VILNIUS UNIVERSITY

FACULTY OF MATHEMATICS AND INFORMATICS

MASTER'S STUDY PROGRAMME
FINANCIAL AND ACTUARIAL MATHEMATICS

From Markowitz's Mean-Variance portfolio optimization to Conditional Value-at-Risk

Nuo Markowitz vidurkio-dispersijos portfelio optimizavimo iki sąlyginės vertės pokyčio rizikos

Master's Thesis

Author: Agnė Marcinevičiūtė

VU e-mail address: agne.marcinkeviciute@mif.stud.vu.lt

Supervisor: Andrius Grigutis, PhD, Associate Professor

Vilnius
2024

Contents

List of Figures

List of Tables

From Markowitz's Mean-Variance portfolio optimization to Conditional Value-at-Risk

Abstract

This master's thesis explores methods of portfolio optimization. Starting from the mean-variance approach Sharpe ratio is calculated and the efficient frontier is plotted to minimize the variance for different returns. To assess the tail risk of the return distribution Conditional Value-at-Risk (CVaR) is estimated. Data includes stock and Bitcoin closing prices and is obtained from Yahoo Finance [12]. The code for the empirical analysis is produced using Matlab programming language [6]. As Bitcoin price is highly volatile and might behave asymmetrically to traditional assets non-parametric methods are used to estimate CVaR. The usage of copulas allows to estimate the distribution of return vectors by modeling marginals individually. To estimate the copula density the kernel density is used. Simulated returns from the estimation are used to optimize CVaR. The contribution to the research lies in the utilization of non-parametric methods to analyze Bitcoin expected return and risk trade-off. Therefore, in this analysis risk perception of investors is analyzed in highly volatile market conditions. The aim of this thesis is to compare both optimization methods for a portfolio consisting of both traditional and non-traditional assets. Results indicate that CVaR creates a more diversified portfolio than mean-variance model. Also, CVaR allocates at least 10% of funds to Bitcoin for all portfolios, as it is viewed as a non-linear return driver. Mean-variance model increases allocation to Bitcoin with the growth of portfolio risk, because it sees volatility directly as risk.

Key words: mean-variance, CVaR, efficient frontier, Bitcoin, copula, non-parametric.

Nuo Markowitz vidurkio – dispersijos portfelio optimizavimo iki sąlyginės vertės pokyčio rizikos

Santrauka

Šis magistro darbas tyrinėja portfelio optimizavimo metodus. Pirmasis pasirinktas vidurkio–dispersijos modelis, suskaičiuotas optimalus Šarpo rodiklis ir nubrėžtas efektyvumo frontas minimizuojantis grąžų dispersiją pasirinktai grąžai. Sąlyginės vertės pokyčio rizikos metodas buvo pasirinktas norint įvertinti skirstinio uodegos riziką. Praktinėje darbo dalyje nagrinėjamos 2018-2023 akcijų ir Bitkoino uždarymo kainos iš Yahoo Finance [12], programavimui naudota Matlab programavimo kalba [6]. Prieš pasirenkant modelį buvo atsižvelgta į Bitkoino kainos nepastovumą ir asimetrinį ryšį su tradiciniais finansiniais instrumentais, todėl tolesnei analizei buvo pasirinktas neparametrinis metodas sąlyginės vertės pokyčio rizikos skaičiavimui. Kopulų naudojimas supaprastina vektorių modeliavimą atskirdamas marginaliuosius skirstinius ir kopulą. Kopulos skaičiavime naudojama branduolio tankio funkcija. Simuliuotos investicijų grąžos naudojamos CVaR optimizavime. Šis darbas prisideda prie egzistuojančios literatūros pritaikydamas neparametrinius metodus Bitkoino tikėtinų grąžų ir rizikos analizėje. Šiame darbe investuotojų rizikos jautrumas analizuojamas į portfelį įtraukus ypač nepastovius instrumentus. Darbo tikslas–palyginti abu optimizavimo modelius portfeliui, kuris sudarytas iš tradicinių ir netradicinių finansinių instrumentų. Rezultatai rodo, kad naudojant CVaR modelį gaunamas labiau diversifikuotas portfelis, nei naudojant variacijos–dispersijos modelį. Taip pat, skaičiuojant CVaR metodu kiekviename iš analizuotų portfelių bent 10 % lėšų investuojama į Bitkoiną, nes dėl šio instrumento netiesiško ryšio su tradicinėmis akcijomis CVaR modelyje jis tampa vienu iš pagrindinių grąžos variklių. Kita vertus, vidurkio–dispersijos modeliu nustatytuose portfeliuose Bitkoino procentinė dalis auga didėjant rizikai, todėl galima teigti, kad šis modelis grąžų nepastovumą laiko tiesioginiu aukštos rizikos požymiu.

Raktiniai žodžiai: vidurkis–dispersija, CVaR, efektyvumo frontas, Bitkoinas, kopula, neparametrinis modelis.

1 Introduction

Portfolio optimization is the core goal in financial management with the aim of constructing a portfolio with the best risk-return trade-off. Markowitz [5] introduced a model that analyzes the mean and variance of investment assets. The aim of this method is to create such portfolio that would maximize return for a given level of risk, or would minimize risk for a given level of return. However, focusing on portfolio mean and defining risk as variance might be an overly simplistic approach to evaluate an investment portfolio. This approach might fail to account for skewness of fat tails of a distribution. This equal treatment of the upside and the downside asset volatility might lead to inaccurate results. These shortcomings are addressed by CVaR model. It focuses on the mitigation of the tail risk and measures worst-case scenario loss beyond a pre-specified confidence level. CVaR method is useful for risk-averse investors who try to have minimal losses during extreme market downturns, especially prominent for cryptocurrencies like Bitcoin.

In this study I compare portfolio efficient frontiers and weight allocations to certain assets resulting from both mean-variance and CVaR models. Analysis includes traditional stocks that are listed on the exchange and Bitcoin. This thesis highlights how both approaches differ in the treatment of risk and the trade-offs between risk and return. For the mean-variance approach I calculate Sharpe ratio and plot the efficient frontier to minimize variance for different returns. For the second approach, the CVaR optimization, I implement non-parametric methods. I use copulas to estimate the distribution by modeling marginals separately. I use marginal cumulative density functions to transform my data and get marginals uniform over [0,1]. For copula density estimations I use kernel density. I sample from copula and obtain simulated returns which by using the inverse cumulative density functions are transformed back to the original marginal distribution. Lastly, by minimizing CVaR with the simulated returns I obtain the results and plot the efficient frontier.

By implementing both parametric and non-parametric methods this thesis aims to provide comparison of both portfolio optimization models and offers insights for investors with different risk appetites. The structure of this thesis is the following: first section is the introduction, in the second section I analyze relevant literature for mean-variance model and alternative risk methods. The third section outlines sample selection logic and the methodology used for both models. In the fourth section I present the findings of the practical assessment and the comparison between two methods. Lastly, the fifth section concludes the takeaways of this master's thesis and provides thought for further analysis.

2 Relevant literature review

In this section I introduce the concept of Markowitz's mean-variance model and outline its shortcomings. Also, I discuss the implications of including Bitcoin in the portfolio. To assess the limitations of the mean-variance model I analyze literature on the alternative risk models such as Value-at-Risk (VaR), CVaR and compare them.

2.1 Markowitz's Mean-Variance Optimization framework

Markowitz [5] first formulated portfolio selection known as the mean-variance optimization framework. He states that to achieve an optimal portfolio, an investor should take into account a trade-off between expected return and variance. Key assumptions of the model are that the investors are rational and risk-averse, and markets are efficient. The rationality constraint makes sure that an investor would prefer a higher return under the same level of risk. And the risk aversion leads to them wanting a lower risk for the same return level. The mean-variance optimization problem derives the efficient frontier, which is comprised of portfolios with the efficient trade-off between risk and return. The portfolios that lie on the efficient frontier are efficient and will be chosen by rational investors as they want maximum expected return with lowest risk. Markowitz [5] defines risk as standard deviation or variance of returns. The drawback of this strong assumption is that it ignores the skewness and flat tails of distributions. In this case, equally treating upside and downside asset volatility might lead to incorrect findings. For example, sensitivity to estimation error in expected returns using historical averages which might not reflect the future market trends falls as a criticism of the mean-variance model [7]. Merton [7] highlights that errors occur estimating the covariance matrix because of large number of parameters used in the estimation. As with n number of assets we would need to estimate $\frac{n(n-1)}{2}$ covariance terms.

