Computational Finance – Lesson 3

04/10/2024

ginevra.angelini@polimi.it

MPT - Drawbacks

Modern Portfolio Theory - Drawbacks

- Mean-Variance analysis represents one of the most used framework for risk management.

- After estimating the inputs (expected returns and variance-covariance matrix), the minimization algorithm allow us to find the portfolio that with equal expected value, is less risky between the possible portfolios.

- This framework has been criticized from several points of view.

  - MPT states that **variance is a comprehensive estimate of risk** and that **returns** are distributed following a **normal distribution**.

  - Then expected returns and the covariance matrix are not known, they have to be **estimated**.

Modern Portfolio Theory - Drawbacks

- Therefore the inputs of the optimization process suffer from two main problems:

  - **Model error**, the model that generates data is incorrect.

  - **Estimation Error**, the estimation procedure is subject to errors. Portfolios on the frontier are really sensitive to little variations of the inputs.

- Furthermore also the output have some problems. In fact often the portfolios on the frontier are **strongly concentrated** on 2-3-4 assets, failing the diversification principle.

- If we want to build a ***robust allocation***, i.e. an allocation strategy that is designed to be **more resilient** or less sensitive to variations in input parameters or assumptions, we have to find solutions to these problems.

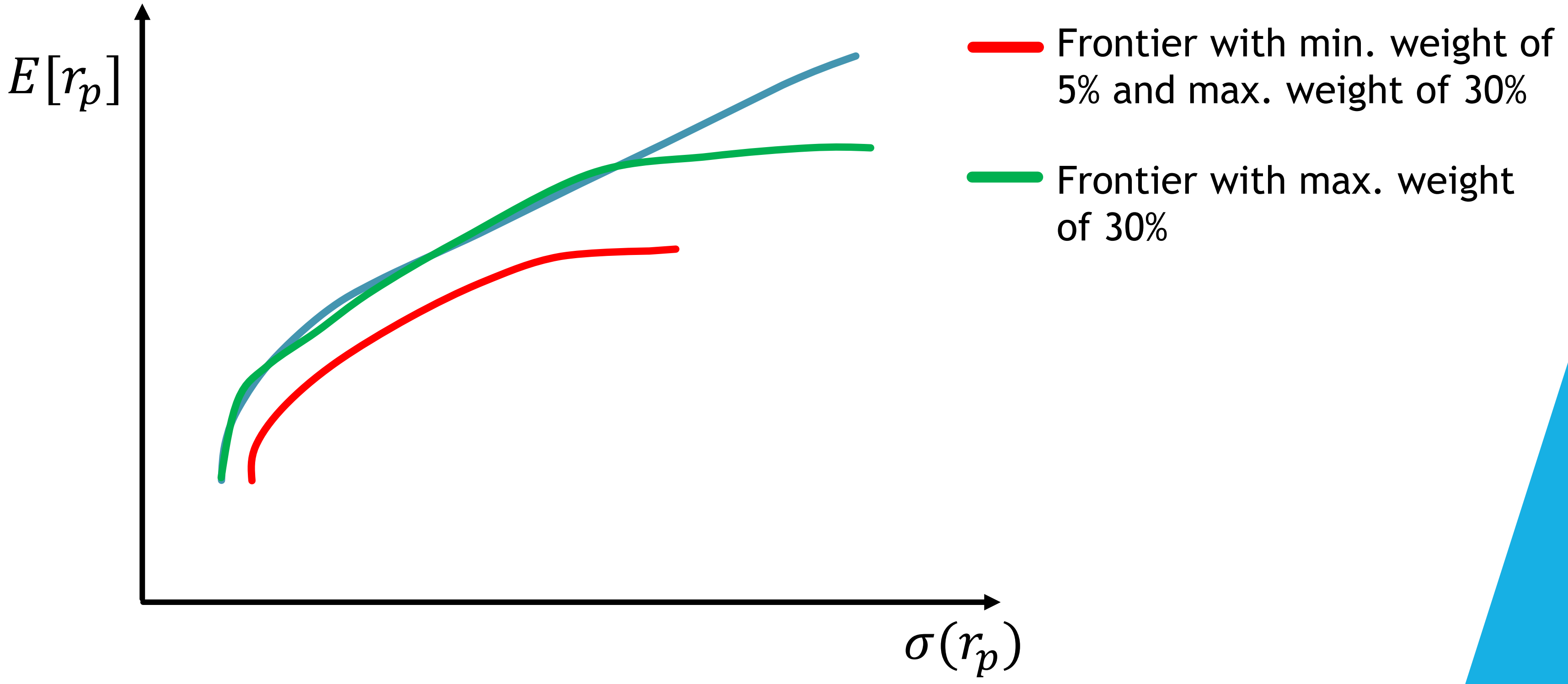
MPT – Robust Allocation

Concentration problem

- A possible solution for the concentration problem is to put a minimum and a maximum limit on the weights of the portfolio.

