Helping You Write Academic Papers in R using Texevier

Nico Katzke[a], John Smith[a,b], John Doe[a,b]

[a] *Prescient Securities, Cape Town, South Africa*

[b] *Some other Institution, Cape Town, South Africa*

Abstract

Abstract to be written here. The abstract should not be too long and should provide the reader with a good understanding what you are writing about. Academic papers are not like novels where you keep the reader in suspense. To be effective in getting others to read your paper, be as open and concise about your findings here as possible. Ideally, upon reading your abstract, the reader should feel he / she must read your paper in entirety.

Keywords: Multivariate GARCH, Kalman Filter, Copula

JEL classification L250, L100

1. Introduction

importance of monte carlo methods

path dependence; first rule of investment management

Due to the aforementioned sensitivity issues surrounding errors in the expected return estimation, this work will only cover so-called risk based portfolios, since these techniques intentionally forego this input. These include the naive equal weight, inverse variance, hierarchical risk parity, equal risk contribution and the minimum variance portfolios. *The theoretical underpinnings of each will be reviewed as well as their relative performance in historical back tests.*

“. Markowitz’s curse is that the more correlated investments are, the greater is the need for a diversified portfolio—and yet the greater are that portfolio’s estimation errors. (De Prado, 2016)”

Corresponding author: Nico Katzke

Email addresses: `nfkatzke@gmail.com` (Nico Katzke), `John@gmail.com` (John Smith), `Joe@gmail.com` (John Doe)

Contributions:

The authors would like to thank no institution for money donated to this project. Thank you sincerely.

2. Aims and Objectives

This work aims to use Monte Carlo Methods to uncover the relationship between a market's covariance structure and the risk return properties of various risk based portfolio algorithms. This will be achieved through the following objectives.

1. Design and create four distinctive *ad hoc* and build one empirical *50 by 50* correlation matrix, each representing a market with a different risk structure. These will range from a structure possessing no clusters to those exhibiting hierarchical clustering.
2. Use the R package *MCmarket* to perform Monte Carlo Simulations, using each of the five correlation matrices from step one as the primary input [REFERENCE MYSELF]. The markets will be built to posses student t multivariate distributions, with 3 degrees of freedom. Meanwhile the individual asset returns will each be normally distributed with a random mean and standard deviation. Each market type will be simulated 10 000 times across 300 periods.
3. Use the simulated market data to calculate the returns obtained from various risk based portfolio's. The first **100** periods will be used estimate an out of sample covariance matrix, this will be used to calculate portfolio weights. These weights will remain for the next 50 periods after which portfolios will be rebalanced by looking back *100* periods, recalculating the covariance matrix and the new portfolio weights. This process is repeated until all *300* simulated periods have been considered. Therefore, each portfolio will end up with a series of *199* returns.
4. The performance of each portfolio will then be compared and contrasted using various portfolio risk/return analytics. Portfolio optimisers will be compared with each other within market types and with themselves across markets types.

3. Litrature Review

3.1. A Review of Portfolio Optimisation Algorithms

3.1.1. Introduction

Since Harry Markovitz's (1952) seminal work on mean-variance portfolios scholars from around the globe have been aspiring to develop a robust algorithm capable of situating a portfolio on the efficient frontier *ex ante* (Markowitz, 1952). There are now a wide array of available alternatives portfolio optimisers. They range from simple heuristic based approaches to advanced mathematical algorithms based on quadratic optimisation, random matrix theory and machine learning methods; with many more are still in the making.

