DANMARKS TEKNISKE UNIVERSITET

42104

Introduction to Financial Engineering

Final project

Manuel Montoya Catalá (s162706)

December 2016

Contents

1 Portfolio optimization

1.1 Diversification

Choose financial stock indices (at least 7) for different markets or sectors. Time interval 20 years, monthly data. When necessary, convert all values in one domestic currency (e.g. USD).

Solution

The financial instruments chosen has been selected form the S&P500 index in order to make the CAPM analysis more meaningful. We have also included the index as a possible comodity of our portfolio. These financial instruments belong to different economic areas and are sumarized as follows:

1. SCHW: The Charles Schwab Corporation. Finance
2. AAPL: Apple Inc. Computers
3. TAP: Molson Coors Brewing Company. Beer.
4. LUV: Southwest Airlines. Transport.
5. CAT: Caterpillar Inc. Construction
6. SPY: Index S&P 500 ETF
7. TXN: Texas Instruments Incorporated. Electronics.

The time period chosen goes from 1-1-1996 to 1-1-2016, 20 years in total of monthly data, 240 samples per asset. We will work with the "Adjusted Close price", that is, the price at the end of the month, discounting dividends and splits. The data has been automatically downloaded from yahoo finance and stored into disk for future use. The next image shows the monthly evolution of the prices across these 20 years. We can observe that in the long run, all of them grow, but each has its own way.

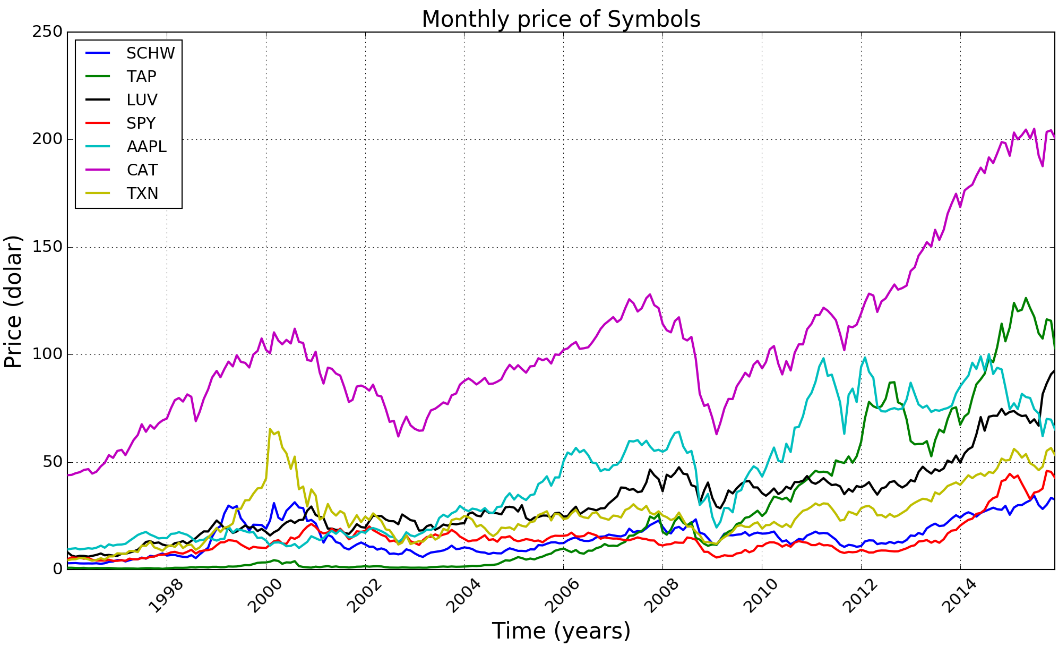

Figure 1: Monthly Adjusted close prices for all Symbols

The next figure shows the individual returns for the same period of time. We can appreciate times of bigger overall variance in the beggining. There are some common relations between all of them, for example, at the end of yeat 2011, all of them dropped. We can appreciate that the signals are not completely gaussian, because there are times with higher values and some small patterns could be seen.

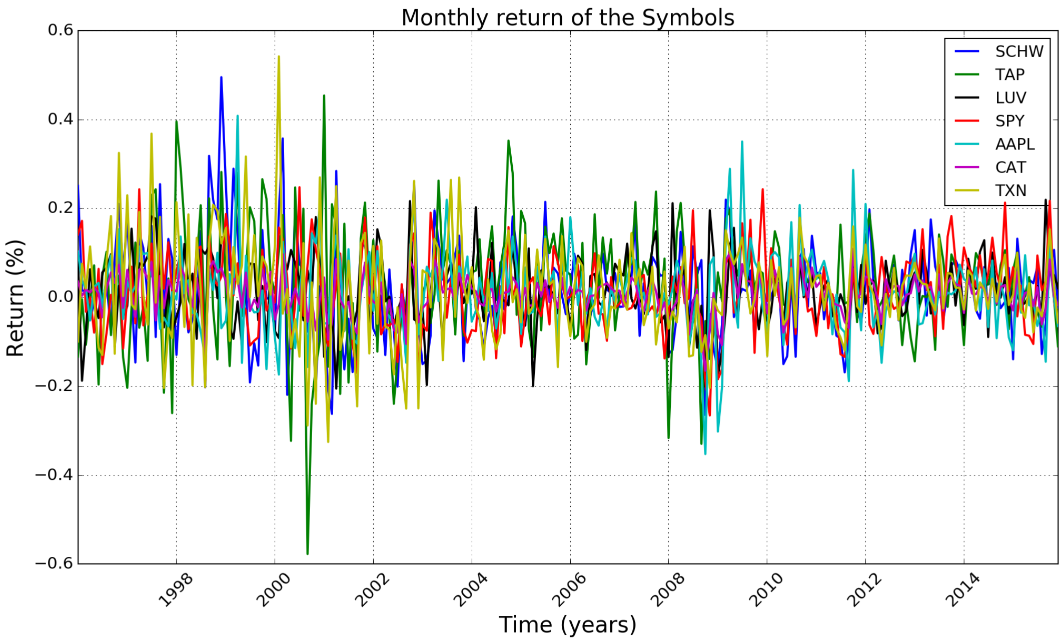

Figure 2: Monthly Adjusted close returns for all Symbols

In order to make the returns more gaussian, we could apply some kind of transformation and operate in a domain where all things are gaussian.

1.1.1 About the logarithm transformation

We want our returns to have a guassian distribution, because everything is analitically beautifly simple in the guassian world, nevertheless, this doesn't usually happen, the data in real life frequently has heavier tails, that is, the distribution has more uncertainty than the gaussian distribution, there are more ocurrences of the signal far from the centroid than the ones that happen in a guassian distribution.

The next graph shows the distribution for the returns of some of the symbols with a guassian distribution fitted to them. We can see some heavy tails, specially, in the left side of the distribution, we can see that it is asymetric and the left part usually has higher negative returns.

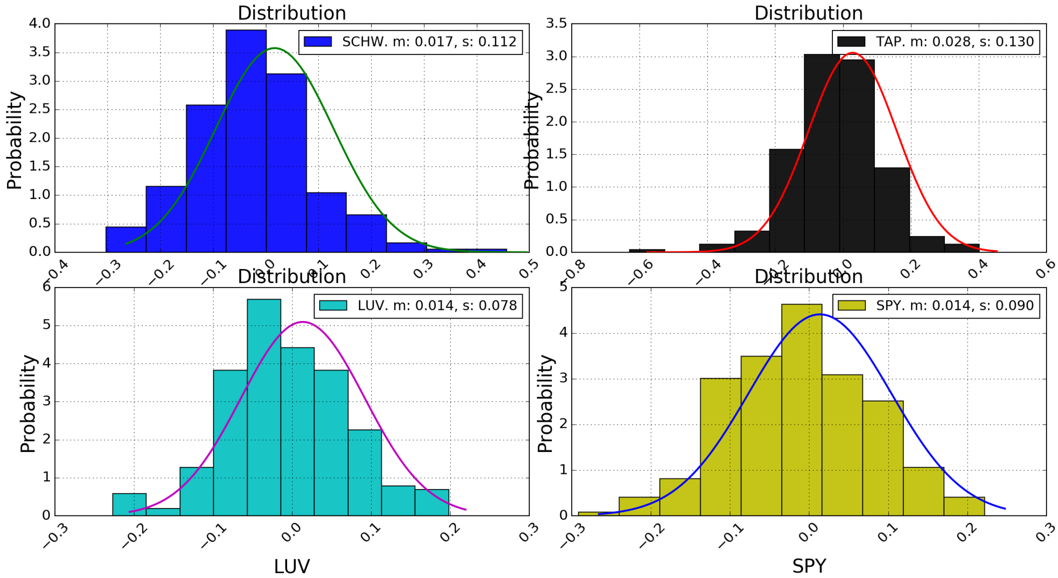

Figure 3: Distribution of the returns for some Symbols

A posible solution is to apply a transformaiton of the data into a new domain where it will be gaussian. Most likely a transformation that penalizes high values of the variable so that it will drag them to the centre of the distribution, making it more guassian-like. One of these examples is the logarithmic transformation.

In the literature it is recomended to perform a transformation of the original signal depending on the properties of its mean-range graph, the so called *Box-Cox transformations*. If the mean and the range of the signal are independent, no transformaiton is needed, but if there is a dependency, then the signal should be transformed. For computing this graph, we choose a window size w and we move it along the signal, in this case the time series of one of the assets, calculating the mean and range of the window.

The next figure shows the graph for different values of the window w. If a relation between the mean and the range is found, then a transformation could be benefitial, this relation means that there are areas in the time series with values higher than usual, which could be explained with a higher variance. Performing this transformation we seek to "saturate" the points with higher variance, so that they are not as high in the new domain. Also, any exponential trend will be transformed into a linear trend that can be easily eliminated by differentiation.

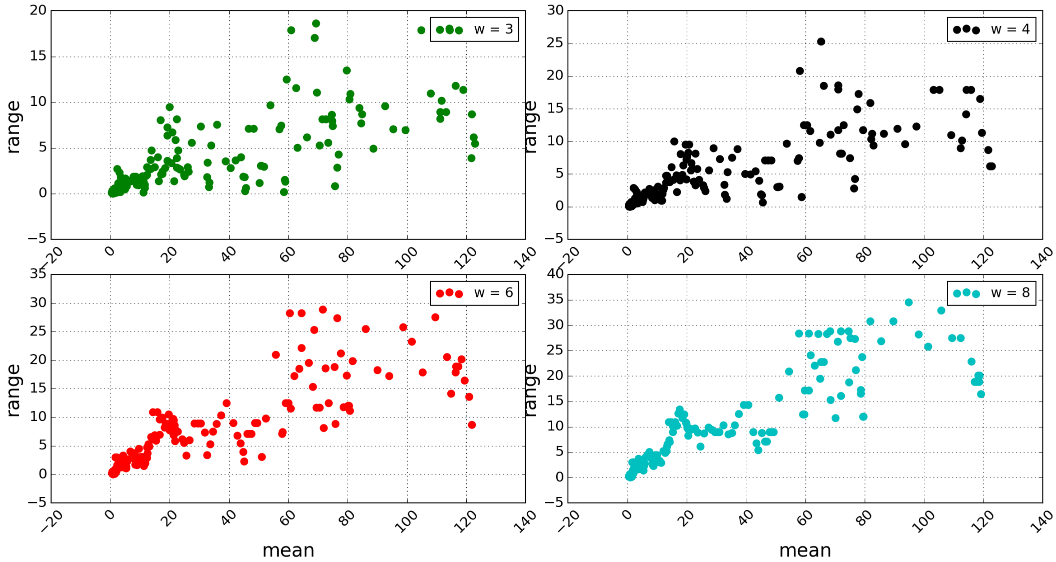

Figure 4: Range-mean graph for several window's size

The proposed transfomation, given the linear relation between the range and the mean, is the logarithmic transformation. The log function and its derivative can be seen in the next figure. As we can observe, the log transformation is most linear at $x = 1$, saturating values higher than it and increasing values that are lower. We can see that its derivative is $1/x$, at $x = 1$ it is 1, for higher values it decreases and for smaller values, it increases. This transformation would lower the tails of the original signal by making the bigger values smaller, and the lower values bigger.

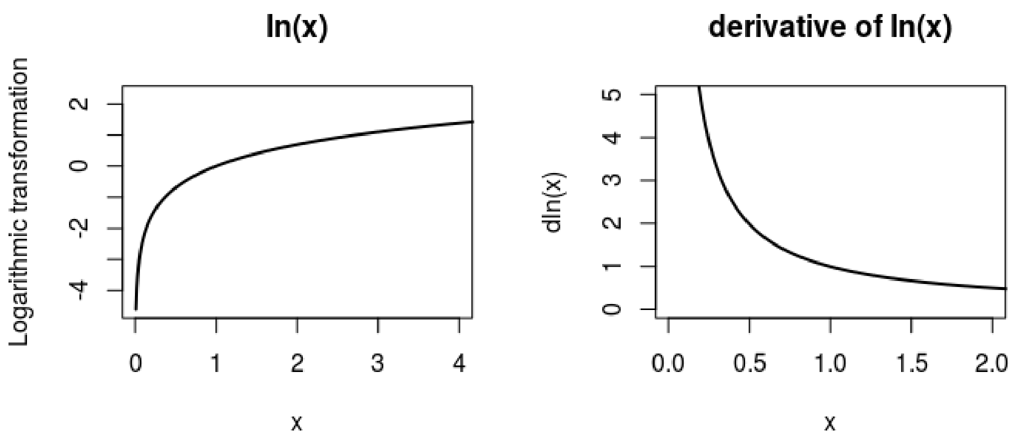

Figure 5: Logarithmic transformation

Of course this transformation could only be applied to a strictly positive signal. Once it is obtained, the output signal is usually differentiated with lag $k = 1$ obtaining the signal:

$$Y_n = log(X_n) - log(X_{n-1}) = log\left(\frac{X_n}{X_{n-1}}\right)$$

A common practice is, instead of doing this, obtain the returns of the original signal, and move them to the linear part of the logarithm transformation adding 1.

$$Y_n = log(1 + r_n) = log\left(1 + \frac{X_n - X_{n-1}}{X_{n-1}}\right) = log\left(\frac{X_n}{X_{n-1}}\right)$$

Which ends up being the same as the previous equation. Once again, we have to hope that the process X_n is either strictly positive or strictly negative, otherwise, the value could be negative or infinity. We will asume that this is the case, if the signal does not meet this requirement, I guess you can always add a constant value to the process so that it does.