- For example we can choose to put a maximum limit exposition of 30% on each asset or put an exposition limit between the 5% and the 30% on each asset.

- But adding constraints lead us to a sub-optimal solution respect to the unbounded solution.



Estimation Error

- Financial data often exhibits characteristics that can violate the assumptions of classical statistical methods. These characteristics include:

  - **Outliers:** Financial data can be prone to outliers.
  - **Non-Normality**: Returns on financial assets often do not follow a normal distribution.
  - **Non-Constant Variance:** The variance of asset returns may not remain constant over time.
  - **Asymmetry:** Financial data can exhibit skewness, meaning it's not symmetrically distributed around the mean.

- Because of these factors, using classical estimators like sample mean and sample covariance matrix can lead to suboptimal portfolio construction.

- In order to minimize the effects of the estimation error we can use two techniques:

  - **Frontier Resampling**

  - **Robust Estimators / Shrinkage Estimators**

Portfolio Frontier - Resampling

- The main idea of this procedure consists in *repeating the optimization procedure* for several times using as inputs (expected returns and the covariance matrix) those estimated on simulated samples.

- In particular the procedure is composed by the following steps:

  i. Estimation of the expected returns e_0 and covariance matrix V_0 from data.

  ii. Generation of M samples of T observations. Each observation is the realization of a random variable independent and identically distributed as a ***Multivariate Normal random variable*** with mean e_0 and covariance matrix V_0.

  iii. Estimation for each sample of the array of Expected Returns e_i and of the Variance-Covariance matrix V_i,

  iv. Calculation for each couple (e_i, V_i) of the i-th portfolio frontier PF_i and the i-th optimal allocation $\overline{w}_i$.

  v. The final frontier is calculated as the ***mean of all the i-th frontiers***

Portfolio Frontier – Robust Estimators

- Before focusing on robust estimators, it is necessary to understand what are **biased estimators**. A biased estimator in statistics is an estimator that, on average, gives a value that differs systematically from the true population parameter it is intended to estimate.

- Let's suppose that we have some data following a probability distribution $P_\theta(x)$ and then we can construct a statistic $\hat{\theta}$ which serves as an estimator of θ based on observed data x. The ***bias*** **of** $\hat{\theta}$ is defined as:

$$\boxed{Bias(\hat{\theta},\theta) = E[\hat{\theta}] - \theta}$$

- An *estimator* is said to be *unbiased* if *its bias is equal to zero* for all values of parameter θ.

- It's important to note *that bias is a property of an estimator*, not of any particular sample. An estimator can be biased even if it produces a very accurate estimate for a specific sample.

- However, it's worth mentioning that **not all biases are necessarily bad**. In some cases, a biased estimator might have other desirable properties (such as lower variance or computational simplicity) that make it a better choice in certain situations. In other cases, unbiased estimators are preferred because they are asymptotically more efficient or have other desirable properties.

Robust Estimators

- A **robust estimator** is a statistical estimator that is ***not strongly affected by outliers*** or departures from assumptions that underlie the estimation procedure. In other words, it provides reliable estimates even when the data deviates from the ideal assumptions.

- Robust estimators are particularly useful in situations where the data may contain outliers, or where the assumptions of classical statistical methods (like normality or constant variance) may not hold. They are designed to be less sensitive to extreme observations.

- By using robust estimators, we can construct portfolios that are more resilient to extreme events, potentially leading to more stable and reliable investment strategies even in the presence of noisy or atypical data.

