Acta Univ. Sapientiae, Economics and Business, **10** (2022) 133–144

DOI 10.2478/auseb-2022-0008

Asset Allocation Strategies Using Covariance Matrix Estimators

László PÁL
Sapientia Hungarian University of Transylvania (Cluj-Napoca, Romania),
Department of Economic Sciences
e-mail: pallaszlo@uni.sapientia.ro

Abstract. The covariance matrix is an important element of many asset allocation strategies. The widely used sample covariance matrix estimator is unstable especially when the number of time observations is small and the number of assets is large or when high-dimensional data is involved in the computation. In this study, we focus on the most important estimators that are applied on a group of Markowitz-type strategies and also on a recently introduced method based on hierarchical tree clustering. The performance tests of the portfolio strategies using different covariance matrix estimators rely on the out-of-sample characteristics of synthetic and real stock data.

Keywords: portfolio optimization, covariance matrix estimators

JEL Classification: G11, C61

1. Introduction

Portfolio optimization is the process of selecting the best possible allocation among certain assets (e.g. individual stocks, asset classes, bonds, cash) according to some specific objective such as risk minimization, return, risk-adjusted return, diversification maximization, and so on. Asset allocation is generally a challenging task due to a lot of factors that can influence the results. The outcome may be influenced by the investment period and asset universe, by the risk tolerance of the investor, and so on.

Henry Markowitz (1952) revolutionized the portfolio optimization by considering the expected return and standard deviation as the key components for quantifying an asset return and risk. This theory formulates portfolio construction as a quadratic optimization problem, where the goal is to maximize return for risk, or equivalently, minimize risk for a given level of expected return. Although the Markowitz model is theoretically sound and has a major impact on portfolio research, its application in real-life situations is challenging. Practically, the researcher or the portfolio manager should estimate the unknown expected value and variance of the returns of the securities in order to apply them in the model. The risk and return are usually calculated inaccurately by using the historical sample, leading to unacceptable solutions with worse out-of-sample performance. Therefore, the portfolios hold a small number of stocks with extreme weights, thus making these portfolios less diversified. Furthermore, small changes in the inputs often cause substantially modified weights in the optimal portfolio. Based on the Markowitz model, in DeMiguel et al. (2009), the constraint variant of the model was investigated. In Scutellà and Recchia (2013), a slightly different approach was proposed based on robust optimization.

Tremendous efforts have been devoted by the researchers to handle estimation errors to improve the performance of the Markowitz model. Considering the estimation errors, the literature on the portfolio selection problem has been extended in several directions. Jorin (1986) proposed shrinkage estimator for expected returns. This kind of estimator usually shrinks the sample estimate towards some average value. Ledoit and Wolf (2003, 2004) presented a transformation procedure of the empirical covariance matrix called shrinkage. Another approach to reducing the risk estimate is to denoise (Bun, Bouchaud, and Potters 2017; López de Prado, 2020) the sample covariance matrix. The denoising procedure eliminates those eigenvalues of the covariance matrix that are affected by noise. Michaud (1998) tries to overcome the uncertainty associated with the estimated parameters using a technique called resampling. A detailed comparison of the Michaud and Markowitz strategies can be followed in Becker, Gürtler, and Hibbeln (2015). The subset resampling (Gillen, 2016; Shen and Wang, 2017) method tries to improve Michaud's algorithm by sampling subset-size portfolios. Recently, the Hierarchical Risk Parity (HRP) by López Prado (2017) has received substantial attention. The method calculates the covariance matrix by hierarchical clustering and avoids the matrix inversion procedure.

In this paper, we analyse the effects of covariance matrix estimators that are applied on a group of Markowitz-type strategies and a recently introduced method based on hierarchical tree clustering. The performance tests of the portfolio strategies using different estimators rely on the out-of-sample characteristics of synthetic and real stock data.

