Alternative Diversification Strategies in the temporal asset allocation.

Juri Hinz and Marc Weibel
School of Mathematical and Physical Sciences
University of Technology Sydney
juri.hinz@uts.edu.au

March 30, 2017

Abstract

In this work, we examine the incremental role of adaptive temporal asset allocation on the performance of diverse traditional portfolio strategies which follow classical investment philosophies. Such strategies are popular in practice and achieve desired statistical properties, (growth stability, mean reversion), but are vulnerable to abrupt changes in the asset dynamics.

We address its issue suggesting a potential improvement through regular revisions of the investment decisions. This concept is realized in terms of a certain time modulation to an appropriate traditional static portfolio.

1 Introduction

The last financial crisis and more recently the dramatic events surrounding Greece has triggered a re-design of portfolio strategies among practioners. Uncertainty about future asset returns within a portfolio optimization framework has led the financial industry to look for new solutions to propose to their clients.

In portfolio management practice, major difficulties originate from problems associated with reliable estimation of statistical model parameters and the sensitivity of the optimal asset allocation with respect to these quantities. Since there are several aspects of these difficulties and an entire range of methods which address them, let us explain our concept focusing on classical dynamic portfolio optimization.

In this context, the major quantitative ingredients are the (conditional) means and the covariances of the so-called asset log-returns. While a reliable and statistically significant estimation of log-return means is virtually impossible, the estimation of covariances may also be extremely difficult in practice, since for a large asset number, the consideration must take into account the asymptotic behavior of the spectrum of random empirical covariance matrices. Given permanent time changes of the price fluctuation intensity and an extreme sensitivity of the optimal portfolio weights, a naive construction of the optimal portfolio is usually has no value for practical applications.

In view of these problems, some practically relevant approaches to portfolio optimization have been suggested in order to overcome or to diminish the dependence on statistical procedures of model identification.

Let us emphasize the *benchmark approach* in this setting. This theory addresses an optimal portfolio selection under minimal theoretical assumptions and presents considerations, justifying the asymptotic optimality of an equally-weighted portfolio. This investment strategy attempts to hold an approximately equal fraction of the entire portfolio wealth in each of the assets, selected for the investment. Since this strategy requires a regular position re-balancing, it is not a static investment, strictly speaking. However, empirical investigations show that for appropriate diversification, even infrequent re-balancing can achieve a reasonable performance.

A similar area of ideas is related to the so-called *risk-parity* approach. In this framework, the investor attempts to build a portfolio choosing portfolio weights such that the marginal contribution from each asset position to an appropriately defined *total portfolio risk* is the same. Such *risk parity* approach is used to build diversified portfolios which do not rely on a return expectations, with the focus on risk management rather than performance. However, the risk parity approach has also been criticized, and some stylized dependence on expected returns has been reintroduced, with extensions in terms of the so-called *minimum-torsion* approach.

A general framework is suggested in [5], the so-called *benchmark approach* which assumes the existence of a numeraire portfolio. This numeraire portfolio displays positive weights and when used as a benchmark renders all benchmarked portfolios to super-martingales. Platen shows that this portfolio is equivalent to the Kelly portfolio which maximizes a logarithmic utility function. This numeraire portfolio cannot be systematically outperformed by any other long-only portfolio. This theoretical numeraire portfolio can be approximated by a worldwide diversified portfolio.

2 Methodology

2.1 Mean-Variance and Naive Diversification

The traditional mean-variance framework presented in [4] is still the most used approach in the financial industry. However, estimation theory has proven that this framework produces optimal portfolio weights that change quite dramatically over time due to the absence of large datasets. Several approaches have been proposed in order to stabilize or shrink the covariance matrix. These minimum-risk approaches do not rely on the estimation of expected returns and only focus on risk.

In several studies, Platen (see [6], [5], [7]) proposes an alternative to Markowitz-based strategies relying on naive diversification. This diversification is deemed to approximate the "ideal" numeraire portfolio, which maximizes the logarithmic utility function of an investor and thus dominates any other strictly positive portfolio over a long period of time. [7] shows that when the number of assets tends towards infinity, the equally-weighted portfolio converges to the numeraire portfolio. This robust approach is consistent and does not imply any asset returns model. Moreover, even after deduction of transaction costs, such portfolios significantly dominate corresponding market-weighted assets, thus empirically proving the asymptotic behavior of the numeraire "proxies".