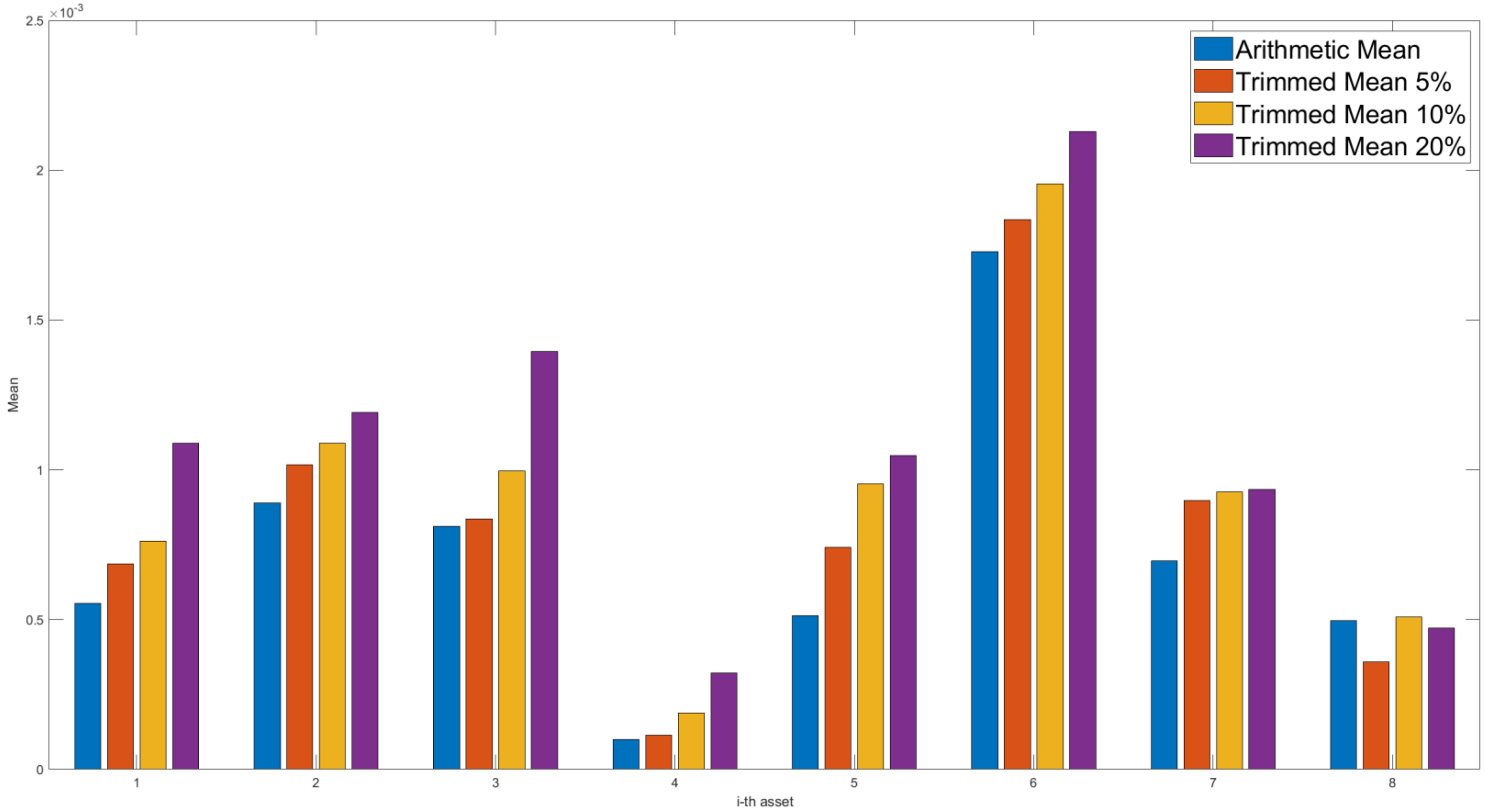
- Some common robust estimators used in portfolio optimization include:

    - ***Robust Mean Estimators***, i.e. trimmed mean

    - ***Robust Covariance Estimators***, i.e. Minimum Covariance Determinant (MCD) estimator

Robust Estimators – Trimmed Mean

- The **trimmed mean**, *i.e. truncated mean*, is a method of averaging that removes a small designated percentage of the largest and smallest values before calculating the mean. After removing the specified outlier observations, the trimmed mean is found using a standard arithmetic averaging formula.



- A trimmed mean is seen as a more realistic representation of a data set as the few erratic outliers have been removed that could otherwise potentially skew the information.

Robust Estimators – MCD

- MCD is a method for estimating the covariance matrix in a way that tries to minimize the influence of anomalies.

- **Objective:** The MCD estimator aims to find a subset of data points that collectively minimizes the determinant of the covariance matrix.

- **Calculation Process:**
  i. Random selection of a subset of data points from the dataset.
  ii. Calculation of the sample covariance matrix of the selected subset.
  iii. Computation of the determinant of this covariance matrix. The goal is to find the subset of points that minimizes this determinant.
  iv. Repeat steps 1,2,3 multiple times, each time selecting a different random subset.
  v. Choose the smallest determinant. The covariance matrix associated with this subset is considered the MCD estimator.