This review

3.1.2. *Common Issues Portfolio Optimizers*

These optimisers are not without their flaws and this section highlights some common issues that tend to worsen their out of sample performance. Mean-variance optimisers, in particular, rely heavily on the accuracy of return forecasts, where small changes in the expected return input can lead to large changes in portfolio weights (De Prado, 2016). Due to this issue the so-called risk based portfolio's that intentionally avoid using expected return forecasts have garnered alot of attention (Maillard, 2010). Furthermore, quadratic programming methods used in many portfolio optimization require the inversion of some positive-definite covariance matrix. This step is particularly susceptible to error if the covariance matrix suffers from a high condition number. A condition number is defined as the absolute value of the ratio between a covariance matrix's largest and smallest eigenvalues (Bailey & Lopez De Prado, 2012; De Prado, 2016). Diagonal matrices have the smallest condition number, of 1, which increases as more correlated variables are added. When working with high conditional number matrices a small change in a single entry's estimated covariance can greatly alter its inverse, which in turn can effect the portfolio weights (De Prado, 2016). This is exacerbated by the fact that covariance matrices themselves are prone to estimation error (Zhou *et al.*, 2019). For a given sample size, larger dimension covariance matrices are prone to more noise in estimation. This is essentially due to a reduction in degrees of freedom as a sample of at least $1/2N(N+1)$ independent and identically distributed (iid) observations are required to estimate an $N \times N$ covariance matrix. Furthermore, financial market covariance structures tend to vary over time and have been know to change rapidly during so-called regime changes (De Prado, 2016). This exacerbates the issue of requiring a large number of observations when estimating the covariance matrix, as there is no guarantee that passed data will be a good refection of the future and looking further into the passed decreases the likelihood of it being so.

3.1.3. *Naive Equal Weight (EW)*

Perhaps the oldest and most simple portfolio diversification heuristic constitutes holding a weight of $1/N$ of the N total assets available to the investor (DeMiguel *et al.*, 2009). This portfolio is commonly called the equal weight or 1/N portfolio, its failure to recognise the importance of covariation between assets has resulted in it also being referred to as the naive portfolio. It commonly used as a benchmark index.

Despite its simplistic nature empirical studies tend to find a statistically insignificant difference in Sharp ratio between the naive portfolio and more advanced portfolio optimisers. This finding was made in DeMiguel *et al.* (2009) who looked at the mean-variance, minimum-variance and Bayes-Stein portfolio's, where EW also performed surprisingly well form a total return perspective.

3.1.4. Minimum Variance (MV)

Portfolio optimisers designed to exhibit the minimum variance have garnered a lot of attention for themselves, largely to their tendency to achieve surprisingly high returns in historical back tests (Clarke *et al.*, 2011). This performance has been attributed to the fact that low volatility stocks tend to earn returns in excess of the market, and high beta stocks tend not to be rewarded by higher returns (Clarke *et al.*, 2011; Fama & French, 1992). These minimum variance portfolios tend to achieve cumulative returns equal to or slightly greater than market capitalization weighted portfolio's whilst maintaining consistently lower variance and achieving a noticeable improvement in downside risk mitigation even during times of financial crisis (Clarke *et al.*, 2011). **The MV portfolio (discussed in this section) is the only portfolio on the efficient frontier that does not depend on expected return forecasts(De Prado, 2016).**

The minimum variance portfolio selects security weights such that the resulting portfolio weights correspond to the portfolio with the lowest possible in sample volatility. Therefore, it has the lowest expected volatility and is, in theory, safest/least risky portfolio (De Carvalho *et al.*, 2012a). Its primary input is a variance covariance matrix, which it uses in its optimisation to overweight low volatility and low correlation securities (De Carvalho *et al.*, 2012a). This approach often works well out of sample, but is known to achieve minor reductions in *ex ante* portfolio volatility by greatly favouring a small number of low volatility/correlation securities (De Prado, 2016 [p. 68]). This tendency to produce highly concentrated portfolio's can cause serious out of sample diversification problem as the sample becomes increasingly susceptible to measurement error.

Let $\sum$ indicate the markets variance covariance matrix and $w = \{w_i, ..., w_N\}$ be a vector of length N containing individual security weights, then the vector containing MV portfolio can be written as (De Carvalho *et al.*, 2012a):

$$w^* = arg\min(w' \sum w) \quad s.t. \ \sum_i^N w_i = 1$$

3.1.5. Inverse-Varience (IV) Weighting

The inverse-variance (IV), equal risk contrition (ERC) and maximum diversification (MD) portfolio's each assume that adequate diversification can be obtained by allocating equal risk to each investable security.