2.2 Random Matrix Theory

2.2.1 Overview

The study of statistical factors inherent to asset returns has a long history in finance. Since the seminal paper of [4], using volatility (or variance) as a risk measure has become standard. The additive property of variance in uncorrelated markets, enables an easily identification of the sources of risk within a portfolio:

$$Var(R_p) = \sum_{i=1}^{N} Var(w_i R_i)$$

where R_p is the portfolio returns and R_i the returns of asset i, with a weight w_i in the portfolio.
However financial assets do display correlation and one often recourse to the principal component analysis in order to extract uncorrelated sources of risk. To this purpose we decompose the covariance matrix Σ of asset returns:

$$\mathbf{E}^{\mathsf{T}} \Sigma \mathbf{E} \equiv \Lambda$$

where $\Lambda \equiv \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$ is a diagonal matrix containing the eigenvalues of Σ and $E \equiv (e_1, \ldots, e_N)$ are the corresponding eigenvectors (column-wise). These eigenvectors define N uncorrelated factors, whose returns are defined by $\bar{\mathbf{R}}_{\mathsf{PCA}} \equiv \mathbf{E}^{-1} R$. The eigenvalues Λ correspond to the variances of these uncorrelated factors.

In [1], the authors showed that the minimum variance portfolio, has proposed by [4], displays the largest weight on the eigenvectors of the correlation matrix with the smallest eigenvalues. An effective empirical estimation of the correlation matrix thus turns out to be a complicated task but plays a major role in portfolio construction.

If we consider N assets with a number of observations T not very large compared to N, one can expect that the estimation of the covariances will be "noisy", meaning that the empirical correlation matrix is to a large extent composed of random entries. We thus have to be careful when using empirical correlations in portfolio construction, above all in minimum-risk strategies. It is of utmost importance to design a procedure allowing to retain real information and removing noise from the eigenvalues and eigenvectors.

2.3 Theory

Random Matrix Theory (RMT) came up in the 50's and are also of interest in a portfolio construction context. Algorithms used in the context of optimal portfolio liquidation, trading off the risk and the impact cost rely on the inversion of the covariance matrix. Thus small or zero eigenvalues, are related to portfolios of assets that have nonzero returns but vanishing or low risk. Small samples or insufficient data lead to estimation errors that impact such portfolios. Random matrix techniques aim at solving this issue of small eigenvalues in the sample covariance matrix.

In their research paper, [3] propose to compare the properties of the empirical correlation matrix to a purely random matrix, retrieved from simulated independent returns. The identification of deviations from the random matrix helps detect the presence of true information.

2.3.1 Random correlation matrices

Let us consider time series of N assets, with T observations. The elements of the empirical correlation matrix C, of size $N \times N$, are given by

$$\mathsf{C}_{ij} = \frac{1}{N} \sum_{t=1}^{T} \tilde{r}_{it} \tilde{r}_{jt}$$

where $\tilde{r}_{it}$ denotes the return of asset i at time t, normalized by volatility such that $\mathbb{V}ar[\tilde{r}_{it}] = 1$. If we use the matrix form the correlation matrix can written as

$$\mathsf{C} = \tilde{R} \tilde{R}^{\mathsf{T}}$$

where $\tilde{R}$ defines the $N \times T$ matrix whose rows correspond to the returns observations for each asset.

Theorem 1 (Marchenko-Pastur theorem)**.** *In random matrix theory, the asymptotic behavior of eigenvalues of large rectangular random matrices are described by the Marchenko-Pastur distribution. Let X denotes a $M \times N$ random matrix whose entries are independently identically distributed random variables with mean 0 and finite variance. $Y_N = N^{-1} X X^T$ and let $\lambda_1, \lambda_2, \ldots, \lambda_M$ be the eigenvalues of Y_N. Consider the random spectral measure*

$$\mu_M(A) = \frac{1}{M} \sum_{j=1}^{M} \delta_{\lambda_j \in A}, \quad A \in \mathbf{R}\,.$$