- **Advantages/Limits:** The MCD estimator is highly resistant to the influence of outliers in the data but it assumes that the data is approximately multivariate normal. In situations where this assumption is severely violated, the MCD estimator may not perform optimally.

Shrinkage Estimators

- **Shrinkage estimators** are a class of statistical estimators that reduce the impact of extreme or noisy observations in a dataset by "shrinking" or pulling the estimates towards a central value. This central value can be a global mean, a Bayesian prior, or some other suitable reference point.

- The motivation behind shrinkage estimators is to improve the precision or accuracy of parameter estimation, especially in situations where there is limited data or where the data may be noisy.

- There are two main types of shrinkage estimators:

  - ***James-Stein Estimator***, a frequentist shrinkage estimator

  - ***Bayesian Shrinkage Estimators***, a bayesian shrinkage estimator

- They both share the idea of improving parameter estimation through *shrinkage*, but they approach it from different statistical philosophies. James-Stein estimator is purely frequentist and uses sample data only, whereas the Bayes-Stein estimator incorporates prior information through Bayesian methods.

James – Stein Estimators

- The ***James-Stein*** is an alternative estimator for the multivariate means. The James-Stein estimator works by "shrinking" individual estimates towards a common average. This means that even if some of the individual estimates are influenced by outliers or extreme values, the resulting James-Stein estimate tends to be less sensitive to such deviations.

- The key idea is that when estimating multiple parameters simultaneously, the individual estimates are typically too variable. By shrinking them towards a common point, we can often achieve a more efficient and reliable overall estimate.

- The James-Stein estimator is defined as follows:

$$\boxed{\hat{\mu}_i^{JS} = e_0 + c_i(\overline{e}_i - e_0)}$$

e_0: is the global sample mean

$\overline{e}_i$: is the sample mean of asset i

σ_i^2: is the variance of asset i

N: is the number of assets $(N \geq 3)$

$$c_i = \max\left[0, 1 - \frac{(N-3)\sigma_i^2}{\sum(\overline{e} - e_0)^2}\right], 0 \leq c_i \leq 1$$

Bayes – Stein Estimators

- ***Bayesian shrinkage estimators*** are a class of statistical estimators that combine prior information with sample data to form estimates. These estimators "shrink" or adjust the estimates towards a central value, striking a balance between the prior information and the observed data.

- Here's how Bayesian shrinkage estimators work:

  - ***Prior Distribution:*** A prior distribution represents what we believe about the parameters before observing any data. It encodes our prior knowledge or beliefs about the parameters.
  - ***Likelihood:*** The likelihood represents the probability of observing the data given the parameters. It captures the information from the sample data.
  - ***Posterior Distribution:*** The Bayesian framework combines the prior information with the likelihood information from the data to form a posterior distribution. This is done using Bayes' theorem. The posterior distribution represents our updated knowledge about the parameters after observing the data.
  - ***Estimation:*** The estimator is derived from the posterior distribution. It's a combination of the prior information and the observed data.

Bayes – Stein Estimators

- So let's suppose that the sample means are distributed as a normal multivariate distribution with mean e_c and variance σ^2, $\bar{e} \sim N(e_c, \sigma^2 \mathcal{I})$ and we have a prior distribution of expected returns such that $e_c \sim N(e_0 1, \tau^2 \mathcal{I})$.

sample mean

- Applying Bayes Theorem, we can find that the generic expected return based on sample observations is:

$$\bar{e}_n^* = \frac{\tau^2}{\tau^2 + \sigma^2}\bar{e}_n + \frac{\sigma^2}{\sigma^2 + \tau^2} e_0 \longrightarrow$$

$\bar{e}_n$: is the sample mean of asset *n*

- This means that the estimates are *shrinked* trough the mean common value and the estimates of the expected returns are less influenced by the extreme observations.

- Thus one of the possible formulations of the Bayes-Stein estimator is:

$$\boxed{\hat{\mu}_{BS} = (1 - \alpha)\bar{e} + \alpha \mathbf{1}\underline{e}}$$

e_n: is the sample mean of asset n, i.e. n-th component of $\bar{e}$

$\underline{e}$: global sample mean

$\alpha = \frac{\lambda}{\lambda + T}$, T is the number of observations

$$\lambda = \frac{(N+2)(T-1)}{(\bar{e} - 1\underline{e})^T V^{-1}(\bar{e} - 1\underline{e})(T - N - 2)}$$