This transformation is highly dependent on the mean of the orginial signal, lets say we add a constant value μ_* to the whole process. This will produce a smoothing on the signal since the division will tend to 1, thus being in the symetric, linear part of the transformation.

$$Y_n = log\left(\frac{X_n + \mu_*}{X_{n-1} + \mu_*}\right)$$

Maybe we can play with this value to obtain the best model for our data. Nevertheless, we have no time to do this in the current report. The next figures show the same information as before but for the log-transformed signals. As we can see:

1. The prices have been smoothed
2. The returns have less variance, but the negative returns are increased and the positive returns are saturated.
3. The mean-range plot is random
4. The distributions are more guassian like.

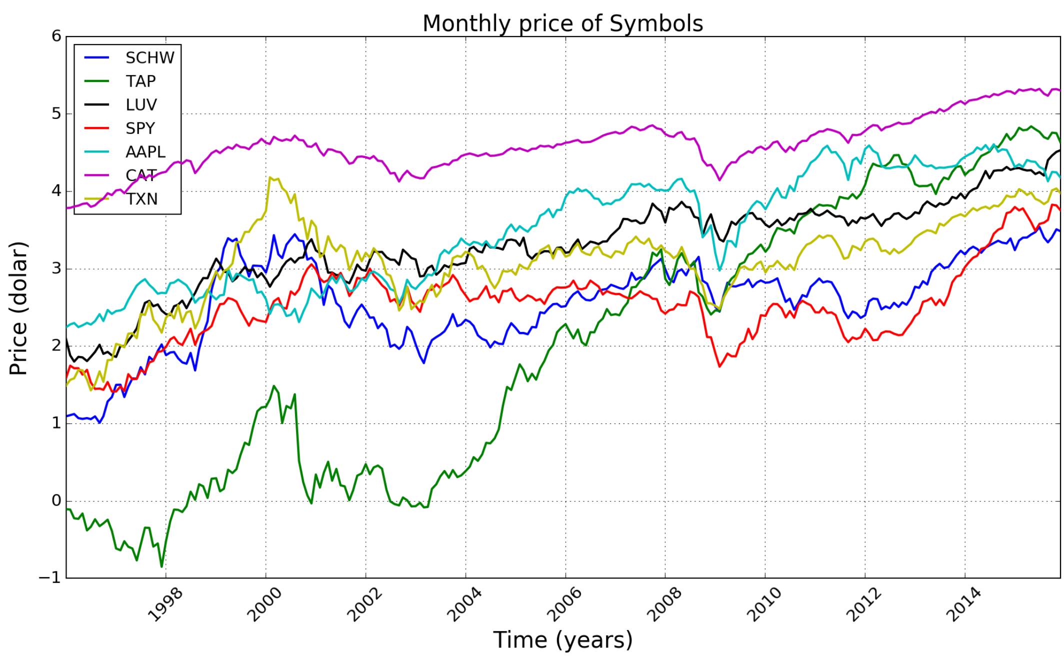

Figure 6: Monthly Adjusted close prices for all Symbols for the log-transformed data

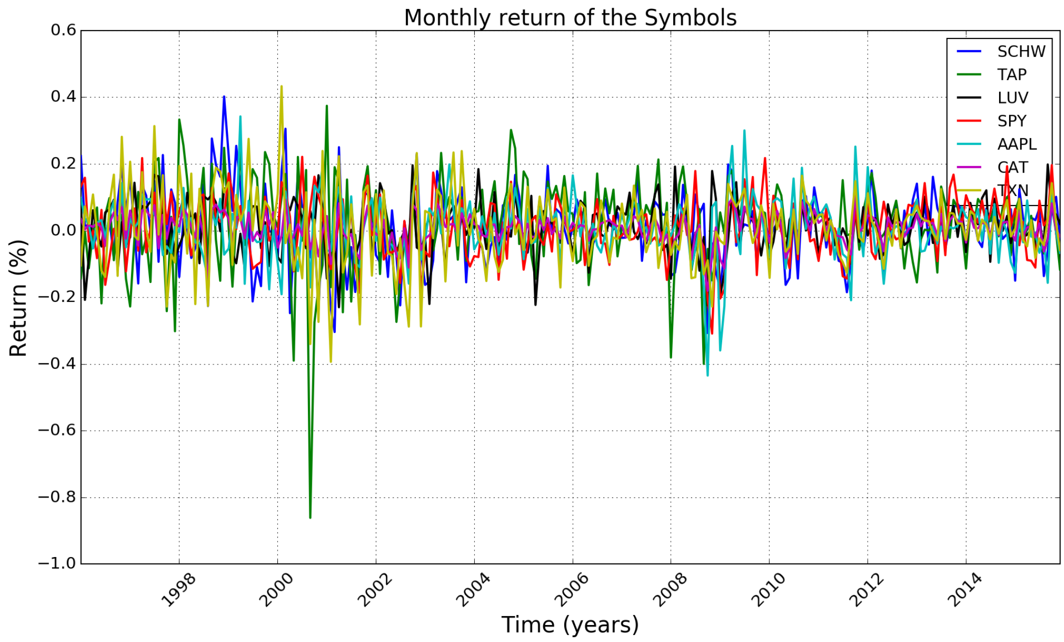

Figure 7: Monthly Adjusted close returns for all Symbols for the log-transformed data

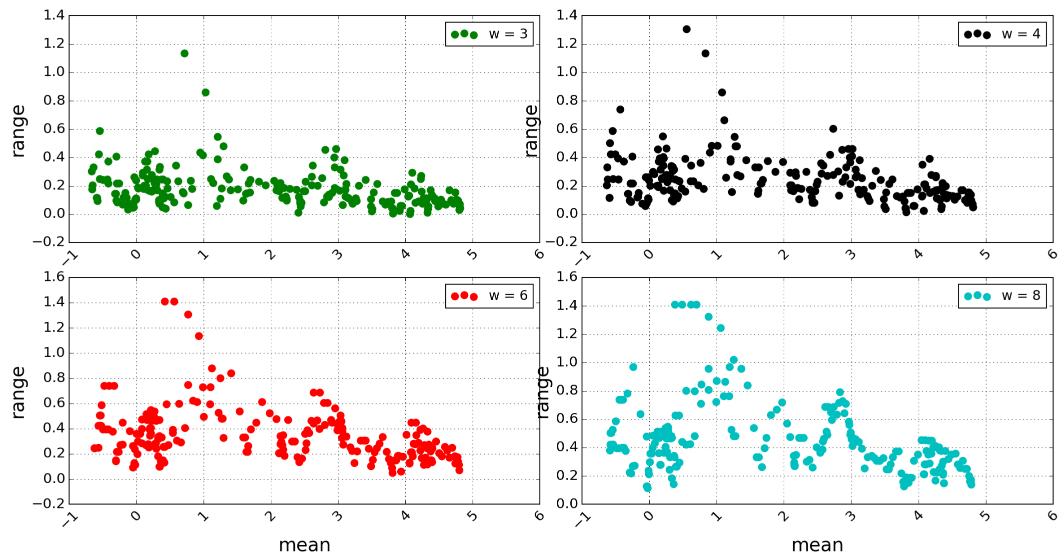

Figure 8: Range-mean graph for several windows with the log-transformed data

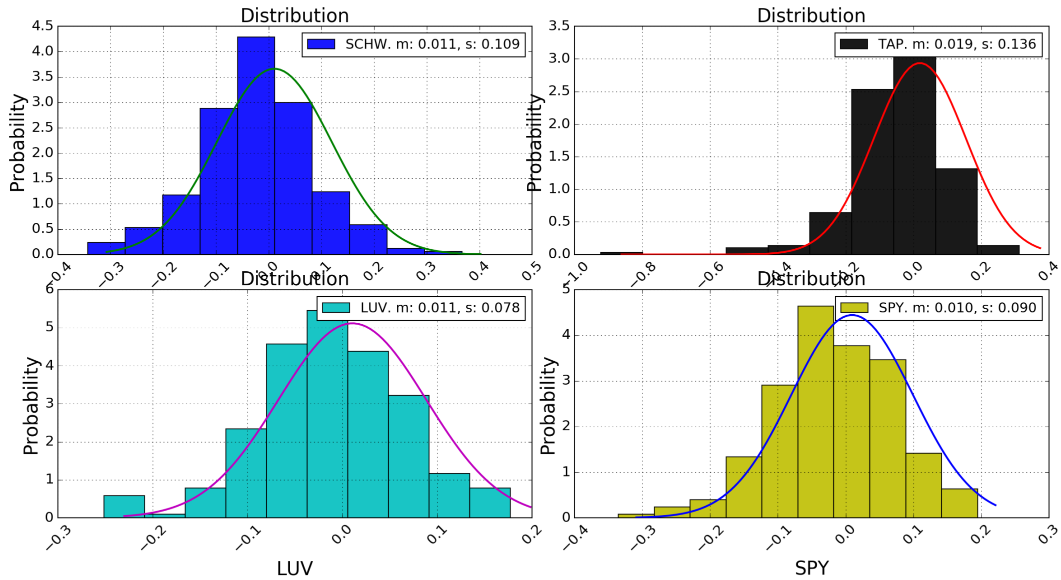

Figure 9: Distribution of the returns with the log-transformed data

So we see how a good transformation of the data could be useful. But we dont feel so confident about it, also, the following graphs and results obtained in the rest of the report are more visual if we do not transform the signals. The fact that this transformation will saturate high positive returns and amplify negative returns does not seem very convincing. Another of the problems with this transformation is, if for example we add the returns of the same signal, then, in the real domain (de-transformed), I am multiplying them, which does really seem too good. Of course there might be tricks and constraints so that we move in a zone of the transformation that will make sense to de-transformate. But we are sticking to the original values of time series and returns.

1.2 Estimate

Expected yearly returns and covariance matrix with a rolling window of 10 years (in annual steps).

Solution

Well, now that we have our data (and we have not transformed it), we calculate the expected yearly returns and covariance matrix form the monthly data. Assuming guassianity, the yearly expected return will be 12 times the monthly expected return and the yearly covariance matrix will be $\sqrt{12}$ the monthly covariance matrix.

The next graph shows the returns asked. We only show the first 6 covariance matrixes but it is representative.

1. Probably the most remarkable thing to say is that the returns and covariance matrices change over time
2. The assets CAT and TXN usually have the highest variance.
3. The TAP and AAPL usually have the highest return.
4. We can also see the crisis started in 2007 in the graph, of course with a little delay since we have a 10 year rolling window.
5. The covariance matrix seems to be always positive, but some of the covariances could be negative.

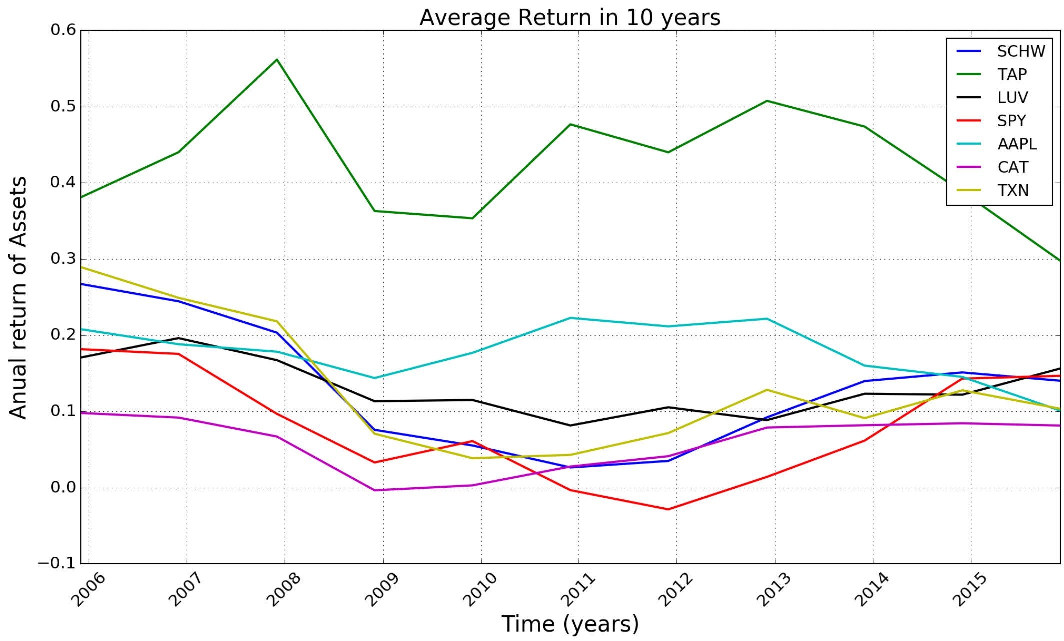

Figure 10: Yearly returns with 10 year rolling window

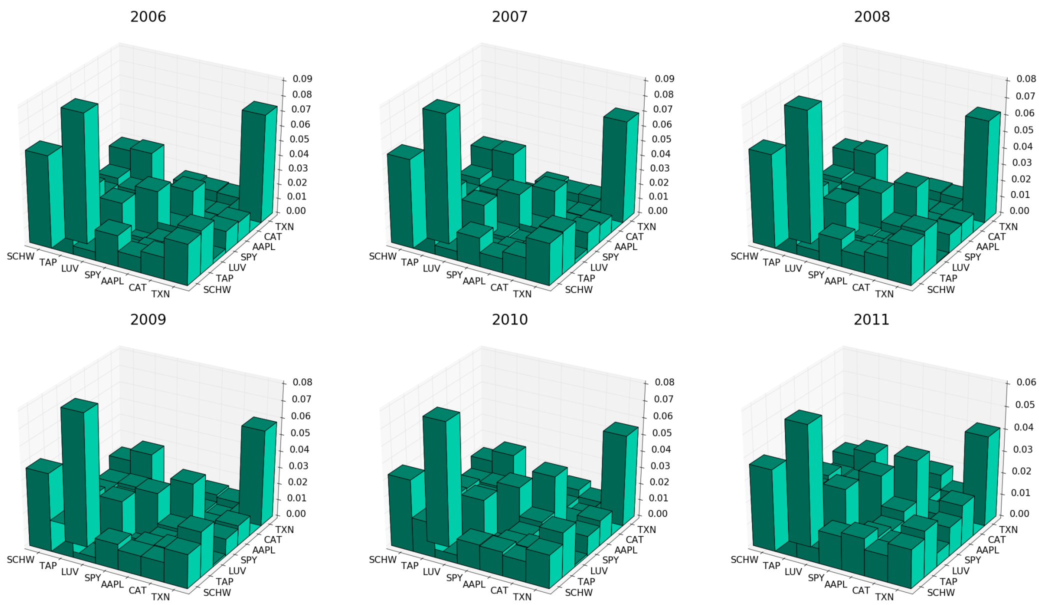

Figure 11: Covariance Matrixes with 10 year rolling window

1.3 Efficient frontier

Show in one or more figures the 11 different efficient frontiers (one for each rolling window) and explain in your report the diversification benefit.