Assume that $M, N \to \infty$ so that the ratio $M/N \to \lambda \in (0, +\infty)$. Then $\mu_M \xrightarrow{d} \mu$, where

$$\mu_A = \begin{cases} (1 - \frac{1}{\lambda})\, \mathbb{I}_{0 \in A} + \nu(A), & \text{if } \lambda > 1 \\ \nu(A), & \text{if } 0 \leq \lambda \leq 1, \end{cases}$$

and

$$d\nu(x) = \frac{1}{2\pi\sigma^2} \frac{\sqrt{\lambda_{max} - x)(x - \lambda_{min})}}{\lambda x} \mathbb{I}_{[\lambda_{min}, \lambda_{max}]}\, dx$$
$$\lambda_{min}^{max} = (1 \pm \sqrt{\lambda})^2 \sigma^2$$

We can now apply the theorem defined in (1) within a portfolio context, Let $\tilde{R}_t \sim \mathcal{N}(m, \mathbb{I}_N)$ denote the independently and normally distributed asset returns[1] and C the empirical correlation matrix.

We denote $\rho_{\mathsf{C}}(\lambda)$ the density of the eigenvalues of the correlation matrix C, defined as:

$$\rho_{\mathsf{C}}(\lambda) = \frac{1}{N} \frac{dn(\lambda)}{d\lambda}$$

where $n(\lambda)$ corresponds to the number of eigenvalues of the correlation matrix C that are less than λ.

In [2], the authors showed, as $N \to \infty$ and $T \to \infty$, with $Q = T/N \geq 1$, $\rho_{\mathsf{C}}(\lambda)$ is exactly known:

$$\rho_{\mathsf{C}}(\lambda) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{\lambda_{\max} - \lambda)(\lambda - \lambda_{\min})}}{\lambda}, \tag{1}$$

$$\lambda_{\min}^{\max} = (1 \pm \sqrt{1/Q})^2 \sigma^2 \tag{2}$$

with $\lambda \in [\lambda_{\min}, \lambda_{\max}]$. As asset returns have been scaled the variance σ^2 is equal to 1.

In the limit Q=1, [3] shows that the distribution of the normalized eigenvalues are given by (1) and that important features can be extracted in the limit $N \to \infty$:

- the lower boundary of the spectrum is strictly positive, except for Q=1; No eigenvalue displays a value between 0 and $\lambda_{\min}$. In the neighborhood of this boundary, the density of eigenvalues exhibits a sharp maximum, except in the limit Q=1, corresponding to $\lambda_{\min} = 0$, where it diverges as $\sim 1/\sqrt{\lambda}$.
- the density of eigenvalues vanishes above $\lambda_{\max}$

Example 2.1. *Let us consider a simple example of $T = 1000$ random asset returns ($N = 200$ assets), with constant variance $\sigma^2 = 1$.*

When N is finite, these particular features displayed in the neighborhood of the the boundaries are not sharp. There is still a small probability of finding eigenvalues below $\lambda_{\min}$ and above $\lambda_{\max}$. This probability vanishes when the number of observation N becomes very large.

[1]Asset returns have been scaled to have $\sigma = 1$.

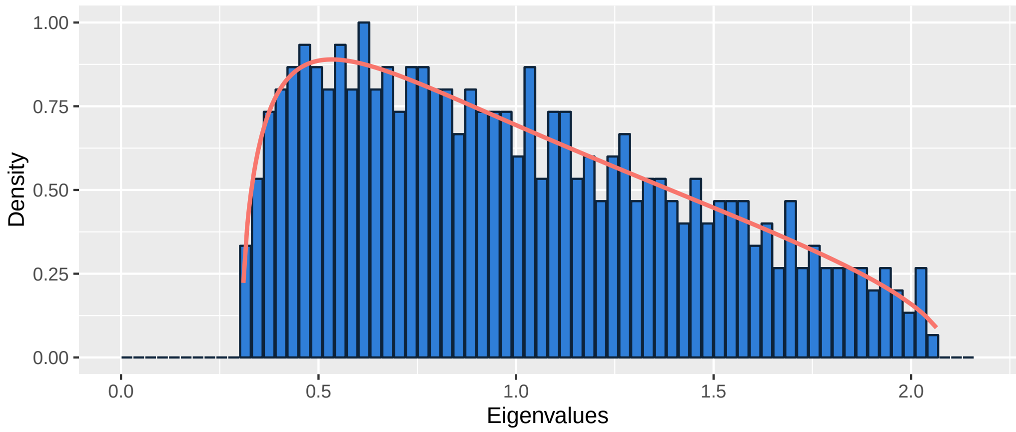

Figure 1: Eigenvalues distribution of 200 random assets

Example 2.2. *We consider 75 stocks (N) in the S&P500 index for which we have 200 observations(T). We thus have* $Q = T/N = 2.66$. *We display the spectrum of eigenvalues and superimpose the Marchenko-Pastur density, with* $Q = 2.66$ *and* $\sigma^2 = 1$.