The introduction of Bitcoin is studied as a diversification tool for a portfolio of stocks [1]. The authors use the mean-variance framework and optimize the portfolio using Monte Carlo simulations. The results suggest that Bitcoin acts as a diversification tool, as almost all portfolios with Bitcoin performed much better compared to portfolios without it. However, they point out that because of high price fluctuations, allocation to Bitcoin should be limited. Some research [11] states that adding Bitcoin to a diversified portfolio minimizes portfolio variance. They find that Bitcoin significantly increases diversification as the correlation between it and other assets is low.

2.2 Alternative risk measures

The shortcomings of the mean-variance model paved the way for the introduction of alternative risk measures. Conditional Value-at-Risk model, also called Expected Shortfall was introduced [9]. It addresses the drawbacks of VaR which only focuses on the loss of a specific percentile. Using CVaR authors [9] calculate the mean loss beyond the VaR threshold, leading to its better performance catching the downside risk. The authors use a linear optimization of CVaR, demonstrating that the computational efficiency can be applied to the case of large portfolios. Results of this article show that CVaR model generates more robust solutions during the crisis periods than the mean-variance or VaR methods.

Moreover, Banihashemi and Navidi [2] analyze Iranian company stock price data from 2015 to 2016 and compare VaR model with the CVaR. They calculate both measures using historical and Monte Carlo simulations. Authors conclude that CVaR model provides more accurate results compared to VaR, especially in capturing the tail risk. Nedela [8] compares the traditional Markowitz model with the CVaR model extending data by including stock prices from the United States, China and the United Kingdom markets. Empirical results do not indicate which model is more suitable for assessing risk during the crisis period. However, investors would have a lower level of risk creating a portfolio with the CVaR model compared to using mean-variance method.

Semenov and Smagulov [10] use copulas for their estimation of the dependencies between financial assets. They estimate both VaR and CVaR using three different copula models. Authors compare the results with historical estimates and find that using any of the copula specifications outperformed historical estimates based on their predictability.

3 Sample selection and research methodology

In this section I present the data and the methodology used for the optimization of both mean-variance and CVaR models. I outline the sample selection criteria used for this analysis. Also, I provide the mathematical formulations for both models as well as the non-parametric approach for the CVaR estimation.

3.1 Sample selection criteria

This thesis seeks to extend Markowitz's mean-variance optimization model paying attention to its shortcomings regarding the downside risk. CVaR model is introduced as an extension to the mean-variance model. I select a sample for both models from 2018.01.01 to 2023.12.31, a six-year sample captures both market downturn and growth cycles. The source of the data is Yahoo Finance [12] and the sample includes both stock and cryptocurrency price data. For this analysis I choose eight assets and obtain their daily closing prices out of which one is Bitcoin, the rest are Apple, Microsoft Corporation, Alphabet, Amazon, Netflix, JPMrogan Chase and Co. and Tesla company stocks all denominated in USD. I add Bitcoin to my portfolio based on the research [4] stating that introducing Bitcoin improves the risk-return trade-off of a traditional asset portfolio. The reason is that even though Bitcoin has high volatility, the correlation with other assets is low.

3.2 Data preparation

Percentage returns are chosen for this analysis as they allow to directly compare different assets regardless of their price level. I calculate daily returns for each asset individually using closing prices, defined as:

$$R_i, t = \frac{P_i, t - P_{i,t-1}}{P_{i,t-1}}, \tag{3.1}$$

where $R_i, t \in \mathbb{R}$ for n assets.

- $R_{i,t}$ is the return of the i-th asset on day t.
- P_i, t: Closing price of the i-th asset at time t.

Portfolio includes n assets, the vector of daily returns $\mathbf{R}_t$ is defined as:

$$\mathbf{R}_t = (R_{1t}, R_{2t}, \dots, R_{nt})$$

- $\boldsymbol{\mu} = \mathbb{E}[\mathbf{R}_t]$: Vector of mean returns.
- $\mathbf{\Sigma} = \mathrm{Cov}(\mathbf{R}_t) = \mathbb{E}[(\mathbf{R}_t - \mu)(\mathbf{R}_t - \mu)^\top]$: Covariance matrix of asset returns.

The vector of expected returns $\boldsymbol{\mu}$ and covariance matrix of asset returns $\mathbf{\Sigma}$ are calculated as:

$$\boldsymbol{\mu} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{R}_t, \quad \mathbf{\Sigma} = \frac{1}{T-1}\sum_{t=1}^{T}(\mathbf{R}_t - \boldsymbol{\mu})(\mathbf{R}_t - \boldsymbol{\mu})^\top.$$

3.3 Markowitz Mean-Variance optimization

The first model I use is the Markowitz mean-variance optimization model introduced by Markowitz [5]. The model assumes risk aversion of the investors and claims that higher risk needs to be compensated by higher expected return. Using the mean-variance optimization framework I create a portfolio that finds combinations of expected return against the risk of these assets. The goal is to perform a minimization problem and achieve the smallest risk for a given return.

Firstly, I determine the weights $\mathbf{w} = [w_1, w_2, \ldots, w_n]^\top$, where each w_i is defined as a percentage of an investment allocated to a security i, and $\sum_{i=1}^{n} w_i = 1$.

The portfolio return R_p represents the actual portfolio return at a specific time and is determined as:

$$R_p = \sum_{i=1}^{n} w_i R_i = \mathbf{w}^\top \mathbf{R}_t, \tag{3.2}$$

where $\mathbf{R}_t$ is the return vector of the individual assets.

Using the mean-variance optimization model for this analysis, I assume that the return proxy is the expected return, while the risk proxy is the variance of the asset returns. Also, a portfolio set should be formed using any combination of the constraints specified above.

Then the portfolio expected return which represents the average portfolio return over time is calculated as follows:

$$E[R_p] = \mathbb{E}[\mathbf{w}^\top \boldsymbol{R_t}] = \mathbf{w}^\top \boldsymbol{\mu}. \tag{3.3}$$

Also, the variance and standard deviation is expressed as:

$$\sigma_p^2 = \mathrm{Var}(R_p) = \mathrm{Var}(\mathbf{w}^\top \mathbf{R}_t) = \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}, \tag{3.4}$$

$$\sigma_p = \sqrt{\sigma_p^2} = \sqrt{\mathrm{Var}(R_p)} = \sqrt{\mathrm{Var}(\mathbf{w}^\top \mathbf{R}_t)} = \sqrt{\mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}}. \tag{3.5}$$

where the covariance matrix of returns is defined as $\boldsymbol{\Sigma}$.

3.4 Sharpe Ratio

Sharpe ratio is a measure comparing the return of the portfolio with its risk. As risk increases by one unit, it shows the additional amount of return received by an investor, therefore, showing how greatly investors are compensated for taking extra risk. According to the Modern Portfolio theory, tangency portfolio that lies on the efficient frontier maximizes the Sharpe ratio.

The Sharpe ratio is defined as:

$$\text{Sharpe Ratio} = \frac{E[R_p] - r_f}{\sigma_p},$$

Where $\mathbb{E}[R_p] - r_f$ is the risk premium and it measures the excess return of the portfolio compared to the risk-free rate. σ_p is the volatility of the portfolio. Moreover, r_f is a risk-free rate of return. This return is a benchmark used to determine excess returns. For this analysis I choose return on Treasury bills as my risk-free rate.

Now, after substituting $E[R_p]$ and σ_p as defined in (3.3) and (3.5) :

$$\text{Sharpe Ratio} = \frac{\mathbf{w}^\top \boldsymbol{\mu} - r_f}{\sqrt{\mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}}} \tag{3.6}$$

To get the optimal portfolio I search for the maximum of the Sharpe ratio:

$$\max_{\mathbf{w}} \frac{\mathbf{w}^\top \boldsymbol{\mu} - r_f}{\sqrt{\mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}}}, \tag{3.7}$$

subject to:

$$\sum_{i=1}^{n} w_i = 1, \quad w_i \geq 0.$$

These conditions imply that the portfolio weights sum to 1. Meaning that the full portfolio is invested without borrowing or any leftover funds. Also, no short selling of the assets is allowed, as weights of the portfolio have to be non-negative. For this analysis the portfolio is long-only, with all capital fully allocated.