The paper is structured as follows. In Section 2, we review the mean-variance model and its special variant, the minimum variance optimization model. Furthermore, we describe the covariance matrix estimators examined in this study and the reference strategies as well. In Section 3, we conduct two experiments: at first, the three Markowitz-type algorithms are assessed using a Monte Carlo experiment with synthetic data, while in the second experiment all the methods are compared based on the S&P 100 stocks data. Finally, Section 5 concludes.

2. The Mean-Variance Optimization Model

Markowitz formulated his portfolio allocation model as a quadratic optimization problem called mean-variance optimization (MVO). According to this formulation for each level of expected return, the portfolio with the smallest variance is preferable. If the return constraints are omitted from the previous model, we get the global minimum variance or simply the minimum variance portfolio (MV). This approach is popular among researchers since the covariance matrix estimation induces a smaller estimation error than the return estimation.

For a given N risky assets, the minimum variance strategy can be defined as:

$$\min_{\mathbf{w} \geq \mathbf{0}} \mathbf{w}^{\mathrm{T}} \sum \boldsymbol{w} \text{ subject to } \boldsymbol{w}^{\mathrm{T}} 1 = 1,$$

where $\mathbf{w} = (w_{t1}, w_{t2}, \ldots, w_{tN})^{\mathrm{T}}$ represents the portfolio weights at time $t = 1, \ldots, T$, while $\sum$ denotes the covariance matrix. Usually, the real covariance matrix is unknown, hence requiring a suitable estimator.

2.1 Covariance Matrix Estimators

The covariance matrix is an important element of many asset allocation strategies. The widely used empirical covariance matrix estimator is unstable, especially when the number of assets (N) is larger than the time observations (T). Furthermore, during the mean-variance optimization, the covariance matrix inverting procedure amplifies the noise and numerical instability. In the following, we present three covariance matrix estimators assessed during our experiments. A more detailed description of covariance matrix estimators can be followed in Senneret et al. (2016) and Choi, Lim, and Choi (2019).

The Sample Covariance Estimate

The sample covariance matrix for N assets with return r_t can be formulated as follows:

$$\widehat{\sum}^{S} := \frac{1}{T} \sum_{t=1}^{T} (\mathbf{r}_t - \boldsymbol{\mu})\,(\mathbf{r}_t - \boldsymbol{\mu})^{\mathrm{T}},$$

where $\boldsymbol{\mu} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{r}_t$ is the expected return.

Ledoit–Wolf Linear Shrinkage Estimator

Shrinkage is a transformation procedure of the sample covariance matrix adopted in order to get a more robust covariance matrix. The linear shrinkage estimator proposed by Ledoit and Wolf (2004) combines the sample covariance matrix with the identity matrix. An important ingredient of shrinkage is the shrinkage intensity, which has a weighting role in the procedure. The Ledoit–Wolf estimator is defined by

$$\widehat{\Sigma}^{LS} := (1-\alpha)\widehat{\Sigma}^{S} + \alpha\frac{\operatorname{tr}[\widehat{\Sigma}^{S}]}{N}\mathbf{I},$$

where the optimal shrinkage intensity $\alpha \in [0,1]$ is chosen to minimize some risk function. A nonlinear variant of the shrinkage estimator is proposed by Ledoit and Wolf (2012). Furthermore, the nonlinear shrinkage estimator is combined with the composite likelihood method and applied to some special models (see Engle, Ledoit, and Wolf, 2019).

Denoising Covariance Matrix

Cleaning large-dimensional covariance matrices is a common task in different research areas as finance, physics, or multivariate statistics. Recently, some interesting works (see, for example: Bouchaud and Potters, 2011; Bun, Bouchaud, and Potters, 2017) have been conducted to find more robust covariance estimators. These experiments are usually relying on tools from Random Matrix Theory (RMT) in order to distinguish the signal part from the noisy part of the covariance matrix. More concretely, the cleaning process relies on correcting the eigenvalues of the covariance matrix by using the Marcenko–Pastur distribution.

