ADVANCED METHODS IN PORTFOLIO OPTIMIZATION FOR TRADING STRATEGIES AND SMART BETA

by

Jason de La Borie de La Batut (CID: 01419438)

Department of Mathematics
Imperial College London
London SW7 2AZ
United Kingdom

Thesis submitted as part of the requirements for the award of the MSc in Mathematics and Finance, Imperial College London, 2017-2018

Declaration

The work contained in this thesis is my own work unless otherwise stated.

Signature and date:

Acknowledgements

The writing of this thesis was for me an essential experience. Hence, I would like to thanks all those who helped me to produce it.Being a student at Imperial College, in the Msc Mathematics and Finance, makes it possible and therefore I am grateful to all the administration, and especially the teachers, who provide each day a wonderful work.

I was pleased to get the opportunity to write this thesis in an Asset Management, hence directly with the financial industry. I would like to thank Emmanuel Boussard and Emmanuel Gavaudan, who gave me an internship, and proposed me a very interesting subject.

I had the extraordinary luck to work with Vincent Haye, Quantitative Analyst at BOUSSARD & GAVAUDAN, who was my intern supervisor during this summer project. Thanks to him, I discovered the systematic trading industry, and the main mathematical theories around it. Without him, I could not have write this paper.

I would like to thank my supervisor for his constant help and support, as well as for his insightful suggestions during the project, and all along the year. I first discoverd portfolio optimization in his course "Convex Optimization".

Contents

Introduction

"A good Portfolio is more than a long list of good stocks and bonds. It is a balanced whole, providing the investor with protections and opportunities with respect to a wide range of contingencies"

Harry Markowitz

In 1952 Harry Markowitz introduced a new mathematical framework in order to build a smart portfolio, called the Modern Portfolio Theory (MPT) [8], for which he later was awarded the Nobel Prize in economics. While Portfolio Management is one of the most relevant part of the work of a Bank or an asset Manager, the MPT is well adapted for systematic Portfolio Construction.

The goal is clear: an investor wants to have the highest return on his investment, but with the lowest possible risk. Unfortunately, this two conditions are not compatible. Indeed we can not imagine an economy in which one can invest in a product which gives almost surely the best yield. In such an economy, everyone would buy this product, and it would become the only available asset on the market. Obviously, the remaining possibility is to offer a way to manage yield versus risk.

Let's take a simple example to illustrate this idea. In a restaurant, imagine a dish (our product) such that if one chooses it, he would never be disappointed neither by the quality, nor by the taste (the asset is riskless). Imagine now that this dish is the cheapest one in the restaurant (we can not find a better yield on the market). Then every single customer would choose this dish, and no one would prefer to choose something else. Given that there are two possibilities for the owner of the restaurant: either to specialize into this dish, or increase its price. As we can not imagine a market with a single asset, the first solution is not applicable in the financial industry, and the price of such a product would increase. Without a perfect product, investors, just as restaurant customers, seek to manage risk versus yield (quality - price ratio).

But what is risk and what is yield? The strongest contribution of Markowitz in his theory was to model these two concepts with an accurate mathematical framework. As we will see in chapter one, the yield of an investment on a stock is, for Markowitz, the mean of the returns on the stock price; and the risk simply its variance. Then, the mathematical problem is easy to understand, and just a bit more difficult to implement. We need to maximize the returns and minimize the variance and find an equilibrium between these two optimization problems. This first kind of optimization (problemS because we can differentiate different goals) is described in chapter two.

At first glance, one could think that Markowitz found the perfect pattern for asset allocation. Unfortunately, or rather, fortunately, it's not the case, and we could quickly improve the MPT model. Indeed, Markowitz's optimization has many drawbacks, in particular, as soon as the size of the portfolio increases. Adding constraints can overcome these drawbacks, but it penalizes the final result. We will try to answer the following question for any situation:
Given a set of tradable assets and constraints, what is the best way to spread our money into these assets ?

In order to answer in a relevant way this question, we will go across all the different kinds of portfolio optimization we can find in the literature. Some of these optimization problems are classical and well known. Most of them must be adapted to some particularities in the data set. Few of them are totally new. The aim of this paper is to give to the reader all the abilities to find accurately the best available portfolio optimization, in function of the purpose. However, it will not present any spectacular new optimization.

During all this project, I tried to develop for BOUSSARD GAVAUDAN Asset Management, the best possible optimization, given a set of constraints which are specified by the company. The reader should keep in mind that all the following had been developed for an industrial purpose, and therefore, some details of the portfolio management will intentionally not be explored. For instance, all the portfolios used in this work are long only, stocks only. We will not invest into a riskless interest rate. Nevertheless, this thesis can be transposed in a more general framework. This can lead to more complicated calculus.

In a first chapter, we will present all the useful mathematical definitions for doing a Portfolio optimization. This includes both the mathematical theory around systematic asset management, but also the various techniques for solving numerically an optimization problem. The first chapter also deals with some technical issues that can make an optimization problem harder to solve. Then, chapter two and three focus on all the optimization problems that have been explored during this project, finishing with the most relevant one. And finally, in the last chapter, we will present all the results we obtained with these optimization problems.

1 Portfolio Management: context and useful tools

First, let us define the general scheme of a portfolio optimization problem.

1.1 Mathematical framework

1.1.1 Definitions

A portfolio is a collection of N assets $(A_1, ..., A_N)$, with a vector of weights $(\omega_1, ..., \omega_N)$. In all the following, we will use the notation for a portfolio $(\mathcal{A}, \Omega)$:
$(\mathcal{A} =: (A_1, ..., A_N); \Omega =: (\omega_1, ..., \omega_N))$,
or if it is not ambiguous: Ω

These assets can be stocks, bonds, derivatives..etc; but in all the following, because of the context of this thesis, we will focus our study on portfolios that contain only stocks on equity. Indeed, as the data we used were only times series on stocks, we did not try to transpose our results for portfolios with other kinds of products.

A weights ω_i is the amount of wealth invested in asset A_i in the portfolio $(\Omega; \mathcal{A})$. In order to be able to compare different portfolios, we will only study the "sum equals 1" portfolios. Each investment is in fact proportional to this standardized portfolio, which is characterized by $\sum_{i=1}^{N} \omega_i = 1$.

Each asset is characterized by its price in time, which is a stochastic process $(S_i(t))_t$. Using these notations, we can analyze the behaviour of a portfolio by studying the following weighting sum of stochastic processes (which is therefore itself a stochastic process):

$$S_t = \sum_{i=1}^{N} \omega_i S_i(t)$$

Remark 1.1: If we study a long only (ie all weights are positive) standardized portfolio, the previous quantity is in fact a convex sum. That point will be very helpful, as the portfolios used in this project are always long only.

However the quantity S_t is not easy to interpret. Let us rather define the rate of return of a portfolio. Assume at time t_0 one invests x_0 in a portfolio $(\Omega, \mathcal{A})$. The quality of this investment between t_0 and t_1 is totally described by its return R, which is the rate such that the wealth at time t_1 is given by $x_1 = Rx_0$. One can also define the rate of returns of this investment, which is r = R-1, such that one can write: $x_1 = (1 + r)x_0$, ie $r = \frac{x_1 - x_0}{x_0}$

When we try to find the best investment, in fact we focus on optimizing the behaviour of the return of this investment (or analogously, the rate of returns).

Define now the return R_i on asset i. x_1 is therefore the sum of each profit for each asset:

$$x_1 = \sum_{i=1}^{N}(\omega_i x_0)R_i = x_0 \sum_{i=1}^{N} \omega_i R_i = x_0 R$$

By identification: $R = \sum_{i=1}^{N} \omega_i R_i$
Hence:

$$r = R - 1 = \sum_{i=1}^{N} \omega_i R_i - \sum_{i=1}^{N} \omega_i = \sum_{i=1}^{N} \omega_i (R_i - 1)$$
$$r = \sum_{i=1}^{N} \omega_i r_i$$

The previous has been obtained between time t_0 and time t_1, but we can do the same for each t, and therefore we obtain the following stochastic process which assess the quality of a portfolio:

$$R_t = \sum_{i=1}^{N} \omega_i R_i^t$$
$$r_t = \sum_{i=1}^{N} \omega_i r_i^t$$

With $r_i^t = \frac{S_i(t+dt)-S_i(t)}{S_i(t)}$, and dt is the accuracy we have in our data set H, which consists in N time series of prices of each stock. In BOUSSARD & GAVAUDAN's database, dt = 1 day. For assessing the quality of an investment on a portfolio, it is therefore essential to model the random behavior of the rate of return.

Fisher Black and Myron Scholes showed [1] that each process $(S_i(t))_t$ is in a first approximation lognormal. They imagined that, for each i, there exists non temporal μ_i and σ_i such that:

$$log(S_i(t)) \sim \mathcal{N}(log(S_i^0) + (\mu_i - \frac{\sigma_i^2}{2})t, \sigma_i\sqrt{t})$$
$$S_i(t) = S_i^0 e^{(\mu_i - \frac{\sigma_i^2}{2})t + \sigma_i \mathcal{W}_i^t}$$
$$S_i^{t_1} = S_i^{t_0} e^{(\mu_i - \frac{\sigma_i^2}{2})(t_1 - t_0) + \sigma_i(\mathcal{W}_i^{t_1} - \mathcal{W}_i^{t_0})}$$
$$dS_i(t) = \mu_i S_i(t)dt + \sigma_i S_i(t) d\mathcal{W}_i^t$$

Where $(\mathcal{W}_i^t)_t$ is a Wiener process, and the last Stochastic Differential Equation has been obtained thanks to Ito's formula.

Let us define now a new quantity L_t which is the weighting sum of the log prices:

$$L_t := \sum_{i=1}^{N} \omega_i L_t^i, \text{ with:}$$
$$L_t^i = log(S_i(t))$$

Define now the infinitesimal log returns for asset i $\mathcal{R}_t$:

$$\mathcal{R}_t^i := dL_t^i$$
$$\mathcal{R}_t^i = d(log(S_i(t)))$$
$$\mathcal{R}_t^i = \frac{dS_i(t)}{S_i(t)}$$
$$\mathcal{R}_t^i = \mu_i dt + \sigma_i d\mathcal{W}_i^t$$
$$\mathcal{R}_t^i \sim \mathcal{N}(\mu_i dt, \sigma_i \sqrt{dt})$$

Hence, the infinitesimal log returns are normal, with constant parameters. Moreover, as $\mathcal{W}_i$ are Brownian motions, its increments are independent. Hence log returns are i.i.d normal laws.

If we only have daily time series on stocks price, one can assume that the previous result is still true if we take the daily returns R_t:

$$R_t^i = \Delta L_t^i$$
$$R_t^i \simeq \frac{\Delta S_i(t)}{S_i(t)}$$
$$R_t^i \sim \mathcal{N}(\mu_i \Delta t, \sigma_i \sqrt{\Delta t})$$

Where Δt = 1 day.

The rate of returns of a Portfolio, which is the weighted sum of its infinitesimal log returns, is hence, a sum of normal laws:

$$\mathcal{R}_t = \sum_{i=1}^{N} \omega_i dL_t^i = \sum_{i=1}^{N} \omega_i [\mu_i dt + \sigma_i d\mathcal{W}_i^t]$$

In a first approximation, we will consider that this rate of returns is equal to daily rate of returns. This quantity is easy to compute with daily time series on stocks price, :

$$
\begin{aligned}
r_t &= \sum_{i=1}^{N} \omega_i \Delta L_t^i \\
&= \sum_{i=1}^{N} \omega_i [\mu_i \Delta t + \sigma_i \Delta \mathcal{W}_i^t]
\end{aligned}
$$

This is assumed to be a sum of normal law, with constant parameters, which simplifies the problem.

Indeed, the mean function of the daily rate of returns can easily be computed using the vector of mean $(E[r_i])_{1 \leq i \leq N}$ of each rate of returns on stock i (this vector of mean is by construction independent of t):

$$E[r_t] = \sum_{i=1}^{N} \omega_i E[r_i^t] \simeq \sum_{i=1}^{N} \omega_i \mu_i \tag{1.1}$$

The variance of the process $(r_t)_t$ is also independent of t:

$$Var[r_t] = \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_i \omega_j Cov[r_i^t, r_j^t] \simeq \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_i \omega_j \sigma_{ij}, \tag{1.2}$$

as $\sigma_{ij} = \sigma_i \sigma_j d\mathcal{W}_i^t d\mathcal{W}_j^t$ is independent of t. If the price of the portfolio is lognormal as well, then the two previous quantities are the log mean and the volatility of the portfolio.

Remark 1.2: From now on, as we can not access the true value of μ_i and σ_{ij}, we will use this notation when we refer to an estimator of this quantity.

1.1.2 Asset Selection

We can divide a portfolio construction in four or three different steps (the first and the second one can be done at the same time).

Firstly, the portfolio Manager has to select an universe of assets. This universe is often as big as possible. It contains all the assets that could enter into the portfolio. An asset A is selected if and only if it satisfy the followings:

- The kind of product is part of the set of products that are traded by the portfolio manager;
- The asset can be traded (it still exists, it is quoted, it is not on the black list...);

As we will see later, this universe will be helpful for computing market signals. Practically, it is relevant to be able to load a database of prices for all the assets of the universe.

Once the universe is built, Portfolio managers apply a filter on the universe in order to select the relevant assets that they want to put into the portfolio. A systematic strategy is a quantitative filter, based on algorithms that select assets according to a decision rule. For instance, two famous decision rules are Trend Following and Mean Reversion. These two steps are called Asset Selections. In this thesis we will not describe any trading strategy.

The next step of the portfolio construction is the optimization. Different Portfolio Optimizations will be presented in the next chapter.

The sizing is the last step of the portfolio construction. It consists in the choice of the amount of wealth that is invested into the portfolio.

The different steps of a portfolio construction are summarized in the following graph:

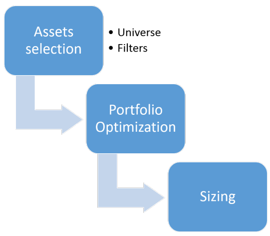

Figure 1: Steps of Portfolio Construction

1.1.3 Estimators

For many purposes, the quantities (1.1) and (1.2) need to be estimated. Indeed, as we will see in the next chapter, these metrics are directly linked with the objective functions in many optimization problems. However, as we can not access to the true value of the mean and the variance of the rates of returns, we need to build robust estimators. According to the first section *Definitions*, the rates of returns are i.i.d normally distributed, with constant parameters. Hence, for each i, we can see each rate of returns for any t, as independent observation of a normal law. Using this observation, we are able to build estimator of the classic metrics of a random variable, such as the mean, the variance, and the covariance matrix.

As the complexity of the problem should not lie in the computation of the estimator, simple estimator such as the sample one must be good enough. Otherwise, we should change the previous model for rate of returns. Recall that the sample mean and the sample covariance matrix, for n variables, and T observation, are given by:

$$\mu_i = \sum_{t=1}^{T} r_i^t$$

$$\Sigma = (\sigma_{ij})_{1\leq i,j\leq n}, \text{ with: } \sigma_{ij} = \frac{\sum\limits_{t=1}^{T}(r_{i,t}-\mu_i)(r_{j,t}-\mu_j)}{T-1}$$

However, even if the model is acceptable we can sometimes face a issue: the lack of data. Indeed, imagine we need to trade a stock which was quoted firstly only a few days ago. Then we will need to compute the covariance matrix of a data set with n stocks, but T<<n days of data. Therefore the covariance matrix will probably be singular, and we need to compute a more robust estimator for this case. The package sklearn.covariance of python is well adapted for this kind of issues. Therefore, as soon as we will have less than 10 days of data, we will use Ledoit and Wolf or OAS covariance matrix, which are shrunk estimator of the covariances matrix. See for more details [2] and [3] .

Remark 1.3 : Bias on the sample estimator.
The stocks are not quoted every day. Indeed, Saturdays, Sundays, and bank holidays introduce gaps into our time series. As week end are the same all over the world we can just skip those days. However, bank holidays may be different, and consequently, sometimes stock prices are the same two days in a row in our data set. For example imagine a portfolio with two stocks: a French one and a English one: FP:FP and RBS:LN (Total and Royal Bank of Scotland). The first of May is a bank holiday in France but not in England. Therefore, the price of one share of FP:FP is the same the 30^{th} of April 1985 and the 1^{st} of May 1985, but not the price of RBS:LN:

Date	FP:FP	RBS:LN
1985-04-29	0.670294	1.099526
1985-04-30	0.674864	1.127212
1985-05-01	0.674864	1.146988
1985-05-02	0.672325	1.139077

Hence there is a rate of returns associated of 0 at this date, which is a mistakes:

Rate of returns FP:FP	Rate of returns RBS:LN
0.68%	2.51 %
0%	1.75%
...	...

