

## PORTFOLIO THEORY – LECTURE NOTES 3

Dr. Andrea Rigamonti

### IMPROVING PORTFOLIO OPTIMIZATION

Let us consider again the portfolio optimization problem. Whatever metrics we will use to evaluate the results, we need to compare them with appropriate **benchmarks** in order to know if the results are satisfying. A benchmark portfolio surprisingly difficult to beat is the one created with a **naïve 1/N rule**.<sup>1</sup> This portfolio simply assigns equal weights to all the assets in all periods, and therefore does not require to estimate any input and to perform any optimization procedure. This is indeed one of its main strengths: it is completely immune to estimation errors. Another advantage is that it has a very low turnover, which translates into very low transaction costs.

Another possible benchmark is the **market**, or more precisely the returns of a large stock market index, like the S&P 500. While it is not possible to buy an index, it is possible to buy ETFs. An **ETF (Exchange-Traded Fund)** is a fund that is traded on the financial markets and which tries to replicate a certain index. By investing in such fund, the investor can invest in a certain index without having to trade all the stocks contained in such index. Stock markets sometimes go through periods of poor performance that can last years, but over the long run they provide good returns (in countries with a solid economy). Therefore, such passive investing solution, which also does not require to perform estimation and optimization procedures, is another reasonable benchmark (over long enough periods of time).

Basic mean-variance optimization with sample estimates often struggles to beat common benchmarks, especially the naïve rule mentioned above. To improve optimization performance we need to reduce the severity of the estimation errors in the inputs, and/or reduce the impact of such errors on portfolio formation. Plenty of solutions have been proposed. We focus on some of them that proved to be effective and not too difficult to apply.

On the side of parameter estimation, we consider techniques developed to improve the estimation of the covariance matrix. While in general the error in the vector of sample means is more severe than the error in the sample covariance matrix, solutions proposed to improve over the latter are simpler and more effective. Moreover, while the number of parameters in  $\boldsymbol{\mu}$  is equal to the number of assets  $N$ , the number of parameters in  $\boldsymbol{\Sigma}$  is equal to  $N^2$ . Therefore, as the portfolio gets larger, estimation error in  $\boldsymbol{\Sigma}$  becomes worse much faster than in  $\boldsymbol{\mu}$ , which might jeopardize the performance benefit that we theoretically get from a larger  $N$  thanks to the greater diversification potential.

The approach arguably more successful in dealing with the estimation of the covariance matrix is given by **shrinkage estimators**. Ledoit & Wolf (2004) define the shrinkage estimator

$$\boldsymbol{\Sigma}_{LW} = \delta \mathbf{I} \mu_{\delta} + (1 - \delta) \boldsymbol{\Sigma}$$

where  $\boldsymbol{\Sigma}$  is the usual sample covariance matrix,  $\mathbf{I}$  is the identity matrix,  $\mu_{\delta}$  is the average sample variance of all the variables (so the product  $\mathbf{I} \mu_{\delta}$  gives a diagonal matrix whose elements on the diagonal are equal to the average sample variance and all the other elements are equal to zero), and  $\delta$  is the shrinkage parameter whose value is between 0 and 1.

It is called “shrinkage” estimator because the sample estimate is shrunk toward a target matrix.  $\boldsymbol{\Sigma}_{LW}$  is basically the result of a weighted average between the sample covariance matrix and the target matrix. The intensity of the shrinkage is controlled by the shrinkage parameter  $\delta$ . Ledoit and Wolf

---

<sup>1</sup> We will refer to this rule/portfolio as “naïve” or “naïve 1/N”, to not confuse it with the 1/N rule we described before. However, it is also commonly called simply “1/N” rule/portfolio in the literature.

(2004) select an optimal value for  $\delta$  with a procedure whose details are beyond our scope here. This shrunk estimator improves over the sample estimator, which translate in better performance of the mean-variance and minimum variance portfolios (and portfolios computed using the covariance matrix as input in general). Moreover, it gives a nonsingular covariance matrix even when the number of periods  $T$  used for estimation is smaller than the number of assets  $N$ . The sample covariance matrix, on the contrary, in such case is singular and therefore not invertible, which makes it impossible to perform the optimization procedures. It is also very easy to use this estimator in common statistical environments like R. Hence, it is now standard to use this estimator instead of the sample covariance in many applications, including portfolio optimization.

We now turn to solutions that mitigate the impact of estimation errors on portfolio formation. One of such solutions relies on a logic similar to the one we just described, and consists in computing **shrinkage portfolios**. The idea, proposed by Tu and Zhou (2011), is that we can improve over the mean-variance portfolio by shrinking it toward the naïve 1/N portfolio. In this way we can combine a portfolio that optimizes the weights but suffers from estimation errors with one that is not optimized but is also immune from estimation errors, obtaining a portfolio that improves over both. The weights of such portfolio are given by

$$\mathbf{w}^* = \delta \mathbf{w}_{NAIVE} + (1 - \delta) \mathbf{w}$$

where  $\mathbf{w}_{NAIVE}$  is the vector of weights of the naïve 1/N portfolio,  $\mathbf{w}$  is the vector of weights of the optimized portfolio (usually the mean-variance portfolio, but can be also other portfolios), and  $\delta$  is again the parameter that controls the shrinkage intensity. The value of  $\delta$  can be chosen using optimization rules, heuristics, or cross-validation.

Another solution to improve over the standard mean-variance portfolio is the **grouping strategy** proposed by Branger et al. (2019). The idea is that since the performance of mean-variance (or minimum variance) optimization suffers more and more as  $N$  increases, due to the growing number of parameters to estimate, one could achieve better performance by grouping together the assets in a certain number of groups. The optimization procedure is then performed between the groups, while within a group the assets are equally weighted. This reduces the dimension of the problem, and therefore the number of parameters to estimate. In other words, it is another strategy that combines the benefits of the naïve 1/N rule (which is applied within groups) and of optimization (which is applied between groups). The stocks can be grouped according to how similar they are in terms of estimated mean, variance or beta, and the number of groups can be chosen using optimization rules, heuristics, or cross-validation. The higher the number of groups, the closer we get to the usual optimization; the smaller the number of groups, the closer we get to the naïve 1/N rule. In the extreme case where we only have one group we obtain the naïve 1/N portfolio; in the extreme case where the number of groups is equal to  $N$ , we get the usual optimized portfolio.