Bouchaud and Potters (2011) proposed a cleaning procedure (eigenvalues clipping) where all eigenvalues below some threshold value are shrunk. Furthermore, Bun, Bouchaud, and Potters (2017) investigated a rotationally invariant estimator with promising results.

In this study, we apply a Targeted Shrinkage Denoising procedure (López de Prado, 2020) to the sample covariance matrix. Basically, this implementation is based on the clipping procedure proposed by Bouchaud and Potters (2011).

2.2 Reference Algorithms

In this subsection, we review the portfolio allocation methods considered in this study. The starting point of our comparisons is represented by the previously presented minimum-variance algorithm. The assessments also include two bootstrapped variants of the minimum-variance method and a recently proposed strategy that follows a slightly different approach than the others. The main subject of the investigations is how these algorithms perform using the different covariance matrix estimators.

Resampling Method

Michaud (1998) proposed the resampling (RES) method in order to overcome the instability of the estimated parameters of the mean-variance optimization method by applying Monte Carlo simulations during the portfolio construction process. The algorithm generates random resamples of asset returns based on the historical data by considering all securities simultaneously. For each of the resample, the optimal weights of mean-variance portfolio are computed, aggregating across all the samples by averaging the optimized weights. As a result, the obtained portfolio has less extreme final compositions and is well-diversified.

Subset Resampling

The subset resampling (SRES) procedure is a recently proposed (Gillen, 2016; Shen and Wang, 2017) variant of the resampling method. Rather than calculating weights for all assets, subset resampling constructs portfolios considering a smaller number of assets. By aggregating a sufficiently large number of subset portfolios, we obtain a well-diversified portfolio. This approach may be useful when there are many securities with short return histories. The performance of the method depends a lot on the two input parameters of the algorithm: the subset size and the number of subsets. In Shen and Wang (2017), the authors conducted detailed experiments and found that the subset resampling procedure is superior to other strategies. Subset resampling has the advantage that usually the subset size is smaller than the size of the observations, hence a more stable estimation of the covariance matrix can be obtained. Applying new estimators, we hope that further improvements can be achieved.

Hierarchical Risk Parity

The Hierarchical Risk Parity (HRP) algorithm was introduced recently by López de Prado (2017) as an alternative asset allocation strategy that alleviates the main pitfalls of the general mean-variance approach. The method performs asset allocation without the need to invert the covariance matrix. Unlike the Ledoit–Wolf and denoising methods, the HRP algorithm simply reorganizes the covariance matrix to place similar assets (in terms of linear co-movements) together and then employs an inverse-variance weighting allocation. The algorithm has three main steps:

- Tree clustering: the procedure transforms recursively the correlation matrix into smaller groups considering some distance metrics.
- Quasi-diagonalization: it is a technique where the covariance matrix is rearranged in order to reflect the similarity of the securities (see *Figure 1*).

– Recursive bisection: split portfolio weights between subsets based on inverse proportion to their aggregated variances.

Portfolios generated by HRP exhibit better out-of-sample performance than other traditional portfolio allocation algorithms (López de Prado, 2017). Although the covariance matrix transformation technique seems to be efficient, a question is arising whether further improvement can be achieved by using Ledoit–Wolf or denoised covariance estimators instead of the sample estimate.

Figure 1. *Quasi-diagonalization (see PortfolioLab)*

3. Empirical Comparison

The main aim of this section is to compare the previously described strategies considering different covariance estimator techniques. At first, the three Markowitz-type algorithms are assessed using a Monte Carlo experiment with synthetic data, while in the second experiment all the methods are compared based on the S&P 100 stock data.