However it is quite inconvenient to look for all the incoherence in the data set while pandas allows us to compute immediately the rate of returns. Therefore there is an bias between the real sample estimator and the one we use. Assume that there is 1 mistakes for 30 days of observed rate of returns. Define $\mu_{i,python}$, $\sigma^2_{i,python}$ the sample mean and the sample variance for rate of returns of stock i obtained with python, and as previously, μ_i and σ_i^2 the true sample mean and variance, and supposed that we have T days of data $(r_{i,k})_k$.
then:

$$\mu_{i,python} = \frac{\sum_{k=1}^{T} r_{i,k}}{T-\frac{1}{30}T} = \frac{30}{29}\mu_i$$

$$\sigma^2_{i,python} = \frac{\sum_{k=1}^{T}(r_{i,k}-\mu_{i,python})^2}{\frac{29}{30}T-1} = \frac{\sum_{k=1}^{T}(r_{i,k}-\mu_i)^2-\frac{1}{29^2}T\mu_i^2}{\frac{29}{30}T-1} = \frac{(T-1)\sigma_i^2}{\frac{29}{30}T-1} - \frac{T\mu_i^2}{\frac{29^3}{30}T-29^2}$$

1.2 Structure of the code

1.2.1 Solving a Convex Program

Previously, we have seen that each investment in a portfolio is proportional to the associated standardized portfolio. The rate of return of the standardized portfolio is therefore a sum of random variables, where the sum of weights is equals to one. Assume that all the weights are positive. Then, this sum is a convex sum. There exists a very complete theory on convex optimization which can be helpful in the construction of optimizer as we will see in the next chapter. Then, there are a lot of advantages in using long only portfolio.

In this subsection, we will sketch how to solve a convex optimization, especially in the quadratic case.

Definition 1.1 : $f : X \to R$ is a convex function if and only if $\forall x_1, x_2 \in X, \forall t \in [0,1]$: $f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$.

Define, for a convex function f, the following program (P):

$$\min f(\omega)$$

$$s.t: \begin{cases} g_j(\omega) \leq 0 & \forall j \leq p \\ h_i(\omega) = 0 & \forall i \leq m \end{cases}$$

Where g_i and h_i are the constraints, and convex functions.
We say that (P) is feasible if there exists ω which satisfies the constraints of the program.

According to Kuhn and Tucker [4], ω^* is a solution of the program if and only if:

- Stationarity: $\nabla(\mathcal{L})(\omega^*) = 0$;
- Primal feasibility: ω^* is feasible;
- Dual feasibility: All the Lagrange multipliers for inequality constraints γ_i are positive or null;
- Complementary slackness: $\gamma_i.g_i(\omega^*)$ = 0;

With:
$\mathcal{L}(\omega) = f(\omega) + \sum\limits_{i=1}^{p} \gamma_i g_i(\omega) + \sum\limits_{i=1}^{m} \mu_i h_i(\omega)$

The package CVXOPT [5] of python provides a numerical solution for the previous program when the constraints are linear, and the objective function f is quadratic, ie when the program can be written:

$$\min \tfrac{1}{2}\omega^T.P.\omega + q^T.\omega$$

$$s.t: \begin{cases} G.\omega \leq h \\ A.\omega = b \end{cases}$$

With $\omega \in \mathrm{R}^N$, $G \in \mathrm{R}^{N*m}$, $h \in \mathrm{R}^m$, $A \in \mathrm{R}^{p*N}$ and $b \in \mathrm{R}^p$, ie there are m inequality linear constraints, and p equalities linear constraints.

Given this numerical solution easy and quick to implement, we will always try to transform our problems into such a program. This numerical solutions are obtained with Cone Programming, and especially Interior-point methods for large-scale cone programming. For more details, see [5], and [6].

Sometimes we will face non quadratic problems (see Equal Risk Contribution). In such a case, we will use for numerical solutions the package scipy.optimize from python. It calls Sequential Quadratic Porgramming algorithms that are more general, but besides, slower to converge. For more details, see [7].

1.2.2 Oriented object Programming

The industrial aim of the project at BOUSSARD & GAVAUDAN is to build a solver that could then be used for any purpose and for any portfolio. Given this issue, we will try to sketch in this section how the code can be engineered.

Because of the previous points, Python appears to be a good coding solution for Portfolio Optimization purposes. We chose to use an oriented object programmation.

Define a class Portfolio which contains all the useful tools for the characterization of a portfolio. This object instantiate the current date, the list of assets in the Portfolio, the prices and rate of returns data set, different estimators of the mean, of the covariance matrix and of the correlation matrix. While the main program compute the back test, an optimizer perform the optimization. Finally an other function assesses the quality of the backtest by producing risk metrics from the $P\&L$.

2 Traditional Portfolio Optimizations

A portfolio optimization can occur in many different contexts, and each optimization is unique. As we can not define a perfect goal for the optimization, it is therefore very difficult to produce a user guide of the construction of a portfolio given a set of assets. It is firstly hard to choose which target function we want to optimize. We list in this section different possible targets, but the last one, the diversification ratio, has been considered when using the optimizer as the best one.

Secondly, a portfolio optimization largely depends on the choice of the constraints. Except during the confrontation between Markowitz Optimization and Most Diversified Portfolio, we will use the constraints $\mathcal{C}$ that are described in the previous chapter. Recall $\mathcal{C}$:

- All weights are positive; (constraints (1))
- The sum of the weights is 1; (constraints (2))
- The weights are not too far apart, ie the maximum weight does not exceed 10 times the minimum weight; (constraints (3))

Consequently, in all the following, we will describe a way of building a portfolio optimization in this particular context which is the set of constraints used at BOUSSARD & GAVAUDAN, but it may vary a bit from one asset manager to another. Nevertheless, as we have seen in chapter one, this is a logical choice of set of constraints.

For each of the following optimization problems we use the same notations as previously:
Σ : Estimator of the covariance matrix of the *returns* ;
μ : Estimator of the vector mean of the *returns* ;
σ : Estimator of the standard deviation of the *returns* ;
ω : Vector weight of the portfolio ;

2.1 Classical Portfolio Optimization: The modern Portfolio Theory

As we have seen in introduction, Markowitz was the first one to build a mathematical framework in order to optimize a portfolio with more than two assets. This portfolio must satisfy a certain yield while limiting the associated risk. In this section we will develop the modern portfolio theory as it is described by H.Markowitz in the article Portfolio Selection [8]

Markowitz measured the yield of its portfolio with an estimator of the expected return $E[R] = \sum_{k=1}^{n} \omega_k E[r_k]$, and the risk with an estimator of the variance of the portfolio $Var(R) = \sum_{i,j}^{n} \omega_i \omega_j Cov[r_i, r_j]$. Using the notation of chapter 1, these estimators can be written with estimators of the means vector and the covariance matrix . Respectively: $\omega^T.\mu$ and $\omega^T.\Sigma.\omega$.

2.1.1 Minimum Variance

Let's first assume that we are only trying to minimize the risk associated to the portfolio. Given the constraints, the construction comes down to the following problem:

$$\min Var(R)$$
$$s.t: \begin{cases} \sum_{i=1}^{n} \omega_i = 1 \\ \omega_i \geq 0 \\ \omega_{max} \leq 10.\omega_{min} \end{cases}$$

Using estimators instead of the real value of the variance, this is equivalent to:

$$\min \omega^T.\Sigma.\omega$$
$$s.t: \begin{cases} \sum_{i=1}^{n} \omega_i = 1 \\ \omega_i \geq 0 \\ \omega_{max} \leq 10.\omega_{min} \end{cases}$$

and can be reduced by adding two variables s_1 and s_2 corresponding to ω_{max} and ω_{min} to the following convex problem:

$$\min \omega^T.\Sigma.\omega$$
$$s.t: \begin{cases} \sum_{i=1}^{n} \omega_i = 1 \\ \omega_i \geq 0 \\ \omega_i \geq s_1 \\ \omega_i \leq s_2 \\ s_2 \leq 10.s_1 \end{cases}$$

Ie, with a writing adapted to the numerical resolution of convex problem with linear constraints (via CVXOPT, as we have seen in the previous chapter):

(With $x = [\omega_1 \ldots \omega_N \, s_1 \, s_2]^T$, $\Sigma' = \begin{bmatrix} \Sigma\,0\,0 \\ 0\,0\,0 \\ 0\,0\,0 \end{bmatrix}$)

$$\min x^T.\Sigma'.x$$
$$s.t: \begin{cases} \begin{bmatrix} 1_N^T & 0 & 0 \end{bmatrix}.x = [1] \\ \begin{bmatrix} -I_N & 0 & 0 \end{bmatrix}.x \leq [0] \\ \begin{bmatrix} -I_N & 1 & 0 \\ I_N & 0 & -1 \end{bmatrix}.x \leq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0_N & -10 & 1 \end{bmatrix}.x \leq [0] \end{cases}$$

2.1.2 Maximum returns

On the other hand, if we only focus on the *maximize returns* condition, then the problem is written:

$$\max E[R]$$
$$s.t: \begin{cases} \sum_{i=1}^{n} \omega_i = 1 \\ \omega_i \geq 0 \\ \omega_{max} \leq 10.\omega_{min} \end{cases}$$

which is mathematically equivalent to:

$$\max \omega^T.\mu$$
$$s.t: \begin{cases} \sum_{i=1}^{n} \omega_i = 1 \\ \omega_i \geq 0 \\ \omega_{max} \leq 10.\omega_{min} \end{cases}$$

and can be reduced by adding two variables s_1 and s_2 corresponding to ω_{max} and ω_{min} to the following convex problem:

$$\max \omega^T.\mu$$
$$s.t: \begin{cases} \sum_{i=1}^{n} \omega_i = 1 \\ \omega_i \geq 0 \\ \omega_i \geq s_1 \\ \omega_i \leq s_2 \\ s_2 \leq 10.s_1 \end{cases}$$

Hence:
(With $x = [\omega_1 \ldots \omega_N \; s_1 \; s_2]^T$, $\mu' = \begin{bmatrix} \mu \, 0 \, 0 \end{bmatrix}$)

$$\max x^T.\mu'$$

$$s.t: \begin{cases} \begin{bmatrix} 1_N^T & 0 & 0 \end{bmatrix}.x = [1] \\ \begin{bmatrix} -I_N & 0 & 0 \end{bmatrix}.x \leq [0] \\ \begin{bmatrix} -I_N & 1 & 0 \\ I_N & 0 & -1 \end{bmatrix}.x \leq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0_N & -10 & 1 \end{bmatrix}.x \leq [0] \end{cases}$$

2.1.3 Mean - Variance

The modern portfolio theory of *H.Markowitz* is based on the combination of these two conditions. A reasonable investor does not just want to maximize the return on his portfolio (because he is exposed to too much risk) or just minimize the associated risk (because he may have too much poor performance). There are two main techniques for combining these conditions, as stated by Markowitz. The first idea is to optimize a combination of the two conditions. When one wants to maximize a function f_1 and to minimize another f_2, one can reduce the problem to the minimization of a weighted difference between f_2 and f_1. The weighting factor is used to control the optimization strategy (if we focus more on limiting the risk, or maximizing the yield). We obtain then a parametric optimization.

The problem is then:

$$\min Var(R) - \lambda.E[R]$$

$$s.t: \begin{cases} \sum_{i=1}^{n} \omega_i = 1 \\ \omega_i \geq 0 \\ \omega_{max} \leq 10.\omega_{min} \end{cases}$$

Or:

$$\min \omega^T.\Sigma.\omega - \lambda.\omega^T.\mu$$

$$s.t: \begin{cases} \sum_{i=1}^{n} \omega_i = 1 \\ \omega_i \geq 0 \\ \omega_{max} \leq 10.\omega_{min} \end{cases}$$

and can be reduced by adding two variables s_1 and s_2 corresponding to ω_{max} and ω_{min} to the following convex problem:

$$\min \omega^T.\Sigma.\omega - \lambda.\omega^T.\mu$$

$$s.t: \begin{cases} \sum_{i=1}^{n} \omega_i = 1 \\ \omega_i \geq 0 \\ \omega_i \geq s_1 \\ \omega_i \leq s_2 \\ s_2 \leq 10.s_1 \end{cases}$$

Hence:

(With $x = [\omega_1 \ldots \omega_N \; s_1 \; s_2]^T$, $\mu' = \begin{bmatrix} \mu \, 0 \, 0 \end{bmatrix}$, and $\Sigma' = \begin{bmatrix} \Sigma \, 0 \, 0 \\ 0 \, 0 \, 0 \\ 0 \, 0 \, 0 \end{bmatrix}$)

$$\min x^T.\Sigma'.x - \lambda.x^T.\mu'$$

$$s.t: \begin{cases} \begin{bmatrix} 1_N^T & 0 & 0 \end{bmatrix}.x = [1] \\ \begin{bmatrix} -I_N & 0 & 0 \end{bmatrix}.x \leq [0] \\ \begin{bmatrix} -I_N & 1 & 0 \\ I_N & 0 & -1 \end{bmatrix}.x \leq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0_N & -10 & 1 \end{bmatrix}.x \leq [0] \end{cases}$$

This parametric technique is very close to the next one and it has a very powerful geometric interpretation. If you plot on a graph all possible combinations of *Mean / Variance* for a given portfolio, then there will be a limit when risk decreases and yield increases. This limit is called *Efficient Frontier*. Each point on this boundary is a risk minimization for a given yield level.

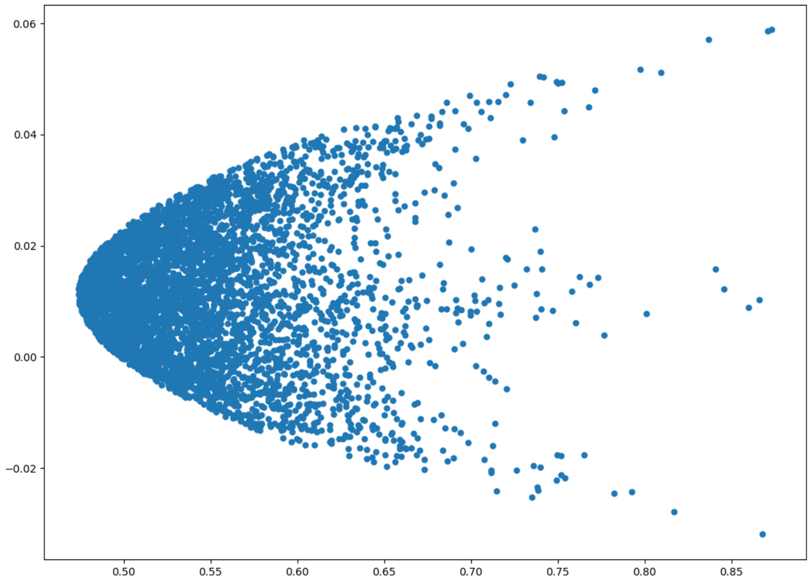

Figure 2: Theoretical Efficient Frontier, *returns* reduced normal centered law. 50,000 random portfolios

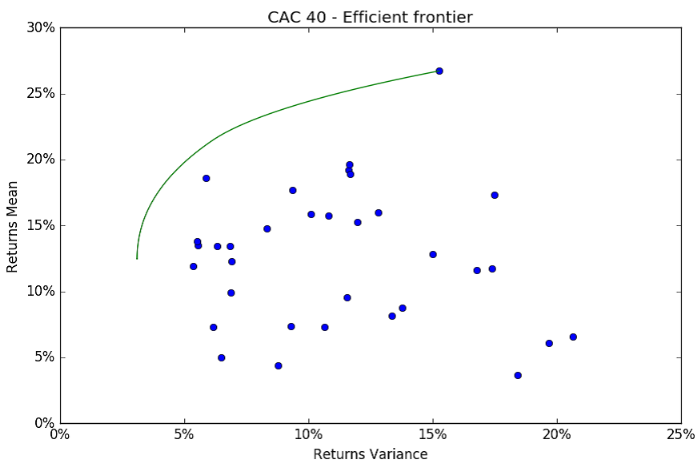

Figure 3: Efficient Frontier. CAC 40 stocks textit Advanced Topics in Operations Research Scenario-Based Portfolio Optimization by Edouard BERTHE [9]

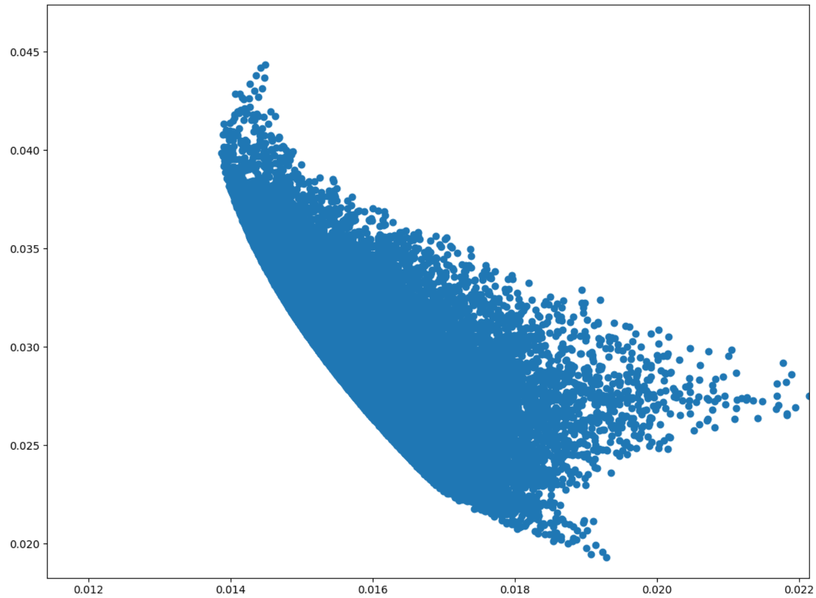

Figure 4: Portfolio of 4 assets from the CAC40 ; 50,000 random portfolios

A parametric equation of this frontier so-called "Efficient Frontier" is the previous optimization. However there exists a more explicit way to characterize this frontier. We can rewrite the previous optimization as follow: minimizing the risk for a given level of expected returns.

