbf{A}$  if it satisfies  $\mathbf{B}\mathbf{A} = \mathbf{I}$ , which we usually denote  $\mathbf{B} = \mathbf{A}^{-1}$ . Using the relation between eigenvalues and their eigenvectors we can derive the eigenvalue decomposition of a symmetric matrix

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top. \quad (2.1.1)$$

A necessary condition for a matrix to be positive definite can be directly obtained from the eigenvalue decomposition. Let  $\mathbf{z} \in \mathbb{R}^p$  and  $\mathbf{y} := \mathbf{U}^\top \mathbf{z} \in \mathbb{R}^p$ , then  $\mathbf{z}^\top \mathbf{A} \mathbf{z} = \mathbf{z}^\top \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\top \mathbf{z} = \mathbf{y}^\top \mathbf{\Lambda} \mathbf{y} = \sum_i \lambda_i y_i^2$  which is a second degree polynomial. If the eigenvalues are all positive then necessarily the matrix is positive definite. If there are some eigenvalues which are zero then the matrix is semi-positive definite.

In all papers of this thesis we assume that the true covariance matrix is positive definite. The assumption has quite a deep economical interpretation. If one (or more) eigenvalue(s) are zero then there is a possibility to construct a portfolio which does not contain any risk with a potentially positive return. An opportunity which should not exist unless the elements of  $\boldsymbol{\mu}$  are all zero. Assume  $\lambda_p = 0$ , let  $\mathbf{u}_p$  be its eigenvector and set  $\mathbf{w} = \mathbf{u}_p / \sum_i u_{ip}$ . The variance of that portfolio is zero since all eigenvectors are orthonormal and its mean is  $\mathbf{w}^\top \boldsymbol{\mu}$  which can be non-zero unless the elements of  $\boldsymbol{\mu}$  are all zero. If the true covariance matrix is

not positive definite there might exist arbitrage opportunities, e.g. the possibility of making profit without taking any risk.

The eigenvalue decomposition is very useful. First, it provides a simple way to construct inverses, which is very important for MPT as seen in (2.0.4). We claim that  $\mathbf{B} = \mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}^\top$  is a valid inverse which is easy to verify since  $\mathbf{B}\mathbf{A} = \mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}^\top\mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$ . To study the inverse we need to study the inverse of the eigenvalues. Secondly, it contains a lot of information that might not be available at first glance. If  $\mathbf{A}$  is a covariance matrix then it contains variances and covariances, describing relations between random variables. The eigenvectors are rotations that try to capture as much variation as possible along its axis. The eigenvalues is the variation along the eigenvectors axis. They describes how the system behaves and not the individual elements and their inverse values describe how the precision matrix behaves. **comment on latent factors**

## Chapter 3

# Statistical models and estimation

To *practically* use the portfolios described by (2.0.4) we have to specify  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$ .

To state his portfolio allocation problem we first introduce the relevant quantities. In this thesis we never use the asset prices themselves but their a transformation of the relative differences, that is, their simple returns. The simple return is defined as  $r_{i,t} := (y_{i,t} - y_{i,t-1})/y_{i,t-1}$  and the log return is then defined as  $x_{i,t} := \log(r_{i,t} + 1)$  where  $y_{i,t}$  is the price of the  $i$ th asset at time  $t$ . We model a portfolio with  $p$  assets as  $\sum_{i=1}^p w_i x_{i,t} = \mathbf{w}^\top \mathbf{x}_t$  where  $\mathbf{w} = (w_1, \dots, w_p)$  are the portfolio weights and  $\mathbf{x}_t = (x_{1,t}, x_{2,t}, \dots, x_{p,t})$  are the log returns. Notice that this is an approximation. In reality we would want to work with  $\sum_{i=1}^p w_i r_{i,t}$  since it is additive in the number of assets. However, logarithmic returns are additive in time which can be desirable. Compounding returns results in simple addition. The difference between the two approaches is very small if the returns are small, which is often true for financial assets. The models we work with in this thesis often rely on the log returns and not the returns. We will omit the time index  $t$  to keep the notation tidy unless otherwise stated.

- Multivariate Normal
- Wishart and inverse wishart distribution, properties
- Matrix variate location and scale.

**Definition 3.0.1.** A random vector  $\mathbf{x} \in \mathbb{R}^p$  follows a multivariate normal distribution with mean vector  $\boldsymbol{\mu} \in \mathbb{R}^p$  and positive definite covariance matrix  $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$  if its density is given by

$$\frac{|\boldsymbol{\Sigma}|^{-1/2}}{2\pi} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \quad (3.0.1)$$

where  $|\mathbf{B}|$  is the determinant of the matrix  $\mathbf{B}$ . We usually use the notation  $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  to indicate the above.