Solution

Well, the efficient frontier as a concept is the set of portfolios that maximize the return E_p for any given risk $Risk_p$, under some constraints. The measure of risk is usually the standard deviation of the returns of the portfolio. This maximization can have different constraints, thus changing the shape of the frontier. Three important types of efficient frontiers are:

- Markowitz: The sum of the allocation weights is 1 and no short sales are allowed.
- Normal: The sum of the allocation weights is 1 and short sales are allowed.
- Lintner: The sum of the absolute value of the allocation weights is 1 and short sales are allowed. The reason for this one is that, in practice, when you short sale, you don't get the money, but rather pay a margin instead, as it is done when you go long.

The way to compute them is different, due to the different constraints, Markowitz and Lintner need a convex optimizer, but the Normal frontier can be computed analytically from the mean returns $\vec{R}$ and covariance matrix S of the assets.

$$w^* = \delta \vec{R} \cdot S^{-1}$$

Where δ is the price of risk (risk adversion), the price people are willing to pay for the risk they are taking. Bigger values of δ mean that the risk is penalized more and we want lower risk portfolios. This equation comes from maximizing the utility function:

$$U = w^{'}\vec{R} - \delta w^{'} S w$$

Since this function is convex with respect to w, to maximize it we just compute the derivative and equal to 0. In our case, we want the optimal portfolio that uses all the money, that is, the sum of its weights is 1, so we are not taking the δ into consideration, it gets normalized away.

$$w^* = \frac{\delta \vec{R} \cdot S^{-1}}{\sum_{i=1}^{N} \delta \vec{R} \cdot S^{-1}[i]} = \frac{\vec{R} \cdot S^{-1}}{\sum_{i=1}^{N} \vec{R} \cdot S^{-1}[i]}$$

To calculate the δ, we can just compute the returns and risk of the portfolio and use the equation:

$$\delta = \frac{R_p - R_f}{\sigma_p^2}$$

As we multiply our portfolio by a constant k, the return will increase by k and the variance by k^2, so the value of δ will diminish. The more risk we are willing to take, the less risk adversion δ we have.

Also, given two points (porfolios) on the Normal efficient frontier, represented by their allocation weights W_1^* and W_2^*, we can obtain any other point in the frontier as a convex linear combination of these two. This is due to the convexity properties of the utility function. It can be proven that all the points in the efficient frontier form a hyperbola. This can be seen from the linear combination of the assets, for a high enough λ, the effect of the other portfolio becomes insignificant and thus we have an asintotic linear progression (in the mean-risk space) of the efficient frontier points.

$$W_n^* = \lambda W_1^* + (1 - \lambda) W_2^*$$

We have tried to play a little with them, so let's look at an example where we see all of them together for a subset of the assets. The next graph shows the different efficient frontiers along with 100000 random portfolios that meet their restrictions. In this kind of graph, we can plot any univariate ramdom variable that we wish (usually gaussian because they have mean and std as suficient statistics) with a single point where the x-axis is its standard deviation (risk), and the y-axis is the mean (return). Each point in the graph corresponds a portfolio (characterized by its mean return and risk), and of course, each portfolio has a set of allocation weights w associated to it.

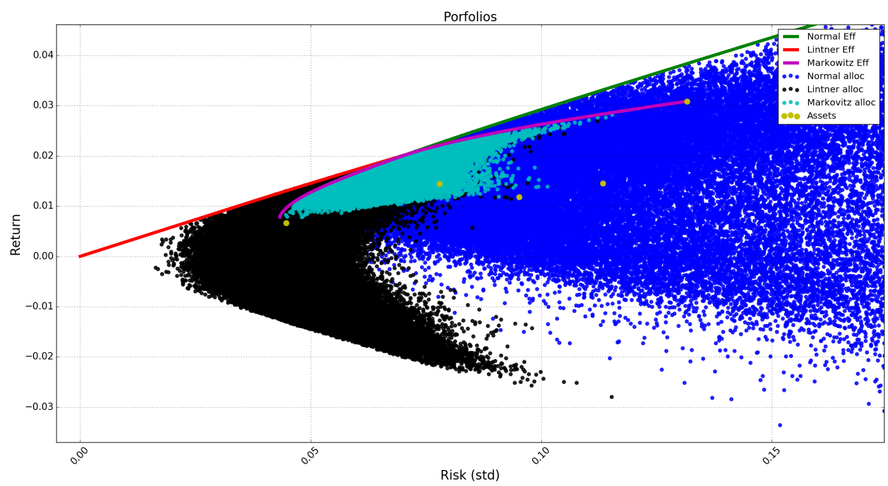

Figure 12: Different allocations and efficient frontier for a subset of assets

The mean and std of the portfolios can be computed of 2 different manners in the library. The first form is just performing the linear combination of the returns sample by sample and then calculate the mean return and variance of the return. The second form is calculating the mean return and covariance matrix of the assets and then use the linear combination properties of gaussian variables to get the resulting mean and variance. Both methods yield the same result.

Some observations about the previous graph can be made:

- First we must clarify that for the Lintern frontier, we were not really able to to do it properly, the frontier drawn ensures that the sum of the absolute weights is lower than 1, not exactly 1. So, in goes all the way to the (0,0) origin, still a good aproximation, the random allocations are right though.
- Obviously, the most restrictive case is the Markowits frontier, then the Lintnet frontier, and then the Normal frontier, that has the previous two as particular cases.
- For more clarity, we can plot the 3 different types of frontiers differently, and then, also, there is a small region where they converge, maybe this region would be a good candidate for portfolio allocation. The following is not the best graph ever, but we hope it helps the overall picture.

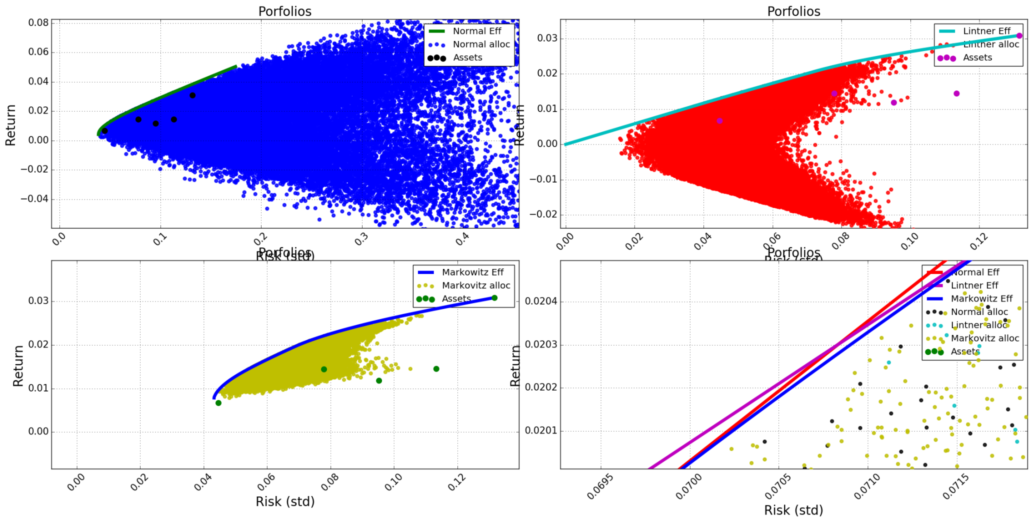

Figure 13: Different allocations and efficient frontier for a subset of assets 2

Moving on to what was really asked in the question, we decide to use the Normal efficient frontier, since it is the less restrictive and easier to compute. For the rest of the report, it is the only one that we will use. The next graph shows the 11 yearly efficient frontiers calculated with the 10 year rolling window. As we can see, the efficient frontier changes with time, since it depends on the returns and covariance matrix that surprisingly enough, change with time.

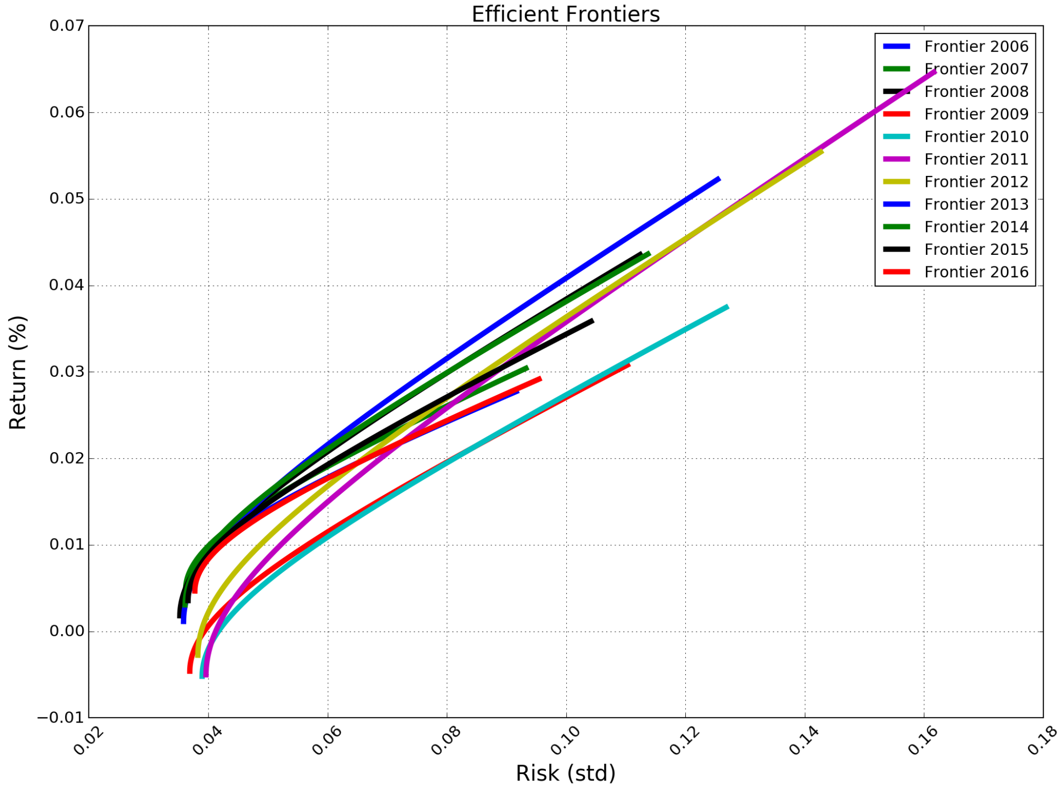

Figure 14: Efficient frontiers for the rolling windows

About the diversification benefit, diversification allows us to find a portfolio (linear combination of assets) whose return sortino ratio is better than any of the separate assets. In this whole model we are assuming that the returns of the stocks follow a multivariate gaussian distribution, with mean R and covariance matrix S, the return and variance of the portolio can be obtained as:

$$R_p = w^{'} R$$

$$Var_p = w^{'} S w$$

The fact that we can find a portfolio with a lower variance is thanks to the covariances between the assets. Given 2 different correlated assets X and Y, the mean (return) and variance (risk) of a portfolio (linear combination) of them is:

$$E(w_1 X + w_2 Y) = w_1 E(X) + w_2 E(Y)$$

$$Var(w_1 X + w_2 Y) = w_1^2 Var(X) + w_2^2 Var(Y) + 2 w_1 w_2 Cov(X, Y)$$

As it can be seen, as long as the covariance is less than 1, we will always obtain a lower variance than performing the same combination for a single asset. This function can generalized for N assets and the same principle holds.

$$Var(w_1 X_1 + \dots w_n X_n) = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + 2 \sum_{i=1}^{n} \sum_{j>i}^{n} w_i w_j \ \mathrm{Cov}(X_i, X_j)$$

In the literature, it is also explained that the return of an asset is composed by two different kinds of risk:

- Systematic Risk: The one given by the environment, like the country, or the economic sector. The risk inherent to the entire market or the entire market segment.
- Non-Systematic Risk: The one given by the specific situation of the asset (a specific company).

The Non-Systematic Risk can be reduced with diversification by investing in sectors that are negatively correlated respect to news.

1.4 Tobin separation

Repeat the assignment c), now including a risk-free rate of 1%. What are the consequences for the portfolio management in such a world?

Solution

The efficient frontier calculation does not depend on the risk free rate. It just gives you the maximum return portfolio for a given risk level, when we are using all our money. Which is the same thing as saying that maximizes the Sortino Ratio. The constraints imply that the sum of the weights of the allocation are 1, so we are using all our money, so we do not either borrow or lend money.

Many times what it is done, is to substract to each asset, the risk-free rate and then recompute everything. This will provoque all assets to go down in the y-axis R_f units, puting the previous R_f value in the origin (0,0). If this is done, then this new efficient frontier is equal to the previous one displaced R_f units downwards. But then we add the risk free rate back to the return of our portfolio and everything is alright.

$$R_p = R_f + w_{ap} \cdot (E_a - R_f)$$

Where w_{ap} is the allocation weights of our portfolio, and $(E_a - R_f)$ is the vector of mean returns minus the risk-free return.

Maybe this is done because it is assumed that assets should yield that return as a bias, in order to even think to start investing in them. Why would you invest in something risky if you can obtain it with zero risk ?

Summarizing, when we have risk free rate R_f we can borrow or lend money at this rate. It is like another asset that we can combine with our portfolio. It is a riskless asset, so the combination of this asset with the portfolio will lie in a straight line in the return-risk graph. The equation of the return of the total portfolio is:

$$R_p = (1 - X)R_f + X \cdot R_{ap}$$

Where R_{ap} is the return of the portfolio composed only by assets. X is the percentaje of our money that we are investing in it, computed as the sum of the weights of the portfolio R_{ap}. And yes, this equation is actually the same as the previous one.

