Aalto University
School of Science

Project plan

MS-E2177, Seminar on Case Studies in Operations Research
March 30, 2022

Lauri Jokinen	lauri.jokinen@aalto.fi	(project manager)
Christian Segercrantz	christian.segercrantz@aalto.fi	
Ilmari Vauhkonen	ilmari.vauhkonen@aalto.fi	

Client: Aktia Life Insurance

Contents

1 Background

The Markowitz model assumes that investor seeks to maximize mean returns and minimize risk. The model shows the expected returns and volatility (standard deviation of the returns, used to quantify the risk) for different portfolios. The optimal portfolio is then chosen, e.g., by fixing the expected returns and minimizing the risk.

The Markowitz theory does not account for external factors that are not visible in the asset sphere included in the model, as it assumes constant expected returns, volatility and cross-asset correlation. For example, changing inflation or economic activity cannot be expressed in the model just by changing parameters, but every single asset parameters should be adjusted. For example, inflation treats various industries in a different way (e.g. raw material producers versus end product manufacturers), so changes in inflation could affect not only the expected returns and volatilities of the assets, but also correlations between asset returns. Also, the optimal portfolio in Markowitz model is very sensitive to input data: with tiny changes in data, the model may suggest reinvesting large amounts of money for a tiny improvement in returns.

An advancement in portfolio optimization was the Black-Litterman model which deals with so called *views* [1]. The views are typically relative relations between two assets, and they have uncertainty factor assigned to them. Now, the market data is processed and filtered in a way that the views hold and we optimize the portfolio based on this data. Black-Litterman model also deals with mean and variance of returns, and the uncertainties of the views are assumed to be normally distributed. The Black-Litterman model has a closed form solution for the optimal portfolio, and is very popular today.

The two models are criticized because they assume that the portfolio's risk is the variance of returns. This means, e.g., that both greater and lower returns are marked as risks, while only the lower returns should be avoided. [2] Thus, an investor would benefit from being able to accommodate their definitions of risk. For more precise portfolio optimization we should consider non-parametric distributions. A non-parametric distribution captures a more precise picture of the market and it also handles outliers better.

In his paper, Meucci proposes a new method, called *entropy pooling*, which is based on the Black-Litterman method, but generalizes it to non-parametric distributions. [3] The paper walks through the steps of forming a posterior distribution from the prior market distribution, that takes the views into account, but disrupts the prior distribution as little as possible. This is accomplished by minimizing the entropy between the two distributions. The optimal portfolio is solved numerically by Bayesian optimization and Monte Carlo method.

By varying the views of the entropy pooling model, it can be used to stress test portfolios in different scenarios. Because the model is non-parametric, the simulation of even relatively rare historical events should be quite accurate.

However, the model has is downsides. For example, with extremely rare events, the posterior distribution is formed with a small amount of data which ultimately leads to inaccurate results. All the models discussed here are based on historical data, i.e., they do not forecast the market.

Portfolio optimization is a very much studied topic, and so it is very difficult to create portfolios which would outperform the existing ones. However, we may have information (views) which help build better performing portfolios. It is important to address the importance of the methods how the views are gathered from experts. With certain types of questions, humans are prone to show strong biases. Also, experts have different systematic biases that need to be tested and calibrated for.

2 Objectives

The objective of the project is to implement a Python code library for Meucci's Entropy pooling method. That is, the library will consist of Python files that contain functions used in data processing and computation, as well as scripts for running the procedure and analyzing the results. Some of the script files might be Jupyter notebooks, as it allows running code and printing outputs (numbers, tables, graphs) in a single file that can be converted to a PDF one. So, this could serve as a basis for a possible future repeated (and even automated) portfolio optimization procedure.

The library will have several separate sections that correspond to the phases of the entropy pooling and optimization procedures. First, code for reading and pre-processing the data needs to be written. The data will then be used to find the entropy-pooled parameters for the assets. Finally, using these newly obtained parameters, the Markowitz portfolio optimization procedure is performed. The results will be displayed, visualized and analyzed in ways that will be defined in detail once the project has advanced to such a point.

As requested by our client, Aktia Life Insurance, the Python code should be flexible and commented well for future development. This means that the procedure of each function is clarified either by writing sufficient comments in the code, or writing a separate documentation. Also, the superstructure of the code (which functions and in which order the script calls) must be presented in the documentation, as well as the versions of Python and all libraries used in the project.