**Definition 3.0.2.** The random matrix  $\mathbf{S}$  of size  $p \times p$  follows a  $p \times p$  dimensional Wishart distribution with  $n$  degrees of freedom,  $n > p$ , if its density is given by

$$\frac{|\mathbf{S}|^{(n-p-1)/2} |\boldsymbol{\Sigma}|^{-n/2}}{2^{pn/2} \Gamma_p(n/2)} \exp \left\{ -\frac{1}{2} \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{S}) \right\} \quad (3.0.2)$$

where  $\Gamma_p(\cdot)$  is the multivariate gamma function and  $\text{tr}(\cdot)$  is the trace operator, i.e. the sum of the diagonal elements and  $\boldsymbol{\Sigma}, \mathbf{S}$  are both positive definitie. We use the notation  $\mathbf{S} \sim W_p(n, \boldsymbol{\Sigma})$  to indicate that  $\mathbf{S}$  follows a Wishart distribution with the given parameters.

**Definition 3.0.3.** A positive definitie random matrix  $\mathbf{A}$  is said to be distributed according to a  $p \times p$  dimensional inverse Wishart distribution with  $n$  degrees of freedom and positive definitie parameter matrix  $\mathbf{V}$  if its density is given by

$$\frac{2^{-(n-p-1)p/2} |\boldsymbol{\Sigma}|^{(n-p-1)/2}}{\Gamma_p((n-p-1)/2) |\mathbf{A}|^{n/2}} \exp \left\{ -\frac{1}{2} \mathbf{A} \mathbf{V} \right\}, \quad n > 2p \quad (3.0.3)$$

which we denote  $\mathbf{A} \sim W_p^{-1}(n, \mathbf{V})$  to indicate that  $\mathbf{V}$  follows an Inverse Wishart distribution with the given parameters.

**Lemma 3.0.4.**  $\mathbf{x}$  is a  $k$  dimensional elliptically contoured distributed random variable with mean  $\boldsymbol{\mu}$  and dispersion matrix  $\boldsymbol{\Sigma}$  where  $\text{rk}(\boldsymbol{\Sigma}) = p$ , if and only if

$$\mathbf{x} \stackrel{d}{=} \boldsymbol{\mu} + r\mathbf{V}\mathbf{u}$$

where  $r > 0$  is a random variable,  $\mathbf{u}$  is a random vector which is uniformly distributed on the sphere, in  $\mathbb{R}^k$ ,  $\mathbf{V}\mathbf{V}^\top = \boldsymbol{\Sigma}$ , and  $r$  and  $\mathbf{u}$  are independent.

### 3.1 Estimation - the name of the game

- What are the implications of using  $\mathbf{S}$  instead of  $\boldsymbol{\Sigma}$ ?
- why is  $\mathbf{S}$  always an admissible estimator? (MM)
- Other types of estimators and why they might be better than  $\mathbf{S}$ .

*Remark 3.1.1.* Unconditional and conditional covariance estimation. Prediction is very hard and constructing viable models. Are returns predictable? Should we even try?

### 3.2 Simulations, inverses and why stochastic representations are valuable

- Motivating simulations and the issue with inversions.
- Simulation of multi- or matrixvariate distributions can be very computationally consuming.
- ...

## Chapter 4

# Higher dimensions or portfolios with many assets

In the previous chapter we presented different ways of estimating the covariance matrix. Under certain conditions and/or statistical models, the sample covariance matrix inherited certain properties. If we have a lot of data on the assets that we are trying to invest in then we can most often be certain that we will hold the correct portfolio. Our estimated portfolio will be consistent, e.g. it estimates the correct object of interest. However, if we believe in diversification then  $p$  should be big as well. It should, in theory, decrease the risk (variance) of the portfolio. That is not always the case. By introducing one new asset to our portfolio of size  $p$  we need to estimate the variance and all covariances in the sample covariance matrix. They constitute an additional  $p + 1$  quantities. The sample covariance matrix suffers from the curse of dimensionality. In terms of estimation uncertainty, this does not scale well.

- **state the implications of large dimensions for portfolio theory**, Consistency of  $\mathbf{S}$  but not of  $\mathbf{S}^{-1}$ .
- Give small example of the different quantities on how they behave
- In terms of simulation, this becomes even worse.

### 4.1 Random Matrix Theory and Stieltjes transforms

The subject of Random matrix theory (RMT) has many applications. It was originally developed in the context of statistical physics by WIGNER. The theory has developed and is now used in many fields such as combinatorics, computational biology, wireless communication and finance. This section gives a brief introduction and will then cover a small part of the field more indepth, namely high dimensional limits involving the inverse sample covariance matrix and transformations thereof.