- If X is 1 we do not borrow or lend money. We do not care about it.
- If X is lower than 1, we are not using all our money so we are lending the rest at the risk free rate R_f
- If X is higher than 1, then we need to borrow money at the risk free rate.

Intuitively, if R_f is high, we will be more clever to lend more money, since we have a fair enough return with no risk, if R_f is low and we find a good portfolio allocation, we might as well borrow money and invest more in it, since we like a little more risk.

Now, let us see the **market line**. The market line is the set of optimal portfolios that we should use. It is the set of portfolios that, combining the assets and the risk-free rate, has the highest return for a given risk. Let's consider the case in which there is no risk free rate, $R_f = 0$. All the portfolios in the efficient frontier are optimal respect to the fact that, if we have to use all our money,$\sum w = 1$, then these points have the best Sortino Ratio that we can get for a given amount of risk. But among them, the is one that has the best sortino ratio, all of them have a different Sharpe Ratio since the efficient frontier is not a line. This is the optimal portfolio to invest in if we use all our money. The Sortino Ratio equals the slope of the portfolio in the return-risk graph.

$$SR = \frac{R_{ab} - R_f}{\sigma_{ap}}$$

Where R_{ap} is the return of the portfolio only composed by assets and σ_{ap} is the variance of the portfolio only composed by assets. The next image shows the market line of the portfolio with no risk free rate. As we can see, all the portfolios in that line are optimal, they are better than any other portfolio we could come up with.

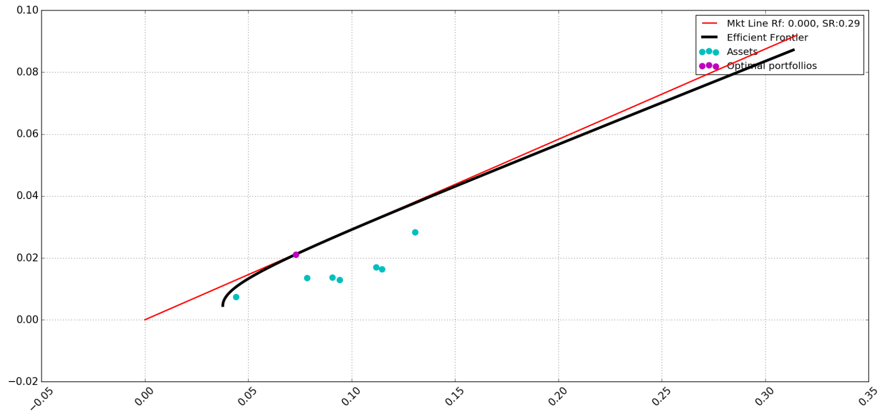

Figure 15: Market line for no risk free rate

If there exists a risk free rate R_f, then, the optimal portfolio in the efficient frontier and market line changes. Now, the market like would hit the y-axis at R_f, and then touch the first portfolio in the efficient frontier that it encounters. From the equation, we can see that the Sortino Ratio decreases, the slope of the market line decreases. So, in a situation where we have a higher R_f:

- We have a new optimal porfolio in the efficient frontier to combine it with.
- We would not choose the optimal previous portfolio in the frontier line because we can obtain a portfolio in the new market line that is above that one.
- The set of portfolios of the new market line will be better for portfolios in which we do not use all our money. Because now we can combine them with the higher R_f that gives more money.
- The set of portfolios of the new market line will be worse for portolios in which we borrow money. Because now the price of borrowing money is higher

The next figure shows the market line and optimal portfolios for different values of risk free rates where we can observe all the previous statements. There are even negative risk free values.

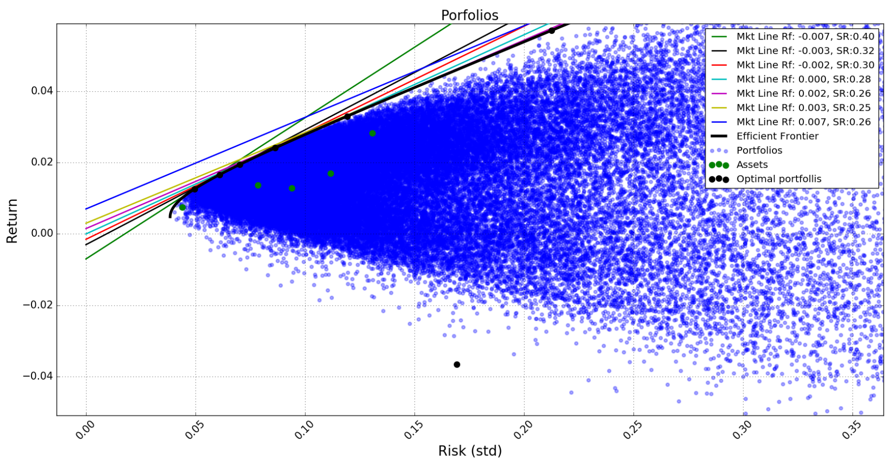

Figure 16: Market lines for different risk free rates

There is still one more thing to realize, look at the lonely black dot in the lower part of the previous image. This is also an optimal portfolio in the efficient frontier that we combine with the risk free rate to get ther maket line, in this case the blue matket line, the one that has the biggest risk free rate. Why is this happening to me ?

Well, what happens is the following, first of all, the frontier line is a hyperbola, so there might not exist a tangent point to it that crosses the y-axis in the given risk-free return, as it happens in this case. The optimal positive porfolio where we can invest all our money will not exist. The optimal line crosses the negative part of the efficient frontier instead. It should be seen that, for every porfolio P, there exists its oposite portfolio $P^- = -P$ which simply has the negated weights as the original portfolio. In the graph, this portolio is the reflexion of the original around the R_f point. In this negated porfolio, we are spending the oposite amount of money, if we spent X in the orginal one, then in this one we are spending $-X$, that is, we are borrowing X.

So, in this case, the optimal thing to do is to take this optimal negative portfolio and negate it. This way, we have borrowed 100% of our money. When $R_f = 0$, never mind, we just use the negative portfolio and that is it. We are obtaining from the market 100% of the money we initially have and we are not spending ours. So we can lend 2 times our original money in free - risk rate ! In this inverted world, to move accross the market line to the right is to sell, and to move back is to buy. When it is negative, I do not invest in the given portfolio, but in the opposite one. The equation of the line still valid, but the std should be inverted. This is what we do to get the real blue market line in the example.

1.5 Asset allocation

Choose a constant required return and calculate the optimal asset allocation for the case of the Tobin separation efficient frontier in the different years. How high is the portfolio turnover on average each year?

Solution

In this exercise, for every rolling window of 10 years, we calculate the efficient frontier and the market line. Then, we move thorough the market line to calculate the portfolio that will give us the desired return, to do so, we just have to multiply the optimal portfolio in the efficient frontier by the right number (which is the same process as moving through the market line). Of course this desired return is only the return that we would have obtained if we hold this portolio during those previous 10 years.

The next graph shows the market line and the portfolio chosen in the return-risk graph for the 11 rolling windows. As we can see, all of them have the same desired return, but with a different risk. We can also see in the legend, the value X, this is the sum of the weights of the portfolio needed to obtain such return.

$$X = \frac{(R_{des} - R_f)}{(R_{opt} - R_f)}$$

Where R_{des} is the desired return and R_{opt} the return of the optimal portfolio in the efficient frontier. In our case, the desired optimal monthly return is $R_{des} = 0.003$ as it can be observed in the picture.

As we can appreciate in the graph, some of this X are negative, which means we actually had to be short on the optimal porfolio because it was a very poor year, the market went down, we saw this in the previous question. This is also the sum of the weights of the desired portolio, since it is the factor that multiplies the optimal porfolio whose sum of weights is 1.

$$W_{des} = X \cdot w^*$$

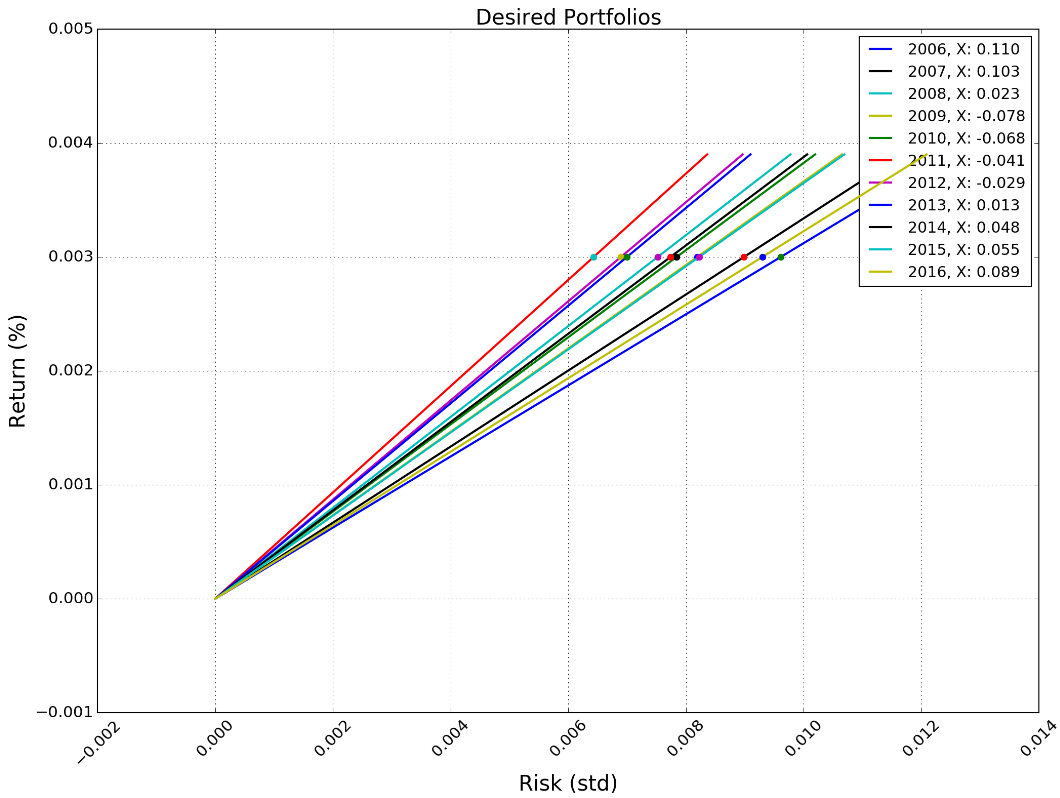

Figure 17: Desired portfolios and their market lines

So, now we compute the turnover for every year, the sum of absolute differences in allocation between the assets of two consecutive years. As it is logical, we cannot compute it for the year 2006 because we do not have the allocation of 2005.

$$Turnover_i = |W_{des} - W_{des}|$$

The next graph shows the turnovers for the different rolling windows. There are 3 bar charts:

- The first is the absolute turn-over at each year in terms of our initial money.
- The second one is the sum of the absolute weights of the desired porfolio each year. This is the real value that tell us each year, how much percentaje of our money we played into the market, and payed trading costs for.
- The third one is the turnover pondered by the absolute value of the previous year. This might me more informative since it tells us how much percentaje of our previous allocation are we moving.

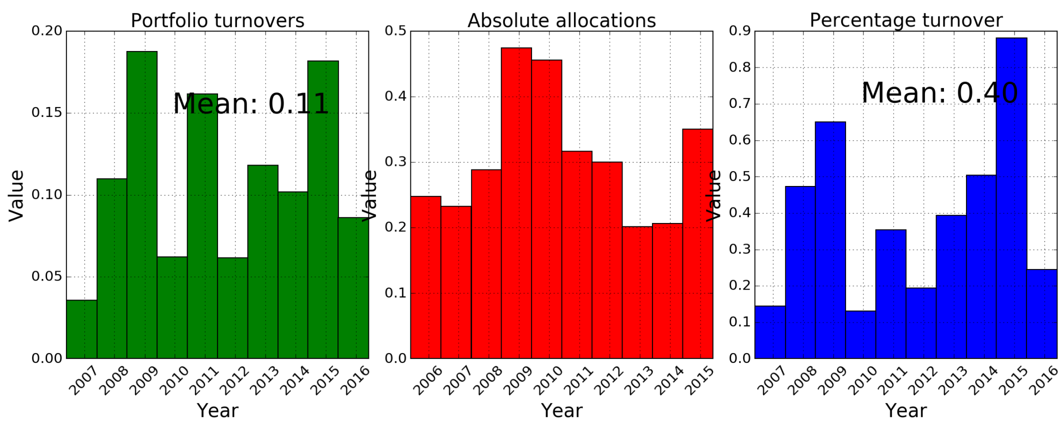

Figure 18: Turnovers

We cannot really appreciate a pattern. But we can see that two big turnovers are in years 2009 and 2013 which are years where the X changes sign. The normalized turnovers might be a more reliable indicator of periods of change in the properties of the market.

1.6 Backtest

Conduct a backtest (ex-post) for the optimized portfolio (with annual re-balancing, see point e). Calculate the average return and the standard deviation for your ex-post portfolio out of sample? (i.e. your portfolio is created by calculating the mean and covariance in-sample to find your optimal portfolio, BUT then you hold your portfolio for one year out-of-sample and calculate the return you have had during this period. Then you repeat the same process again and again for each period.)

Solution

In this question, we follow the same process of the previous one, we select a desired return, calculate the optimal portfolio using the market line and then hold it for a year. In this testing year, we will obtain 12 returns from which we can obtain the mean return and standard deviation. Of course, since the statistical properties of the return in that year might have nothing to do with the statistical properties of the previous window of 10 years, the obtained returns can be anything.

First things first, the next graph shows the asked average return and standard deviation of the portolio returns. It also includes the desired return, as a reference point. To see how well the portfolio returns compare to the index returns, we also plot them.