To solve the optimization problem I use the Lagrange multipliers:

$$\mathcal{L}(\mathbf{w}, \lambda) = \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w} - \lambda_1(\mathbf{w}^\top \boldsymbol{\mu} - r_{\text{target}}) - \lambda_2(\mathbf{1}^\top \mathbf{w} - 1),$$

where r_{target} is the desired return of the portfolio and $\mathbf{1}$ is a vector of ones. Solving $\frac{\partial \mathcal{L}}{\partial w} = 0$ gives the result as follows:

$$\mathbf{w} = \boldsymbol{\Sigma}^{-1}(\lambda_1 \boldsymbol{\mu} + \lambda_2 \mathbf{1}).$$

Lastly, I plot the efficient frontier that minimizes the variance for different target returns, defined as:

$$\min_{\mathbf{w}} \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}, \tag{3.8}$$

subject to:

$$\mathbf{w}^\top \boldsymbol{\mu} = r_{\text{target}}, \quad \sum_{i=1}^{n} w_i = 1, \quad w_i \geq 0.$$

3.5 Conditional Value-at-Risk (CVaR) optimization

Introduced by Rockafellar and Uryasev [9], Conditional Value-at-Risk assesses the tail risk of the distribution of an investment portfolio. It takes the weighted mean of the losses in the tail of the distribution, in turn measuring the expected loss that goes beyond the Value-at-Risk (VaR_α) point at confidence level α. The use of CVaR instead of VaR benefits the most when the asset or an asset class is less stable over time. When dealing with high-volatility assets, VaR might not be able to catch all the volatility as it is indifferent to anything past its risk threshold. Therefore, the worst-case loss scenario at a time period is represented by VaR, the CVaR represents the expected loss if we cross that worst-case bound. In this analysis, I add cryptocurrencies into the portfolio and to account for the high volatility of their returns I employ the CVaR method.

Firstly, for confidence level α, the VaR_α is defined as the α-quantile of the return distribution $F(R_p)$:

$$\text{VaR}_\alpha = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}. \tag{3.9}$$

where $F(x) = P(X \leq x)$ is the cumulative distribution function of the random variable X.

I choose α, the confidence level, to be 95% (or 0.95).

CVaR measures the expected returns conditional on it being less than or equal to VaR_α, defined as:

$$\text{CVaR}_\alpha = \mathbb{E}[R_p \mid R_p \leq \text{VaR}_\alpha].$$

Using the formulation introduced in [9], it is calculated as:

$$\mathrm{CVaR}_\alpha = \nu + \frac{1}{1-\alpha}\mathbb{E}\left[\max(0, \nu - R_p)\right], \tag{3.10}$$

where $\nu = \mathrm{VaR}_\alpha$.

3.6 Copula estimation method

Copulas are defined as multivariate cumulative distribution functions where marginal probability distribution of a variable is uniform on [0,1]. Copula models the dependence between random variables separately from the marginal distributions.

For clarity analyzing a bivariate case, according to Sklar's theorem there exists a unique copula C for any bivariate distribution function $F_{X,Y}(x, y)$ with continuous marginals $F_X(x)$ and $F_Y(y)$, defined as:

$$F_{X,Y}(x, y) = C(F_X(x), F_Y(y)). \tag{3.11}$$

Formally, following [3], I use the definition of copula $C(u, v)$

$$[0, 1]^2 \to [0, 1]$$

and it maps the vector on the unit square and has the following properties:

1. $C(u, v)$ in an non-decreasing function in each u and v.
2. $C(1, v) = v$ and $C(u, 1) = u$, for all $u, v \in [0, 1]$.
3. For $a_u \leq b_u$ and $a_v \leq b_v$, the probability $P(U \in [a_u, b_u], V \in [a_v, b_v])$ has to be non-negative. The third property leads to the rectangle inequality:

$$\sum_{i=1}^{2}\sum_{j=1}^{2}(-1)^{i+j}C(u_i, v_j) \geq 0,$$

with $u_1 = a_u$, $u_2 = b_u$, $v_1 = a_v$, and $v_2 = b_v$.

Furthermore, the copula density $c(u, v)$ is defined as:

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v}. \tag{3.12}$$

- X, Y: returns of the securities.
- $F_X(x), F_Y(y)$: Marginal cumulative density functions (CDFs) of asset returns.
- u, v: variables calculated from the margins, uniform over $[0, 1]$.

For this analysis I use copulas as they allow me to estimate the distribution of asset return vectors by modeling marginals separately. It is useful for my portfolio with stocks and cryptocurrencies, as Bitcoin has a non-linear correlation with traditional stocks. Other correlation measures might fail to estimate the tail dependence, especially, the co-movements during the market crisis periods. Also, cryptocurrencies introduce higher asymmetry, Bitcoin and other stocks might correlate differently depending on market conditions (bullish vs. bearish markets). To capture these portfolio characteristics I continue with the copula estimation model and extend it using non-parametric kernel density estimation.

3.7 Kernel density

Kernel-based estimation is a non-parametric model that estimates the probability density function of a variable (in this case asset return). The copula density $c(u, v)$ is estimated as a function of the joint density $f(x, y)$ and the marginal densities $f_X(x)$ and $f_Y(y)$, equal to:

$$c(u, v) = \frac{f(F_X^{-1}(u), F_Y^{-1}(v))}{f_X(F_X^{-1}(u)) f_Y(F_Y^{-1}(v))}. \tag{3.13}$$

where

- $f(F_X^{-1}(u), F_Y^{-1}(v))$ is the joint density $f(x, y)$.
- $f_X(F_X^{-1}(u))$, $f_X(F_Y^{-1}(v))$ are the marginal densities $f_X(x)$ and $f_Y(y)$.
- $F_X^{-1}(u)$ and $F_Y^{-1}(v)$ are inverse cumulative distribution functions mapping u and v from $[0, 1]$ to the original X and Y space.

Joint density $f(x, y)$ is approximated using kernel density as:

$$\hat{f}(x, y) = \frac{1}{n h_X h_Y} \sum_{i=1}^{N} K\left(\frac{x - x_i}{h_X}\right) K\left(\frac{y - y_i}{h_Y}\right), \tag{3.14}$$

with marginal densities $f_X(x)$ and $f_Y(y)$ estimated as:

$$\hat{f}_X(x) = \frac{1}{n h_X} \sum_{i=1}^{N} K\left(\frac{x - x_i}{h_X}\right).$$

- n: Number of observations.
- h_X, h_Y: Bandwidth parameters used to smoothen the data.
- $K(\cdot)$: Kernel function, calculated as:

$$K(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}.$$

- $\hat{f}(x, y)$: Joint kernel density estimate for X and Y.
- $\hat{f}_X(x), \hat{f}_Y(y)$: Marginal kernel density estimates for X and Y.

Using the mathematical framework outlined in sections 3.5, 3.6 and 3.7 I estimate CVaR following these steps:

1. I use CDFs $F_X(x)$ and $F_Y(y)$ to transform historical asset returns to marginals that are uniform over [0,1], defined as:

$$u_i = F_X(x_i), \quad v_i = F_Y(y_i), \quad \text{where } u_i, v_i \in [0, 1].$$

2. To estimate the copula density $c(u, v)$ I apply kernel density estimation, leading to copula density being defined as:

$$\hat{c}(u, v) = \frac{\hat{f}(F_X^{-1}(u), F_Y^{-1}(v))}{\hat{f}_X(F_X^{-1}(u)) \hat{f}_Y(F_Y^{-1}(v))},$$

where $\hat{f}(x, y)$ is the joint density, and $\hat{f}_X(x)$, $\hat{f}_Y(y)$ are marginal densities.

3. By sampling from the copula I obtain simulated returns R_s. I use the inverse CDFs and transform each pair of sampled (u, v) to the original marginal distribution and define the portfolio return:
$$X = F_X^{-1}(u), \quad Y = F_Y^{-1}(v).$$
$$R_s = w_X X + w_Y Y,$$
where w_X and w_Y are the portfolio weights of assets X and Y.

4. I approximate equation (3.10) using simulated portfolio returns $R_{s,i}$
$$\mathrm{CVaR}_\alpha = \nu + \frac{1}{(1-\alpha)N} \sum_{i=1}^{N} \max(0, \nu - R_{s,i}), \tag{3.15}$$

5. Lastly, I solve the CVaR optimization problem where I minimize CVaR:
$$\min_{\mathbf{w},\nu} \nu + \frac{1}{(1-\alpha)N} \sum_{i=1}^{N} \max(0, \nu - \mathbf{w}^\top \mathbf{R}_i), \tag{3.16}$$
subject to:
$$\mathbf{w}^\top \boldsymbol{\mu} \geq r_{\text{target}}, \quad \sum_{j=1}^{n} w_j = 1, \quad w_j \geq 0.$$

4 Data analysis and results

The last section of the main part includes the data and results of the analysis. First, I plot the efficient frontier and its weights for the mean-variance model. Secondly, I perform the same procedure for the CVaR model. Lastly, I compare both models, outline the differences in results and discuss the reasons for these differences. I perform all calculations using Matlab programming language [6].