- The Marcenko-Pastur (MP) law?
- Short on Stieltjes transforms.
- Theorem 1 of Rubio and Mestre (2011).





# Chapter 5

## Summary of Papers

The papers presented here are among a total of ... papers produced. These are selected based their common theme.

### Paper 1 - Sampling ...

The paper investigates a fundamental question in modern portfolio theory. What are the implications to using the sample covariance matrix  $\mathbf{S}$  and the sample mean  $\bar{\mathbf{x}}$  instead of the true covariance matrix  $\mathbf{\Sigma}$  and  $\boldsymbol{\mu}$ ? The paper does so when returns follow a multivariate normal distribution.

### Paper 2 - Tangency portfolio

In this paper we investigate the another portfolio which contains a risk-free asset. The return or rate of the risk-free asset is denoted  $r_f$ . To introduce this option into our portfolio we add the risk-free rate as part of the portfolio  $w_0 r_f + \mathbf{w}^\top \mathbf{x}$  and optimize over  $w_0$  as well. The portfolio is obtained from the quadratic utility function, that is, it originates from the following portfolio allocation

$$\min_{w_0, \mathbf{w}} w_0 r_f + \mathbf{w}^\top \boldsymbol{\mu} - \frac{1}{2\gamma} \mathbf{w}^\top \mathbf{\Sigma} \mathbf{w} \quad (5.0.1)$$

$$\text{s.t. } w_0 + \mathbf{w}^\top \mathbf{1}_p = 1 \quad (5.0.2)$$

However, since  $w_0 + \mathbf{w}^\top \mathbf{1} = 1$  we substitute  $w_0 = 1 - \mathbf{w}^\top \mathbf{1}$  and solve the unconstrained optimization problem instead. The portfolio has many interesting properties. If there is a risk-free asset we can actually increase the return and decrease the risk of our position in the market. This is most easily explained by the efficient frontier, displayed in Figure 2.1, and what happens with it when the risk-free rate is introduced. In Figure ?? we illustrate what happens. Since we have the opportunity to invest in a risk-free asset we can decrease the risk we take in comparison to earlier. If we do not want to take any market risk, then we can place all our cash in a bond with rate  $r_f$ . Taking any combination between will result in less risk and more return.

The paper investigates what the distribution is of the tangency portfolio, or in a more broader context, the capital market line. We use an extension to the multivariate Gaussian Model from Paper 1, the Closed Skew-Normal Matrixvariate Model. This model can include skewness in the asset returns a trait returns usually exhibit (see e.g. Cont (2001)). We investigate what implications the model has on the estimated tangency portfolio.

### Paper 3

This paper deals with the fact that taking limits changes estimates. Assuming that the investor is interested in daily returns then if he/she invest in the GMV portfolio today and wait a week the estimates of the portfolio estimate will have changed. Rebalancing every day will induce a cost. Assuming that we are ok