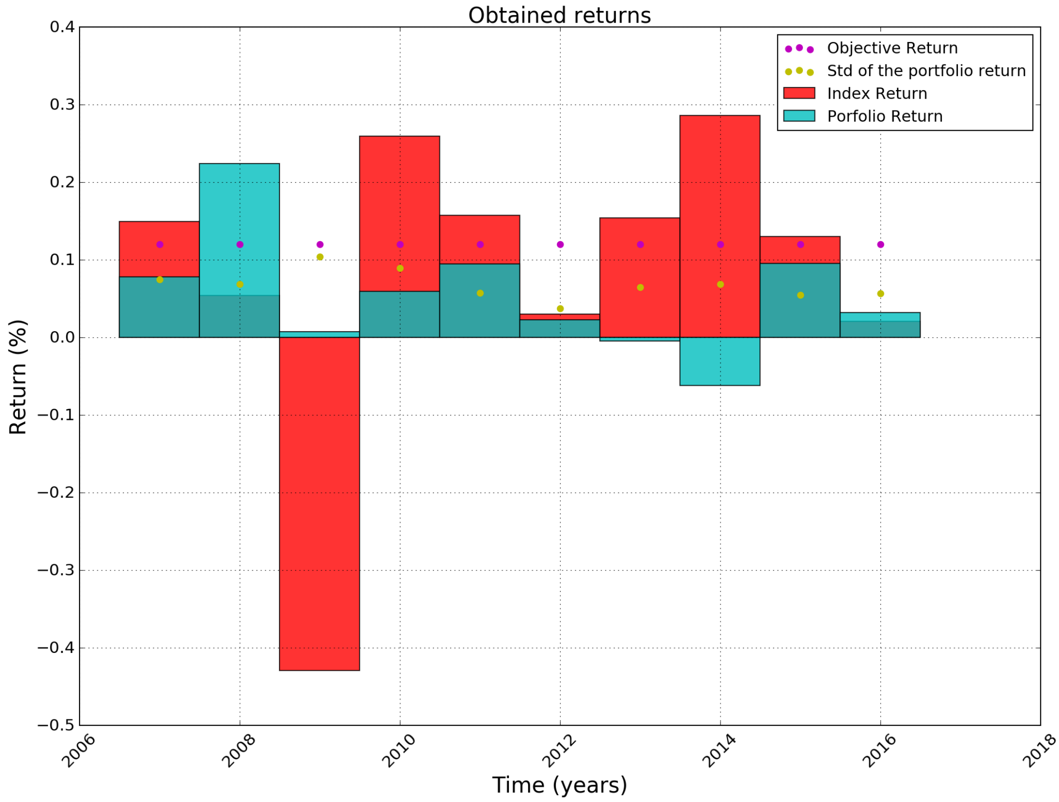

Figure 19: Backtest returns

As it is observed, nothing reliable is obtained, the values of return are within the confidence interval, they could be noise, well, 2008 was great, and we got lucky in the great depresion of 2009, where the index suffered a lot. We do not reach the expected returns but something is something, usually our returns are positive.

We can also try yo visualize what happens with the evolution of the returns during the year we are holding it. The next image shows the cumulative return obtained for the 10 different realizations of the backtest. The x-axis are the months past after holding the portfolio. A pattern observed is that they usually converge at the fifth month.

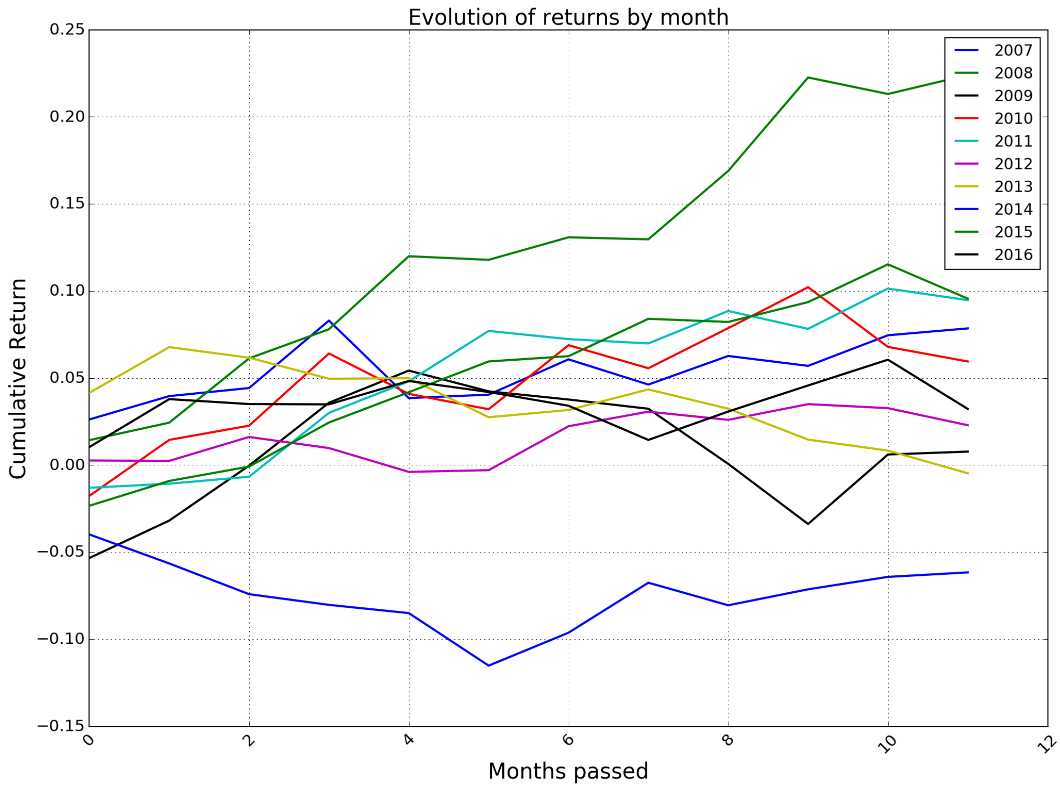

Figure 20: Monthly study of the backtest returns

Maybe we should have picked another point in the efficient frontier to draw the market line, maybe there is an area of the efficient frontier that is more robust to changes. For this purpose we have made the following graph, that shows for some of the test periods, the efficient frontier calculated with the 10 year rolling window, and the efficient frontier calculated with the returns of the test year. Even though it only contains 12 samples, so it is not reliable. In practice, no conclusions were drawn. But we can see that usually, a higher risk original porfolio, translates into a higher risk test portfolio.

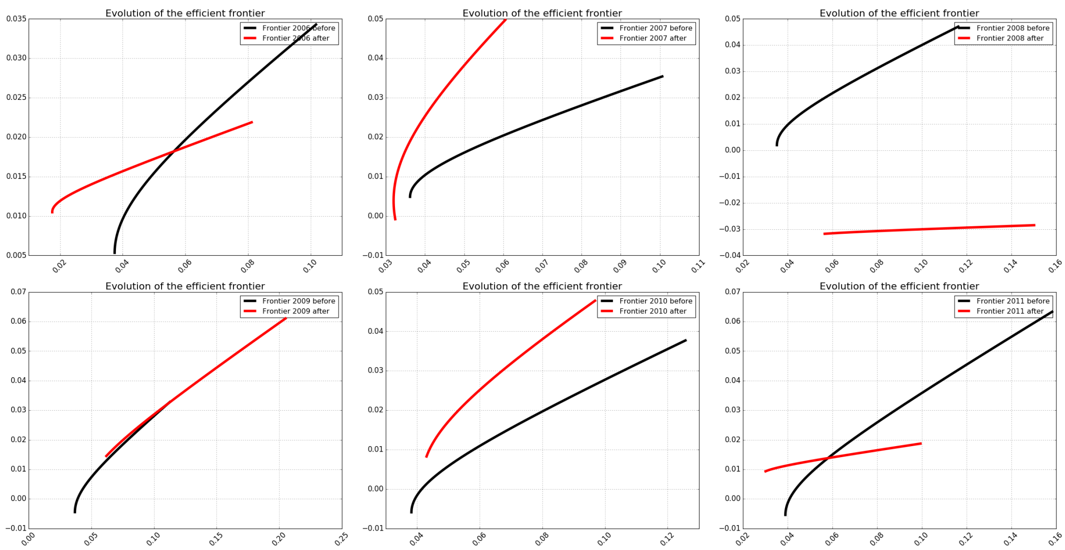

Figure 21: Transformation of the efficient frontier in the backtest

1.7 beta

Use a broad stock index to test, whether our portfolio is in line with the CAPM prediction.

Has your portfolio created "alpha"? Is the "alpha" significant? (see Jensen's alpha)

Solution

Let's see, let us use our 20 years of data and calculate the mean returns, covariance matrix and efficient frontier. Our index to be used is the SPY index of course, since all our symbols have been obtained from that index.

In the CAPM model, we asume that all the our assets can be expressed in terms of the index as a linear regression. It assumes that any symbol and the index form a gaussian bivariate distribution. This model is equivalent to the model in the efficient frontier in that they both assume guassianity of the stocks. The returns of the $i-th$ asset can be expressed as:

$$R_i(t) = \alpha_i(t) + \beta \cdot (R_{market}(t))$$

If there is a risk free return:

$$R_i(t) = R_f + \alpha_i(t) + \beta \cdot (R_{market}(t) - R_f)$$

Where in this report we assume the R_{market} to be the returns of the index. So, in the CAPM model, we reduce every asset to a set of 2 values, alpha and beta, (α_i, β_i), in this case α_i is the mean value of the random vairable $\alpha_i(t)$. This, being the linear regression values between the symbol and the index. There is a relation between the β and the covariance matrix of the symbol and the asset.

$$\beta = \frac{\sigma_{iM}}{\sigma_M^2}$$

where σ_{iM} is the covariance between the symbol i and the market, and σ_M^2 is the variance of the market, in this case, the variance of the index.

The next image shows the (α, β) model and the gaussian model between some of the stocks, and the index.

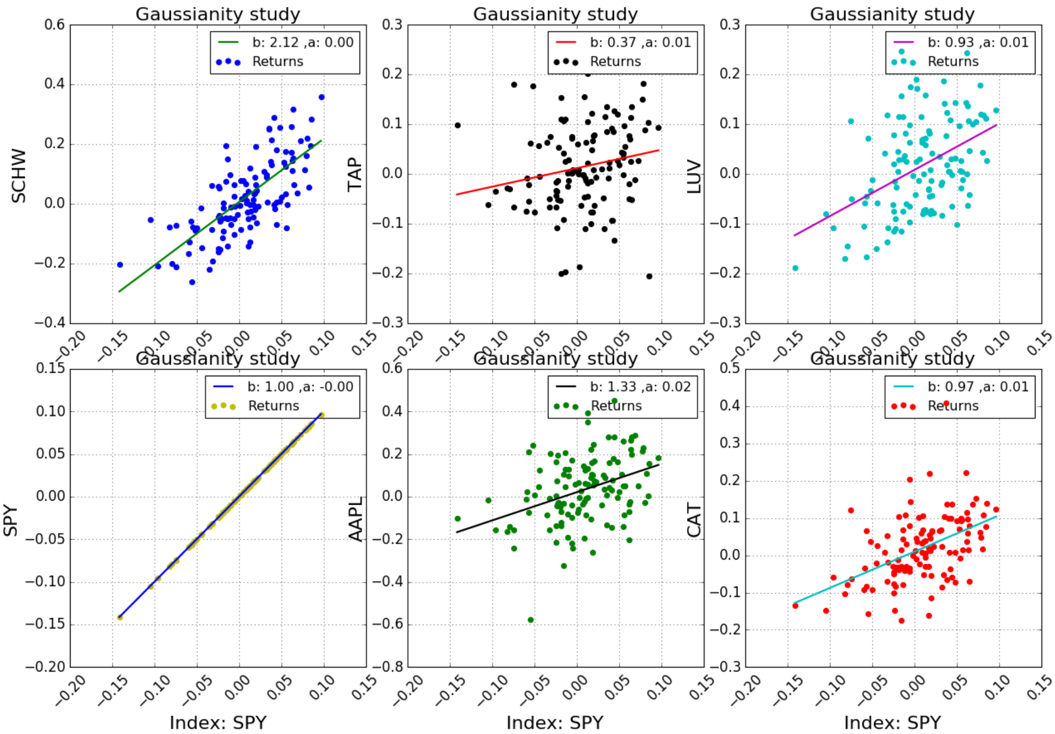

Figure 22: Alpha beta regression of the symbols

And the next image shows the histogram of Jensen's alpha of these symbols, that is their $\alpha_i(t)$, obtained easily from the equation. Geometrically, this is the distance to every point in the scatter plot to the line, plus the mean alpha value since the line does not cross in the origin. This random variable is assumed guassian by the CAPM model but as we can see in the histograms, not so much. The histogram for the SPY, the 4th one, should actually be 0, but it is fun to see the numerical errors python does, they are in the order of 10^{-17}.

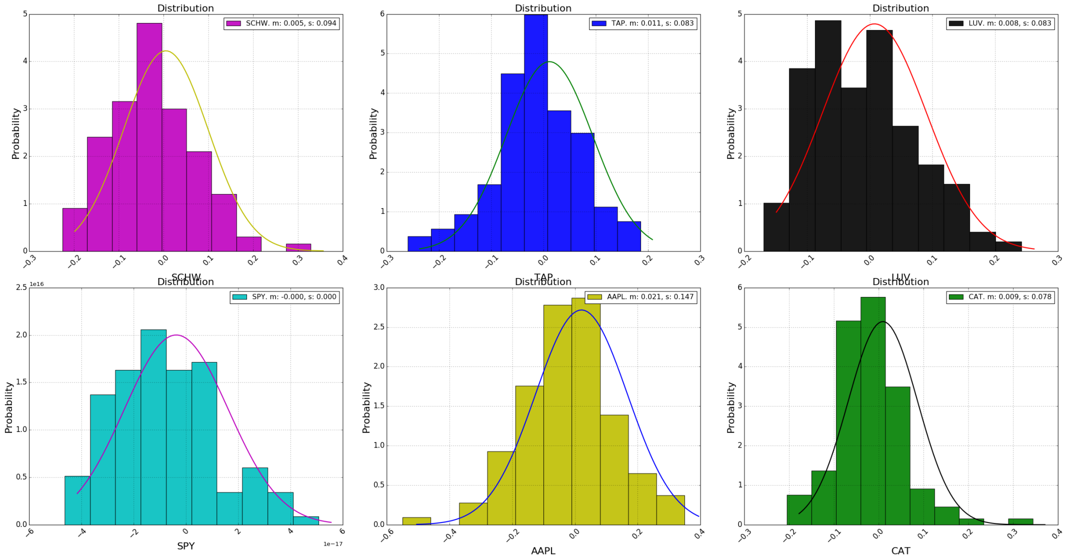

Figure 23: Jensen Alpha of the symbols

The objective of the Efficient Frontier is not directly to maximize α and minimize β (to lower exposure to the market), but to maximize the total return, caused by both α and the market (β), for a given level of risk, caused also for both α and the market (β). So there is no guarantee that the optimal efficient frontier will perform well in terms on this model, although at some level it should.