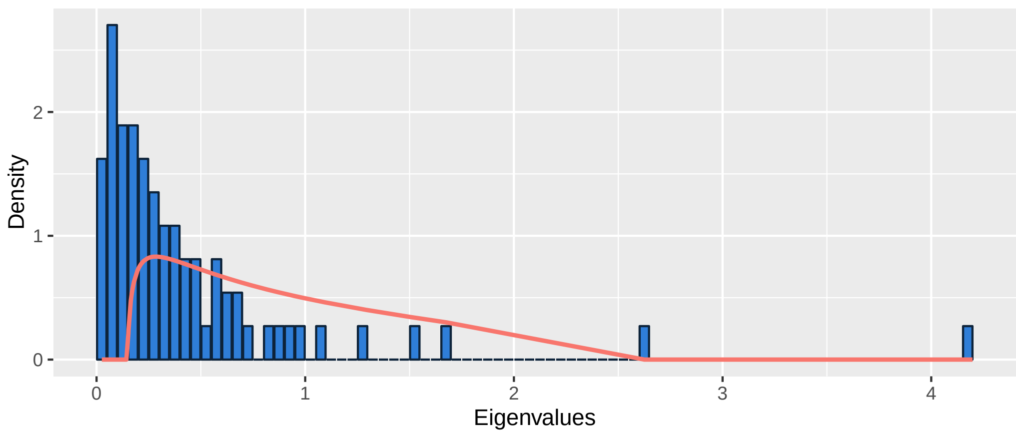

Figure 2: Eigenvalues spectrum of 75 stocks in the S&P500

We can observe that the highest eigenvalues are massively off the higher bound defined by the upper bound $\lambda_{\max}$ and that the overall fit is not really satisfying.

When we look at the corresponding eigenvectors, as expected, we can notice that all components are roughly equal on all stocks, thus proving that the first component corresponds to a proxy for the market itself. We can reject the hypothesis of "pure noise" for the first principal component. Another conjecture would be to assume that the other principal components, that are de facto orthogonal to the market proxy, are pure noise.

To this purpose, we can subtract the contribution of $\lambda_{\max}$ from the nominal value for $\sigma^2 = 1$, leading to $\sigma^2 = 1 - \lambda_{\max}/N = 0.94$. Figure 3 displays the empirical distribution with this better fit for σ^2 (cyan line).

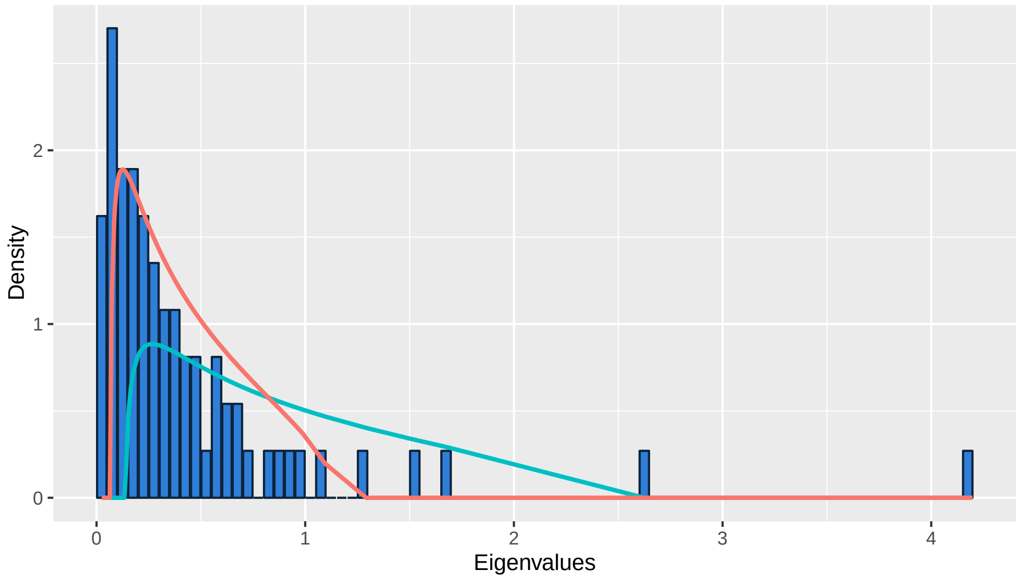

Figure 3: S&P500: Marchenko-Pastur density (best fit)