The IV portfolio, also known as the equal-risk budget (ERB), portfolio aims to allocate an equal risk budget to each investable security (De Carvalho *et al.*, 2012b). Where the risk budget is defined as the the product of a the security's weight and volatility. Therefore, if we define σ_i as security i's volatility, then marginal volatility is equally distributed across N securities by setting security weights as such:

$w_{iv} = (\frac{1/\sigma_1}{\sum_{j=1}^{N} 1/\sigma}, ..., \frac{1/\sigma_N}{\sum_{j=1}^{N} 1/\sigma})$

3.1.6. Equal Risk Contribution (ERC)

The ERC portfolio is similar to the IV, but also takes covariance into account (De Carvalho *et al.*, 2012b). The basic idea behind the ERC is to weight the portfolio such that each security contributes equally to risk, which in turn maximises risk diversification (Maillard, 2010). Generally speaking the ERC acts similar to a weight constrained minimum variance portfolio, with constraints ensuring that adequate diversification is maintained. The weights of the ERC portfolio $x = (x_1, x_2, ..., x_n)$ consisting of n assets is calculated as follows.

let σ_i^2 resemble asset i's variance, σ_{ij} the covariance between asset i and j and $\sum$ be the markets variance covariance matrix. Portfolio risk can now be written as $sigma(x) = \sqrt{x^T \sum x} = \sum_i \sum_{j \neq i} x_i x_j \sigma_{ij}$ (Maillard, 2010). The marginal risk contribution $\partial_{x_i} \sigma(x)$ can then be defined as follows (Maillard, 2010):

$\partial_{x_i} \sigma(x) = \frac{\partial \sigma(x)}{\partial x_i} = \frac{x_i \sigma_i^2 + \sum_{j \neq i} x_j \sigma_{ij}}{\sigma(x)}$

Therefore, $\partial_{x_i} \sigma(x)$ refers to the change in portfolio volatility resulting from a small change in asset i's weight. ERC uses this definition to guide its algorithms central objective to equate the risk contribution for each asset in the portfolio *ex ante*. If we define $(\sum x)_i$ as the i^{th} row resulting from the product of $\sum$ with x and note that $\partial_{x_i} \sigma(x) = (\sum x)_i$, then the optimal ERC weight can be written as (Maillard, 2010):

$x^* = \{x \; \epsilon [0, 1]^n : \sum x_i = 1, x_i \times (\sum x)_i = x_j \times (\sum x)_j \; \forall \; i, j\}$

3.1.7. Maximum Diversification (MD)

Choueifaty & Coignard (2008) originally designed the MD portfolio to maximize the diversification ratio (DR), which they defined as the as the sum of each securities risk bucket divided by portfolio volatility (De Carvalho *et al.*, 2012b). If we define $w = (w_1, ...w_N)^T$ as a vector of portfolio weights, V as a vector of asset volatilities and $\sum$ as the covariance matrix. Then the DR can be expresses as:

$DR = \frac{w'.V}{\sqrt{w'Vw}}$

Therefore, much like the IV and ERC portfolio's, the MD portfolio attempts to diversify the portfolio by allocating equal risk to each security (Choueifaty & Coignard, 2008). The MD portfolio accomplishes this by over-weighting low volatility securities and those that are less correlated with other stocks (De Carvalho *et al.*, 2012b). For detail regarding the theoretical results and properties of the MD portfolio see Choueifaty & Coignard (2008: 33–35).

'The MD strategy, introduced by Choueifaty and Coignard [2008], invests in the portfolio that maximizes a diversification ratio. The ratio is the sum of the risk budget allocated to each stock in the portfolio divided by the portfolio volatility. This strategy should invest in stocks that are less correlated to other stocks.' (copy pasted)

3.2. Empirical Backtests

Choueifaty *et al.* (2013) conducted an empirical back test comparing the relative performance if numerous portfolio optimisers between 1999 and 2010. They used historical data from the MSCI World world index and considered the largest 50% of assets at each semi-annual rebalance date. At each rebalance date the covariance matrices, used as inputs in the portfolio optimisers, were estimated using the passed years worth of data (Choueifaty *et al.*, 2013). This was done to reduce noise in estimation. All portfolio's were restricted to long only. The MV portfolio achieved an annual return of 6.7% and outperformed the ERC and EW portfolio's who returned 6.3% and 5.8% respectively. Unsurprisingly, the MV portfolio possessed the lowest daily volatility (10%) followed by the ERC and then the EW portfolio's (with 12.9% and 16.4% respectively). Accordingly the MV portfolio scored the highest sharp ratio (0.36) followed by the ERC and EW portfolio's (0.24 and 0.16 respectively).