The next two graphs show the scatter plot of the returns for 6 different random portfolios and the histogram of their Jensen's alpha. As we can see, everything is quite random, good portfolios are not so easy to do. The total β of the portfolio is weighted sum β of its assets, and the same thing goes with the α. The total Jensen's alpha is also the weightes sum of the Jensen's alpha. If the individual Jensen's alpha where uncorrelated we will be able to diversify risk. Also, the total Jensen's alpha will be more gaussian-like due to the Central Limit Theorem, everything tends to be gaussian.

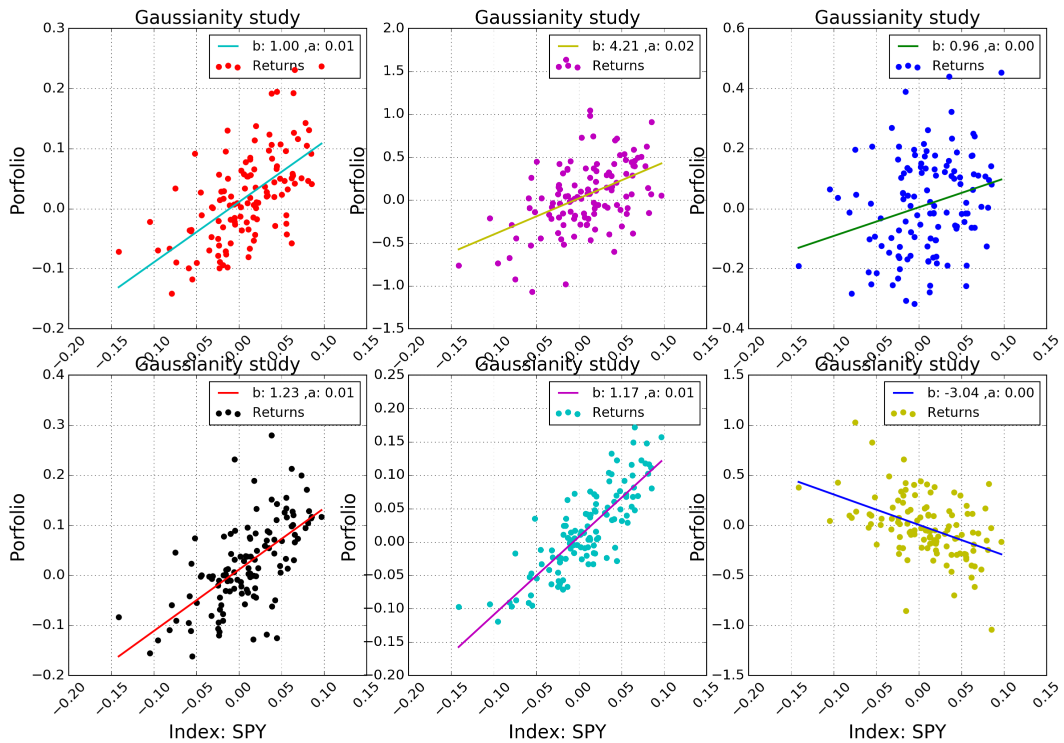

Figure 24: Alpha beta regression of random portfolios

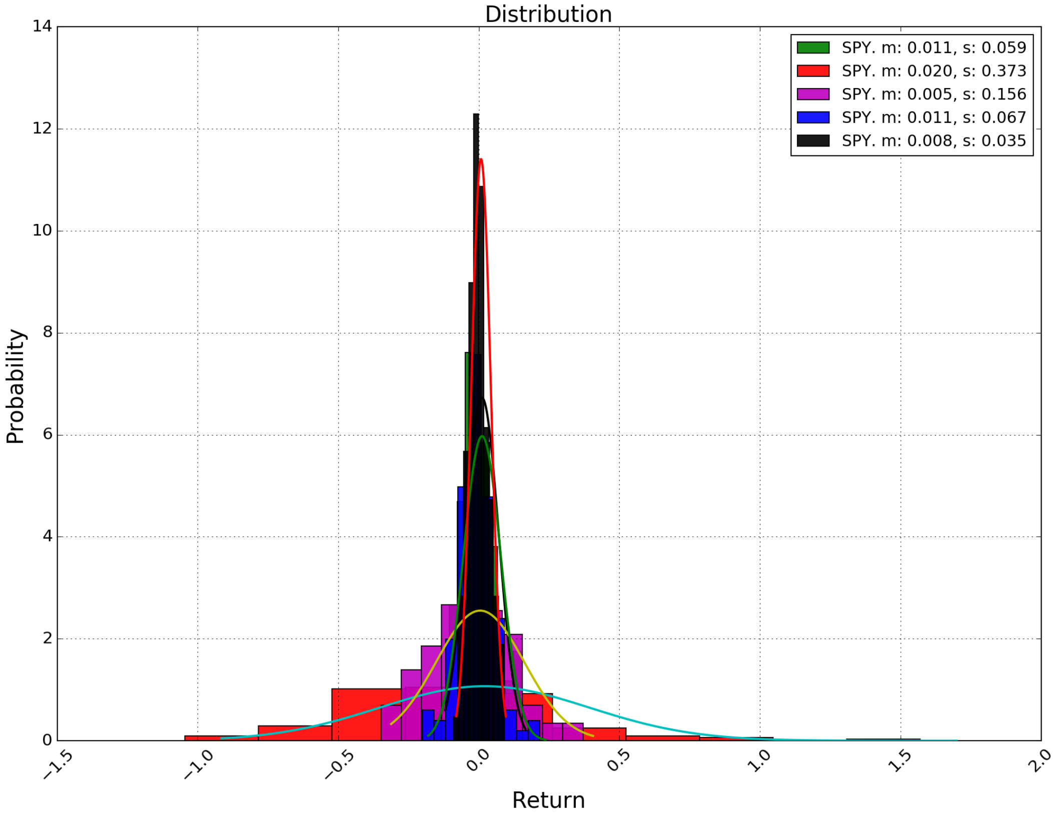

Figure 25: Jensen's alpha of random portfolios

Now we will see the alpha beta regression scatter plot for 3 different portfolios that are in the efficient frontier. As we can see, the bigger the return, the bigger the risk, this was very much expected from the efficient frontier shape. We can also see that, in order to increase the return, the beta is increased, so the returns of the market must be positive. The Jensen's alphas are somewhat gaussian but not really.

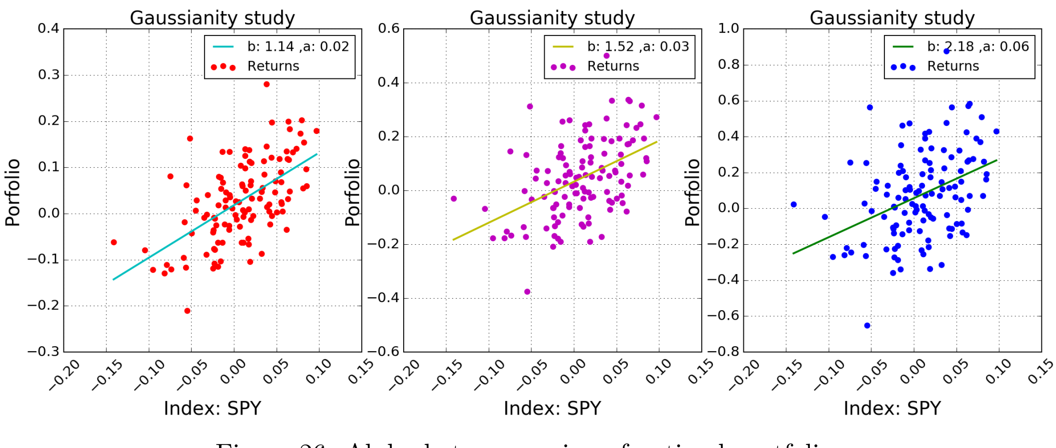

Figure 26: Alpha beta regression of optimal portfolios

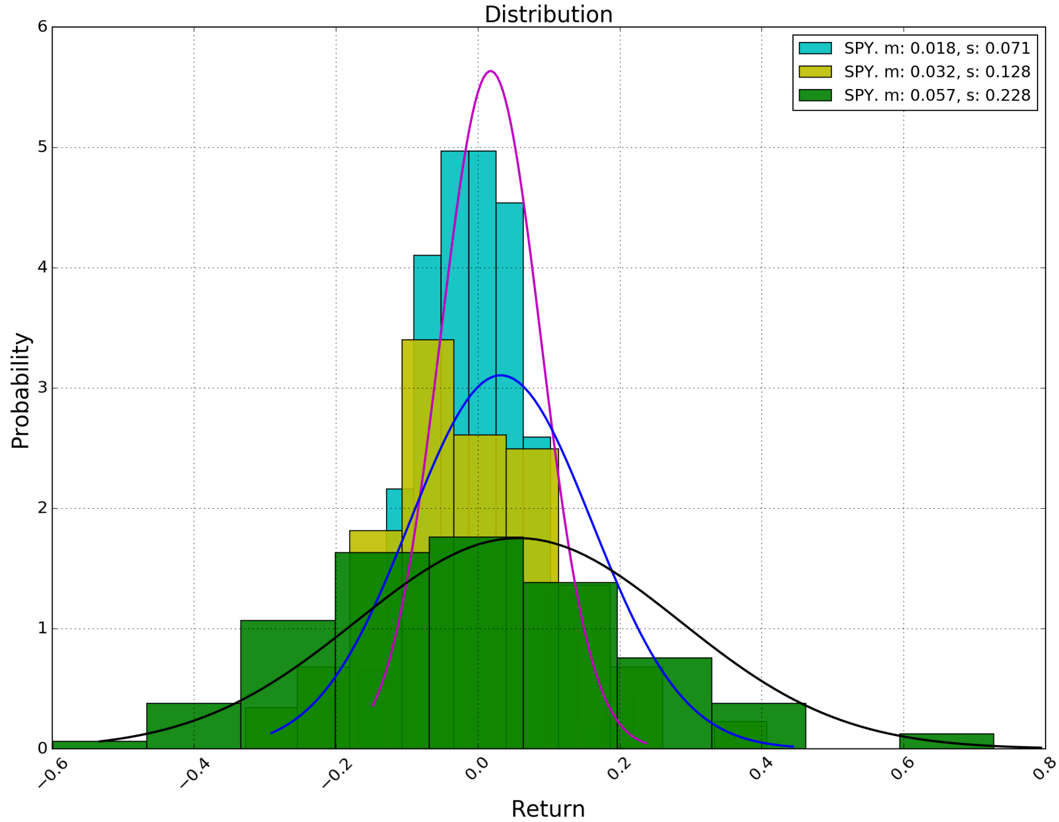

Figure 27: Jensen's alpha of optimal portfolios

We ran a set of statistical tests to show that the distribution is not really gaussian. Also tested if the α is significant, that is, a test to check if the mean of the distribution is non-zero due to chance, or actually we have enough samples, mean and low variance to say that the mean α is significant. The p-value for the means of these 3 portfolios are $0.00635, 0.00678, 0.00740$ respectively. These are the probablities of obtaining a set of samples more "extreme" than the ones we got if we assume the null hypotesis, that the mean of the distribution is zero. Since this probability is very low that means that obtaining this data points if the null hypotesis is true is very unilikely so we can reject the null hypotesis, so we the mean of the distribution is actually non-zero, so these values are significant. Usually the threshold for rejecting the null hypotesis is 0.05. We can also see how the p-value actually increases as the from the lowest return portfolio to the highest return. This is because, even though the mean is higher, the uncertainty grows more than what the mean does. We can also observe that, as always, the negative returns tail is heavier than the positive returns tail.

1.8 Black-Litterman (BL)

How would the asset allocation and your backtest change if you use the model implied expected returns of the Black-Litterman approach?

Solution

So, we have implemented the Black-Litterman model in Python. This BL is a bayesian model in which we can cambine the allocation weights of our initial portfolio and our views of what is going to happen in the market. Instead of completely trusting the porfolio obtained from the historic prices, we will introduce some modifications in the allocation weights according to information that we believe in (prior views) of the form: "The asset A will perform better that asset B by V, and I am Q certain about it".

In this bayesian model, we asume that:

- The returns are jointly gaussian.
- The optimal allocation weights can be obtained from the returns as:

$$w^* = \delta \vec{R} \cdot S^{-1}$$

Where $\vec{R}$ is the vector of mean returns and S^{-1} is the inverse convariance matrix of the returns.

- Our prior views are also gaussian and are based on relative differences between the different assets.

Once we fulfill all of this, we are good to go, for this model we need:

- **A prior portfolio**, called the equilibrium portfolio, composed by the variables $\{w, S, \delta\}$ which are the ones explained before. This previous porfolio could be any, for example we could just use the optimal portfolio given by the efficient frontier, where we can easily obtain the real δ or we can just put another one, like the one based on the relative market capitalization of the assets. That is, the money we should invest in each asset is proportional to the money that it currently has. In this second scenario, we have to come up with risk tolerance (price of risk) δ ourselves.

- **Prior Views**: These are our personal views on how good will a symbol perform compared to other symbols. They are of the type, the returns of $S1$ with be N units bigger than the ones of $S2$ ans I am Q sure about it. Matematically speaking this can be writen as $S1$ - $S2 = N$ where N is Gausian Noise with mean Q and variance Ω given by our confidence on our view. The Gaussian Noise is our prior and has mean given in Q and variance given by Ω (how confident we are in our prior). The matrix P is the selection Matrix, which multiplied by the symbols vector will give the right hand side of the equation $S1$ - $S2$. A important variable is τ, which is the coefficient of uncertainty in the prior estimate of the mean Π, it regulates how uncertain we are about our prior portfolio, which is the same as saying how certain we are about our prior views. It is a parameter to optimize. It is usually 1 or 0.25. If we are sure about our prior views we put a higher τ.