4.1 Portfolio return data

To examine the data, firstly, I plot asset returns of all eight securities. It includes: Apple (AAPL), Microsoft Corporation (MSFT), Alphabet (GOOGL), Amazon (AMZN), Netflix (NFLX), JPMrogan Chase and Co. (JPM), Tesla (TSLA), Bitcoin (BTC). From a first look at Figure 1 most of the distributions look mostly symmetric around the mean, which suggests a normal distribution as a good fit. However, some assets show deviations and asymmetry. All distributions display high kurtosis as they have sharp peaks and most of the data is concentrated around the mean. Returns with high kurtosis tend to have heavy tails and more outliers. BTC, TSLA and NFLX distributions have the heaviest tails, it suggests that they might suffer from extreme returns, both positive and negative. The most volatile asset is BTC with the range of returns from -0.4 to 0.2. The second most volatile is NFLX with returns ranging from -0.3 to 0.1. Other assets have narrower ranges of returns and falling somewhere in the range from -0.2 to 0.2.

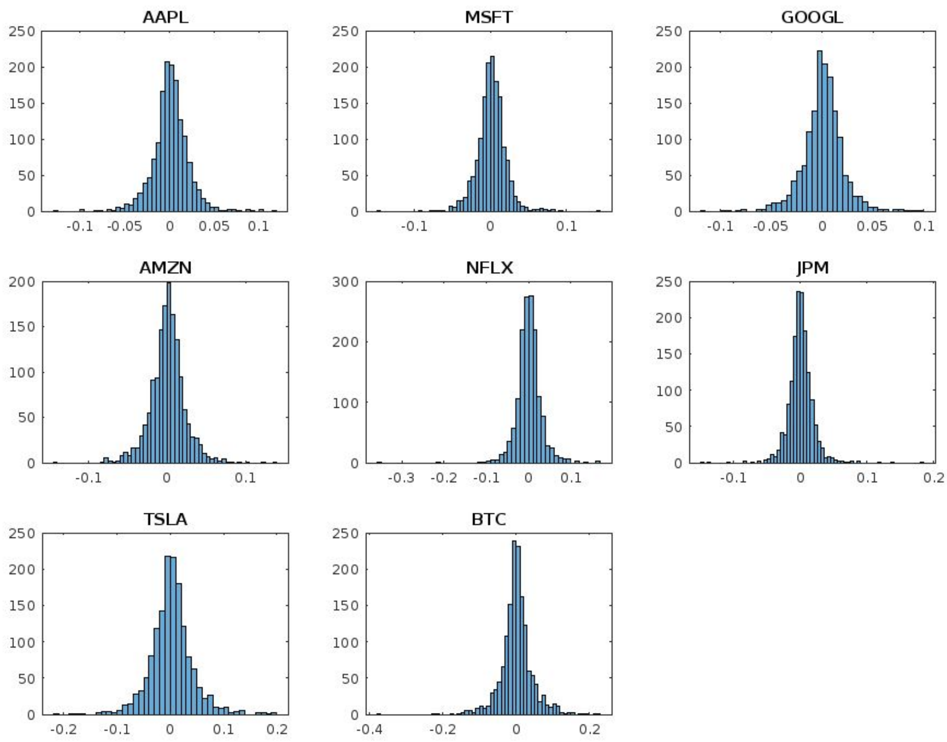

Figure 1: Histograms of asset returns

According to the returns data and the asset lifetime, I specify three asset groups based on the perceived risk profile. First includes stable stocks, such as APPL, MSFT, GOOGL, JPM exhibiting the most stable returns around the mean. The second group includes growth assets and includes AMZN and NFLX. Lastly, I define the third group as higher-risk assets and it contains TSLA and BTC which are newer assets then the rest.

4.2 Mean-Variance optimization results

Using Matlab built-in functions for mean-variance optimization, two conditions are met implicitly. First, it ensures a valid finite mean vector of returns. Second, the feasible portfolio set is a non-empty and compact set.

After the mean-variance optimization I plot the efficient frontier presented in Figure 2. The X-axis shows the risk measured by standard deviation and it ranges from around 1.5% to 4%. The Y-axis represents the expected return of a portfolio. As risk increases the mean return can rise from 0.8% to 2.4%. The efficient frontier is the steepest in the low risk segment up to standard deviation of 2%. In this segment returns begin from around 0.8% and can increase to 1.5% with standard deviation below 2%. This segment includes conservative, risk-averse investors, who choose portfolios with low volatility. The moderate risk segment starts where standard deviation is above 2%. It represents most of the investors with higher risk tolerance, as it finds a balance between moderate risk and higher return. Average return for this segment is from the range of 1,5% to 2%. The last segment is the high-risk segment, where the return increases but at a diminishing rate compared to the standard deviation. Portfolios from this segment might allocate more funds to the third group of assets. An investor would have to double their risk to 4% to get a higher return by only 0.4% compared to the moderate segment. For example, increasing the standard deviation from 1.5% to 2.5% would lead to an increase in return by 0.9 p.p., but rising risk from 3% to 4% would result in the growth of the mean return by just 0.4%.

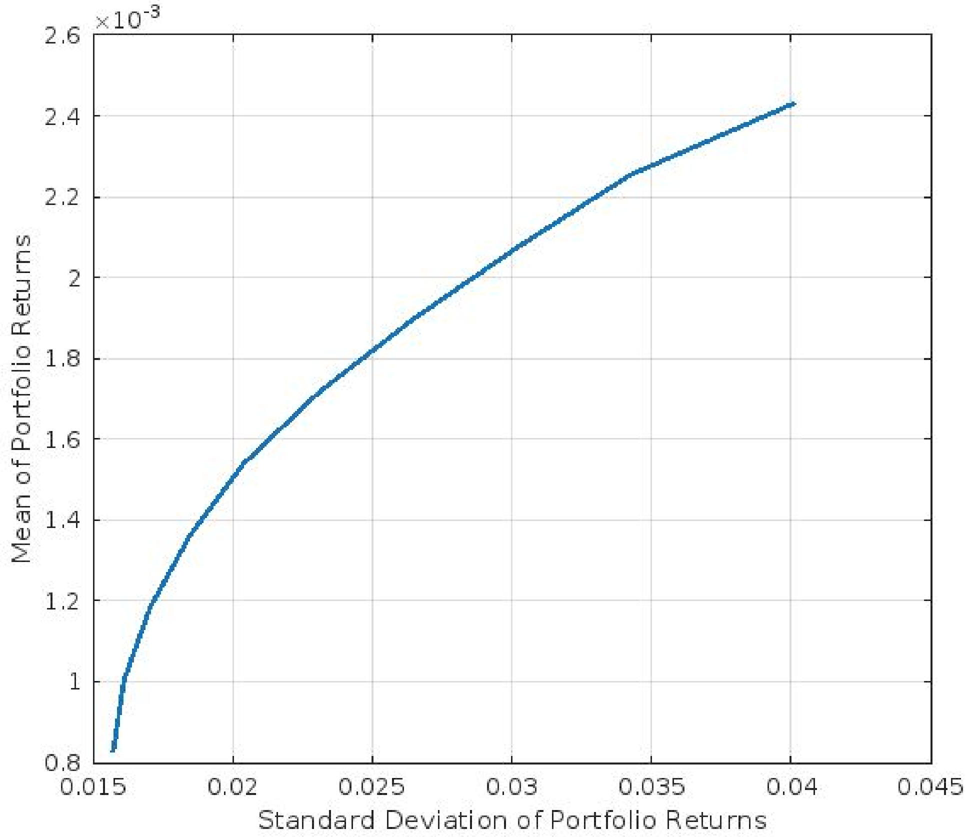

Figure 2: Efficient frontier for Mean-Variance estimation

Another way to analyze the efficient frontier is to plot ten portfolios that lie on the efficient frontier and compare their weights. The area plot in Figure 3 shows a mix of different assets in each of the plotted ten portfolios. The X-axis of this area plot represents each of the ten portfolios, from one to ten, while the Y-axis shows the weight assigned to an asset. Each color represents one of our eight assets and is indicated by the legend. Larger area shows that more funds are allocated to that asset. For example, if I want to analyze the second portfolio I find number two on the X-axis and see that blue area at the bottom (APPL) has a weight of 20% (Y-axis). This means that the second portfolio allocates 20% of funds to APPL. Portfolios are listed from the least risky (number 1) to the riskiest (number 10). First three portfolios are the most diversified and stable assets have the highest weight. For instance, the first portfolio includes all assets, focusing on stable stocks and allocating almost no funds to the riskier assets such as BTC and TSLA. Moreover, AAPL, MSFT and JPM account for around 80% of the total allocation for the third portfolio, indicating portfolio stability. Starting from the sixth portfolio the investment strategy turns to aggressive. Over 60% of the funds are allocated to TSLA and the rest to BTC. These high weights indicate the high return potential of the higher-risk asset group.