3.1. Monte Carlo Experiment

This experiment is based on a novel approach (Gautier Marti, 2020) for sampling realistic financial correlation matrices. The method (called CorrGAN) relies on a generative adversarial network. CorrGAN generates correlation matrices that have many “stylized facts” seen in empirical correlation matrices based on asset returns. We simulate data using a 0-mean multivariate Gaussian parameterized by CorrGAN-generated correlation matrices with 80 stocks. An in-sample and an out-of-sample dataset of daily returns are generated based on the computed distributions. We use different observation numbers: $T = 50$, 100, and 200. The comparison methodology is based on in-sample and out-of-sample results as follows: the empirical covariance matrix is estimated by using the in-sample dataset. Then the asset allocation strategy computes the optimal weights which are further used to find the portfolio returns and their associated volatilities. The main performance metric is the portfolio volatility in- and out-of-sample.

The presented methodology is applied to the MV, RES, and SRES Markowitz-type asset allocation strategies, each of them using the following covariance estimation methods: sample estimate, Ledoit–Wolf estimator, and denoising covariance. In the case of RES, the results are averaged over 1,000 draws, while SRES used 21 as the size of the subsample and 3,000 for the subsample number. A single experiment is repeated for 1,000 generated correlation matrices, and the portfolio volatilities (in-sample and out-of-sample) are computed.

Table 1 contains the root-mean-square errors (RMSE) of volatilities of the examined methods using different covariance matrix estimators and observation numbers. Based on the results, SRES provides the best RMSE values indifferent of the sample size and covariance estimator. Obviously, the differences become smaller as the sample size increases, because the covariance matrix becomes more stable. Considering the effects of covariance estimators on the individual strategies, it can be stated that the Ledoit–Wolf approach managed in all cases to improve the results obtained by the sample estimate. On the other hand, the denoised variant provides better values than the sample estimate just using MV and RES for $T = 50$.

The distributions of the volatilities with 50 observations can be followed in *figures 2–4*. The figures also reflect the fact that the SRES is more stable than the MV and RES strategies independently of the covariance estimator.

Table 1. *RMSE for different covariance estimators and sample sizes*

Covariance estimation methods	T = 50			T = 100			T = 150		
	MV	**RES**	**SRES**	**MV**	**RES**	**SRES**	**MV**	RES	SRES
Sample estimate	0.1659	0.1449	**0.1313**	0.0987	0.0896	**0.0801**	0.0726	0.0680	**0.0622**
Ledoit–Wolf shrinkage	0.1559	0.1397	**0.1256**	0.0955	0.0878	**0.0786**	0.0713	0.0672	**0.0616**
Denoised covariance	0.1592	0.1429	**0.1339**	0.0986	0.0902	**0.0829**	0.0747	0.0698	**0.0647**

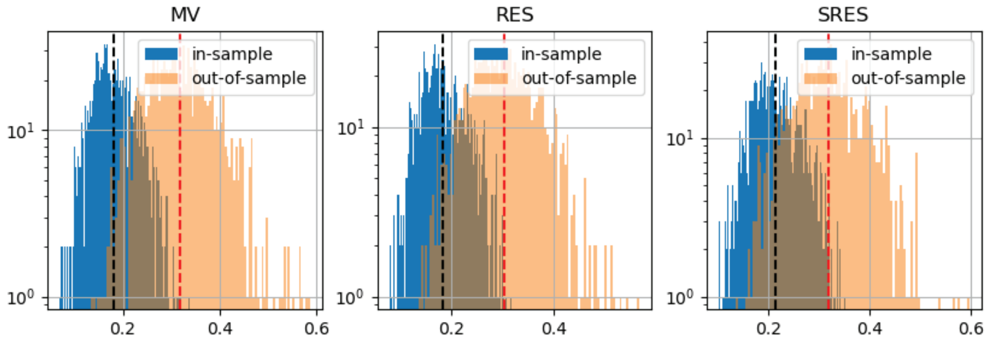

Figure 2. *Distribution of portfolio variance using the sample covariance estimator*