Let us write an equivalent way of leading a mean variance optimisation:

$$\min Var(R)$$

$$s.t: \begin{cases} \sum_{i=1}^{n} \omega_i = 1 \\ \omega^T.\mu = R \\ \omega_i \geq 0 \\ \omega_{max} \leq 10.\omega_{min} \end{cases}$$

which is mathematically equivalent to:

$$\min \omega^T.\Sigma.\omega$$

$$s.t: \begin{cases} \sum_{i=1}^{n} \omega_i = 1 \\ \omega^T.\mu = R \\ \omega_i \geq 0 \\ \omega_{max} \leq 10.\omega_{min} \end{cases}$$

and can be reduced by adding two variables s_1 and s_2 corresponding to ω_{max} and ω_{min} to the following convex problem:

$$\min \omega^T.\Sigma.\omega - \lambda.\omega^T.\mu$$

$$s.t: \begin{cases} \sum_{i=1}^{n} \omega_i = 1 \\ \omega^T.\mu = R \\ \omega_i \geq 0 \\ \omega_i \geq s_1 \\ \omega_i \leq s_2 \\ s_2 \leq 10.s_1 \end{cases}$$

Hence, we obtain the following convex problem. If we release all the constraints $\mathcal{C}$, the solution is the plot of the Efficient frontier with R in ordinate.

(With $x = [\omega_1 \dots \omega_N \; s_1 \; s_2]^T$, and $\Sigma' = \begin{bmatrix} \Sigma\,0\,0 \\ 0\,0\,0 \\ 0\,0\,0 \end{bmatrix}$)

$$\min x^T.\Sigma'.x$$

$$s.t: \begin{cases} \begin{bmatrix} 1_N^T & 0 & 0 \end{bmatrix}.x = [1] \\ \begin{bmatrix} \mu & 0 & 0 \end{bmatrix}.x = [R] \\ \begin{bmatrix} -I_N & 0 & 0 \end{bmatrix}.x \leq [0] \\ \begin{bmatrix} -I_N & 1 & 0 \\ I_N & 0 & -1 \end{bmatrix}.x \leq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0_N & -10 & 1 \end{bmatrix}.x \leq [0] \end{cases}$$

3 Non-Traditional Portfolio Optimizations

3.1 Some Useful optimization problems

Although elegant and easy to implement, previous optimization problems suffer many disadvantages. First, we quickly realized that if we release the constraint between ω_{max} and ω_{min}, the result is absolutely not the same. Markowitz's unconstrained optimisations tend to put almost all allocations on a small subset of assets. This subset is all the more limited when the number of assets increases (see [10], which shows that the combination of constraints (1) and (3) is impossible if you want to conduct an effective optimization, as soon as the number of assets N $>>$ 50).

For instance, see below the plot of the allocations of the portfolio for a mean - variance optimization, long only, convex (assets has been taken randomly in an universe of size 5,000), which can be summed up in the following sentence: Mean Variance optimization for portfolios with many assets kills the diversification of the portfolio, which is always been a strength of a successful portfolio management.

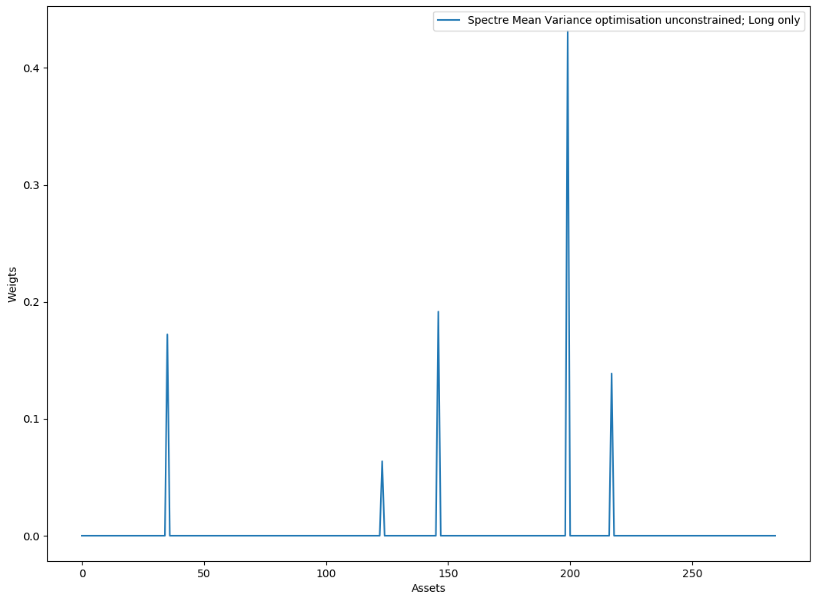

Figure 5: Portfolio of 250 random assets

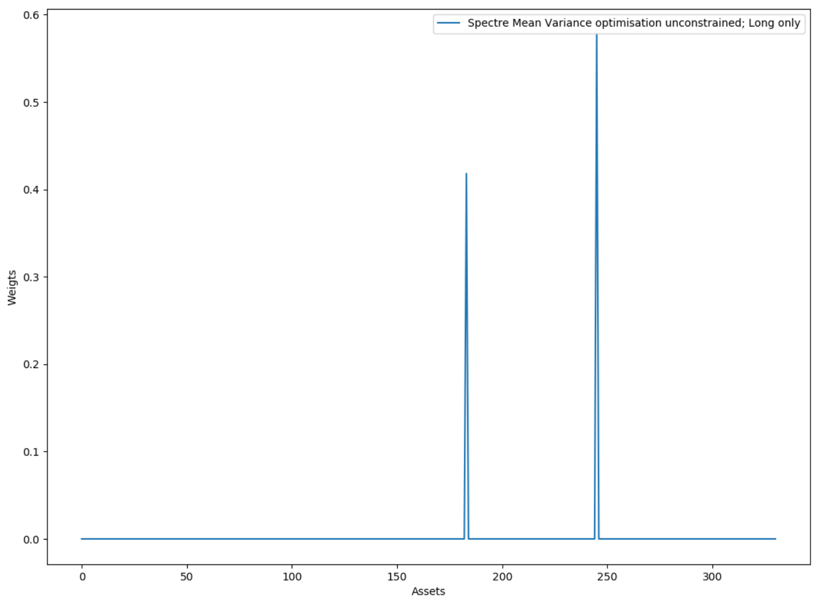

Figure 6: Portfolio of 300 random assets

3.1.1 Equal Risk Contributions

To remove this trend, a first idea is to go closer to the portfolio *Equally Weighted* (EWP), but it does not take into account the risk associated with each allocation. The following solution had been established first by Sebastien Maillard, Thierry Roncalli and Jerome Teiletche [11].

The Equal Risk Contribution Portfolio (ERC) is an alternative to the EWP, which seeks to match not the weights, but the risk associated to each weight. If we assume that the standard deviation of the portfolio is an homogeneous function of the weights ω, then, the Euler's theorem for homogeneous functions states that:

$$\sigma(\omega) = \sum_{i=1}^{n} \omega_i \frac{\partial\sigma(\omega)}{\partial\omega_i} = \sum_{i=1}^{n} \sigma_i(\omega)$$

Ie each asset i, weighted ω_i, whose *returns* are r_i, makes a contribution to the risk:

$$\sigma_i(\omega) = \omega_i \frac{\partial\sigma(\omega)}{\partial\omega_i} = \omega_i \frac{\omega_i\sigma_{ii} + \Sigma_{j\neq i}\omega_j\sigma_{ij}}{\sigma(\omega)}$$

$\sigma_i(\omega)$ is called the i^{th} risk contribution for a portfolio $\Omega : \omega = (\omega_i)_{1\leq i\leq n}$. The goal of this optimization is to find the vector ω such that for all $i \neq j$ $\sigma_i(\omega) = \sigma_j(\omega)$. One can find by calculation a closed-form solution in some particular cases (n = 2 or when the correlation between the assets is constant). On the other hand, for any number of assets n > 2, the solutions can only be expressed via endogenous formulas, ie function of the solution.

We are looking for: $\omega^* = \{\omega_k \in [0;1]^n : \sum_{i=1}^{n} \omega_i = 1,\, \omega_i \frac{\partial\sigma(\omega)}{\omega_i} = \omega_j \frac{\partial\sigma(\omega)}{\omega_j} \quad \forall i,j\,\}$

$$\begin{aligned}
\sigma(\omega) &= \sqrt{\omega^T.\Sigma.\omega} \\
\frac{\partial\sigma}{\partial\omega} &= \frac{1}{2\sqrt{\omega^T.\Sigma.\omega}}.2\Sigma.\omega = \frac{\Sigma.\omega}{\sqrt{\omega^T.\Sigma.\omega}} = \frac{1}{\sigma(\omega)}(\Sigma.\omega) \\
\sigma_i(\omega) &= \omega_i \frac{\partial\sigma}{\partial\omega_i} = \frac{\omega_i}{\sigma(\omega)} \sum_{j=1}^{n} \omega_j\sigma_{ij} = \frac{\sum_{j=1}^{n} \omega_i\omega_j\sigma_{ij}}{\sigma(\omega)}
\end{aligned}$$

Remark 3.1: $(\Sigma\omega)_i \propto \frac{\partial\sigma}{\partial\omega_i}$.

Hence $\omega^* = \{\omega_k \in [0;1]^n : \sum_{i=1}^{n} \omega_i = 1,\, \omega_i(\Sigma\omega)_i = \omega_j(\Sigma\omega)_j \quad \forall i,j\,\}$

CASE n = 2:

$\Omega : \omega = (\lambda, 1-\lambda)$

$$\begin{pmatrix}\sigma_1(\omega)\\ \sigma_2(\omega)\end{pmatrix} \propto \begin{pmatrix}\lambda(\Sigma\omega)_1\\ (1-\lambda)(\Sigma\omega)_2\end{pmatrix}$$

$$\begin{aligned}(\Sigma\omega)_1 &= (\Sigma\omega)_2\\ \sigma_1^2\omega + \sigma_{12}(1-\lambda) &= \sigma_{12}\omega + \sigma_2^2(1-\lambda)\\ \sigma_1^2\lambda^2 &= (1-\lambda)^2\sigma_2^2\\ \sigma_1\lambda &= (1-\lambda)\sigma_2\\ \lambda &= \frac{\sigma_2}{\sigma_1+\sigma_2}\end{aligned}$$

$$\omega^* = (\frac{\sigma_2}{\sigma_1+\sigma_2}, \frac{\sigma_1}{\sigma_1+\sigma_2}) \tag{3.1}$$

CASE CONSTANT CORRELATION:

Assume now that the correlation between every asset in the Portfolio is constant equals to ρ. Then $\forall$ (i, k):

$$\begin{aligned}(\Sigma\omega)_i &= (\Sigma\omega)_k\\ \sum_{j=1}^{n}\omega_i\omega_j\sigma_{ij} &= \sum_{p=1}^{n}\omega_k\omega_p\sigma_{kp}\\ \omega_i^2\sigma_i^2 + \sum_{j\neq i}^{n}\rho\omega_i\omega_j\sigma_i\sigma_j &= \omega_k^2\sigma_k^2 + \sum_{j\neq k}^{n}\rho\omega_k\omega_j\sigma_k\sigma_j\\ \omega_i^2\sigma_i^2 + \rho\omega_i\omega_k\sigma_i\sigma_k &= \omega_k^2\sigma_k^2 + \rho\omega_i\omega_k\sigma_i\sigma_k\\ \omega_i\sigma_i &= \omega_k\sigma_k\\ \omega_i &= \omega_k\frac{\sigma_k}{\sigma_i}\\ 1 &= \omega_k\sigma_k\sum_{i=1}^{n}\frac{1}{\sigma_i}\\ \omega_k &= \frac{\sigma_k^{-1}}{\sum_{i=1}^{n}\sigma_i^{-1}}\end{aligned}$$

$$\omega^* = \left(\frac{\sigma_k^{-1}}{\sum_{i=1}^{n}\sigma_i^{-1}}, k\in[0,n]\right) \tag{3.2}$$

CASE CONSTANT VARIANCE:

Assume now that the variance of each asset in the Portfolio is constant equals to $a > 0$. Then $\forall$ (i, k):

$$
\begin{aligned}
(\Sigma\omega)_i &= (\Sigma\omega)_k \\
\sum_{j=1}^{n} \omega_i \omega_j \rho_{ij} a &= \sum_{p=1}^{n} \omega_p \omega_k \rho_{pk} a \\
w_i &= \frac{\sum\limits_{p=1}^{n} \omega_p \omega_k \rho_{pk}}{\sum\limits_{j=1}^{n} \omega_j \rho_{ij}} \\
1 &= \sum_{i=1}^{n} \frac{\omega_k \sum\limits_{p=1}^{n} \omega_p \rho_{pk}}{\sum\limits_{j=1}^{n} \omega_j \rho_{ij}} \\
\omega_k &= \frac{(\sum\limits_{p=1}^{n} \omega_p \rho_{pk})^{-1}}{\sum\limits_{i=1}^{j} (\sum\limits_{j=1}^{n} \omega_j \rho_{ij})^{-1}}
\end{aligned}
$$

$$
\omega^* = \left(\frac{(\sum_{p=1}^{n} \omega_p \rho_{pk})^{-1}}{\sum_{i=1}^{n} (\sum_{j=1}^{n} \omega_j \rho_{ij})^{-1}}, k \in [0, n] \right) \tag{3.3}
$$

Remark 3.2: We note that this formula for the weights of the portfolio is endogenous, and so cannot be directly used for the optimization.

GENERAL CASE:

In the general case, we will see that the formula we obtain is once again endogenous, and so connot be used. Nevertheless, let us define first:

$\sigma_{i\omega} = Cov(r_i, \sum_{j=1}^{n} \omega_j r_j) = \sum_{j=1}^{n} \omega_j \sigma_{ij}$

$\beta_i = \frac{\sigma_{i\omega}}{\sigma^2(\omega)}$ Beta of stock i against the portfolio.

As all risk contributions are the same, then we have:

$$
\begin{aligned}
\sigma_i(\omega) &= \frac{\sigma(\omega)}{n} \\
&= \frac{\omega_i \sigma_{i\omega}}{\sigma(\omega)} \\
&= \sigma(\omega)\omega_i \beta_i = \frac{\sigma(\omega)}{n} \\
\omega_i &= \frac{\beta_i^{-1}}{n}
\end{aligned}
$$

$$
\omega^* = \left(\frac{\beta_i^{-1}}{n}, i \in [0, n] \right) \tag{3.4}
$$

This two endogenous formulas are useful if and only if we are able to estimate the betas against the portfolio or the weighted sum of the correlations. In any other case, we need to develop a numerical solution. Sebastien Maillard, Thierry Roncalli and Jerome Teiletche showed [11] that the equal risk contribution portfolio exists for all matrix H of data, and hence used the minimization of the following quartic function:

$$f(\omega) = \sum_{i=1}^{n}\sum_{j=1}^{n}(\omega_i(\Sigma\omega)_i - \omega_j(\Sigma\omega)_j)^2$$

This problem is then:

$$\min \sum_{i=1}^{n}\sum_{j=1}^{n}(\omega_i(\Sigma\omega)_i - \omega_j(\Sigma\omega)_j)^2$$

$$s.t: \begin{cases} \sum_{i=1}^{n}\omega_i = 1 \\ \omega_i \geq 0 \\ \omega_{max} \leq 10.\omega_{min} \end{cases}$$

This form is the most adapted to study this quartic system. That will be done with the package scipy.optimize of python, as CVXOPT is not usable for quartic problems.