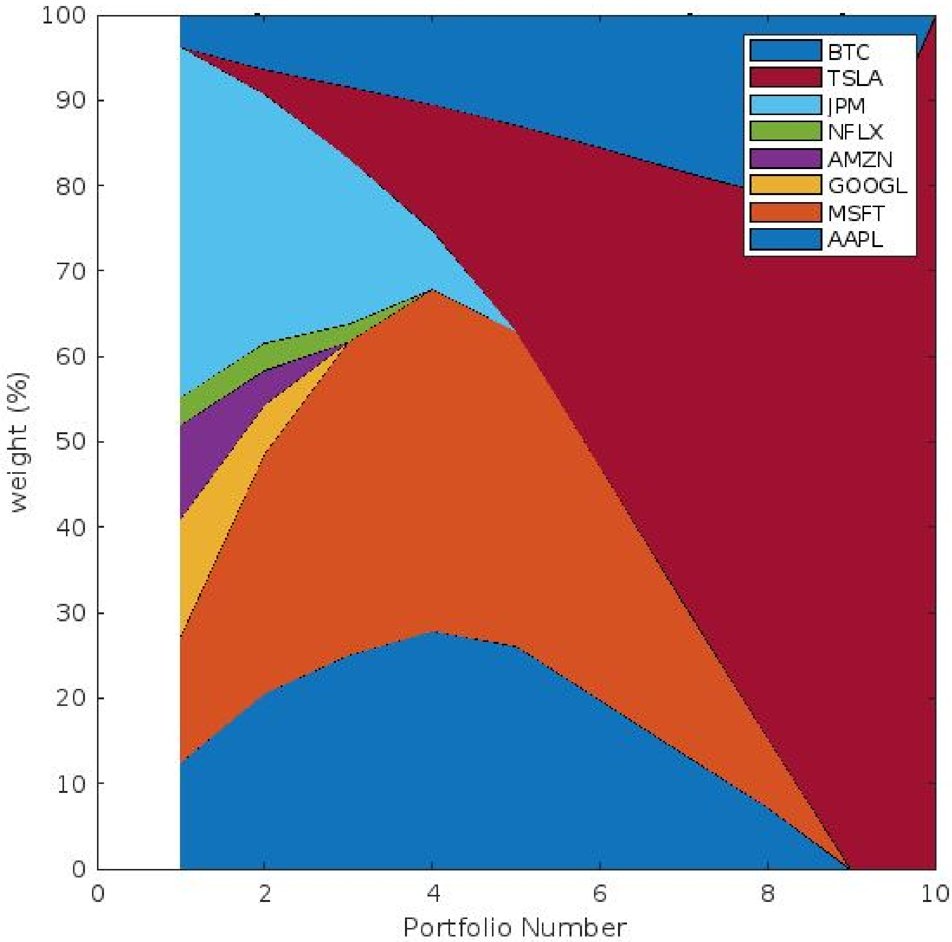

Figure 3: Weights for Mean-Variance portfolio

To find the highest risk-return combination among all portfolios on the frontier I estimate the tangency portfolio, for my risk-free rate I use Treasury bill rate. Tangency portfolio is an optimal portfolio that lies on the frontier where the efficient frontier is tangent to the capital allocation line. It has the highest Sharpe ratio, representing the optimal risky allocation, combining risky assets with the risk-free Treasury bills. This tangency portfolio is represented by a star in Figure 4 and has asset weights as indicated in Table 1. The standard deviation of the tangency portfolio is 2,1% with the expected return of 1,6%. From Table 1 we can see that this portfolio is comprised of four assets. Most of the weight is concentrated in the stable stocks, such as AAPL and MSFT. However, the tangency portfolio allocates over a tenth of the funds to Bitcoin, highlighting high potential future returns.

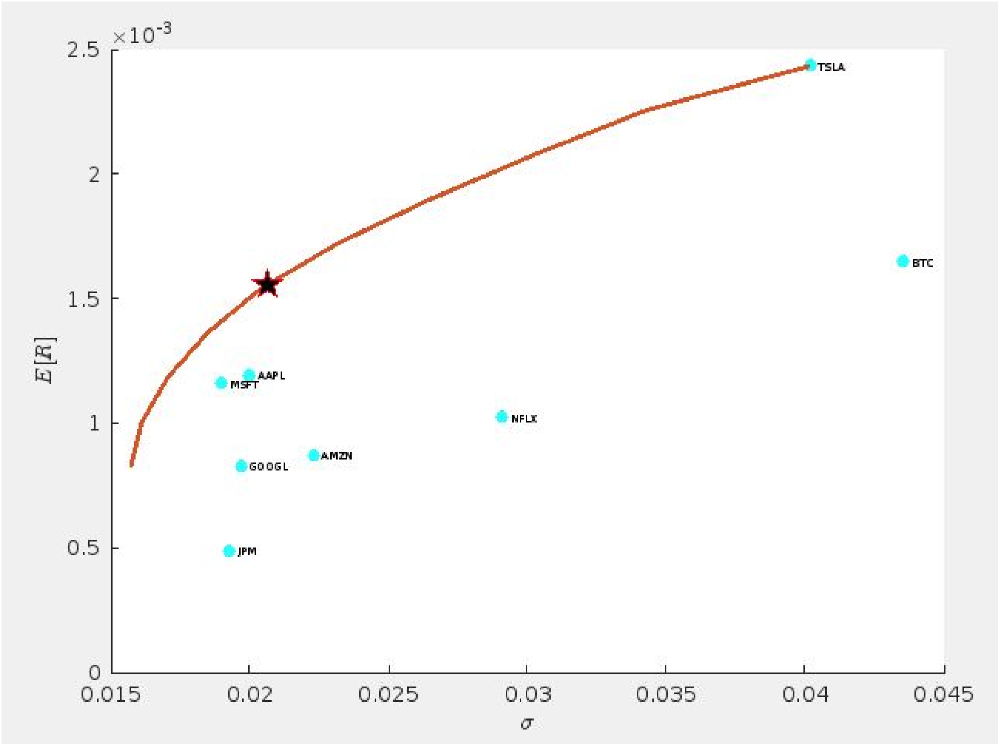

Figure 4: Tangent portfolio for Mean-Variance estimation

Asset	**Weight (%)**
AAPL	25.4
MSFT	35.8
GOOGL	0
AMZN	0
NFLX	0
JPM	0
TSLA	25.7
BTC	13.1

Table 1: Target portfolio weights

4.3 CVaR optimization results

Mean-variance optimization relies on the assumption of linearity and it fails to account for a more complex, non-linear dependence. Standard deviation might fail to account for the asymmetries and the tail risk of the more extreme events. As I include Bitcoin into my portfolio and its price fluctuates considerably, I need a more flexible approach to estimate the potential risk of high losses. That is why I implement a non-parametric CVaR return optimization model. Using copula estimation method with kernel density specification I create an empirical distribution and generate 1000 simulated scenarios. I plot simulated returns for each asset in Figure 5. Comparing the distributions from Figure 1 to Figure 5 they appear similar. Both historical and simulated returns exhibit comparable means, variances and the shapes of distributions. It indicates that the simulations accurately replicated the historical trends in the data. Importantly, the tails of the distributions align closely, as the CVaR focuses on the tail risk. Histograms suggest that using the simulated data CVaR should catch extreme events effectively.