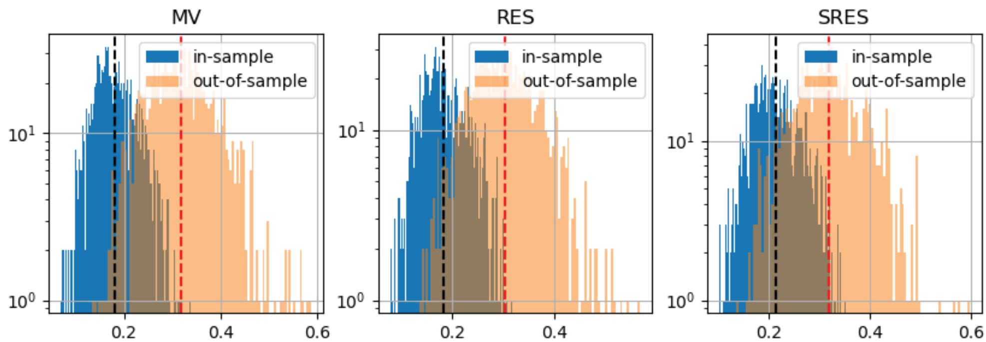

Figure 3. *Distribution of portfolio variance using the Ledoit–Wolf covariance estimator*

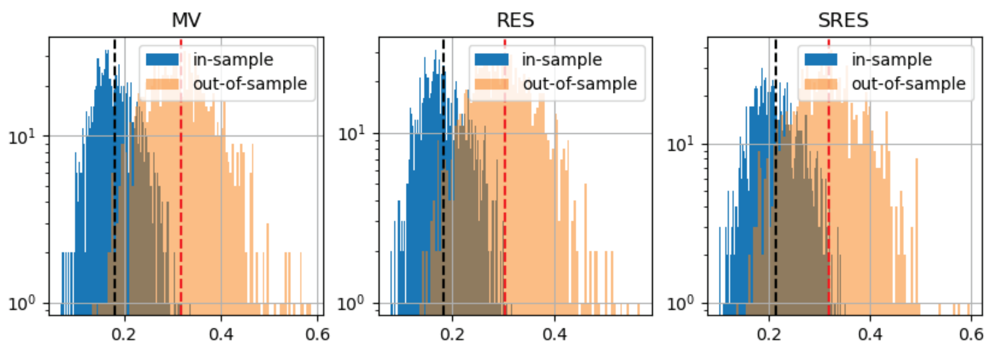

Figure 4. *Distribution of portfolio variance using the denoised covariance*

3.2. Experiments on Real Stock Data

In the following experiment, the MV, RES, SRES, and HRP asset allocation strategies are compared using different covariance matrix estimators. We provide a rolling window approach of the competitiveness of the strategies and analyse their performance on real stock data. The rolling window approach (DeMiguel, Garlappi, and Uppal, 2009) is a frequently applied procedure for performance analyses of asset allocation strategies. For a given T long dataset of asset returns, a rolling window (estimation window) of size M is chosen. The data of the first rolling window are applied to find the optimal allocation for each of the strategies. After that, the obtained weights are applied to the next period (out-of-sample) to calculate the returns. We continue this process by moving the rolling window towards the end of the dataset, getting a series of $T - M$ out-of-sample returns.

The presented procedure is applied for weekly returns of the securities of the S&P 100 from 1 January 2005 to 1 January 2020. Hence, the total observations consist of $T = 1{,}024$ weeks (15 years), and we considered those stocks without missing data resulting in 87 assets. We choose in-sample (rolling window) sizes of $M = 50$, 100, and 150 weeks and an out-of-sample size of 25 weeks.

In order to compare the performance of the selected allocation strategies, we use the following measures to evaluate the out-of-sample characteristics: the Sharpe ratio, maximum drawdown, annual growth rate (CAGR), annual volatility, and turnover rate.

Tables 2–4 report the results of the presented indicators for the MV, RES, SRES, and HRP methods using the sample covariance estimate, the Ledoit–Wolf estimator, and the denoised covariance. The Hierarchical Risk Parity strategy performs the best considering the CAGR and turnover rate indicators independently of the applied covariance estimators. This means that HRP usually provides more diversified portfolios than its counterparts. As we have expected, the impacts of the covariance estimators on the HRP method do not differ significantly. In the case of the Ledoit–Wolf estimator, HRP has the lowest turnover rates, while the denoised covariance matrix provides better CAGR values. Except for a few cases, the subsampling variant of the minimum-variance method achieves the best Sharpe ratios and annual volatilities. SRES is much more preferable in terms of diversification (turnover rate) than the MV and RES methods and achieves similar values as HRP as the estimation window size increases (see for M = 150).