3.1.2 Diverse metrics to optimize: Tracking Error, Alpha/Beta approach, Max Drawdown

Tracking Error:

Suppose that we have data on a benchmark, a reference signal, as it is the case for example in contexts of overcome an index. Define the R_f signal of *returns* of this benchmark. The tracking error captures the proximity of the portfolio's *returns* signal to the benchmark. It is the mean quadratic distance of the two signals:

$$TE = \sqrt{\sum_{t=0}^{T}(R_{f_t} - R_t)^2}$$

With T the data period considered, R_t the *returns* at time t, so $R_t = \omega^T.r_t$, with $r_t = (r_t^i)_i$ the vector of *returns* of the assets of the portfolio at time t.

$$
\begin{aligned}
TE &= \sqrt{\sum_{t=0}^{T}(R_{f_t} - R_t)^2} \\
TE^2 &= \sum_{t=0}^{T}(R_{f_t} - R_t)^2 \\
&= \sum_{t=0}^{T}(R_{f_t} - \omega^T.r_t).(R_{f_t}^T - r_t^T\omega)
\end{aligned}
$$

Define $A = (r_t)_t = \omega.H$, $B = (R_{f_t})_t$, with H the temporal matrix of *returns*.
Then:

$$
\begin{aligned}
TE^2 &= (A^T - B^T).(A - B) \\
&= (\omega^T.H^T - B^T).(H.\omega - b) \\
&= \omega^T.H^T.H.\omega - \omega^T.H^T.B - B^T.H.\omega \\
&= \omega^T.(H^T.H).\omega - 2\omega^T.(H^T.B)
\end{aligned}
$$

Suppose we want to replicate the performance of the benchmark. So our optimization can be written:

$$\min \omega^T.(H^T.H).\omega - 2\omega^T.(H^T.B) \text{ s.t: } \mathcal{C}$$

which, under mathematical writing, is a convex problem adapted to CVXOPT:

$$\min x^T.(H^T.H).x - 2x^T.(H^T.B)$$

$$
s.t: \begin{cases}
\begin{bmatrix} 1_N^T & 0 & 0 \end{bmatrix}.x = [1] \\
\begin{bmatrix} -I_N & 0 & 0 \end{bmatrix}.x \leq [0] \\
\begin{bmatrix} -I_N & 1 & 0 \\ I_N & 0 & -1 \end{bmatrix}.x \leq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\begin{bmatrix} 0_N & -10 & 1 \end{bmatrix}.x \leq [0]
\end{cases}
$$

Alpha/Beta Approach:

Suppose once again that we have the *returns* signal of the benchmark, and that the latter has a generally positive return. Finally, suppose that this benchmark is sufficiently correlated with the assets of our portfolio, so that we can perform a relevant linear regression between the two signals of *returns*. Then in these particular conditions, if we maximize the slope (β) of the regression

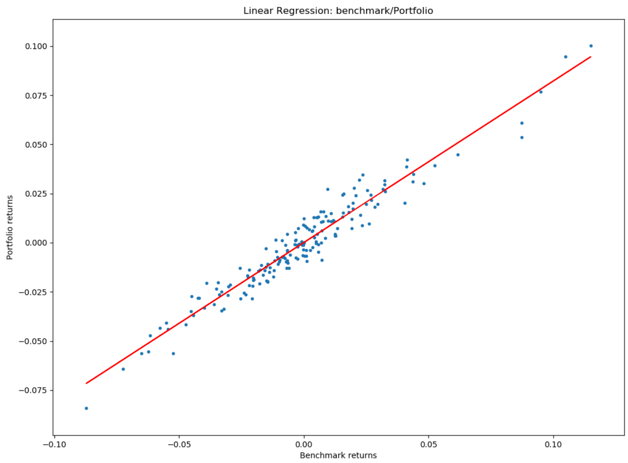

Figure 7: Linear Regression: Benchmark / Portfolio, CAC40 assets

and the intercept (α), we almost surely outperform the benchmark.
The problem becomes:

$$\max \alpha + \beta$$

$$s.t: \begin{cases} \sum_{i=1}^{n} \omega_i = 1 \\ \omega_i \geq 0 \\ \omega_{max} \leq 10.\omega_{min} \end{cases}$$

Max Drawdown:

Finally, min max Drawdown optimization is a more qualitative alternative to min variance. Indeed the max *drawdown* is a good indicator of the downward volatility. We can minimize this max *drawdown*. So we avoid minimizing the volatility upwards.

Definition 1.2: For a signal, the drawdown corresponds to the amplitude of the decline after a major peak.

The max drawdown of a (S_t) signal can be calculated numerically with the following algorithm:

. maw drawdown = 0

. peak=S_0

. For all t:

. If $S_t >$peak:

. peak = S_t

. drawdown = $\frac{peak - S_t}{peak}$

. maw drawdown = max(drawdown, max drawdown)

3.2 Most Diversified Portfolio

3.2.1 Definitions

In 2008 Yves Choueifaty and Yves Coignard introduced a new kind of Portfolio Construction. They tried to find an elegant alternative to the MPT. The starting point of their article Toward Maximum diversification [12] deals with the estimation of Markowitz's Parameters. According to them, while it seems relevant to use the sample covariance matrix to capture the risk, it is too ambitious to assume that mean returns are constant. Therefore, model it with the sample mean is unsuitable. They wanted to free the mean dependence of a portfolio optimization. Knowing that, they quantified an other characteristic of a portfolio, the diversification, with a metric that depends only on a covariance matrix estimator.

A quantitative diversification metric should capture how much a portfolio is built with "very different kind of" assets. Yves Choueifaty proposed in *TOWARD MAXIMUM DIVERSIFICATION* [12] that the diversification of a portfolio is all the more important as the volatility is far from the volatility if all the assets were uncorrelated. Hence, this metrics can be written (with S = $(\sigma_i)_{1 \leq i \leq n}$):

$$DR(\omega) = \frac{\omega^T.S}{\sqrt{\omega^T.\Sigma.\omega}}$$

The strategy followed when doing Most Diversified Portfolio (MDP) optimization is to maximize this ratio.
ie:

$$\max \frac{\omega^T.S}{\sqrt{\omega^T.\Sigma.\omega}}$$

$$s.t: \begin{cases} \sum\limits_{i=1}^{n} \omega_i = 1 \\ \omega_i \geq 0 \, \forall i \\ \omega_{max} \leq 10.\omega_{min} \end{cases}$$

Unfortunately, this problem is not quadratic convex, and as it is, we can not solve it quickly thanks to CVXOPT as we did earlier in the MPT context. In *Toward Maximum Diversification*, Yves Choueifaty and Yves Coignard choose to use the analytical solution. Indeed the diversification ratio is easily differentiable, and the maximization is reduced to a simple matrix equation. However this analytical solution is not adapted to the constraints $\mathcal{C}$, and to the structure of the data. First, the anlytic solution constantly violates the constraint (3). In addition, since there is no large data window, it is impossible to use the majority of the covariance matrix estimators. Indeed, the latter are regularly non-invertible, whereas the analytical solution involves only this inverse.
However, with some calculus, we can make this problem quadratic convex, and therefore numerically solvable in all cases.

Indeed, define: $y = \frac{\omega}{k}$, with $k \geq 0$ such that $\frac{\omega^T}{k}.\sigma = 1$.

Then:

$$\begin{aligned} DR(\omega) &= \frac{\omega^T.S}{\sqrt{\omega^T.\Sigma.\omega}} \\ &= \frac{k}{k}.\frac{\omega^T.S}{\sqrt{\omega^T.\Sigma.\omega}} \\ &= \frac{\frac{\omega^T.S}{k}}{\sqrt{\frac{\omega^T}{k}.\Sigma.\frac{\omega}{k}}} \\ &= \frac{1}{\sqrt{y^T.\Sigma.y}} \end{aligned}$$

and so:
max$\{\frac{\omega^T.S}{\sqrt{\omega^T.\Sigma.\omega}}; \omega \in \mathcal{C}\} =$ min$\{\sqrt{y^T.\Sigma.y};\ k \geq 0,\ y^T.S = 1,\ \frac{y}{k} \in \mathcal{C}\}$
argmax$\{\frac{\omega^T.S}{\sqrt{\omega^T.\Sigma.\omega}}; \omega \in \mathcal{C}\} \propto$ argmin$\{y^T.\Sigma.y;\ k \geq 0,\ y^T.S = 1,\ \frac{y}{k} \in \mathcal{C}\}$,

with $\mathcal{C}$ the set of constraints.

It is therefore sufficient to add a variable k to obtain the following convex problem:

$$\min y^T.\Sigma.y$$

$$s.t: \begin{cases} \omega = \frac{y}{k} \\ k \geq 0 \\ y^T.S = 1 \\ \sum_{i=1}^{n} \omega_i = 1 \\ \omega_i \geq 0 \\ \omega_{max} \leq 10.\omega_{min} \end{cases}$$

which is equivalent to:

$$\min y^T.\Sigma.y$$

$$s.t: \begin{cases} \omega = \frac{y}{k} \\ k \geq 0 \\ \sum_{i=1}^{n} y_i = k \\ y^T.S = 1 \\ s_1 = y_{min} \geq 0 \\ s_2 = y_{max} \leq 10.y_{min} \end{cases}$$

So finally:

(With $x = [y_1 \ldots y_N \, k \, s_1 \, s_2]^T = [k.\omega_1 \ldots k.\omega_N \, k \, s_1 \, s_2]^T$, and $\Sigma' = \begin{bmatrix} \Sigma\,0\,0\,0 \\ 0\,0\,0\,0 \\ 0\,0\,0\,0 \\ 0\,0\,0\,0 \end{bmatrix}$)

$$\min x^T.\Sigma'.x$$

$$s.t: \begin{cases} \begin{bmatrix} S^T & 0 & 0 & 0 \end{bmatrix}.x = [1] \\ \begin{bmatrix} 1_N^T & -1 & 0 & 0 \end{bmatrix}.x = [0] \\ \begin{bmatrix} 0_N & -1 & 0 & 0 \end{bmatrix}.x \leq [0] \\ \begin{bmatrix} 0_N & 0 & -1 & 0 \end{bmatrix}.x \leq [0] \\ \begin{bmatrix} -I_N & 0 & 1 & 0 \\ I_N & 0 & 0 & -1 \end{bmatrix}.x \leq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0_N & 0 & -10 & 1 \end{bmatrix}.x \leq [0] \end{cases}$$

In addition to Choueifaty and Coignard's approach, we will propose an explanation of why even in a Markowitz approach, the Most Diversified Portfolio (MDP) can be seen as the best Mean versus Variance Portfolio.

*

* *

From now and in all the rest of this section, let us consider the most diversified portfolio only in the unconstrained case, ie we reduce the problem to the case where the only constraint is $\sum_{i=1}^{n} \omega_i = 1$

In this general case, let us now derive the solution to the MDP problem
Recall that $DR(\omega) = \frac{\omega^T S}{\sqrt{\omega^T \Sigma \omega}}$.

$$\begin{aligned}
\frac{\partial DR}{\partial \omega} &= \frac{\sqrt{\omega^T.\Sigma.\omega}.S - \omega^T.S\frac{2\Sigma.\omega}{2\sqrt{\omega^T.\Sigma.\omega}}}{\omega^T.\Sigma.\omega} \\
\sqrt{\omega^T.\Sigma.\omega}S &= \frac{\omega^T.S}{\sqrt{\omega^T.\Sigma.\omega}}\Sigma.\omega \\
(\omega^T.\Sigma.\omega).S &= (\omega^T.E).\Sigma.\omega \\
\omega &= \frac{\omega^T.\Sigma.\omega}{\omega^T.S}\Sigma^{-1}.S = \alpha(\omega)\Sigma^{-1}.S
\end{aligned}$$

With:

$$\begin{aligned}
\alpha(\omega) &= \frac{\omega^T.\Sigma.\omega}{\omega^T.S} \\
S^T.\Sigma^{-1}.1_N^T.\alpha(\omega) &= \frac{S^T.\omega}{\omega^T.S} = 1 \\
\alpha(\omega) &= \frac{1}{1_N^T.\Sigma^{-1}.s}
\end{aligned}$$

$$\omega = \frac{\sigma^{-1}.S}{1_N^T.\Sigma^{-1}.S} \tag{3.5}$$

3.2.2 MDP vs MPT

We are now going to prove that the MDP is not only an efficient alternative to the MPT approach. Indeed, if one puts it into the mean versus variance mathematical framework, then, the MDP appears to give a good solution to the following problem:
What is the best E such that my Portfolio satisfies the program (P):

$$\begin{aligned}
&\min \tfrac{1}{2}\sigma^2 = \tfrac{1}{2}\omega^T.\Sigma.\omega \\
&s.t : \left\{ \sum_{i=1}^{n} \omega_i \mu_i = E \right.
\end{aligned}$$

The problem (P) is a characterization of the efficient frontier. Robert Merton [13] first proved the following.

Assume first that we have a matrix of returns H, which is a n×m matrix $H = (r_{ij})_{1\leq i\leq n, 1\leq j\leq m}$, where r_{ij} is the returns of asset i at times j. Assume also that we have functions f_1 and f_2 of only H that build an estimator Σ invertible (the real covariance matrix is always invertible) of the covariance matrix $\Sigma = (\sigma_{ij})_{1\leq i\leq n, 1\leq j\leq n} = f_1(H)$ and an estimator μ of the mean of returns vector $\mu = (\mu_i)_{1\leq i\leq n} = f_2(H)$, with the same notation as in the previous sections. Then, there is no randomness in the following works as everything is deterministic, and depends only on the data set H, and the accuracy of the estimator function f_1 and f_2.

Derivation of the efficient Frontier:
For a data set H, define the Portfolio Mean versus Variance space (Mean, Variance), where each existing point (E, σ^2) is a portfolio Ω with mean E, and variance σ^2, ie $E = \sum_{i=1}^{n} \omega_i\mu_i$, and $\sigma^2 = \omega^T.\Sigma.\omega = \sum_{j=1}^{n}\sum_{i=1}^{n} \omega_i\omega_j\sigma_{ij}$. For Mean versus Variance Portfolio analysis, people often use the (σ^2, E) space, but using (E, σ^2) is better in this case.

Al the points (E, σ^2) of the efficient frontier satisfy the following problem:

$$min_{\sigma,\omega}\ \frac{1}{2}\sigma^2$$

$$s.t: \begin{cases} \sigma^2 = \frac{1}{2}\omega^T.\Sigma.\omega \\ \sum_{i=1}^{n} \omega_i\mu_i = E \\ \sum_{i=1}^{n} \omega_i = 1 \end{cases}$$

Which is equivalent to solving the following (P) for each E:

$$min_{\omega}\ \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n} \omega_i\omega_j\sigma_{ij}$$

$$s.t: \begin{cases} E - \sum_{i=1}^{n} \omega_i\mu_i = 0 \\ 1 - \sum_{i=1}^{n} \omega_i = 0 \end{cases}$$

and then plug ω into $\sum_{j=1}^{n}\sum_{i=1}^{n} \omega_i\omega_j\sigma_{ij}$ in order to obtain σ^2.

As we saw in chapter 1, if ω satisfies (P), then, the K.K.T condition give:

- $\nabla(\mathcal{L})(\omega) = 0$;
- ω is feasible;

- All the Lagrange multipliers for inegality constraints γ_i are positive;
- $\gamma_i.g_i(\omega) = 0$, where g_i are the inequality constraints;

Hence, for (P), ω is a solution if:

- Condition 1: $\forall i, \nabla(\mathcal{L})(\omega)_i = \sum_{j=1}^{n}(\omega_j\sigma_{ij}) - \gamma_1\omega_i\mu_i - \gamma_2 = 0$; (a)
- Condition 2: $E = \sum_{i=1}^{n}\omega_i\mu_i$ and $1 = \sum_{i=1}^{n}\omega_i$; (b1), (b2)
- Condition 3: None;
- Condition 4: None;

With $\mathcal{L}(\omega) = \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}\omega_i\omega_j\sigma_{ij} + \gamma_1(E - \sum_{i=1}^{n}\omega_i\mu_i) + \gamma_2(1 - \sum_{i=1}^{n}\omega_i)$.
(a) $\iff \Sigma.\omega - \gamma_1\mu - \gamma_2 1_N = 0$
(a) $\iff \omega = \gamma_1\Sigma^{-1}.\mu + \gamma_2\Sigma^{-1}.1_N$

Define $(v_{ij})_{ij} := \Sigma^{-1}$

$$\forall k, \qquad \omega_k = \gamma_1\sum_{j=1}^{n}v_{kj}\mu_j + \gamma_1\sum_{j=1}^{n}v_{kj} \tag{3.6}$$

And then by multiplying by μ_k and summing:
(b) $\iff$ E = $\sum_{k=1}^{n}\omega_k\mu_k = \gamma_1\sum_{k=1}^{n}\sum_{j=1}^{n}v_{kj}\mu_j\mu_k + \gamma_2\sum_{k=1}^{n}\sum_{j=1}^{n}v_{kj}\mu_k$
(b) $\iff$ 1 = $\sum_{k=1}^{n}\omega_k = \gamma_1\sum_{k=1}^{n}\sum_{j=1}^{n}v_{kj}\mu_j + \gamma_2\sum_{k=1}^{n}\sum_{j=1}^{n}v_{kj}$

Define:
A := $\sum_{k=1}^{n}\sum_{j=1}^{n}v_{kj}\mu_j = 1_N^T.\Sigma^T.\mu$
B := $\sum_{k=1}^{n}\sum_{j=1}^{n}v_{kj}\mu_j\mu_k = \mu^T.\Sigma^T.\mu$
C := $\sum_{k=1}^{n}\sum_{j=1}^{n}v_{kj} = 1_N^T.\Sigma^T.1_N$