We see that some eigenvalues are still above $\lambda_{\max}$ and can thus be considered as information and reduces the random part of the correlation matrix. σ^2 can be considered as a parameter that we can adjust to optimize the fit. The best fit is obtained for $\sigma^2 = 0.44$ and corresponds to the red line in Figure 3. It accounts for roughly 95% of the spectrum, while the remaining highest eigenvalues are still well above the upper threshold $\lambda_{\max}$.

If we now randomize our asset returns, by shuffling the returns of each assets we obtain a spectrum of the eigenvalues that in the limit follows a Marchenko-Pastur density, as shown in Example 2.3.

Example 2.3. *We consider the same stocks but shuffle the time series of each asset and compute the corresponding asset values. We repeat this procedure 1'000 times and average the results. We can observe that the random returns are very well explained by the theoretical density.*

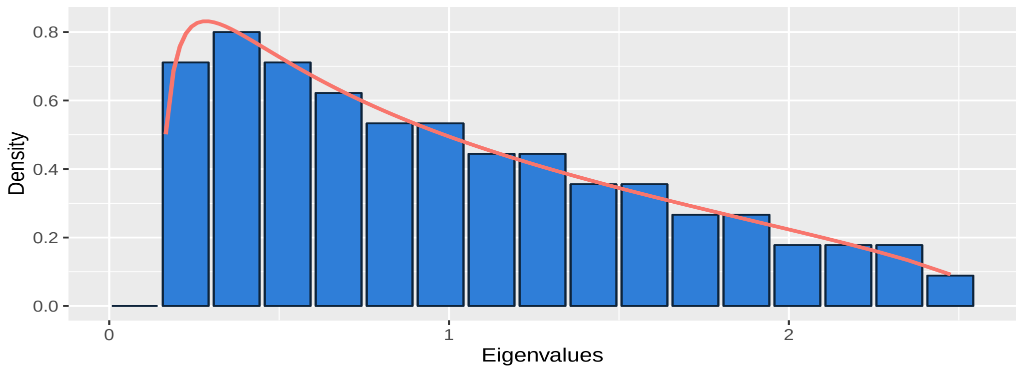

Figure 4: S&P500: Reshuffled assets)

2.4 Minimum-Torsion

Let us assume the price evolution $(S(t) = (S_1(t), \dots, S_N(t))_{t\in\mathbb{N}}$ of given $N \in \mathbb{N}$ financial assets follows an adapted stochastic process taking values in $\mathbb{R}^N$ realized on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in\mathbb{N}})$. If we denote by $\pi(t) = (\pi_1(t), \dots, \pi_N(t)$ the vector of fractions of the wealth invested in the assets $i = 1, \dots, N$ at time $t = 0, 1, 2 \dots,$, then, following the self-financed strategy determined by $\pi = (\pi(t))_{t\in\mathbb{N}}$, the wealth $(S^\pi(t))_{t\in\mathbb{N}}$ evolves as

$$S^\pi(t+1) = S^\pi(t)(1 + \sum_{i=1}^{N} \pi_i(t) R_i(t+1)), \qquad t = 0, 1, 2 \dots$$

with the so-called returns

$$R_i(t+1) = (S_i(t+1) - S_i(t))/S_i(t), \qquad t \in \mathbb{N}\, i = 1, \dots, N$$

of the assets $i = 1, \dots, N$. For instance the so-called *equally-weighted portfolio* suggests holding the same fraction of the wealth in each asset at any time

$$\pi_i(t) = \frac{1}{N}, \quad t \in \mathbb{N},\, i = 1, \dots, N.$$

The traditional mean-variance considerations on portfolio optimization is based on the assumption that the mean vector and covariance matrix of the returns $(R(t) = (R_i(t))_{i=1}^N)_{t\in\mathbb{N}}$ do not change with time $t \in \mathbb{N}$. Let us agree that $R := R_1$ represents the distribution of the returns. The main ingredients for the calculation of optimal portfolio in the spirit of Markowitz are the return covariances

$$\Sigma = \mathrm{Cov}(R), \qquad \sigma^2 = \mathrm{Var}(R) = \mathrm{diag}^{-1}(\Sigma), \tag{3}$$