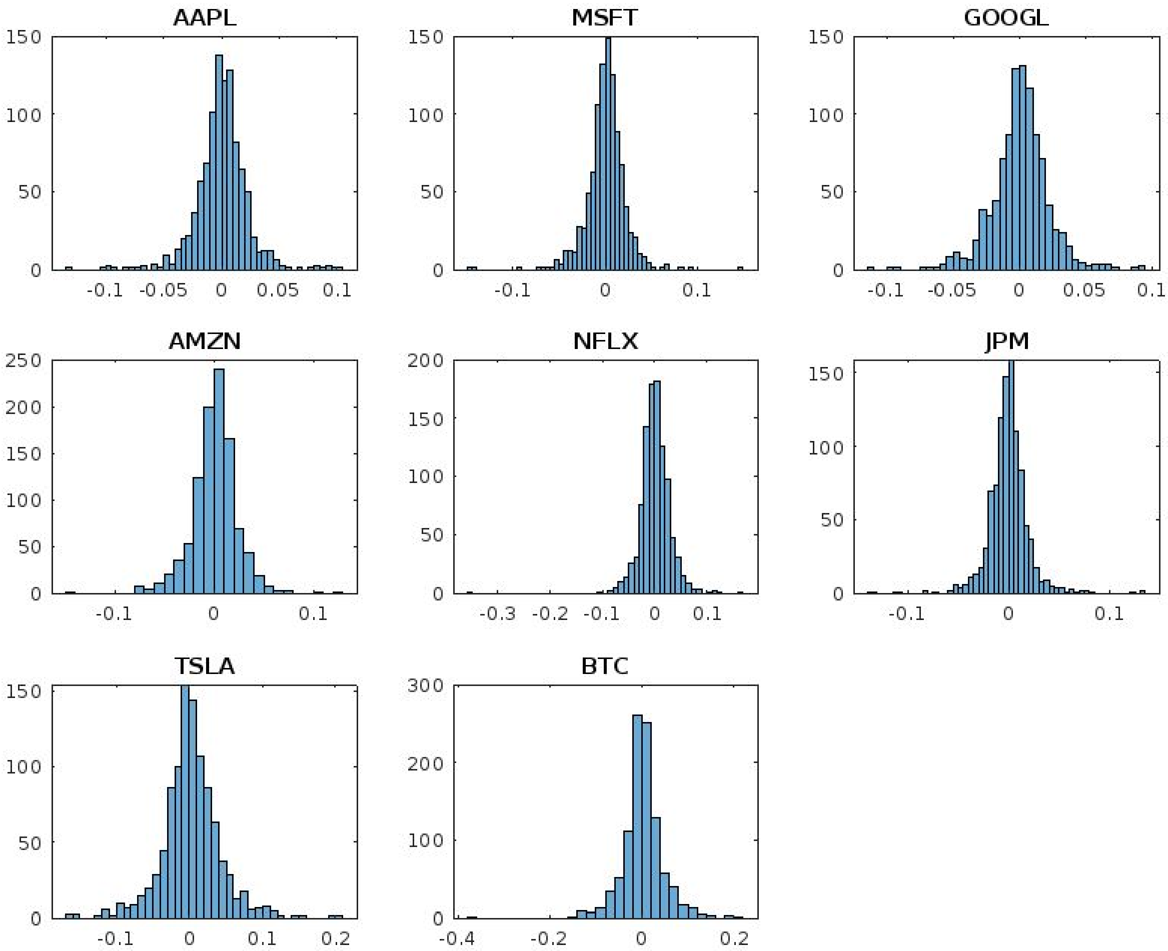

Figure 5: Histogram of returns for asset scenarios

After analyzing simulated return data I plot the efficient frontier using CVaR portfolio optimization. The X-axis shows portfolio CVaR focusing on the loss beyond the 95% confidence level and it ranges from around 4% to 8.5%. The Y-axis represents the expected return of a portfolio. With increase of risk the expected return can rise from 0.4% to 2.5%. Similarly to the mean-variance portfolio, CVaR efficient frontier exhibits diminishing marginal returns. For instance, increasing CVaR from 4% to 5% lead to a higher return by 0.8 p.p. However, moving from 7% to 8% return increases by 0.4 p.p. Highest return to CVaR ratio is achieved at around CVaR of 0.5%. The benefit of the CVaR model is that it minimizes the impact of market crisis, which is an especially important quality considering investing in high-volatility assets like Bitcoin.

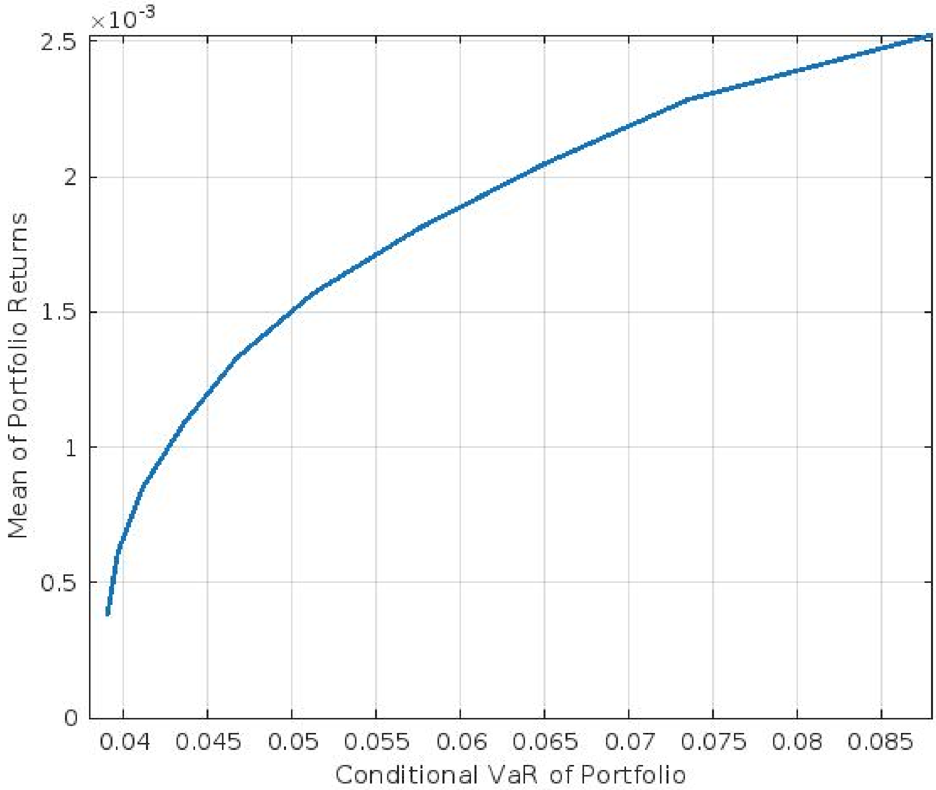

Figure 6: Efficient frontier for CVaR estimation

To ensure that the results of both models are comparable I also estimate ten portfolios for the CVaR optimization problem. These portfolio lie on the efficient frontier and I plot each of their weights in Figure 7. As before, the X-axis shows the portfolio number, while the Y-axis indicates its weight. Eight portfolios out of ten exhibit high diversification, some investing in all assets. On the far left I plot more conservative portfolios with around 85% of their funds allocated to the stable stock group. Each of those uses around 30% of their total funds to acquire GOOGL stock. Interestingly, even the conservative investors allocate some of their capital to Bitcoin, average Bitcoin weight is 10%. Risk-seeking investors on the far right do not diversify their portfolio and moderately increase Bitcoin weight to 25% and massively increase investment size to TSLA.