Then:

$$\begin{cases} E = \gamma_1 B + \gamma_2 A \\ 1 = \gamma_1 A + \gamma_2 C \end{cases} \tag{3.7}$$

$$\begin{aligned}\gamma_1 &= \frac{1}{B}(E - \gamma_2 A)\\ 1 &= \frac{A}{B}(E - \gamma_2 A) + \gamma_2 C\\ \gamma_2(C - \frac{A^2}{B}) &= 1 - \frac{A}{B}E\\ \gamma_2 &= \frac{1 - \frac{A}{B}E}{C - \frac{A^2}{B}}\\ &= \frac{B - AE}{BC - A^2}\end{aligned}$$

Define D := BC - A^2

$$\gamma_2 = \frac{B - AE}{D} \tag{3.8}$$

$$\begin{aligned}\gamma_1 &= \frac{1}{B}(E - \frac{AB - A^2E}{D})\\ &= \frac{ED - AB + A^2E}{BD}\\ &= \frac{E(D + A^2 - AB}{BD}\\ &= \frac{EBC - AB}{BD}\end{aligned}$$

$$\gamma_1 = \frac{EC - A}{D} \tag{3.9}$$

Hence, the Lagrange multipliers, and the weights ω are given by:

$$\begin{aligned}\gamma_1 &= \tfrac{EC-A}{D}\\ \gamma_2 &= \tfrac{B-AE}{D}\\ \omega &= \tfrac{EC-A}{D}\Sigma^{-1}.\mu + \tfrac{B-AE}{D}.\Sigma^{-1}.1_N\end{aligned}$$

Using (2.1):

$$\begin{aligned}\omega_k &= \frac{\sum_{j=1}^{n}(EC - A)v_{kj}\mu_j + \sum_{j=1}^{n} v_{kj}(B - AE)}{D}\\ &= \frac{1}{D}(\sum_{j=1}^{n}(EC\mu_j - A\mu_j + B - AE))\end{aligned}$$

$$\omega_k = \frac{1}{D}(E\sum_{j=1}^{n} v_{kj}(C\mu_j - A) + \sum_{j=1}^{n} v_{kj}(B - A\mu_j)) \tag{3.10}$$

Then, using (a):

$$\begin{aligned}
\sum_{j=1}^{n} \omega_j \sigma_{ij} &= \gamma_1 \mu_1 + \gamma_2 \\
\sum_{i=1}^{n} \sum_{j=1}^{n} \omega_i \omega_j \sigma_{ij} &= \gamma_1 \sum_{i=1}^{n} \mu_i \omega_i + \gamma_2 \sum_{i=1}^{n} \omega_i \\
\sigma^2 &= \gamma_1 E + \gamma_2 \\
&= \frac{E^2 C - AE + B - AE}{D} \\
\sigma^2 &= \frac{CE^2 - 2AE + B}{D}
\end{aligned}$$

$$\sigma^2 = f(E), \text{ such that } f \text{ is a parabola} \tag{3.11}$$

It is a convex parabola, hence it has a minimum. Indeed:
$\frac{\partial^2 f}{\partial^2 E} = 2.\frac{C}{D} \geq 0$, hence f is convex.
$\frac{\partial f}{\partial E} = \frac{2CE - 2A}{D} = 2\frac{CE-A}{D} = 0 \iff \boxed{E = \frac{A}{C}}$

By the way, we can note that:
$f(\frac{A}{C}) = \frac{1}{C} = \frac{1}{\sum\limits_{k=1}^{n} \sum\limits_{j=1}^{n} v_{kj}} = \frac{1}{1_N^T.\Sigma^T.1_N}$

And with (2.5), it corresponds to:

$$\omega_k = \frac{\sum\limits_{j=1}^{n} v_{kj}}{C} = \frac{\Sigma^{-1}.1_N}{1_N^T.\Sigma^T.1_N} \tag{3.12}$$

Then, the final Merton's result is the following. The efficient frontier is characterized by the equations:

$$E = \frac{A}{C} \pm \frac{1}{C}\sqrt{DC(\sigma^2 - \frac{1}{C})} \tag{3.13}$$

And finally, the useful equation of the Mean versus Variance efficient frontier is :

$$\boxed{\sigma^2_{MPT}(E) = \frac{1_N^T.\Sigma^{-1}.1_N.E^2 - 2.1_N^T.\Sigma^{-1}.\mu.E + \mu^T.\Sigma^{-1}.\mu}{\mu^T.\Sigma^{-1}.\mu.1_N^T.\Sigma^{-1}.1_N^T - (1_N^T.\Sigma^{-1}.\mu)^2}} \tag{3.14}$$

*

* *

Choueifaty and Coignard did not try to see the MDP as a particular case of the MPT. However, the following is a tentative to derive in a same way as Merton did, the different existing curves in the MDP problem.

Unfortunately, the solutions use many different constants, and so it is not really pleasant to read.

Once again, let us first assume that Σ and μ are known matrix and vector, functions only of the data set H. Define now $S = (\sqrt{\sigma_{ii}} = \sigma_i)_{1 \leq i \leq n}$, the vector of standard deviation of H. Hence σ is also a function of only the data set H. There is still no randomness in this problem.

Equation of the MDP frontiers:

We will define in this section, two frontiers for the MDP. The first one, analogously to the MPT efficient frontier, is characterize by the equation $\sigma^2_{MDP} = f(E_{MDP})$, where E_{MDP} and σ^2_{MDP} are the mean and the variance of any portfolio which maximize the diversification ratio (DR) for a given level of mean $E = E_{MDP}$. The second frontier is characterized by all the portfolios which maximize the diversification ratio for a given level of mean $E = E_{MDP}$, in the $(E, DR(E))$ space. One can generate 5,000 random portfolios and plot the second frontier as we did for the efficient frontier. See the graph below.

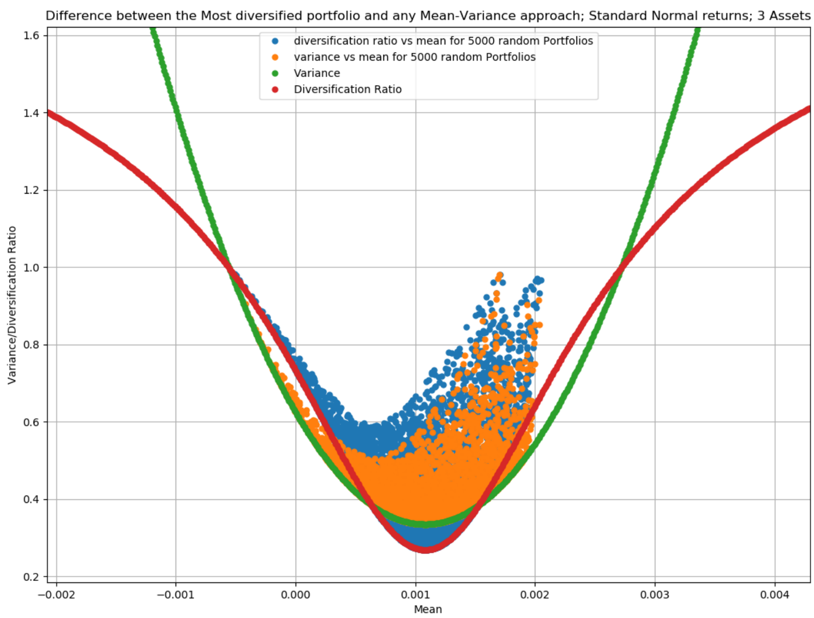

Figure 8: MDP DR frontier and MPT efficient frontier. Standard Normal returns. 5,000 Portfolios

The points of this second frontier are those who satisfy for a given E:

$$\omega^* = argmax_\omega \ \frac{\omega.S}{\sqrt{\omega^T.\Sigma.\omega}}$$

$$s.t: \begin{cases} E - \sum_{i=1}^{n} \omega_i \mu_i = 0 \\ 1 - \sum_{i=1}^{n} \omega_i = 0 \end{cases}$$

with $DR(E) = \frac{\omega^*.S}{\sqrt{\omega^{*T}.\Sigma.\omega^*}}$

As we have seen previously, this is equivalent to (P):

$$y^* = argmin_{y,k} \ \frac{1}{2} y^T.\Sigma.y = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \sigma_{ij}$$

$$s.t: \begin{cases} Ek - \sum_{i=1}^{n} y_i \mu_i = 0 \\ k - \sum_{i=1}^{n} y_i = 0 \\ -k \leq 0 \end{cases}$$

with:

$\omega_i^* = \frac{y_i^*}{k}$

$DR(E) = \frac{\omega^*.S}{\sqrt{\omega^{*T}.\Sigma.\omega^*}}$

If (y, k) is a solution of (P), then it satisfies the K.K.T conditions:

- $\mathcal{L}(y,k) = \frac{1}{2} \sum_{j=1}^{n} \sum_{i=1}^{n} y_i y_j \sigma_{ij} + \gamma_1 (k - \sum_{i=1}^{n} y_i) + \gamma_2 (Ek - \sum_{i=1}^{n} y_i \mu_i) + \gamma_3 (1 - \sum_{i=1}^{n} y_i \sigma_i) - \gamma_4 k$;
- Condition 1:
  - $\forall i, \sum_{j=1}^{n} (y_j \sigma_{ij}) - \gamma_1 - \gamma_2 \mu_i - \gamma_3 \sigma_i = 0$; (a_1)
  - $\gamma_1 + E\gamma_2 - \gamma_4 = 0$; (a_2)
- Condition 2:
  - $k = \sum_{i=1}^{n} y_i$ (b_1)
  - $E.k = \sum_{i=1}^{n} y_i \mu_i$ (b_2)
  - $1 = \sum_{i=1}^{n} y_i \sigma_i$ (b_3)
  - $-k \leq 0$ (b_4)
- Condition 3: $\gamma_1 \geq 0$, $\gamma_2 \geq 0$, $\gamma_3 \geq 0$ and $\gamma_4 \geq 0$; (d)
- Condition 4: $-k.\gamma_4 = 0$; (e)

(e) gives us immediately that either k = 0 or $\gamma_4 = 0$. But as $k > 0$ as $k = \sum_{i=1}^{N} y_i$, hence, $\gamma_4 = 0$. So with (a_2):

$$\gamma_1 = -\frac{1}{E}\gamma_2 \tag{3.15}$$

(a_1) gives us:

$$\begin{aligned} \Sigma.y &= \gamma_1 1_N + \gamma_2\mu + \gamma_3 S \\ y &= \gamma_2\Sigma^{-1}\mu + \gamma_1\Sigma^{-1}1_N + \gamma_3\Sigma^{-1}S \end{aligned}$$

$$\forall k, \qquad y_k = \gamma_2 \sum_{j=1}^{n} v_{kj}\mu_j + \gamma_1 \sum_{j=1}^{n} v_{kj} + \gamma_3 \sum_{j=1}^{n} v_{kj}\sigma_j \tag{3.16}$$

By summing over k, then multiplying by μ_k and summing over k, and finally multiplying by σ_k and summing over k:

$$\begin{aligned} k &= \sum_{k=1}^{n} y_k \\ &= \gamma_2 \sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj}\mu_j + \gamma_1 \sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj} + \gamma_3 \sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj}\sigma_j \\ E.k &= \sum_{k=1}^{n} y_k\mu_k \\ &= \gamma_2 \sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj}\mu_j\mu_k + \gamma_1 \sum_{k=1}^{n}\sum_{j=1}^{n} \mu_k v_{kj} + \gamma_3 \sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj}\sigma_j\mu_k \\ 1 &= \sum_{k=1}^{n} y_k\sigma_k \\ &= \gamma_2 \sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj}\mu_j\sigma_k + \gamma_1 \sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj}\sigma_k + \gamma_3 \sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj}\sigma_j\sigma_k \end{aligned}$$

Define and recall:
A := $\sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj}\mu_j = 1_N^T.\Sigma^{-1}.\mu$
B := $\sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj}\mu_j\mu_k = \mu^T.\Sigma^{-1}.\mu$
C := $\sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj} = 1_N^T.\Sigma^{-1}.1_N$
D := $BC - A^2$
F := $\sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj}\sigma_j = 1_N^T.\Sigma^{-1}.S$

G := $\sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj}\sigma_j\mu_k = \mu^T.\Sigma^{-1}.S$
H := $\sum_{k=1}^{n}\sum_{j=1}^{n} v_{kj}\sigma_j\sigma_k = \sigma^T.\Sigma^{-1}.S$

$$k = \gamma_2 A + \gamma_1 C + \gamma_3 F = \gamma_2(A - \frac{C}{E}) + \gamma_3 F$$
$$E.k = \gamma_2 B + \gamma_1 A + \gamma_3 G = \gamma_2(B - \frac{A}{E}) + \gamma_3 G$$
$$1 = \gamma_2 G + \gamma_1 F + \gamma_3 H = \gamma_2(G - \frac{F}{E}) + \gamma_3 H$$

These three equations can be written as the following matrix single equation:

$$\begin{bmatrix} 1 & \frac{C}{E} - A & -F \\ E & \frac{A}{E} - B & -G \\ 0 & \frac{F}{E} - G & -H \end{bmatrix} . \begin{bmatrix} k \\ \gamma_2 \\ \gamma_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

$$Q. \begin{bmatrix} k \\ \gamma_2 \\ \gamma_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \tag{3.17}$$

$$\begin{bmatrix} k \\ \gamma_2 \\ \gamma_3 \end{bmatrix} = Q^{-1}. \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} . \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} k \\ \gamma_2 \\ \gamma_3 \end{bmatrix} = \begin{bmatrix} -c \\ -f \\ -i \end{bmatrix}$$

Using the Cramer's Formula:

$c = \frac{1}{|Q|}(AG - \frac{GC}{E} + \frac{AF}{E} - FB)$
$f = \frac{1}{|Q|}(G - EF)$
$i = \frac{1}{|Q|}(\frac{A}{E} - B - C + AE)$

$k = \frac{1}{|Q|}(\frac{GC - AF + FBE - AGE}{E})$
$\gamma_2 = \frac{1}{|Q|}(EF - G)$
$\gamma_3 = \frac{1}{|Q|}(\frac{BE + CE - AE - A}{E})$

$$
\begin{aligned}
|Q| &= BH - \frac{HA}{E} - FFE(\frac{F}{E} - G) + HE(\frac{C}{E} - A) + G(\frac{F}{E} - G) \\
&= BH - \frac{HA}{E} - F^2 + FGE + HC - AHE + \frac{GF}{E} - G^2 \\
&= \frac{BHE - HA - F^2E + FGE^2 + HCE - AHE^2 + GF - G^2E}{E} \\
&= \frac{1}{E}[(FG - AH)E^2 + (BH + CH - F^2 - G^2)E + GF - HA] \\
|Q| &= \frac{1}{E}P_1(E)
\end{aligned}
$$

Finally, we can compute the three Lagrange multipliers, and the value of k:

$$k = \frac{(FB - AG)E + GC - AF}{P_1(E)} \tag{3.18}$$

$$\gamma_1 = \frac{G - FE}{P_1(E)} \tag{3.19}$$

$$\gamma_2 = \frac{FE^2 - GE}{P_1(E)} \tag{3.20}$$

$$\gamma_3 = \frac{(B + C - A)E - A}{P_1(E)} \tag{3.21}$$

Recall that $DR(E) = \frac{1}{y^T \Sigma y}$. With (a_1):

$$
\begin{aligned}
\Sigma.y &= \gamma_1.1_N + \gamma_2.\mu + \gamma_3.S \\
&= \gamma_2(\mu - \frac{1_N}{E} + \gamma_3 S) \\
y^T.\Sigma.y &= \gamma_2(y^T\mu - \frac{y^T 1_N}{E}) + \gamma_3 y^T S \\
&= \gamma_2(E.k - \frac{k}{E}) + \gamma_3 \\
&= \gamma_2 k \frac{E^2 - 1}{E} + \gamma_3 \\
&= \frac{1}{P_1(E)^2}[(FB - AG)E + (GC - AF)][EF - G][E^2 - 1] + \frac{1}{P_1(E)}[(B + C - A)E - A]
\end{aligned}
$$

Define:

$\alpha := (FB - AG)F(B + C - A)$

$\alpha_1 := \frac{GC - AF}{FB - AG}$

$\alpha_2 := \frac{G}{F}$

$\alpha_3 := \frac{A}{B+C-A}$
And we obtain:

$$y^T \Sigma y = \alpha[\frac{1}{P_1(E)^2}((E+\alpha_1)(E-\alpha_2)(E^2-1)) + \frac{1}{P_1(E)}(E-\alpha_3)]$$
$$= \frac{\alpha}{P_1(E)^2}[(E+\alpha_1)(E-\alpha_2)(E^2-1) + P_1(E)(E-\alpha_3)]$$

Finally, the equation of the MDP diversification ratio frontier is:

$$DR(E) = \frac{P_1(E)}{\sqrt{\alpha}}.\frac{1}{\sqrt{(E+\alpha_1)(E-\alpha_2)(E^2-1) + P_1(E)(E-\alpha_3)}} \quad (3.22)$$

Then, for a given mean target E, maximizing the diversification ratio leads to the following MDP variance frontier equation:

$$\sigma^2_{DR} = \frac{y^T \Sigma y}{k^2} = \gamma_2 \frac{1}{k}(E - \frac{1}{E}) + \frac{\gamma_3}{E}$$

$$\sigma^2_{DR} = \frac{\alpha}{[(FB-AG)E + (GC-AF)]^2}[(E+\alpha_1)(E-\alpha_2)(E^2-1) + P_1(E)(E-\alpha_3)] \quad (3.23)$$

From now on, the following results depend on three hypothesis and one lemma, that are very consistent with the simulated results.