whose reliable estimation has attracted persistent attention in the literature. Given the covariance matrix Σ, the so-called principal component analysis (PCA) is based on diagonalization $D = T\Sigma T^\top$ with entries of the diagonal matrix D given by the eigenvalues of Σ whose orthonormal eigenvectors are rows of the orthogonal matrix T. The so-called principal components are given by $(TR(t))_{t\in\mathbb{N}}$ which can be considered as an approximation of the returns of the synthetic assets $(TS(t))_{t\in\mathbb{N}}$. Such linear transformation of the original price process $(TS(t))_{t\in\mathbb{N}}$ to $(TS(t))_{t\in\mathbb{N}}$ whose components have uncorrelated returns can be utilized in the portfolio optimization. However, the process $(TS(t))_{t\in\mathbb{N}}$ may appear artificial. That is, to reach return uncorrelation, other linear transformations than T may be of interest, thus we address the following question:

$$\left.\begin{array}{l}\text{determine a linear transformation } T^*: \mathbb{R}^N \to \mathbb{R}^N \\ \text{such that } T^*R \text{ are uncorrelated and the} \\ \text{components of } T^*R \text{ are close to those of } R\end{array}\right\} \tag{4}$$

In what follows, we present an approach to this problem, in terms of an algorithm which yields a matrix T^* solving (4), in certain sense.

Given the return covariance matrix Σ and the vector σ^2 of variances as in (

2.4.1 Corrected-Benchmark Portfolio

In the spirit of risk-parity strategies, which take correlations among assets into account we use the torsion matrix in order to correct the equally-weighted portfolio (naive diversification). We calculate the corrected weights by multiplying the portfolio weights $(w_i = \frac{1}{N})_{i=1}^N$ by the torsion matrix T^* computed in (**??**).

This transformation does not necessarily results in a fully invested portfolio. The difference can be either invested in cash or the weights can be scaled to add up to one. This methodology results in a portfolio where highly correlated assets are under-represented. We characterize the corrected-benchmark portfolio in terms of wealth fractions $\pi(t) = (\pi_i(t))_{i=1}^N$, invested in each asset $i = 1, \dots, N$ at time $t \in \mathbb{N}$ as

$$\pi^\top(t) = w^{*\top}T^*/w^{*\top}T^*\vec{1}, \qquad t \in \mathbb{N}. \tag{5}$$

In this formula, $w^{*\top}T^*$ stands for the wealth fractions, invested in risky assets given the un-correlation from minimum-torsion matrix T^*. According to this approach, only a fraction $w^{*\top}T^*\vec{1} \in]0,1[$ of the wealth would be invested (Here $\vec{1}$ stands for the vector whose entries are equal to one). In order to achieve a full investment of all available funds, we scale this portfolio appropriately, obtaining (5). A similar strategy can be obtained if short positions are not feasible,

$$\pi^\top(t) = (w^{*\top} T^*)^+ / (w^{*\top} T^*)^+ \vec{1}, \qquad t \in \mathbb{N}. \tag{6}$$

Here we use $(\cdot)^+$ to denote a component-wise application of positive-part function. We examine the behavior of both (5), (6) portfolio strategies in an empirical study in Section 3.

3 Case Study

We consider 375 stocks (N) in the S&P500 index for which we dispose of 1095 observations (T) covering the period from 2012/01 to 2016/05.

A common study design is to split the sample into a training and an independent testing set, where the former is used to develop the model and the latter to evaluate its performance.
Accordingly we start in a first step by analyzing the empirical histogram of the eigenvalues of the considered stocks over the first half of the period (2012/01-2013/12) and superimpose the theoretical Marchenko-Pastur density provided by the random matrix theory framework detailed in section ??.