Considering the individual covariance estimators, it can be concluded that the Ledoit–Wolf estimator and the denoised covariance usually improve the most important indicators obtained by the sample estimate. As we have previously observed, the differences are not significant for HRP, but we managed to improve the diversification property of the method, especially using the Ledoit–Wolf estimator. This estimator is slightly better than the denoised variant considering the Sharpe ratio and the turnover rate. In the latter case, the reduction is more significant for the MV strategy.

Table 2. *Sample covariance estimate*

Perf. measures	M = 50				M = 100				M = 150			
	MV	RES	SRES	HRP	MV	RES	SRES	HRP	MV	RES	SRES	HRP
Sharpe ratio	0.875	0.883	**0.914**	0.899	0.779	0.840	**0.860**	0.850	0.820	0.842	0.857	**0.861**
Max. drawdown	**37.394**	40.462	38.986	42.209	43.600	42.386	**39.712**	43.511	42.372	42.240	**40.156**	43.155
CAGR	11.083	11.287	11.3749	**12.599**	10.230	10.968	10.888	**12.270**	10.790	11.000	10.914	**12.651**
Annual volatility	13.008	13.119	**12.698**	14.366	13.737	13.493	**13.033**	14.961	13.670	13.499	**13.127**	15.215
Turnover rate	112.279	90.191	68.416	**41.365**	76.042	60.639	42.227	**29.822**	55.036	45.580	30.521	**25.880**

Table 3. *Ledoit–Wolf covariance shrinkage*

Perf. measures	M = 50				M = 100				M = 150			
	MV	RES	SRES	HRP	MV	RES	SRES	HRP	MV	RES	SRES	HRP
Sharpe ratio	0.886	0.911	**0.925**	0.899	0.815	0.862	**0.871**	0.862	0.835	0.856	**0.871**	0.855
Max. drawdown	**38.440**	40.015	40.012	42.543	42.088	41.347	**39.664**	42.300	41.528	41.533	**39.990**	42.966
CAGR	11.205	11.750	11.927	**12.669**	10.598	11.272	11.285	**12.495**	10.947	11.245	11.373	**12.603**
Annual volatility	**12.969**	13.181	13.141	14.455	13.497	13.470	**13.320**	14.981	13.573	13.546	**13.424**	15.266
Turnover rate	89.432	77.241	56.552	**33.976**	61.899	53.190	37.229	**26.160**	47.604	41.349	27.740	23.854

Table 4. *Denoised covariance*

Perf. measures	M = 50				M = 100				M = 150			
	MV	RES	SRES	HRP	MV	RES	SRES	HRP	MV	RES	SRES	HRP
Sharpe ratio	0.891	0.898	**0.925**	0.911	0.821	0.848	0.847	**0.861**	0.836	0.839	0.835	**0.868**
Max. drawdown	**35.576**	39.730	37.216	43.836	41.626	41.298	**38.468**	43.221	42.395	42.374	**39.536**	42.898
CAGR	11.137	11.305	11.243	**12.939**	10.959	10.976	10.480	**12.548**	10.826	10.898	10.467	**12.800**
Annual volatility	12.790	12.875	**12.378**	14.548	13.859	13.358	**12.747**	15.083	13.390	13.434	**12.958**	15.254
Turnover rate	100.273	79.947	59.668	**40.900**	72.172	54.824	37.000	**29.199**	56.159	41.601	27.898	**26.883**