Lemma 3.1: P_1 has no root.
Hypothesis 3.1: Define g_1 such that $\forall E$, $g_1(E) = \sigma^2_{DR}(E)$, and $g_2(E) = \sigma^2_{MPT}(E)$ The minimum of g_1 is "close" to the minimum of g_2. One can even conjecture that the two minimum are the same.
Hypothesis 3.2: In this neighborhood, in most cases $k \leq 1$, so $\frac{1}{k} \geq 1$.In the other cases, k is very close to 1. Hence, given the formula of k, we can assume that $\frac{1}{k}$ is almost linear in this neighborhood. $\frac{1}{k} \simeq \beta.E$
Hypothesis 3.3: The variance of the most diversified portfolio is close to the minimum of g_1, but not the same. We can assume then that the variance of the most diversified portfolio satisfies the equation (2.9).

Proof of lemma .11: P_1 has no root is equivalent to $|Q|$ is never equals to zero, which is equivalent to Q is always invertible, ie for every E, there exists Lagrange multipliers and k such that the diversification ratio has a maximum. Finally, "P_1 has no root" is equivalent to "for a given mean

E, the Diversification Ratio has a maximum". And we have shown with (2.17), that the maximization of DR with a target E has a solution for every E if Σ is invertible, which is one of our first assumptions. Hence, P_1 has no root.

Unfortunately, the three hypothesis are tough to prove, mainly because the derivation of the minimum of g_1 involves the derivation of a ratio of high degree polynomial. So, the exact solution must be impossible to derive without a numerical approach. However, we can find results that are very consistent with these hypothesis. For example the following graphs of k and $\frac{1}{k}$ show that k seems to have no pole and only one zero, and given (2.13), this means that P_1 has no root.

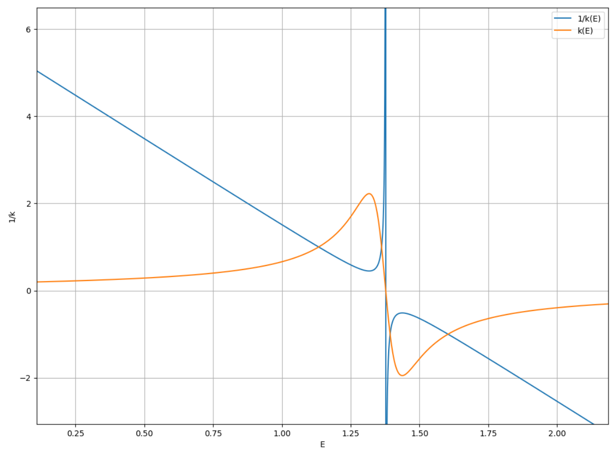

Figure 9: k and $\frac{1}{k}$; 10 random assets with Normal returns

The following graph highlights both hypothesis 1 and hypothesis 2. Indeed we can see that The Variance of the MPT frontier is close to the MDP variance frontier, and especially when the Diversification Ratio is maximized. Moreover, the plot of $\frac{1}{k}$ in function of E shows that k<1 when we are in the neighborhood of the maximum diversification portfolio. This graph has been obtained with 10 random asset from the CAC 40.

Finally, this third graphs shows that it is relevant to assume hypothesis 3, and especially that the maximum diversification portfolio is not the same as the minimum variance portfolio.

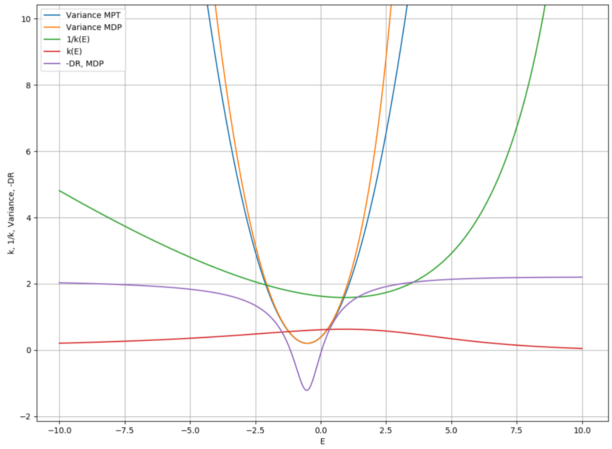

Figure 10: Variances, k, $\frac{1}{k}$ and -DR; 10 random assets with Normal returns

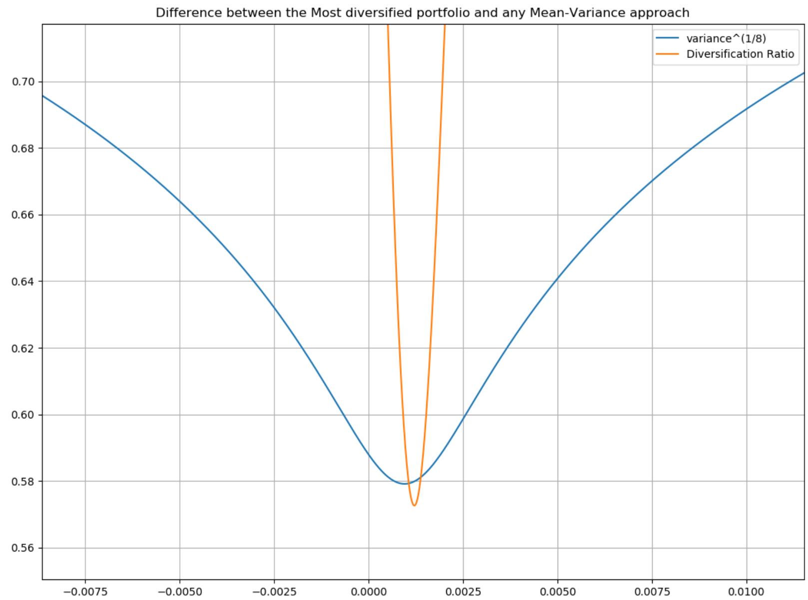

Figure 11: Variance$^{1/8}_{MPT}$ and Diversification ratio ; 10 random assets with Normal returns

Conclusions:

There are many conclusions that follow the previous equation, this lemma and these three hypothesis. Firstly, the lemma implies that the diversification ratio has no root. Moreover, it is seemingly always positive.

Given (2.18), we see that σ_{DR} has one pole of degree two. Moreover, when E tense to infinity, $\sigma_{DR} \sim \frac{E^4}{E^2} = E^2$ and $\sigma_{MPT} \sim E^2$. Hence in the two problems, the variances have the same features when the mean tense to infinity. This point can be highlighted by a plot such as the following, where we see that the ratio is constant when E tense to infinity:

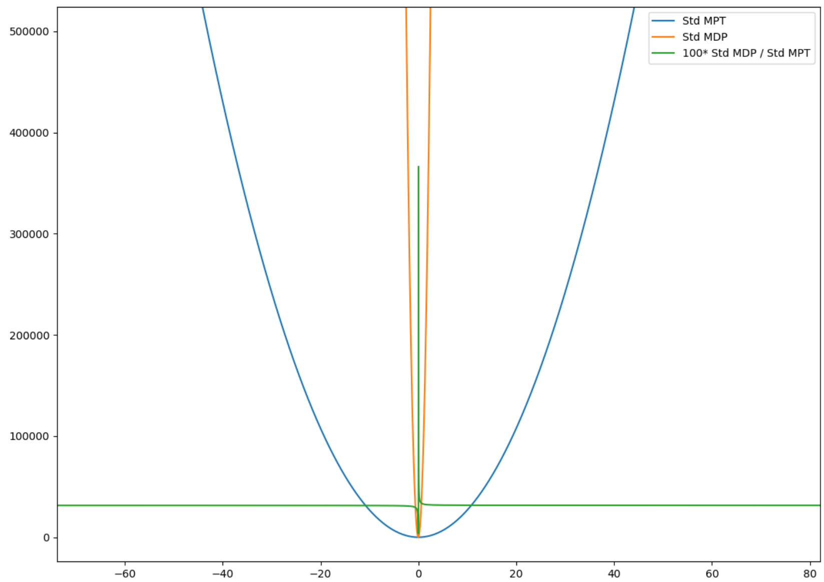

Figure 12: Ratio of variances ; 10 random assets with Normal returns

Recall that the diversification ratio is maximized when $\omega = \frac{\Sigma^{-1}.S}{1_N^T.\Sigma^{-1}.S}$. This correspond then to a mean $E = \frac{G}{F}$. In order to compare the mean in the MPT framework and the mean in the MDP framework, one can compare the mean when minimizing the variance, and the mean when maximizing the diversification ratio, ie compare $\frac{G}{F}$ and $\frac{A}{C}$. However, hypothesis 3 says that $\frac{1}{k} \simeq \beta.E$, and hypothesis 1, allows us to use the dynamic of σ_{MPT} for σ_{DR} in the neighborhood of the solution. So in the neighborhood of it maximum, which is also a neighborhood of the MPT variance, the diversification ratio can be written(with $E_{MPT} = \frac{A}{C}$):

$$
\begin{aligned}
DR(E) &= \frac{\beta E}{\sqrt{\sigma_{DR}^2}} \\
&= \frac{\beta E}{\sqrt{\frac{C}{D}(E - E_{MPT})^2 + \frac{1}{C})}} \\
DR'(E) &= \frac{\beta\sqrt{\frac{C}{D}(E - E_{MPT})^2 + \frac{1}{C}} - \beta E \frac{2\frac{C}{D}(E-E_{MPT})}{2\sqrt{\frac{C}{D}(E-E_{MPT})^2+\frac{1}{C}}}}{\frac{C}{D}(E - E_{MPT})^2 + \frac{1}{C}} = 0 \\
\beta(\frac{C}{D}(E - E_{MPT})^2 + \frac{1}{C}) &= \beta E \frac{C}{D}(E - E_{MPT}) \\
\frac{C}{D}(E - E_{MPT})(E - E_{MPT} - E) &= -\frac{1}{C}
\end{aligned}
$$

Given the three hypothesis, the mean of the most diversified portfolio is:

$$E = E_{MPT} + \frac{D}{C^2 E_{MPT}} > E_{MPT} \tag{3.24}$$

The most diversified portfolio has a higher expected mean of returns.

And finally, given the formulas of σ_{MPT} and DR, the variance around the most diversified portfolio decreases as $E^2 + cst$ with E<1, while the diversification decreases as $\frac{1}{\sqrt{E+cst}}$ with E<1. The the diversification ratio declines much faster.

*

* *

As a conclusion on the most diversified portfolio, we can say that:

- When one maximizes the diversification ratio:
  - The portfolio is well diversified, much more than almost everywhere else ;
  - The variance is low, close to the minimum possible variance ;
  - The expected returns is bigger than the one obtain when minimizing the variance ;
- While when one minimize the variance:
  - The diversification is low ;
  - The variance is just a bit lower;

We could minimize the variance with a chosen target mean, but it could lead to a very low diversification. Maximizing the diversification make it possible to get away from dependence on the estimation of the mean. It acts in fact like a tiny pointers on the best portfolio with respect to the variance, the mean and the diversification.

In the applications that we will see in the chapter 4, we will always try to use first the Most diversified portfolio. We will see that it (almost always) leads to a good result.

4 Applications

The previous optimizers can be helpful in many purposes. They find their utility in the third step of the portfolio construction defined in chapter one. In order to assess the quality of the solutions, we will put it into practice in two different contexts. The first application was the only one that was engineered for real goals at BOUSSARD & GAVAUDAN. We will introduce in this goal some additional constraints in order to build a very robust portfolio.

In all the following applications of the previous optimizers, each covariance matrix estimator and mean estimators are obtained through the sample estimator if we have more than 100 days of data, with the OAS method otherwise.

4.1 Trading Strategy

Assume we get two filters. The first one select between 80 and 200 assets based on trend following strategies. The rebalancing (ie the date of activation of a filter on the universe) is done every three months. Assume we get filters that select assets based on trend following strategies. This gives us between 300 and 600 stocks in which we have to invest a wealth of 1. The filters are applied each 15 days, hence we need to recompute the optimization each 15 days.

As we saw in chapter 3, the Markowitz's optimization does not give satisfying results when the number of assets increases. We then choose to use the Most Diversified Portfolio, and we will see that this optimizer gives better results.

Recall that the most diversified portfolio consists in maximizing the following ratio:

$$\frac{\omega.S}{\sqrt{\omega^T.\Sigma.\omega}}$$

Recall also the given Boussard and Gavaudan's constraints $\mathcal{C}$:

- All weights are positive;

- The sum of the weigths is equal to one;
- the ratio between the lowest weight and the highest is not higher than 10;

As we saw previously, this optimization is equivalent to solving the following convex (quadratic) problem with linear constraints:

$$y^* = argmin_{y,k} \, \tfrac{1}{2} y^T.\Sigma.y = \tfrac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \sigma_{ij}$$

$$s.t: \begin{cases} k - \sum_{i=1}^{n} y_i \\ Ek - \sum_{i=1}^{n} y_i \mu_i = 0 \\ k - \sum_{i=1}^{n} y_i = 0 \\ -k \leq 0 \end{cases}$$

with:
$\omega_i^* = \frac{y_i^*}{k}$
$DR(E) = \frac{\omega^*.S}{\sqrt{\omega^{*T}.\Sigma.\omega^*}}$

This optimization gives us a vector of weights Ω, which allow us to build a real portfolio. Each 15 days, date $(t_k)_k$, there is a rebalancing, ie the filters are reapplied in order to update the lists of assets which satisfies the willing behaviour. If we build such a portfolio, then the $P\&L$ $(P_t)_t$, which is a stochastic process, is given by:

$$P_t = P_{t_k}.\sum_{i=1}^{n} w_i.\frac{S_i(t)}{S_i(t_k)}$$

Indeed if we want to invest $\omega_i\%$ of 1 in asset A_i at time t_k then we need to buy $\frac{w_i}{S_i^{t_k}}$.
This P&L, called "new P&L" in the next graphs, is however not acceptable. Such a position is much too sensitive to the market. Indeed, by building this portfolio, we hope that the selected assets will have a better behaviour than the market, but imagine there is a general fall of all stocks prices, then we risk to loose money. We need to hedge our position. For this purpose, we want to be long our optimized portfolio, and short the market. With such a strategy, in the general case, as our portfolio is supposed to be higher than the market, we win money, and if there is a general fall, as we are short the market, we limit the loss. The only case when we loose money is when the $P\&L$ falls quicker than the market. This case shall not occur if the filters select many diversified assets, ie if there is no bias in a particular factor. This point will be discussed later.

How to be long or short the market ? As we will see in the next section, there exist funds that replicate the behaviour of the well known stock market index, as CAC 40 or FTSE. Our hedging strategy is simply hedging pounds against pounds with future on the associated index. This means that each time we buy 1 of a stock, we short sell 1 of a future on the stock market index where this asset is quoted. For example, imagine we buy 1 of a stock of a french firm, then we will short sell 1 of future on index that replicates the CAC 40 (CF1). The portfolio can be summarized in the following table:

(A, ω)	A$_i$	Hedge (DAX, CAC40, FTSE...)
	ω_i	-ω_i

And the strategy is characterized by the following $P\&L$ stochastic process ($H_i(t)$ is hedge associated to $S_i(t)$)

$$P_t = P_{t_k} + P_{t_k}.\sum_{i=1}^{n} w_i.(\frac{S_i(t)}{S_i(t_k)} - \frac{H_i(t)}{H_i(t_k)})$$

Using this optimization, we obtain the following backtest on 10 years of data, using 300 days look back windows for computing the estimator of the covariance matrix (and therefore the vector of standard deviation). The old $P\&L$ is obtained with an Equally Weighted Portfolio (ie same weight for each asset), the new with the Most Diversified Portfolio. The exact details of the stocks can not be given to the reader, for industrial reasons.