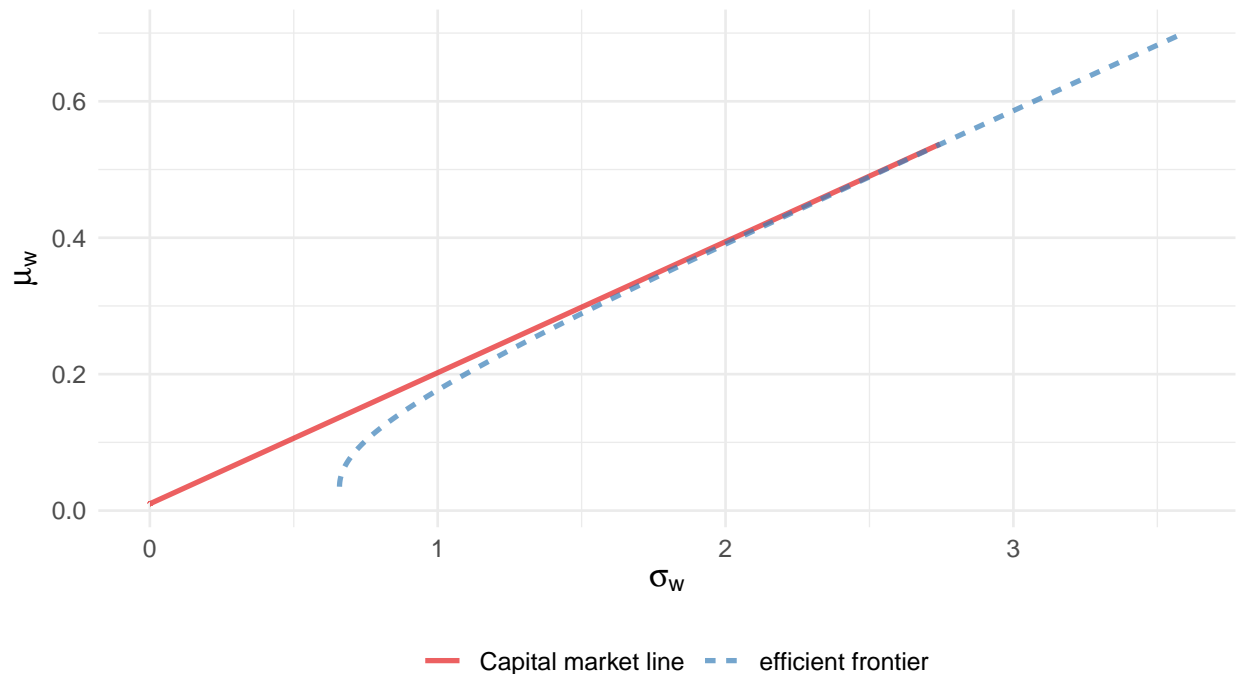


Figure 5.1: Efficient frontier and the capital market line for 30 (randomly) selected stocks from the S&P500

with rebalancing the portfolio at fixed time points, which could be daily, we develop a reweighting scheme the induced cost. That is, minimize the movement between portfolio reallocations.

This paper continues on to use another model. Although more general, its also

The paper is accompanied by a R package, available on CRAN. You are free<sup>1</sup> to install it by `install.packages("DOSPortfolio")`. Below is a short introduction on how to construct the portfolio estimates

```
df <- read_csv("../data/returns.csv")
library(DOSPortfolio)
set.seed(1234)
p <- 350; n <- 400
# Sample p assets
asset_cols <- sample(2:ncol(df), size = p)
# specify reallocation points
reallocation_points <- seq(n, nrow(df), by=n)
# estimate portfolio weights
dos_weights <- df %>%
  select(all_of(asset_cols), -date) %>%
  DOSPortfolio(.,
    reallocation_points = reallocation_points,
    target_portfolio = rep(1, ncol(.))/ncol(.),
    shrinkage_type = "overlapping")
```

## Paper 4

Double shrinkage?

---

<sup>1</sup>or rather encouraged!

# Chapter 6

## Future research

There are many possible extensions and future projects to the thesis at hand.

- Are shrinkage intensities for the sample covariance matrix optimal for the precision or MPT problem?
- One of the most interesting issues of the elliptical distribution and its inverse sample dispersion matrix. What are the moments of  $(\mathbf{Z}\mathbf{R}\mathbf{Z}^\top)^{-1}$ ?
- There are different ways of incorporating estimation uncertainty, one solution is robust optimization. Are there connections to be made? Is robust optimization just Empirical Bayes?
- Sequential reweighting extension to Paper 3. When should we reweight?
- Higher dimensions, other estimators, hard shrinkage.
- BEKK models are usually hard to fit and use for MPT. Their forecasts are not always positive definite. The first issue is simple to cope with if one can formulate the models as Recurrent Neural Networks and use the efficient deep-learning libraries Torch or Tensorflow. By doing so, one would also get the possibility to develop new models by constructing new layers to the networks. It would also be easier to integrate different sources of information in the models.

Although not part of this Phd thesis - Flipped classroom and online learning tools.



# Bibliography

- Bauder, D., Bodnar, T., Parolya, N., and Schmid, W. (2020). Bayesian mean–variance analysis: optimal portfolio selection under parameter uncertainty. *Quantitative Finance*, page to appear.
- Bodnar, T. and Gupta, A. (2009). Construction and inferences of the efficient frontier in elliptical models. *Journal of the Japan Statistical Society*, 39.
- Bodnar, T., Parolya, N., and Schmid, W. (2013). On the equivalence of quadratic optimization problems commonly used in portfolio theory. *European Journal of Operational Research*, 229(3):637–644.
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236.
- Harville, D. A. (1997). *Matrix algebra from a statistician’s perspective*, volume 1. Springer.
- Hult, H., Lindskog, F., Hammarlid, O., and Rehn, C. J. (2012). *Risk and portfolio analysis: Principles and methods*. Springer Science & Business Media.
- Markowitz, H. (1959). *Portfolio selection: efficient diversification of investments*. John Wiley & Sons, Inc and Chapman & Hall, Ltd.
- Rubio, F. and Mestre, X. (2011). Spectral convergence for a general class of random matrices. *Statistics & Probability Letters*, 81(5):592–602.