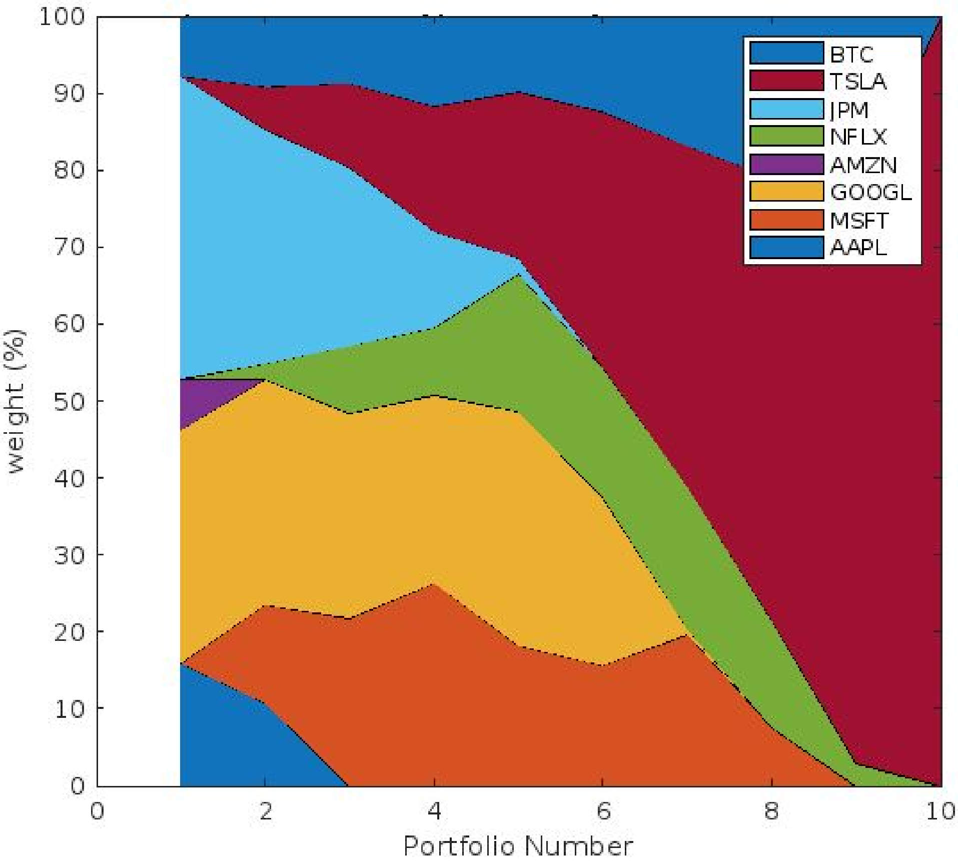

Figure 7: Weights for CVaR portfolio

4.4 Comparison of models

In this section I compare the efficient frontiers and the weight allocation for mean-variance estimation and the CVaR optimization. Both of these approaches optimize portfolio risk-return trade-off. However, they differ in their assumptions and view of risk. Mean-variance model minimizes portfolio variance or standard deviation, while CVaR model focuses on extreme losses and aims to catch the tail risk. I plot both mean-variance and CVaR efficient frontiers in Figure 8. In this figure the Markowitz mean-variance frontier showed in a blue dashed line is flatter than CVaR frontier plotted in a solid red line at lower levels of risk. For mean-variance portfolio an increase in returns would require a higher increase in risk. Markowitz efficient frontier becomes steeper at the far right of the graph, when the increase in risk has a diminishing growth in return. CVaR efficient frontier, on the other hand, is steeper at low levels of CVaR achieving higher returns for the same level of risk. It is in line with the specification of the method, as CVaR aims to provide insight for risk-averse investors and manage tail risk, especially, in the low-risk region. When both efficient frontier lines reach high risk at 8% they converge reflecting similar investment behavior in risky assets. The CVaR model provides an advantage for risk-averse investors by achieving higher expected returns for the same risk. They should be chosen by investors looking for stability and would like to avoid extreme losses. Markowitz mean-variance model performs well at high-risk segments focusing on the maximization of returns.

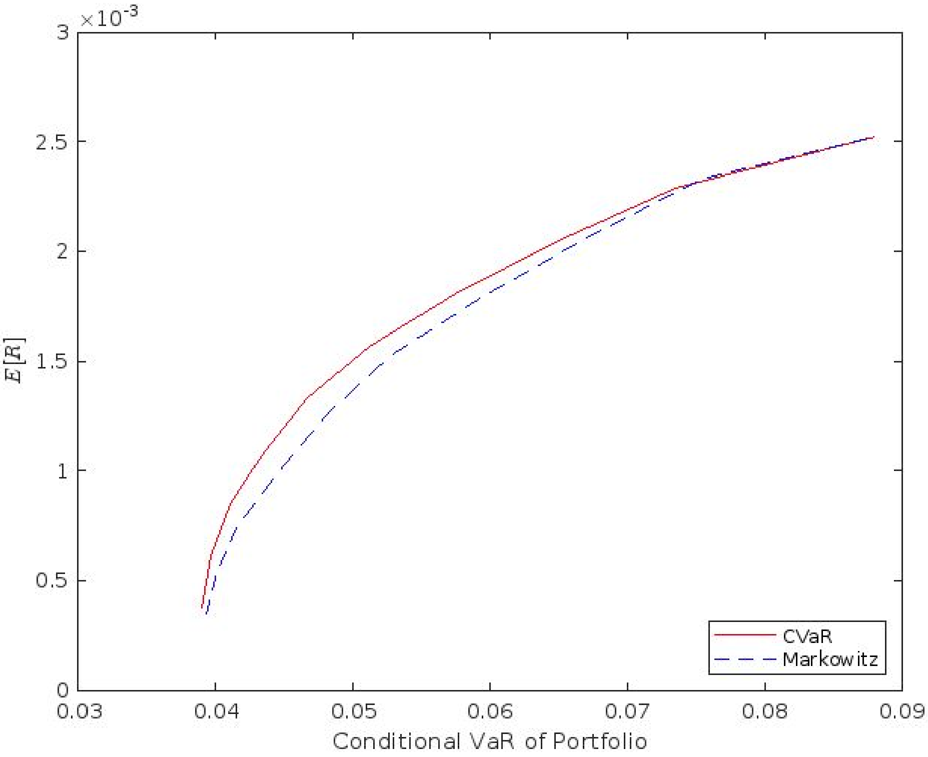

Figure 8: Comparison of efficient frontiers

I examine the weight allocations of assets within the portfolios to gain a deeper understanding about the structure of efficient portfolios for each model.

Comparing Markowitz mean-variance optimization model to CVaR I see an earlier introduction of BTC for CVaR portfolio. This introduction occurs because CVaR aims to mitigate the tail-risk focusing on the return potential and views BTC as a non-linear return driver. For low to moderate risk portfolios BTC stays at a stable rate of 10%, while mean-variance method starts with a minimal allocation to BTC and increases it with the growth of risk. Markowitz portfolios exhibit a sharp reduction in stable assets focusing on only three stable assets and decreasing their share as we move to risky portfolios. On the other hand, CVaR portfolio is well diversified, with 90% of funds allocated to stable or growth stocks. It gradually reduces the share of less risky assets as the risk profile changes from risk-averse to risk-seeking investors. It reflects a strong focus on tail-risk mitigation.

Overall, the weight allocation shows that Markowitz model invests in less number of assets and favors high risk-return trade-offs. On the other hand, CVaR model allows for a more diversified portfolio to avoid the extreme downside risk.

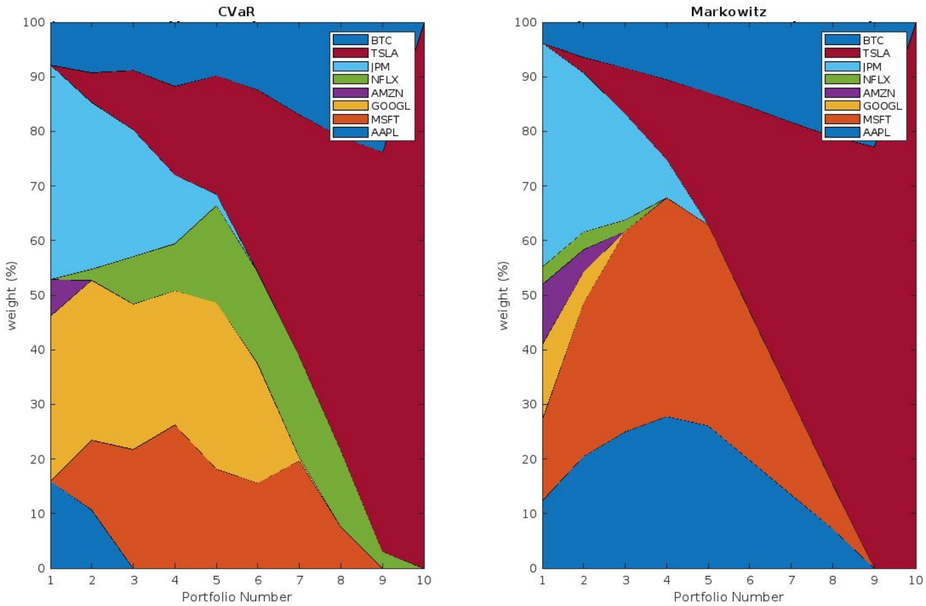

Figure 9: Comparison of weights

5 Conclusion

This thesis compares mean-variance method and the CVaR model used to optimize portfolio return, focusing on their efficient frontiers and portfolio weight allocations. The results show that while both models try to get the best risk-return trade-off, the difference in methodology leads to distinct outputs for investors based on their risk tolerance. Overall, CVaR efficient frontier is higher and steeper than mean-variance frontier for lower-risk portfolios. Markowitz's mean-variance efficient frontier becomes steeper and coincides with the CVaR efficient frontier at higher levels of risk where the increase in risk has diminishing returns. Results indicate that CVaR model aims to reduce tail risk and best suits risk-averse investors as it gets the highest return in low-risk regions. If an investor is looking for stability and minimal losses, then choosing a portfolio from the CVaR efficient frontier would yield the highest return. Mean-variance portfolio performs equally as good as CVaR at higher risk levels as both models allocate heavily to high-risk assets.

Additionally, CVaR model produces more diversified portfolios for low to moderate risk investors. It introduces highly volatile assets like Bitcoin earlier and keeps it at around 10%, as it views Bitcoin as a return diver. The mean-variance model starts with smaller allocation to Bitcoin and increases it when moving to more risky portfolios.