4. Methodology

This work used Monte Carlo simulation methods to investigate the link between a markets correlation structure and the relative performance of the EW, MV, IV, ERC and MD portfolios. The following may become confusing so please note the following terminology. In this study the term market can be thought of as a single observation consisting of the daily returns for a number of assets, where market type refers to an ensemble of markets each designed to posses the same characteristics. Here the ensemble units are referred to as realizations or counterfactuals.

The R package MCmarket was used to simulate five distinctive market types, each corresponding to a unique correlation structure. Four of the correlation matrices were designed *ad hoc*, to posses a unique correlation structure, while the fifth was calculated using S&P 500 data. These correlation matrices range from one exhibiting no correlation (i.e. a diagonal matrix) to one with hierarchical clustering (see 4.1) (**???**).

The long only EW, MV, IV, ERC and MD portfolios were then backtested on the simulated markets using periodic reweighting.

Portfolio analytics are then performed for each portfolio type on each realization within the market types. These portfolio analytics include the standard deviation (sd) of daily returns, downside deviation, value at risk (VaR), conditional VAR (CVaR), Sharp ratio, average drawdown and maximum drawdown. The mean and median of the portfolio metrics are then calculated for each market type.

Finally, the portfolio metrics are compared within portfolio types across market types and within market types across portfolio types.

4.1. Correlation Structures

This section describes the process and motivation behind the creation of the four *ad hoc* correlation matrices (section 4.1.1) and the empirical correlation matrix (section \ref{emp) used as the primary inputs in the Monte Carlo simulations.

4.1.1. Ad Hoc

Figure 4.1 presents the four *ad hoc* correlation matrices used in this study.