4. Conclusions

In this paper, we have investigated different covariance matrix estimators applied to asset allocation strategies. We have compared the minimum variance portfolio with its resampling and subset resampling variants on synthetic data. Based on the conducted Monte Carlo experiment, we can conclude that the Ledoit–Wolf shrinkage estimator achieves the best performance independently of the applied allocation strategy. The RMSE values of portfolio volatilities between in-sample and out-of-sample show that the SRES algorithm is more robust than the MV and RES strategies. In the second experiment, the three Markowitz-type algorithms were compared with a recently proposed hierarchical tree clustering strategy using real financial data. The methods were assessed using a rolling window approach by varying the number of observations. Finally, the out-of-sample characteristics of portfolios were evaluated. The HRP method performs well considering the CAGR and turnover rate indicators. SRES usually outperforms MV and RES, especially for Sharpe ratios and annual volatilities. The diversification ability of SRES is also better than for MV and RES, and it achieves similar values as HRP as the number of observations are getting larger. The Ledoit–Wolf estimator and the denoised covariance usually improve the most important indicators obtained by the sample estimate. Using the newer nonlinear variant of the Ledoit–Wolf estimator or applying other cleaning schemes of covariance matrices, we may achieve further improvements. Another interesting research direction would be to apply clustered selection in the RES and SRES strategies.

References

Becker, F.; Gürtler, M.; Hibbeln, M. (2015), Markowitz versus Michaud: Portfolio optimization strategies reconsidered. *The European Journal of Finance* 21(4): 269–291.

Bouchaud, J.-P.; Potters, M. (2011). Financial applications of random matrix theory: A short review. In: *The Oxford handbook of Random Matrix Theory*. Oxford University Press. 824–850.

Bun, J.; Bouchaud, J.-P.; Potters, M. (2017). Cleaning large correlation matrices: Tools from Random Matrix Theory. *Physics Reports* 666: 1–109.

Choi, Y. G.; Lim, J.; Choi, S. (2019). High-dimensional Markowitz portfolio optimization problem: Empirical comparison of covariance matrix estimators. *Journal of Statistical Computation and Simulation* 89(7): 1278–1300.

DeMiguel, V.; Garlappi, L.; Nogales, F. J.; Uppal, R. (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science* 55: 798–812.

DeMiguel, V.; Garlappi, L.; Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies* 22: 1915–1953.

Engle, R. F.; Ledoit, O.; Wolf, M. (2019). Large dynamic covariance matrices. *Journal of Business & Economic Statistics* 37(2): 363–375.

Gillen, Benjamin J. (2016). Subset optimization for asset allocation. *Social Science Working Papers* 1421. California Institute of Technology, Pasadena.

Ledoit, O.; Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empirical Finance* 10: 603–621.

(2004). Honey, I shrunk the sample covariance matrix. *The Journal of Portfolio Management* 30: 110–119.

(2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices. *Annals of Statistics* 40(2): 1024–1060.

López de Prado, M. (2016). Building diversified portfolios that outperform out-of-sample. *Journal of Portfolio Management* 42: 59–69.

(2020). *Machine learning for asset managers*. Cambridge University Press.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance* 7: 77–91.

Marti, G. (2020), CORRGAN: Sampling realistic financial correlation matrices using generative adversarial networks. *ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain. 8459–8463.

Michaud, R. O. (1989). The Markowitz optimization enigma: Is optimized optimal? *Financ. Anal. J.* 45: 31–42.

PortfolioLab. https://hudson-and-thames-portfoliolab.readthedocs-hosted.com/.

Scutellà, M. G.; Recchia, R. (2013). Robust portfolio asset allocation and risk measures. *Ann. Oper. Res.* 204: 145–169.

Senneret, M.; Malevergne, Y.; Abry, P.; Perrin, G.; Jaffrès, L. (2016). Covariance versus precision matrix estimation for efficient asset allocation. In: *IEEE Journal of Selected Topics in Signal Processing* 10(6): 982–993.

Shen, W.; Wang, J. (2017). Portfolio selection via subset resampling. *AAAI'17: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*. 1517–1523.