Flexibility will be guaranteed by making sure that the code can be modified easily, for example, that some parameter values are not hard-coded and that the dimensions of data structures (input files, correlation matrices and the like) can vary.

In addition to the code library, a research report on the topic will be written. It will motivate the problem, review the literature, explain the approach and methodology and present and discuss the results obtained with the tool.

3 Tasks

The project can be broken down to coding tasks and writing the report. First task, and perhaps the most mathematically laborious task, is to formulate the views to suitable optimization constraints. These constraints and the prior distribution is then used in the second task, which is the entropy minimization algorithm. The minimization algorithm relies on Monte Carlo simulation and Bayesian optimization, and it will give us the posterior distribution.

The posterior distribution can then be used for two separate procedures: we could evaluate an optimal portfolio, or we can compare given portfolios and their returns under the circumstances set by the views. These two cover the third coding task.

The fifth coding task consists of visualizing the different key parts as well as possible. The prior and posterior distributions are important to visualize because they help the user approach and understand the data and results.

As a team, we will need to further specify these tasks, and agree on the overall code structure. We decided to assign the coding tasks as follows:

- Handling of scenario data and views: Lauri
- Implementing the entropy minimization algorithm: Ilmari
- Implementing the Markowitz portfolio optimization approach: Christian

With the code tasks finished, we will test the implementation for bugs, and make the library as user friendly as possible. At this point, we will present the implementation to our client, hear their thoughts about our work, and write the code documentation. Finally, we will summarize the project into a report.

Along the project, we aim to analyze and investigate the model with mathematical tools. A discussion about the model will be added to the report. We will discuss, e.g., about methods of gathering views data and using historical market data for forecasting purposes. In the code, we will add warnings for typical misuses, e.g., when results are based on a small data sample.

4 Schedule

The schedule is as follows:

1. Literature review (completed)
2. Dividing coding tasks and planning a detailed code structure (5 workdays; deadline 11.2.)
3. Working on tasks and writing code (22 workdays; deadline 15.3.)
4. Refining and commenting on code, producing results for the scenarios (10 workdays; deadline 29.3.)
5. Writing the final report and documentation (24 workdays; deadline 30.4.)

The schedule's tasks can overlap.

The scope of the project is well defined, but seems quite narrow with respect to the schedule. We may broaden the scope during the project, when we have a better understanding of the topic and the project's challenges.

5 Resources

The team members have a background in mathematics and operations research. The topic of the project belongs to investment science and the team possesses basic knowledge of it. The contact person from Aktia Life Insurance, Ville Hemmilä, has a long background in economics and investment and will thus be good support. The professor in charge of the project seminar is Ahti Salo, who will give feedback and assist in case we run into problems.

We use example data and views generated by Aktia Life Insurance, similar to that of real portfolios and views. Additional test data can be found online. In terms of tools, Python is a very flexible language with its wide range of packages, of which many will likely come to use. It also has a rather simple syntax and is very popular today.

6 Risks

The risks are listed in Table 1.

Risk	**Likelihood**	**Impact**	**Effect on project**	**How to prevent**
Problems in team collaboration	Low	High	Project schedule will be delayed	Meetings at regular intervals.
Lack of capabilities	Low	High	We're not able to implement the given model	More research, asking assistance and advice from Aktia or course staff
Insufficient communication between team and client	Medium	High	The client is not satisfied with the results	Clear communication between team members, project manager and the client. Also, staying on schedule will help.
The given model does not satisfy client requirements (even though the model is given to us by the client)	Medium	Medium	Project becomes irrelevant to the client	We evaluate the results of the model critically and on a regular basis with the client

Table 1: Risks

References

[1] Fischer Black and Robert Litterman. "Asset allocation: Combining investor views with market equilibrium". In: *Goldman Sachs Fixed Income Research* 115 (1990), pp. 7–18.

[2] Harry Markowitz. *Portfolio Selection. Efficient Diversification of investments*. Cowles Foundation for Research in Economics at Yale University, 1959. URL: https://cowles.yale.edu/sites/default/files/files/pub/mon/m16-all.pdf.

[3] Attilio Meucci. "Fully flexible views: Theory and practice". In: *Fully Flexible Views: Theory and Practice, Risk* 21.10 (2008), pp. 97–102.