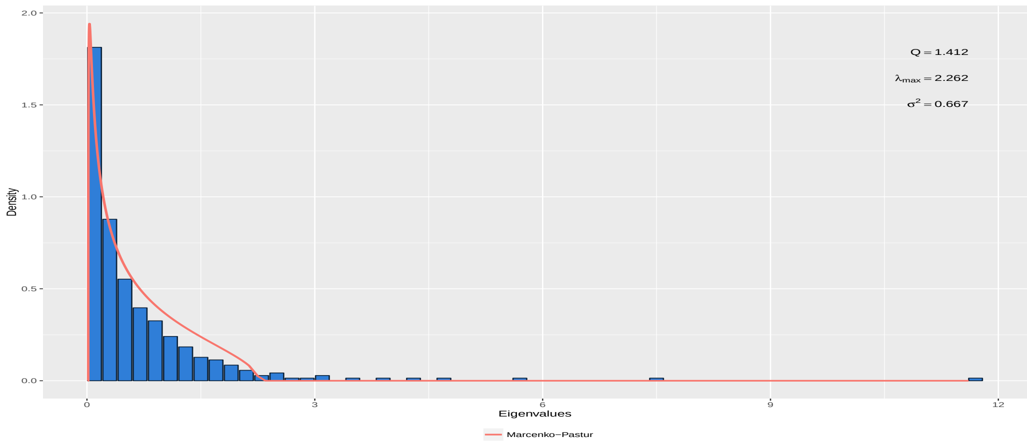

Figure 5: Eigenvalues spectrum of the S&P500 stocks

Based on this analysis (Figure 5), we retain eigenvalues above $\lambda_{\max} = 2.262$ that are assumed to contain "information" and shrink the remaining ones that correspond to "noise". For the shrinkage procedure we follow [3] and replace the noisy eigenvalues with average value such that the trace of the covariance matrix remains unchanged. This results in a "denoised" covariance matrix that will be used to compute the *correction matrix*, given by the minimum-torsion methodology explained in Section 2.4.

The following R code details the minimum-torsion algorithm:

```
# Compute the Minimum-Torsion Matrix
minimum_torsion <- function(cov.matrix)
{ # returns a matrix T such that sum (Var(TR - R)/Var(R))
  # is minimal where R is a random vector with cov(R)=cov.matrix
  # subject to entries of TR uncorrelated

  sigmas <- diag(cov.matrix)^(0.5)
  # Correlation matrix
  C2 <- diag(1/sigmas)%*%cov.matrix%*%diag(1/sigmas)

  E <- eigen(C2)
  # Square root of C2
  C <- E$vectors %*% diag(E$values^(0.5)) %*% t(E$vectors)
  # Inverse of C2
  Cinv <- E$vectors %*% diag(E$values^(-0.5)) %*%
    t(E$vectors)
  # Requirements for break conditions
  PIold <- C
  # Initialization
  D <- diag(x = 1,
            nrow = nrow(C),
            ncol = ncol(C))

  repeat {
    DC <- D %*% C
    E <- eigen(DC %*% t(DC))
    U <- E$vectors %*% diag(E$values^(-0.5)) %*%
      t(E$vectors) %*% DC
    diagonal <- pmax(0, diag(U %*% C))
    D <- diag(x = diagonal)
    PI <- D %*% U
    tolerance <- max(abs(PI - PIold))
    PIold <- PI
    V <- PI %*% Cinv

    # Convergence check: must be decreasing
    print(
      sum(
        diag(V %*% C2 %*% t(V) - V %*% C2 - C2 %*% V + C2)
      )
    )
    if (tolerance < 0.00001)
      break
  }

  result <- diag(sigmas)%*%V%*%diag(1/sigmas)
  return(result)
}
```

In a second step, we used the torsion-matrix calibrated to the training sample and apply the corrected-benchmark strategy defined in (2.4.1) on the test sample covering the period 2014/01-2016/05. Figure 6 compares the cumulative performance of the *naive strategy* (equally-weighted) to the *corrected-benchmark* approach. We also consider an alternative portfolio, where we impose a long-only constraint in order to deliver a fair comparison to the equally-weighted strategy.