This thesis contributes to the existing literature by comparing mean-variance and CVaR models and analyzing not only traditional stocks, but also Bitcoin. This thesis highlights that Bitcoin as a high-return asset based on this research should be included even in low-risk portfolios. However, allocating most of the funds to it would still be a high-risk position. Further research could extend this analysis and evaluate how portfolio weights would change based on the market conditions, estimating separate efficient frontiers for bull and bear markets. Also, this CVaR model could be used to focus on only cryptocurrencies, such as Bitcoin, Ethereum, other emerging tokens, to assess the tail-risk during different market conditions.

References

[1] W. Bakry, A. R. Khaki, S. Al-Mohamad, and N. El-Kanj, *Bitcoin and Portfolio Diversification: A Portfolio Optimization Approach*, Journal of Risk and Financial Management, vol. 14, no. 7, 2021, article 282. `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3614606`

[2] S. Banihashemi and S. Navidi, *Portfolio Performance Evaluation in Mean-CVaR Framework: A Comparison with Non-Parametric Methods Value at Risk in Mean-VaR Analysis*, Operations Research Perspectives, vol. 4, 2017, pp. 21–28. `https://www.econstor.eu/bitstream/10419/178274/1/1-s2.0-S2214716016300665-main.pdf`

[3] R. De Matteis, *Fitting Copulas to Data*, Diploma Thesis, Institute of Mathematics, University of Zurich, June 2001. `https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=1c55760c6fda7782f7fc13f666d5e79343b41656`

[4] A. Eisl, S. M. Gasser, and K. Weinmayer, *Caveat Emptor: Does Bitcoin Improve Portfolio Diversification?*, SSRN Electronic Journal, 2015. `https://doi.org/10.2139/ssrn.2408997`

[5] H. Markowitz, *Portfolio Selection*, The Journal of Finance, vol. 7, no. 1, 1952, pp. 77–91. `https://www.jstor.org/stable/2975974`

[6] MathWorks Quant Team, *CVaR Portfolio Optimization*. MATLAB Central File Exchange, 2024. `https://www.mathworks.com/matlabcentral/fileexchange/38288-cvar-portfolio-optimization`. Accessed on December 5, 2024.

[7] R. C. Merton, *On Estimating the Expected Return on the Market: An Exploratory Investigation*, Journal of Financial Economics, vol. 8, no. 4, 1980, pp. 323–361. `https://www.sciencedirect.com/science/article/abs/pii/0304405X80900070`

[8] D. Neděla, *Comparison of Selected Portfolio Approaches with Benchmark*, Proceedings of the 38th International Conference on Mathematical Methods in Economics, Brno, Czech Republic, 2020, pp. 1–8. `https://www.researchgate.net/publication/344370105_Comparison_of_Selected_Portfolio_Approaches_with_Benchmark`

[9] R. T. Rockafellar and S. Uryasev, *Optimization of Conditional Value-at-Risk*, Journal of Risk, vol. 2, no. 3, 2000, pp. 21–41. `https://www.ise.ufl.edu/uryasev/files/2011/11/CVaR1_JOR.pdf`

[10] M. Semenov and D. Smagulov, *Portfolio Risk Assessment Using Copula Models*, arXiv preprint arXiv:1707.03516, 2017. `https://www.researchgate.net/publication/318392526_Portfolio_Risk_Assessment_using_Copula_Models`

[11] E. Symitsi and K. J. Chalvatzis, *The Economic Value of Bitcoin: A Portfolio Analysis of Currencies, Gold, Oil, and Stocks*, Research in International Business and Finance, vol. 48, 2019, pp. 97–110. `https://ideas.repec.org/a/eee/riibaf/v48y2019icp97-110.html`

[12] Yahoo Finance, *Historical Price Data*. Accessed on December 5, 2024. `https://finance.yahoo.com`

A Appendix

Below is the Matlab code used for this analysis:

```
datamystock = readtable('stock_closing_prices_2.xlsx');
sstocksandbtc = {'AAPL', 'MSFT', 'GOOGL', 'AMZN', 'NFLX', 'JPM', '
    TSLA', 'BTC'};
atrnr = numel(sstocksandbtc);
rett = tick2ret(datamystock{:, sstocksandbtc});
porfo = Portfolio('AssetList', sstocksandbtc, 'RiskFreeRate', 0.01/
    252);
porfo = estimateAssetMoments(porfo, rett);
porfo = setDefaultConstraints(porfo);
weigmystock = estimateMaxSharpeRatio(porfo);
[riskyone, returnone] = estimatePortMoments(porfo, weigmystock);
figure;
tobe = uitabgroup;
tobe1 = uitab(tobe, 'Title', 'EF');
ex = axes('Parent', tobe1);
[vid, kov] = getAssetMoments(porfo);
scatter(ex, sqrt(diag(kov)), vid, 'oc', 'filled');
text(sqrt(diag(kov)) + 0.0003, vid, sstocksandbtc, 'FontSize', 5);
hold on
[riskytwo, returntwo] = plotFrontier(porfo, 10);
plot(riskyone, returnone, 'p', 'MarkerSize', 15, 'MarkerEdgeColor', '
    r', 'MarkerFaceColor', 'k');
hold off;
figure;
plotAssetHist(sstocksandbtc, rett);
nrber = 1000;
astscen = simEmpirical(rett, nrber);
pppp = PortfolioCVaR('Scenarios', astscen);
pppp = setDefaultConstraints(pppp);
pppp = setProbabilityLevel(pppp, 0.95);
figure;
plotAssetHist(sstocksandbtc, astscen);
figure;
wghtone = estimateFrontier(pppp);
plotFrontier(pppp, wghtone);
plotWeight(wghtone, sstocksandbtc, '');
Number = 6;
plotCVaRHist(pppp, wghtone, rett, Number, 50)
portfwoogh = Portfolio;
portfwoogh = setAssetList(portfwoogh, sstocksandbtc);
portfwoogh = estimateAssetMoments(portfwoogh, rett);
portfwoogh = setDefaultConstraints(portfwoogh);
whgttwos = estimateFrontier(portfwoogh);
plotFrontier(portfwoogh,whgttwos);
```

```
plotWeight(whgttwos, sstocksandbtc, '');
prtone = estimatePortReturn(pppp,wghtone);
prttwo = estimatePortReturn(pppp,whgttwos);
priskone = estimatePortRisk(pppp,wghtone);
prisktwo = estimatePortRisk(pppp,whgttwos);
figure
plot(priskone,prtone,'-r',prisktwo, prttwo,'--b')
plotWeight2(wghtone, whgttwos, sstocksandbtc)
function scenstock = simEmpirical(ret,nScenario)
[no,nstock] = size(ret);
g = zeros(no,nstock);
for o = 1:nstock
    g(:,o) = ksdensity(ret(:,o),ret(:,o),'function','cdf');
end
[ro, dfff] = copulafit('t',g);
ri = copularnd('t',ro,dfff,nScenario);
scenstock = zeros(nScenario,nstock);
for o = 1:nstock
    scenstock(:,o) = ksdensity(ret(:,o),ri(:,o),'function','icdf');
end
end
function plotAssetHist(symbol,ret)
figure
stocknr = numel(symbol);
clr = 3;
rows = ceil(stocknr/clr);
for k = 1:stocknr
    subplot(rows,clr,k);
    histogram(ret(:,k));
    title(symbol{k});
end
end
function plotCVaRHist(p, w, ret, portNum, nBin)
portfoliostockrt = ret*w(:,portNum);
valueatrisk = estimatePortVaR(p,w(:,portNum));
condvalueatrisk = estimatePortRisk(p,w(:,portNum));
valueatrisk = -valueatrisk;
condvalueatrisk = -condvalueatrisk;
figure;
go = histogram(portfoliostockrt,nBin);
hold on;
ed = go.BinEdges;
calcl = go.Values.*(ed(1:end-1) < valueatrisk);
htwo = histogram('BinEdges',ed,'BinCounts',calcl);
htwo.FaceColor = 'r';
plot([condvalueatrisk;condvalueatrisk],[0;max(go.BinCounts)*0.80],'--
   r')
```

```
text(ed(1), max(go.BinCounts)*0.85,['CVAR = ' num2str(round(-
    condvalueatrisk,4))])
hold off;
end
function plotWeight(wght, smb, title)
figure;
wght = round(wght'*100,1);
area(wght);
title(title);
legend(smb);
end
function plotWeight2(weigonestock, wighstock2, smb)
figure;
weigonestock = round(weigonestock'*100,1);
area(weigonestock);
legend(smb);
wighstock2 = round(wighstock2'*100,1);
area(wighstock2);
legend(smb);
end
```