The following will be a qualitative study (done with real values anyway) of the effect of the Black Litterman model on the posterior returns and results. The following image shows the weights and returns obtained for the BlackLitterman model for a subset of the assets. We can see 3 sets of weights and returns:

- The optimal prior weights and returns. These are the ones obtained for the optimal portfolio (using the efficient frontier) for a window of 10 years. In this case we can actually calculate the risk adversion δ from the data.

- The prior weights and returns used in the Black Litterman model. These weights are calculated from the market capitalization of the assets, so they can only be positive. The justification to do this is that the amount of money in an asset is a measure of how condifent people are to invest there. The prior returns Π can be calculated from these weights, the covariance matrix of the returns and a risk adversion δ that we should estimate ourselves.

- The posterios weights and returns used in the Black Litterman model. These weights and returns are calculated by combining our prior views with the Black Litterman equilibrium portfolio. In this case, our prior views were that:

  - The first asset will perform better than the second one by 0.0002
  - The second asset will perform better than the third one by 0.0001

  The prior covariance matrix is obtained from the assets covariance matrix and the certainty of our prior views is $\tau = 0.3$.

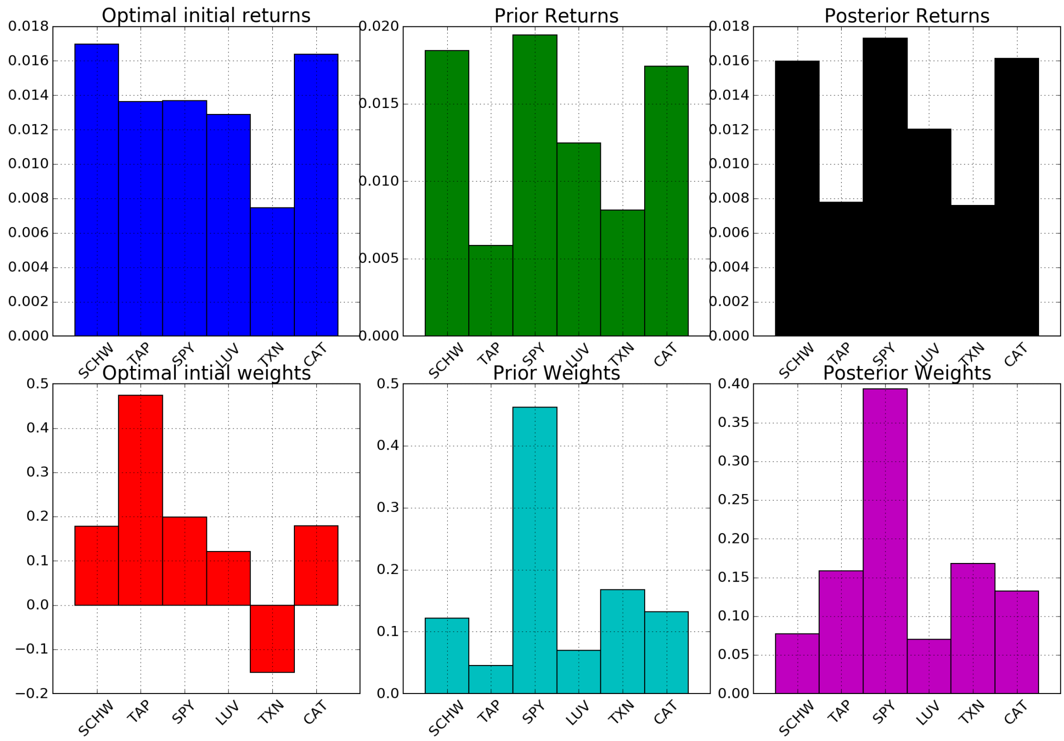

Figure 28: BL weights and returns

Some observations are:

- The BlackLitterman prior gives a lot of weight to the SPY because it has a high market capitalization, but the efficient frontier gives it less weight due to its variance and correlations with other symbols.
- The efficient portfolio prior can have negative values, but the one calculated with the market portfolio cannot.
- We can see how the covariance matrix has shaped the prior returns from the prior weights.
- The 3 first prior returns have been modified in the posterior. The first asset has been lowered because the first prior view assumed a lower differece with the second one. The second asset has increased as well because in the prior we say that it should perform better than the third one, so it is increased. The posterior weights are changed accordingly.
- Having the right model priors does not mean that we will get a higher returns, just returns that are hopefully more likely to happen.

Let's play a little with τ, if we increase τ then we are more sure of our prior views. We set it to a ridiculosly high value $\tau = 10$, this way, the posterior returns will fit our prior views because we are soo sure of our priors. The next image shows this information

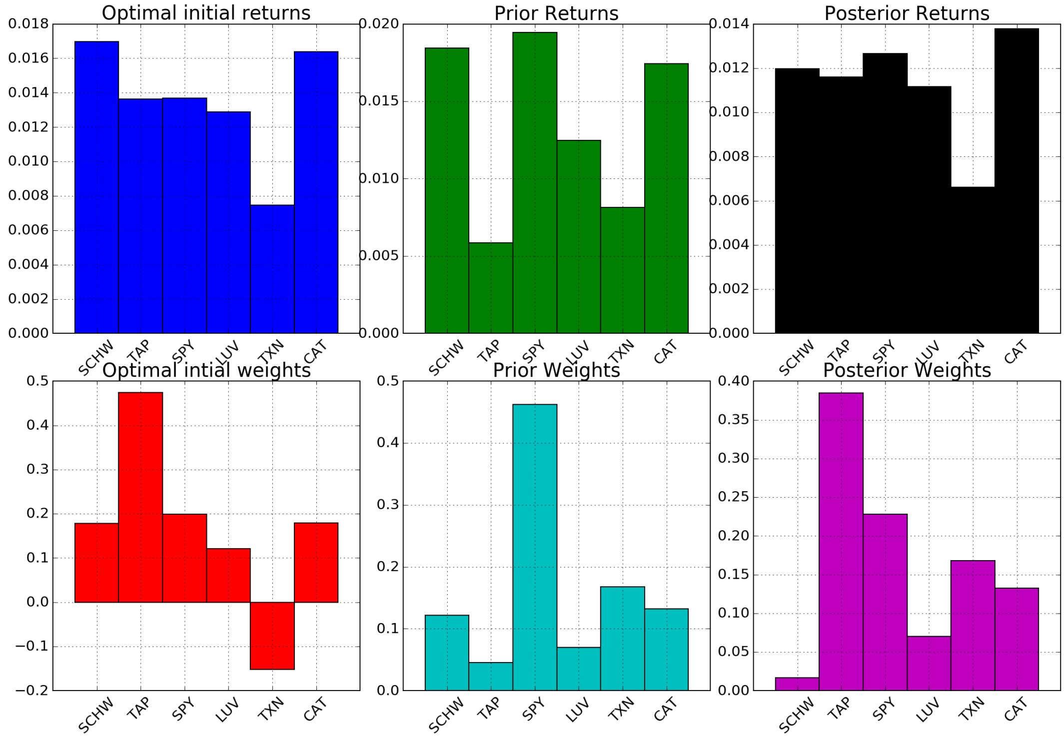

Figure 29: BL weights and returns 2

As we can see, the posterior returns fit now much better the relations that we imposed in our prior views. The posterior weights take into account also the covariance matrix of the returns so drawing conclusion on why they changed like this would require a more deep study.

Maybe it would be interesting to apply the BL using as prior portfolio the one obtained with the market capitalization, and as prior views, the relations between the returns obtained in the optimal portfolio. It would be like mixing long term information and shorter term information. But it is not done in this report because we do not have enough time.

1.9 Timing

Calculate the Treynor-Mazuy measure and give an interpretation to these results (portfolio returns without BL)

Solution

The Treynor-Mazuy measure is a way to see if we have been clever in managing our portfolio. In the CAPM model, if we do not change our portfolio during the time, our portfolio will be expresed as a linear regression on the index returns (alpha, beta). Being these values the weighted linear combination of the values of our assets in the portfolio.

If we have been clever, then we moved our portolio around respect to β, so:

- If the returns of market (index) are going to be high, we should have a portfolio with a high β so that we earn more than the market.
- If the returns of market (index) are going to be low (negative), we should have a small β so that we do not lose money as the market is doing. Or even we can have negative β to earn money when the market is going down.

So, in the scatter plot of index-returns, to market-returns, instead of fitting a straigh line (given by α, β), we could fit a curve, a quadratic curve for example, and if it is convex (positive quadratic term) we did good, and if it concave, maybe we should study philosophy instead.

As an initial example, let's calculate the Treynor-Mazuy for the return of the optimal portfolio for the 20 years, with no management, just use the optimal fixed portfolio during the 20 years. We obtain the following graph. As we can see, the cuadratic component is almost non-existing and we just have a line fitting the regression given by the CAPM model.

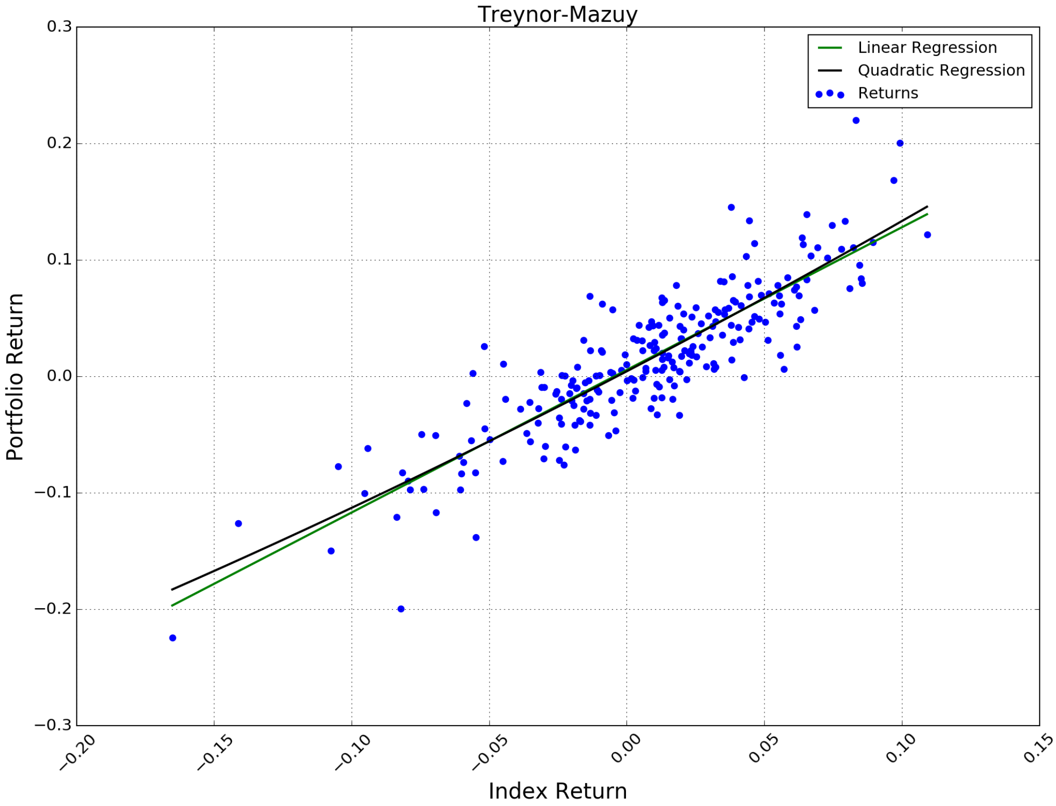

Figure 30: Treynor-Mazuy of the whole dataset a posteriori

Now, instead of calculting the optial portfolio with the 20 years and then obtain the returns of them. Let's only use the returns obtained in the backtest where we compute the optimal portfolio with a rolling window of 10 years and then hold it and compute the returns for the next year. In this case we only hace 12*10 = 120 points. As we can see in the next figure, surprinsingly enough, it did not work. We barely have a beta or alpha, just randomness.

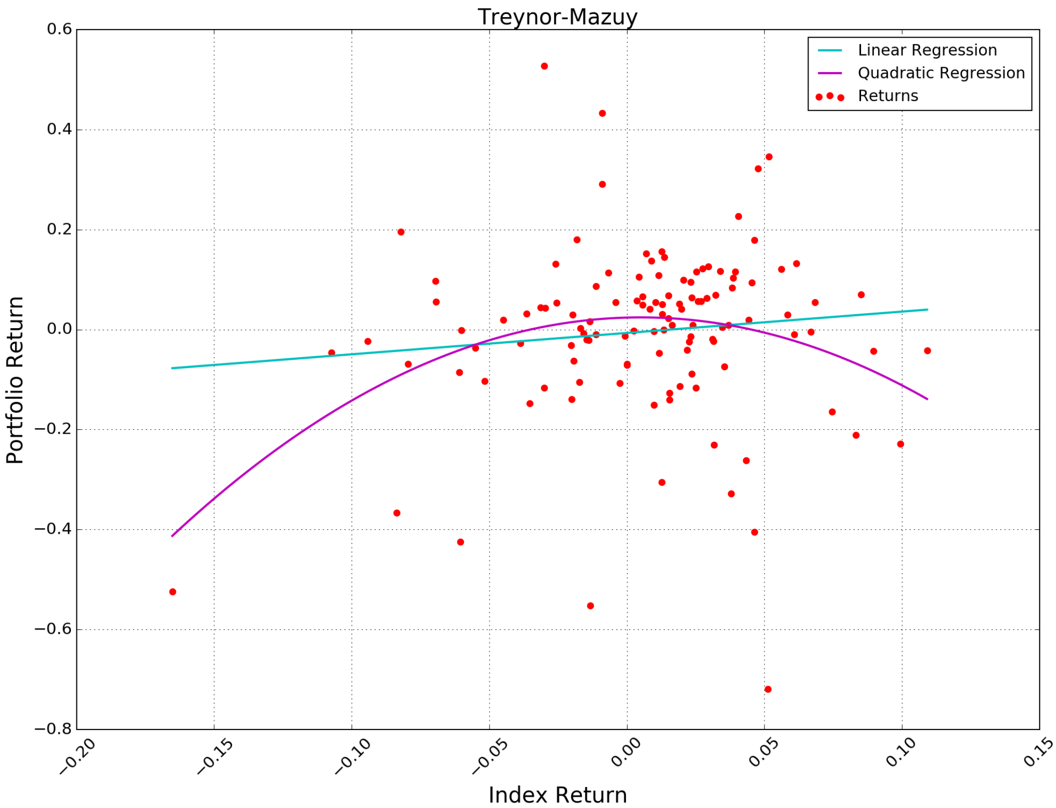

Figure 31: Treynor-Mazuy of Backtest

2 Bonds

2.1 a

Use five different "real" bonds and calculate for these bonds the yield to maturity, duration and convexity.