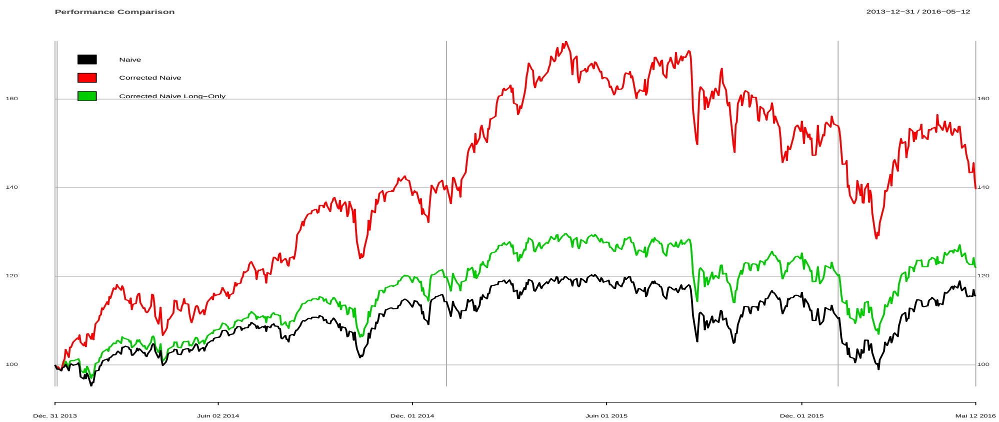

Figure 6: Cumulative performance

We can observe a similar dynamics of the three strategies, the unconstrained corrected-benchmark portfolio, where short positions are allowed, displays a better performance together with higher volatility. The constrained portfolio also slightly outperforms the naive diversification approach with comparable volatility.

In order to validate the multi-period approach, where we allow to rebalance the portfolio dynamically over time, we compare in Figure 7 the three strategies vs their buy-and-hold equivalent.
An astonishing pattern emerges from this figure, both the dynamic naive diversification approaches and the long-only corrected-benchmark strategy display an underperformance with regards to the buy-and-hold equivalent, whereas the corrected-benchmark strategy (constrained) moves from a cumulative outperformance of more than 6% in mid June 2015 towards an underperformance of 2% at the end of the testing period. We also note that all strategies suffered a drawdown in the second half of 2015, with a dramatic collapse in performance of the unconstrained strategy.

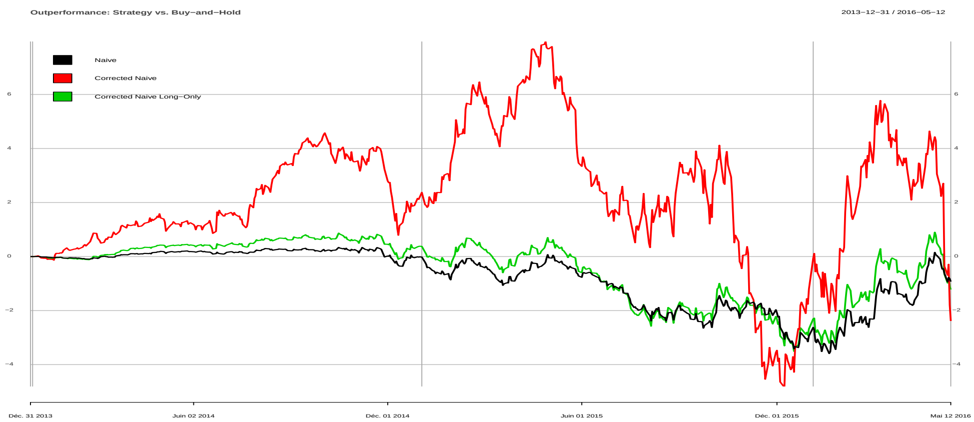

Figure 7: Outperformance vs. buy-and-hold portfolio

References

[1] J.P. Bouchaud and M. Potters. *Theory of Financial Risk*. Aleea-Saclay, Eyrolles, Paris, 1997.

[2] A. Edelman. Eigenvalues and condition numbers of random matrices. *SIAM J. Matrix Analy. Appl.*, 9(4):543, 1988.

[3] Laurent Laloux, Pierre Cizeau, Jean-Philippe Bouchaud, and Marc Potters. Random matrix theory and financial correlations. Science Finance (CFM) working paper archive 500053, Science Finance, Capital Fund Management, 1999. URL `http://EconPapers.repec.org/RePEc:sfi:sfiwpa:500053`.

[4] H. Markowitz. Portfolio selection. *The Journal of Finance*, 7:77–91, 1952.

[5] Eckhard Platen. A Benchmark Approach to Investing and Pricing. Research Paper Series 253, Quantitative Finance Research Centre, University of Technology, Sydney, August 2009.

[6] Eckhard. Platen and David. Heath. *A benchmark approach to quantitative finance / Eckhard Platen, David Heath*. Springer Berlin, 2006. ISBN 9783540262121 3540262121 9783540262121.

[7] Eckhard Platen and Renata Rendek. Approximating the numeraire portfolio by naive diversification. Research Paper Series 281, Quantitative Finance Research Centre, University of Technology, Sydney, 2010.