However, the net asset value of the portfolio is not enough to asses the quality of the optimiza-

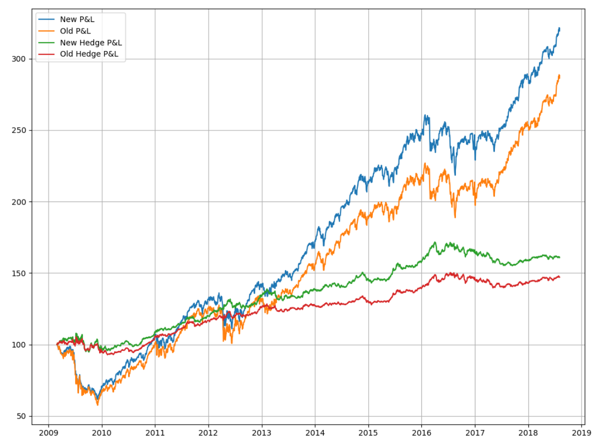

Figure 13: $P\&L$s before (orange) and after (blue) optimization. Portfolio with 300 to 600 assets, rebalancing each 15 days

tion. Therefore we compute few risk metrics for the two $P\&Ls$ and compare them. We decided to use the following metrics:

- The max drawdown;
- The volatility;
- The Sharpe Ratio: the sharpe ratio had been introduced by William Sharpe Sharpe, [14]. Mathematically, the sharpe ratio is $S = \frac{E[r_p - r_b}{\sigma_b}]$, where r_p, r_b and σ_b are respectively the annual rate of returns of our portfolio, the annual rate of returns of the market (ie the hedge), and the volatility of the market. Qualitatively, it examines if an investment have a sufficiently higher returns than the market compare to the additional risk we take. For a good investment, the sharpe ratio should be close to 1 (the higher the sharpe ratio is, the best the investment is). If we assume that each realized rate of returns is an observation of a normally ditributed law $\mathcal{N}(\mu.t, \sigma.\sqrt{\Delta t})$, for computing the sharpe ratio, we take the mean and the standard deviation of rate of returns on the realized $P\&L$ of the hedge and of the portfolio, divided by square root of the number of years;
- The drawdown: The value "drawdown" at each step, in the algorithm we show in part 3.1.2;
- The time to recovery: Duration between two zeros of the previous value;
- Ratio new hedged $P\&L$ versus new hedged $P\&L$

For the previous Portfolio, the risk metrics, the drawdown and the ratio are the following:

Risk metrics	Equally weighted portfolio	MDP
Max Drawdown	0,1195642631	0,122335565
Starting date max drawdown	2009-07-08 00:00:00	2009-07-08 00:00:00
Ending max drawdown	2010-02-08 00:00:00	2009-10-02 00:00:00
Sharpe Ratio	0,6864942773	0,7147881804
Time to recovery maximum	385	384
Time to recovery mean	32,05555556	30,14545455
Volatility	0,05705023047	0,06736278115

Given the previous back test results, the MDP is reliable, and we can implement it in real time. The previous portfolio was re balanced every 15 days, but each time, only a small number of assets was removed from the portfolio, and a small number of names enter into the portfolio. And the portfolio contains a very high number of assets. But we can do exactly the opposite, ie make an optimization problem on portfolios with between 80 and 300 stocks, which is re balanced each 3 months, with a big number of change.

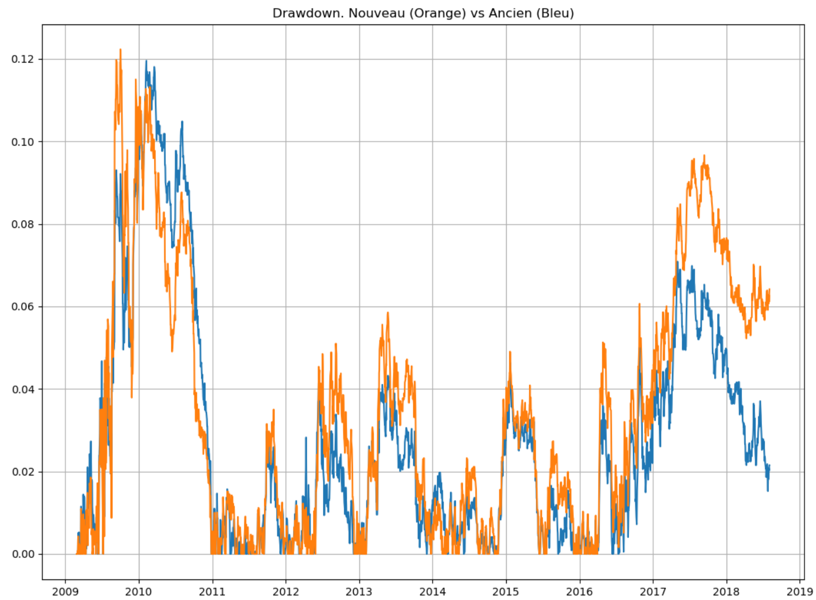

Figure 14: Drawdowns 10 years of realized $P\&Ls$; before optimization (bleu) and after (orange)

Figure 15: Ratio new hedged $P\&L$ versus new hedged $P\&L$

An other strategy from Boussard and Gavaudan, based on mean reversion, applies filters such that this kind of portfolio is obtained. This Portfolio is much more sensitive to the optimization. Firstly, we applied the Most Diversified Portfolio Optimizer, with the constraints $\mathcal{C}$ on a backtest period of 10 years, with a look back windows of 300 days for computing the estimators:

Once again, to assess the quality of the optimization we can examine the realized $P\&Ls$ with

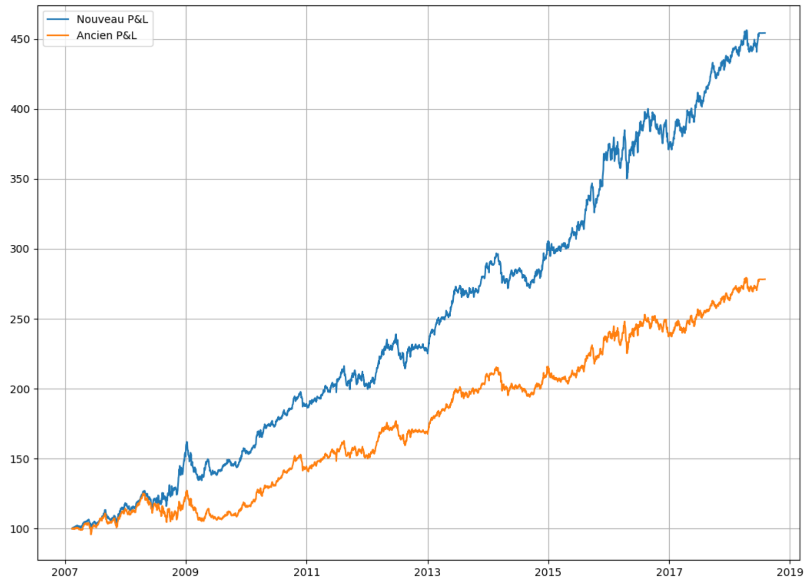

Figure 16: $P\&Ls$ before (orange) and after (blue) optimization. Portfolio with 80 to 300 assets, rebalancing each 3 months

risk metrics:

Risk metrics	Equally weighted portfolio	MDP
Max Drawdown	0,1723112529	0,1698323082
Starting date max drawdown	2009-01-09 00:00:00	2009-01-08 00:00:00
Ending max drawdown	2009-04-07 00:00:00	2009-04-07 00:00:00
Sharpe Ratio	0,8401890185	1,199580859
Time to recovery maximum	304	298
Time to recovery mean	34,68421053	30,19277108
Volatility	0,1017409651	0,1053957081

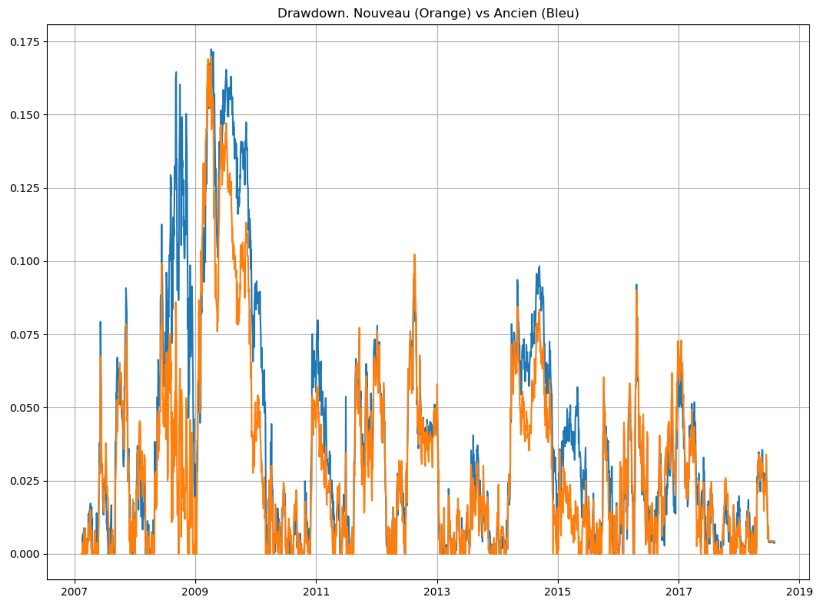

Figure 17: Drawdowns 10 years of realized $P\&Ls$; before optimization (blue) and after (orange)

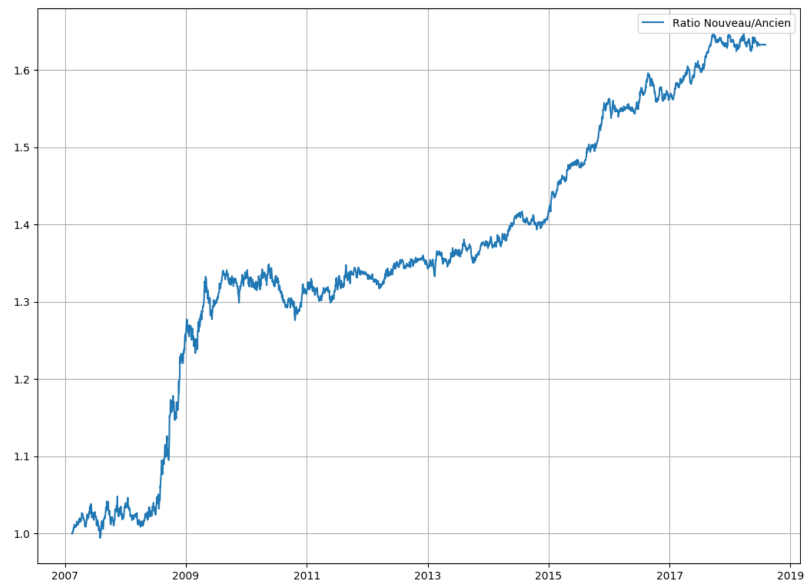

Figure 18: Ratio new hedged $P\&L$ versus new hedged $P\&L$

As this portfolio seems much more sensitive to the optimization, we will now compare the results with the two other optimization problems we have seen: the ERC and the Min Variance portfolio:
The risks metrics for these two optimization problems can be found in appendix.

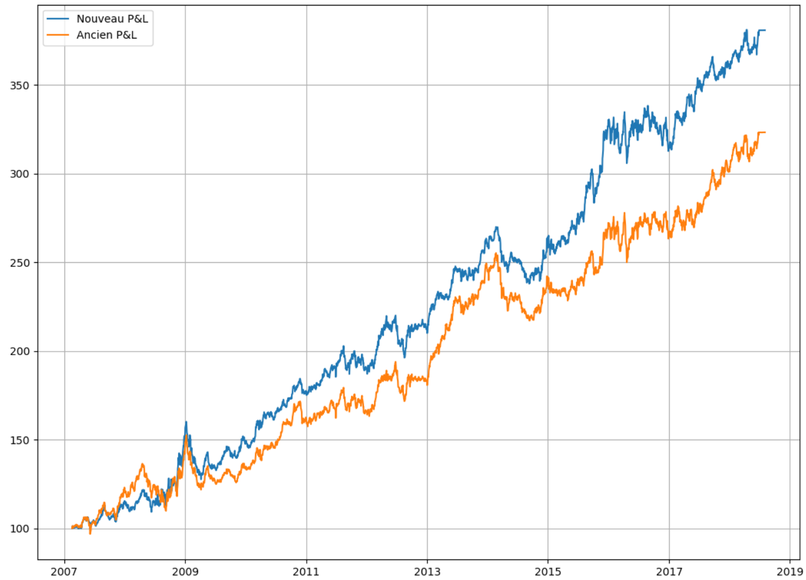

Figure 19: MDP (blue) versus Min Variance (orange)

Obviously, there is a clear advantage in performing the third step of the portfolio construction, as there is a huge increase of the net asset value (NAV) which does not come with any damage of the risks metrics,. Moreover the most diversified portfolio, with constraints $\mathcal{C}$ seems to be the best solution to this problem.

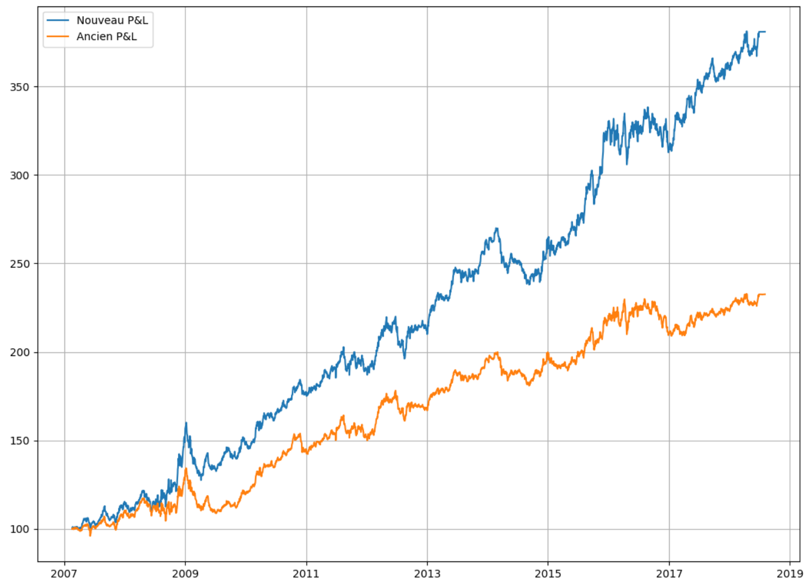

Figure 20: MDP (blue) versus ERC (orange)

*

* *

The convex structure of the MDP optimizer with linear constraints allows us to add new constraints that have not been taken into account until now. However, as we increase the number of constraints, we reduce the size of the feasible set (which is the set of weights which satisfy the constraints), and the feasible set can finally be far away from the optimal unconstrained value. Nevertheless, a good optimization is not only a good objective function, it is also a smart choice of constraints. There are three additional constraints which can be added to the MDP optimizer These constraints are often useless, but ignoring them can be lethal in extrem cases.

The lack of liquidity:

When one applies the previous optimization in the reality one can faces a major problem: the lack of liquidity. Indeed the filters could select stocks which are not easily purchased in the market. If the optimization asks to buy x of stock A, but only $k < x$ is on sale, then the portfolio can not be launched. Moreover, buying too much shares of stock A can affect the other trading strategies.

We made the choice to not trade more than 10% of the daily volume. The daily volume is the quantity of asset i traded each day. It can be estimated by taking the median of the realized daily volume over the past 3 months. Therefore, the new constraint is:

$$\omega_i \leq L_i, \text{ with } L = (L_i)_{1 \leq i \leq N} \text{the vector of daily volume.}$$

In the MDP framework, the previous constraint become:

$$\frac{y_i}{k} \leq L_i$$
$$y_i - L_i k \leq 0$$
$$\begin{bmatrix} I_N & -L & 0 & 0 \end{bmatrix}.x \leq 0$$

Bias:
As we said previously, with our hedging strategy, the only way to loose money is being long a portfolio with a decreasing $P\&L$ while the market increases.

This could occur as soon as there exists a bias in the portfolio, ie being long the portfolio is in fact equivalent to being long a certain factor. This factor can be:

- A country: The portfolio is very sensitive to the market of a particular country
- A sector: The portfolio is very sensitive to a particular industry

A perfect portfolio would have no bias to a particular factor.

For calculating these bias, we retained the following solution. The reader must be aware that this solution is an idea, and there exists no strong theory around it.

First of all we calculate (estimate) the covariance matrix of all the assets in the universe. Then we do a PCA (Principal Components analysis) over this covariance matrix, and take the first vector of the PCA.
Definition 4.1: *Wikipedia: Principal component analysis* [15]
Principal component analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. If there are n observations with p variables, then the number of distinct principal components is $\min(n-1, p)$. This transformation is defined in such a way that the first principal component has the largest possible variance (that

is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. The resulting vectors are an uncorrelated orthogonal basis set.

Our idea is to consider that the normalized first vector ω^* of the PCA gives a portfolio $(\Omega', \mathcal{A}')$ that explains the market, where $\mathcal{A}'$ is the universe.