Solution

A bond is a financial instrument that, once bought, yields a given amount of money c a number f of times per year, plus a final amount of money p at the end of the bond's lifespan. They rely in the fact that, the value of money decreses with time, due to inflation, the value of a dolar today will be only worth $\frac{1}{(1+r_t)}$ the next year, where r_t is the inflation rate. So, every time the bonds yields money, that money is less valuable than our initial invesment due to inflation.

The proper weighted price of a bond taking into account the depretiation of money is:

$$P = \sum_{i=1}^{N} \frac{C(i)}{(1 - y/f)^i}$$

where: P is the price of the bond, $C(i)$ is the money that it yields the $i - th$ time, f is the frequency of the bond, the number of times per year that yields money. Most importantly we have the y or yield to maturity. It is the yearly "inflation rate" assumed by the bond, it is usually higher than the real one, because, who would invest in it otherwise ? The time when the bond expires is called maturity and the time until that happens is called time to maturity T, it is usually express in terms of how many periods of frequency f are left for the bond to die, that is the number N in the equation.

The yield to maturity y that entities apply to the bonds they sell usually increases with the time to maturity, since the more time you hold the bond, the more risk you are exposed to. As time passes, the yield to maturity y of a bond decreases due to this. Also, the higher y, the less the price of the bond, that means, if we assume that the inflation is higher, the money that we will get in the future will be worth less and so the total price decreases.

Usually, a bond yields a fixed amount of money c called cupon, at the given yearly frequency f and it yields a higher amount at its death, called par value or face value p. These are the type of bonds we implemented in out library. The equation of these kind of bonds is:

$$P = \sum_{i=1}^{N} \frac{C(t)}{(1 - y/f)^i} = \sum_{i=1}^{N-1} \frac{c}{(1 - y/f)^i} + \frac{c + p}{(1 - y/f)^N}$$

Given the structure of the bond (how it pays the cupons,face value and time to maturity $\{f, c, p, N\}$), now we can calculate the price P given the yield to maturity in an easy way, just apply the formula. We can also obtain the yield to maturity y if we have the price, but for this we have to solve a polynomial formula of very high degree, usually a function solver is used, in our case, we use the Newton's method to calculate it in Python.

The yearly frequency of return f is important if we take into account the "compound return" of an investing. This is the return we would obtain in an investing if we reinvested the return each time it is given to us. If we received the returns in a continous way, we would reach and exponential boundary, called the spot rate. We have generated the following graph to express this point. It plots the total return obtained for 50 years, for different frequencies of yielding (and thus, reinvesting), when we have a yearly return of 5%.

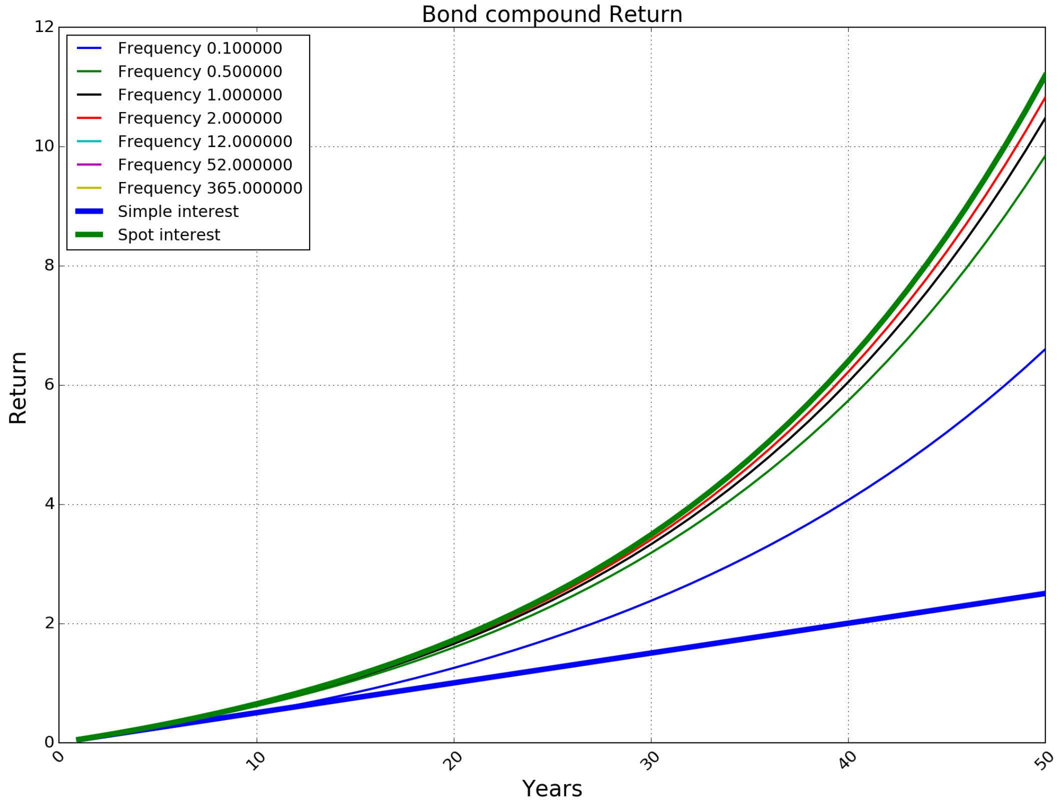

Figure 32: Compound return

Now we go for the modified duration and convexity, these are fancy names for very simple concept. As we said earlier, the price of a bond P decreases as the yield to maturity y increases, intuitively because the money we will get in the future will be worth less, and mathematically it can easily be seen in the equation. The modified duration and convexity are values that chatecterize the curse and can be used to estimate it in a straight forward manner.

Before going into them, we should know that we can estimate any function around a point P using a Taylor series aproximation of order n, for this we only need the derivatives up to order n of the function at the point P and we can approximate it as:

$$\hat{f}(t+\delta t) = f(t) + \frac{1}{1!}\frac{df(t)}{dt}(\delta y) + \frac{1}{2!}\frac{df(t)^2}{dt^2}(\delta y)^2 + \frac{1}{3!}\frac{df(t)^3}{dt^3}(\delta y)^3 + ...$$

The modified duration D is the first derivative of the yield curve $P(y)$ normalized by the price. And then negated because the curve is decreasing as the yield increases and we want a positive value to show to our clients.

$$D(y) = -\frac{dP(y)}{dy}\frac{1}{P(y)}$$

The convexity is the second derivative of the yield curve $P(y)$ normalized by the price.

$$C(y) = \frac{dP(y)^2}{dy^2}\frac{1}{P(y)}$$

So we can approximate the price of a bond given the modified duration and convexity as:

$$\hat{P}(y+\delta y) = P(y) + P(y)[-D(\delta y) + \frac{1}{2}C(\delta y)^2]$$

Notice that we multiply the modified duration and convexity by P(y) so that we transform them back to the derivatives. As we have seen, the D and C can be interpreted as sensibility measures that tell us how much a change in yield δ will affect the price.

The next graph is an example of a bond for which we calculated the yield curve, the modified duration, the convexity and the absolute error of the estimation using the duration and convexity.

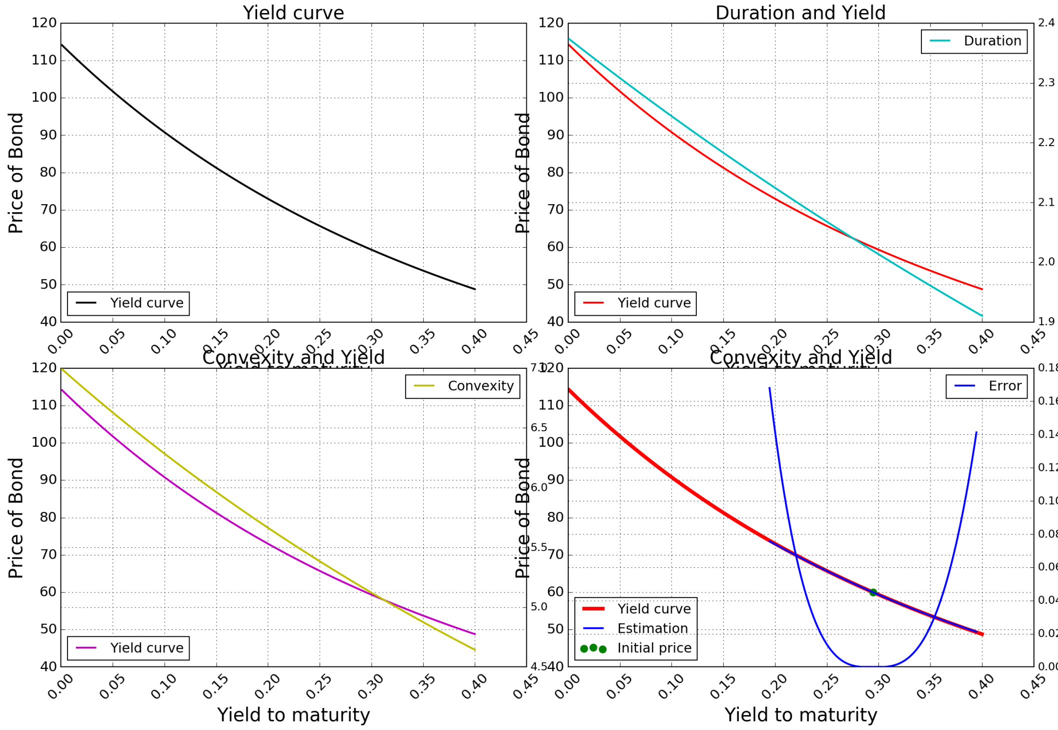

Figure 33: Yield curve, duration, convexity and estimation error

Things we can observe from the graph are:

- The duration decreases as the yield increases. That means that the rate of change of price with respect to the yield decreases, so we are less exposed (less sensitivity).
- The convexity decreases as the yield increases. That means that the rate of change of the rate of change of price with respect to the yield becomes decreases as the yield increases, so we are less exposed less quickly as the yield increases (less sensitivity).
- The aproximation error is very accurate because the function can very much be approximated with an order 2 polinomial. We can see how the error grows faster in left side, as the rate of change of the price is higher, when the yield is lower.

Finally, answering the initial question, the modified duration of our bonds, obtained from screener yahoo finance ranges between 3.0 and 2.2. The convexity ranges between 6.5 and 8.1. The yield to maturity ranges between 2.1 and 3.0.

2.2 b

Calculate the duration and convexity of a portfolio of these bonds, if EUR 100.000,– is invested in each of them.

Solution

So, if we invest 100.000 in each of them, that means we have a portfolio with N bonds whose total price is:

$$P_p = \sum_{i=1}^{N} Pinv_i = Pinv_1 + Pinv_2 + Pinv_3 + ...$$

Where $Pinv_i$ is the total amount of money invested in the bond i, in this case 100.000, if we can only buy a discrete number of bonds, at a price P_i, for the $i-th$ bond, then, the price of our portfolio will be expressed as:

$$P_p = \sum_{i=1}^{N} k_i \cdot P_i = k_1 \cdot P_1 + k_2 \cdot P_2 + ...$$

Where k_i is the number of bonds of type i that we are buying.

Since the modified duration and convexity are linear operators (derivatives) divided by the price of the portfolio, then we can calculate them in linear terms of the modified duration and convexity of the individual bonds. The duration of the portfolio is:

$$D(P_p) = -\frac{dP_p}{dy}\frac{1}{P_p} = -\frac{1}{P_p}\frac{d}{dy}\left(\sum_{i=1}^{N} k_i \cdot P_i\right) = -\frac{1}{P_p}\Big[k_1 \cdot \frac{d}{dy}(P_1) + k_2 \cdot \frac{d}{dy}(P_2) + ... = -\frac{1}{P_p}\Big[k_1 \cdot \frac{d}{dy}(P_1) + k_2 \cdot \frac{d}{dy}(P_2) + ...$$

We replace the derivative using the modified duration equation for the individual bonds, obtaining.

$$D(P_p) = \frac{1}{P_p}\Big[k_1 \cdot D(P_1) \cdot P_1 + k_2 \cdot D(P_2) \cdot P_2 + ...$$

The same reasoning can be applied to the convexity, obtaining:

$$C(P_p) = C\left(\sum_{i=1}^{N} k_i \cdot P_i\right) = \frac{1}{P_p}\Big[k_1 \cdot C(P_1) \cdot P_1 + k_2 \cdot C(P_2) \cdot P_2 + ...$$

And with these, we can approximate the yield curve of our portfolio around a fixed point.

The duration of the portfolio is 2.4 and the convexity is 6.9.

2.3 c

Estimate the potential decline in the market value of your portfolio, if the yield increases by 150 basis points.

Solution

As we have seen, we can make an estimation of the yield-price curve using the modified duration and convexity in the polinomial Taylor series approximation of degree 2.

$$\hat{P}_p(y + \delta y) = P_p(y) + P_p(y)[-D(P_p(y))(\delta y) + \frac{1}{2}C(P_p(y))(\delta y)^2]$$

The estimated decline is 4.1 % of the original price, it declines more than 20.000 EUR.

3 Python code

This assignment has been entirely coded in Python by the authors of this document. For this purpose, different libraries has been developed for obtaining and mantaining the data, performing mathematical operations, and plotting graphs. Then the assignment is solved with such a library, of course it was like 90% of time development of the library and then 10% of time using it, and it can do several more things than the ones exposed in this report, but we do not have as much time to include them, mainly in the plotting of several graphs together, but we did not have as much time. There is of course a lot of room for improvement, The main python document to use is called *Final_Assignment.py* in which all the libraries are loaded, the dataset is loaded and initialized and some trials of the library are done. To execute the individual parts of this assignment, one should scroll down to the end of the file and uncomment the corresponding answer to visualize.