Using 1000 randomly selected subperiods of 500 days, we need to show the stability/instability of the weights of the optimal portfolio when different estimates of the parameters mu and sigma are used i.e the estimate of the subsample or the estimate of the entire population (around 2000 days).

Here we limited our boxplots to the first 20 stocks in order to give the reader a clearer view of the results, it has to be stressed that our calculations aren’t just limited to 20 stocks but they include all the stocks composing the S&P500.

What we learned while doing this exercise is the importance of the number of days of observations we need in order to have a stable (here meaning invertible) covariance matrix that allows us to solve the Markowitz algorithm without excessively amplifying the magnitude of the optimal weights. An efficient and time-saving method to calculate the covariance matrix while using a computer is to use eigenvalues and eigenvectors to apply the spectral decomposition theorem. Applying the spectral theorem means building two matrices:

* The first one is a diagonal matrix with eigenvalues on the diagonal and zeros elsewhere
* The second one is built combining eigenvectors by column

Notice that the inverse of a diagonal matrix with the same allure of the first one just described is the same matrix but with coefficients on the diagonal in the form of 1/cij where cii denotes the coefficients of the original matrix. When cii are close to zero 1/cii tends to infinity and the correctness/stability of any linear transformation resulting from this matrix collapses.

When too few observations are used to calculate the covariance matrix, the risk of having very small coefficient on the diagonal grows. The unwanted result reflected on weights instability while applying the Markowitz optimization.

In the first series of graphs we’re interested in determining whether the optimal weights variability is due to the sensibility of the estimate of mu or sigma.

In the first boxplot we determine the optimal weights using both mu and sigma estimates of the same subsample.

In the second boxplot we determine the optimal weights using the estimate of mu of the same subsample and the estimate of sigma of the entire sample.

In the third boxplot we determine the optimal weights using the estimate of sigma of the same subsample and the estimate of sigma of the entire sample.

A screenshot of a cell phone

Description automatically generatedAs a result we see that weights instability and variability is primarily due to the estimation of sigma.

To contain instability of the weights we have to work on the covariance matrix in order to ”artificially” make it always useful and stable i.e invertible. One technique that allows us to always come up with an invertible matrix is known as matrix shrinkage. It works by applying a convex combination of a definite positive matrix with a positive semi definite matrix, the result is a definite positive matrix which, by fact, is invertible. Here sigma is going to play the role of our definite semi positive matrix while F is going to be a definite one, namely:

* The identity
* The constant correlation matrix

We then apply an arbitrary weight (here lambda=0,2) to the transformation and we obtain a stable matrix.

In the first case we use the identity matrix and lambda=0,2. We transform our estimate of sigma and recompute the same problem as before in order to find the optimal weights. We determine the optimal weights using both mu and sigma estimates of the same subsample but we use the shrinked covariance matrix. The left boxplot show our result and the one on the left is the same as exercise one while using the estimates of mu and lambda of the subsample (Exercise 1). As we clearly see, the effect of a bad estimate of sigma is drastically reduced as wanted.

A screenshot of a cell phone

Description automatically generated

In the second case we use the constant correlation matrix and lambda=0,2. We transform our estimate of sigma and recompute the same problem as before in order to find the optimal weights. The left boxplot show our result and the one on the left is the same as exercise one while using the estimates of mu and lambda of the subsample (Exercise 1). As we clearly see, the effect of a bad estimate of sigma is drastically reduced as wanted.

A screenshot of a cell phone

Description automatically generated

The obvious practical problem is which value to choose for the shrinkage constant. Any choice of δ strictly between 0 and 1 would yield a compromise between S and F. But this results in infinitely many possibilities. Intuitively, there is an ‘optimal’ shrinkage constant. It is the one that minimizes the expected distance between the shrinkage estimator and the true covariance matrix.

Another limit of shrinkage while using the constant correlation matrix is that the model would be inappropriate if assets came from different asset classes, such as stocks and bonds.