The portfolio has no bias for a particular factor f if the sum of the stocks which belongs to this factor f does not exceed 1.5 times the sum in the market portfolio. For example the constraints for the factor UK can be written:

$$\sum_{A_i \in UK} \omega_i \leq 1.5 \sum_{A'_i \in UK} \omega'_i$$

All of this constraints are linear as they can be written:

$$E\omega \leq 1.5Bias$$
$$Bias = (\sum_{A'_i \in factor f} \omega'_i)_f$$
$$E_{if} = \delta_{stock\, i \in factor\, f}$$

In the MDP framework, it becomes:

$$E\omega \leq 1.5Bias$$
$$E\frac{y}{k} \leq 1.5Bias$$
$$E.y - 1.5Bias.k \leq 0$$
$$[\quad E \quad -1.5Bias \quad 0 \quad 0 \quad]x \leq 0$$

We implemented this method for the previous portfolio with the MDP optimizer. Fortunately, it deteriorates not the NAV, and add a security in case of huge fall of any factor. See below for example the bias in countries:

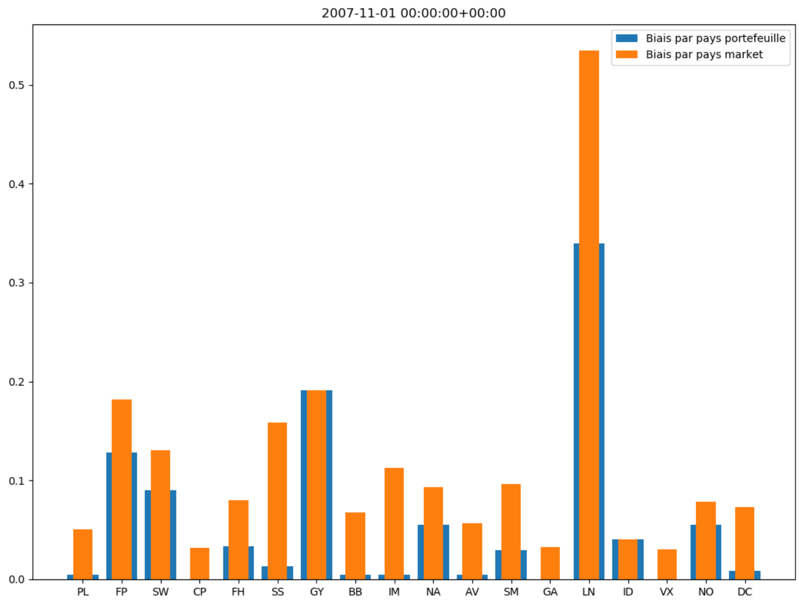

Figure 21: Bias in the portfolio with MDP (blue) versus Market bias (orange)

Beta:

Finally, as we do not want the portfolio to be highly correlated to the hedge, we want to fix its beta smaller than one. We compute the beta to its hedge for each stock thanks to the formula $\beta_i = \frac{Cov(r_i, r_{hedge})}{Var(r_{hedge})}$ (or by doing a linear regression), with r the temporal vector of rate of returns. The nex constraints is then:

$$\sum_{i=2}^{N} \omega_i(\beta_i - 1) \leq 0$$

$$(\beta - 1)\omega \leq 0$$

$$(\beta - 1)y \leq 0$$

$$[\begin{array}{cccc} \beta - 1 & 0 & 0 & 0 \end{array}]x \leq 0$$

With $\beta = (\beta_i)_{1 \leq i \leq N}$

The three previous linear constraints do not deteriorate the NAV and add a security. See below the $P\&Ls$ with and without these three constraints:

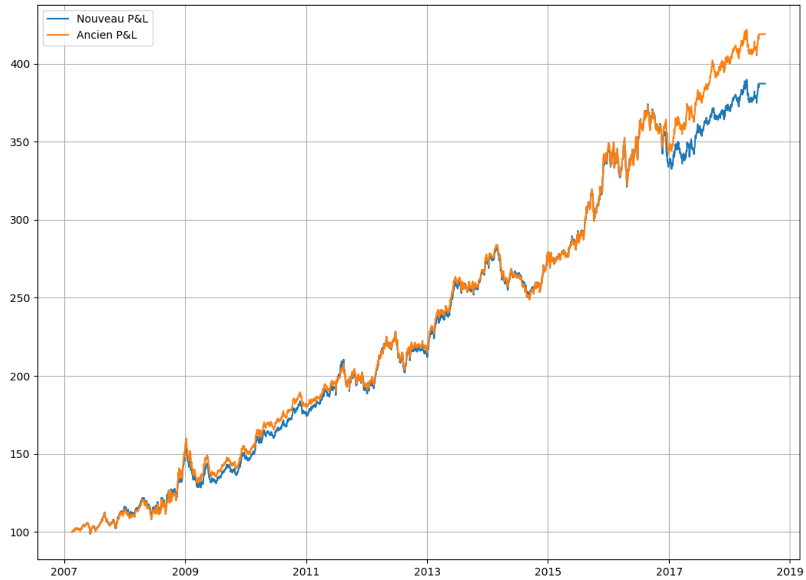

Figure 22: MDP with additional constraints (blue) versus MDP without additional constraints (orange)

Remarks 4.1: In all the previous graphs we added trading costs a posteriori. Each trade implies trading costs. Especially if the stocks is not really liquid and we trade a huge quantity, then there is a high price impact. We model this impact by adding a penalty of 20 bps, ie the true $P\&L$ is given by the following stochastic process:

$$P_t = P_{t_k} + P_{t_k} \sum_{i=1}^{N} \omega_i \left(\frac{S_i(t)}{S_i(t_k)} - \frac{H_i(t)}{H_i(t_k)} \right) - 0.20T$$

$$T = \sum_{i=1}^{N} |\omega_i - \omega_i^{old}|$$

With ω^{old} and ω the weights before and after each rebalancing.
Unfortunately all our tempatives to include trading costs in the optimizers failed, therefore we will not go in depth in trading costs in this thesis. But hopefully trading costs seems to never affect $P\&L$.

4.2 Robust strategies for Smart Beta

Since the financial crisis, investor have had an increasing confidence in simple products, and a decreasing willing to invest into highly sophisticated products. This explains the development of ETF fund.

Definition 4.2: From *Investopedia.com*

An ETF, or exchange-traded fund, is a marketable security that tracks an index, a commodity, bonds, or a basket of assets like an index fund. Unlike mutual funds, an ETF trades like a common stock on a stock exchange. ETFs experience price changes throughout the day as they are bought and sold. ETFs typically have higher daily liquidity and lower fees than mutual fund shares, making them an attractive alternative for individual investors.

The idea behind smart beta is to track an index as an ETF fund does, and then optimize the allocation of weights into the new product. The goal is simply do to build a more profitable index than the original one. This aim seems to be well adapted to our optimizer, and as we did previously, starting from the list of stocks in the DAX index GX1, we obtain the following MDP index:
With the risk metrics:

Figure 23: MDP DAX in blue versus GX1 index (DAX) in orange

Risk metrics	Equally weighted portfolio	MDP
Max Drawdown	0,512008891	0,524153448
Starting date max drawdown	2008-01-08 00:00:00	2008-01-08 00:00:00
Ending max drawdown	2009-03-06 00:00:00	2009-03-09 00:00:00
Sharpe Ratio	0,1819538599	0,3868824587
Time to recovery maximum	733	578
Time to recovery mean	74,03030303	58,92682927
Volatility	0,2303282692	0,206753243

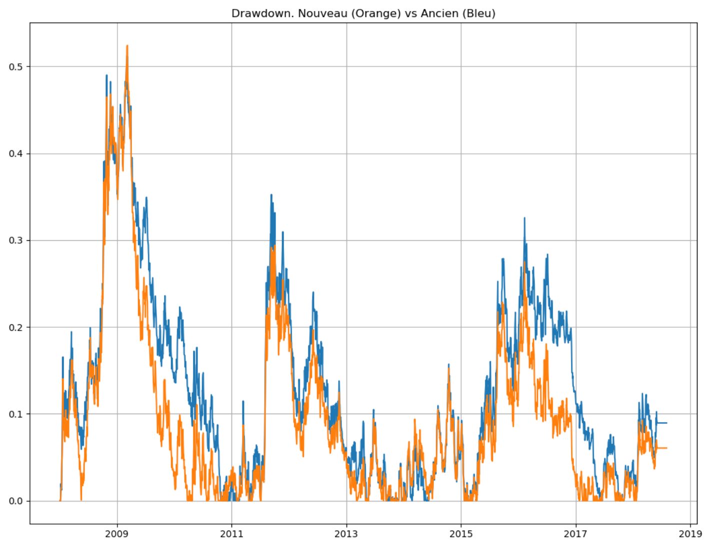

Figure 24: MDP DAX drawdown in orange versus GX1 index drawdown (DAX) in blue

Of course the previous analysis can be done for other index, and leads each time to the same kind of satisfying result.

Conclusion

In the complex path which leads to the construction of a portfolio, we focused in this thesis in the third one: the portfolio optimization. We have seen that this step can produce very different $P\&Ls$. Hence it is useful to study in depth the solutions that are already available in the literature, and then improve it. Hopefully, the convex optimization theory offers us powerful tools, such that we can easily build optimizer that are adaptable, and ease the decision.

We have built many optimizers in this thesis, but we studied in depth the most diversified portfolio (MDP), the Equal Risk Contribution (ERC), and the Markowitz Optimization (MPT), under the Boussard & Gavaudan's constraints . Both theoretically and in practice, the Most Diversified Portfolio appears to be the best solution to the issue of this thesis (recall the problematic in introduction: *Given a set of tradable assets and constraints, what is the best way to spread our money into these assets?*).

The final optimizer, which is currently used at Boussard & Gavaudan is mainly based on the Most diversified Portfolio theory. Once the result of optimization problem seems stable, we add further constraints, which are just safeties against extrem events. We obtained finally a robust optimizer, adaptable to many purposes.

A Appendix

Appendix A.1: ERC and Min Variance Risks Metrics

For the second portfolio, if we analyse the $P\&L$ with an ERC versus a MDP optimization for ten years, we obtain the following risk metrics:

Risk metrics	ERC	MDP
Max Drawdown	0,1899758564	0,2029667151
Starting date max drawdown	2009-01-07 00:00:00	2009-01-08 00:00:00
Ending max drawdown	2009-07-07 00:00:00	2009-04-07 00:00:00
Sharpe Ratio	0,7473438023	1,000278229
Time to recovery maximum	392	326
Time to recovery mean	44,13114754	34,61333333
Volatility	0,09444328992	0,1118017278

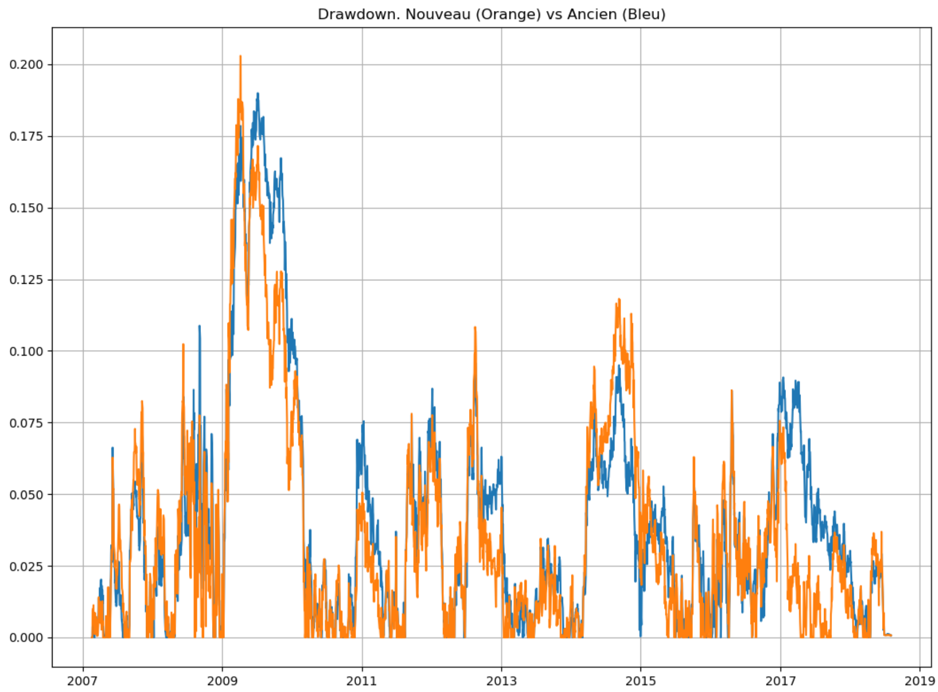

Figure 11: Drawdown MDP (orange) versus ERC (blue)

Risk metrics	Minimum Variance	MDP
Max Drawdown	0,2014053906	0,2029667151
Starting date max drawdown	2009-01-09 00:00:00	2009-01-08 00:00:0000
Ending max drawdown	2009-04-07 00:00:00	2009-04-07 00:00:00
Sharpe Ratio	0,8720384494	1,000278229
Time to recovery maximum	409	326
Time to recovery mean	40,93939394	34,61333333
Volatility	0,1125295415	0,11180172788

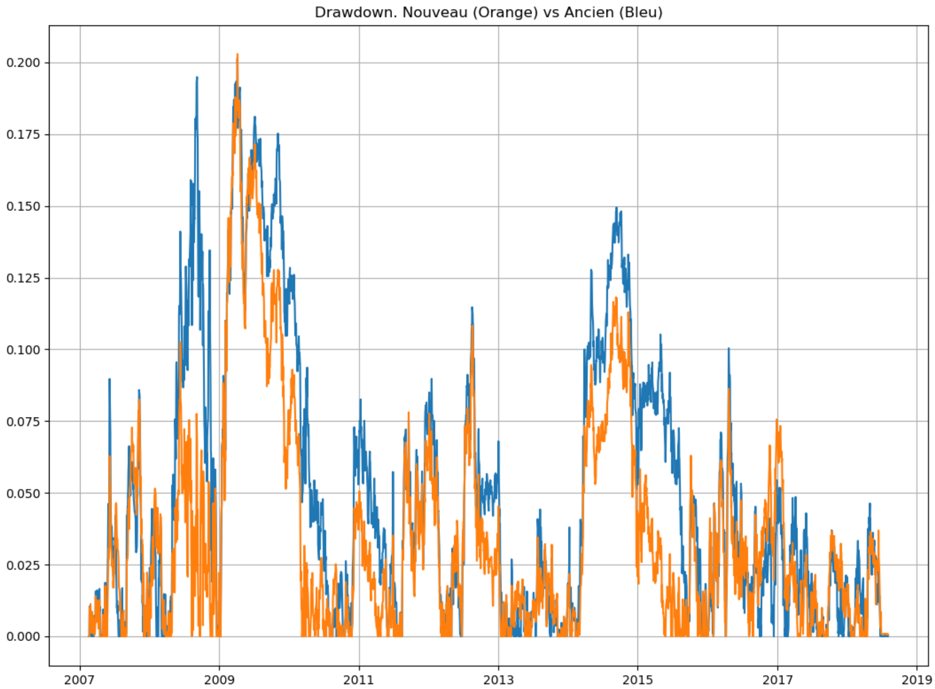

Figure 11: Drawdown MDP (orange) versus Minimum Variance (blue)

References

[1] F. Black and M. Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3): 637-659, 1973.

[2] O. Ledoit and M. Wolf, A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices, *Journal of Multivariate Analysis*, Volume 88, Issue 2, 26 February 2001.

[3] Chen et al., Shrinkage Algorithms for MMSE Covariance Estimation, *IEEE Trans. on Sign. Proc.*, Volume 58, Issue 10, October 2010.

[4] Kuhn, H. W.; Tucker, A. W. Nonlinear programming. *Proceedings of 2nd Berkeley Symposium.* Berkeley: University of California Press. pp. 481 492 1951.

[5] *http://cvxopt.org, http://cvxopt.org/userguide/coneprog.htmlcvxopt.solvers.conelp*

[6] Martin Anderson, Joachim Dahl, Zhang Liu and Lieven Vandenberghe, Interior-point methods for large-scale cone programming.

[7] *https://en.wikipedia.org/wiki/Sequential_quadratic_programming.*

[8] Portfolio Selection, *Journal of Finance*, Harry Markowitz 1952

[9] Edouard Berthe, Advanced Topics in Operations ResearchScenario-Based Portfolio Optimization Friday 21th October, 2016.

[10] Moshe Levy and Yaacov Ritov, PORTFOLIO OPTIMIZATION WITH MANY ASSETS: THE IMPORTANCE OF SHORT-SELLING *Anderson Graduate School of Management, UC Los Angeles*, 2001.

[11] Sebastien Maillard, Thierry Roncalli and Jerome Teiletche *On the properties of Equally Weighted Portfolio*, June 2008.

[12] Yves Choueifaty and Yves Coignard, Toward Maximum Diversification, *TOBAM*, 2008.

[13] Robert Merton, An Analytic Derivation of the Efficient Portfolio Frontier, *The Journal of Financial and Quantitative Analysis* 1973.

[14] William F. (1994). The Sharpe Ratio. *The Journal of Portfolio Management.* 21 (1): 49âĂŞ58. Retrieved June 12, 2012.

[15] $https://en.wikipedia.org/wiki/Principal\_component\_analysis$

[16] Margaret Archibald and Cwerado Martinez, The Hiring Problem and Permutation, *FPSAC*, 2009