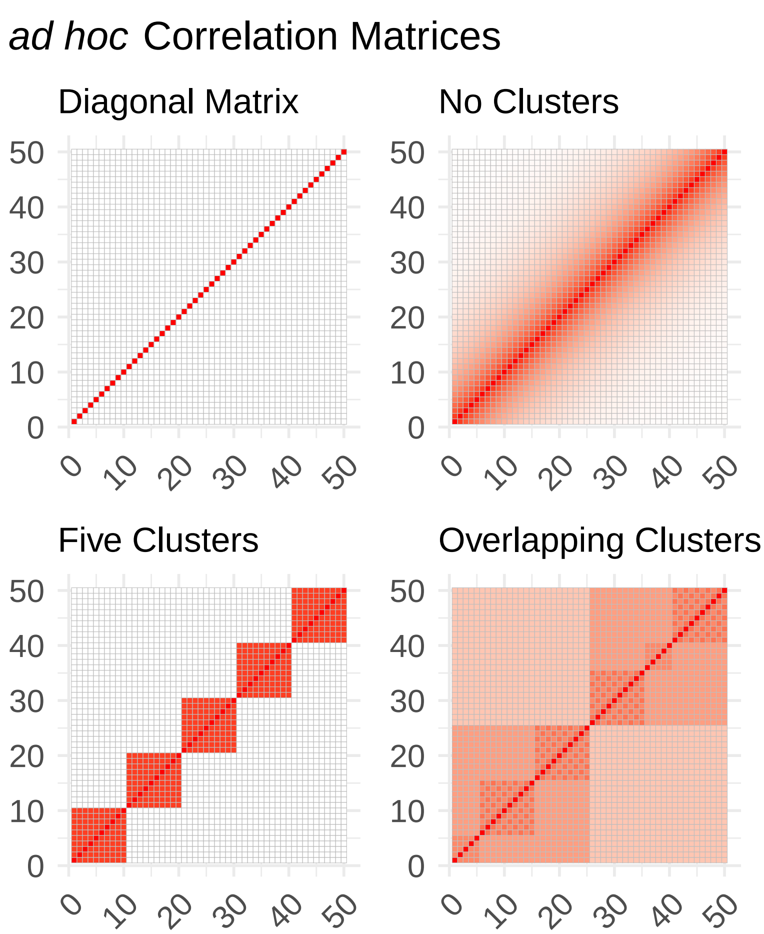

Figure 4.1: Correlation Matricies

4.1.2. Emperical

The empirical correlation matrix used in this study was calculated by randomly selecting the daily returns for 50 of the largest (market capitalization) 100 S&P 500 stocks between 1 January 2016 and 1 January 2021. Markets capitalization?? were measured on 1 January 2021. The covariance matrix was then calculated using the R package fitHeavyTail. This package uses ML estimation to fit a multivariate t-distribution according to the following methodology.

Chuanhai Liu and Donald B. Rubin, "ML estimation of the t-distribution using EM and its extensions, ECM and ECME," Statistica Sinica (5), pp. 19-39, 1995.

4.2. Monte Carlo

This work uses the R package MCmarket to carry out its Monte Carlo simulations. This package uses a generalized Monte Carlo framework from (**???**).

4.3. Back Tests

To remain in line with the literature, and the mandate for the majority of investors, the long-only constraint was applied to all portfolio's. In addition, a constraint limiting the maximum weight for a single security to 20%. This additional constraint is intended to prevent portfolio's that tend to building excessively highly concentrated holdings, while remaining flexible enough to punish those who do so. Therefore, the constraints are effectively providing a fair playing ground for the portfolio's to compete.

The first 50 periods were used to estimate the co-variance matrix, used as the primary input in portfolio weight calculations for the respective portfolios. These weights were then used in conjunction with the simulated returns to calculate the returns for the respective portfolios. d

4.4. Portfolio Analytics

5. Results and Discussion

6. Conclusion

I hope you find this template useful. Remember, stackoverflow is your friend - use it to find answers to questions. Feel free to write me a mail if you have any questions regarding the use of this package. To cite this package, simply type citation("Texevier") in Rstudio to get the citation for Katzke (2017) (Note that united references in your bibtex file will not be included in References).

References

Bailey, D.H. & Lopez De Prado, M. 2012. Balanced baskets: A new approach to trading and hedging risks. *Journal of Investment Strategies (Risk Journals)*. 1(4).

Choueifaty, Y. & Coignard, Y. 2008. Toward maximum diversification. *The Journal of Portfolio Management*. 35(1):40–51.

Choueifaty, Y., Froidure, T. & Reynier, J. 2013. Properties of the most diversified portfolio. *Journal of investment strategies*. 2(2):49–70.

Clarke, R., De Silva, H. & Thorley, S. 2011. Minimum-variance portfolio composition. *The Journal of Portfolio Management*. 37(2):31–45.

De Carvalho, R.L., Lu, X. & Moulin, P. 2012a. Demystifying equity risk–based strategies: A simple alpha plus beta description. *The Journal of Portfolio Management*. 38(3):56–70.

De Carvalho, R.L., Lu, X. & Moulin, P. 2012b. Demystifying equity risk–based strategies: A simple alpha plus beta description. *The Journal of Portfolio Management*. 38(3):56–70.

DeMiguel, V., Garlappi, L. & Uppal, R. 2009. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The review of Financial studies*. 22(5):1915–1953.

De Prado, M.L. 2016. Building diversified portfolios that outperform out of sample. *The Journal of Portfolio Management*. 42(4):59–69.

Fama, E.F. & French, K.R. 1992. The cross-section of expected stock returns. *the Journal of Finance*. 47(2):427–465.

Katzke, N.F. 2017. *Texevier: Package to create elsevier templates for rmarkdown*. ed. Stellenbosch, South Africa: Bureau for Economic Research.

Maillard, T., Roncalli. 2010. The properties of equally weighted risk contribution portfolios. *The Journal of Portfolio Management*. 36(4):60–70.

Markowitz, H. 1952. Portfolio selection. *The Journal of Finance*. 7(1):77–91.

Zhou, R., Liu, J., Kumar, S. & Palomar, D.P. 2019. Robust factor analysis parameter estimation. In ed. Springer *International conference on computer aided systems theory*. 3–11.

Appendix

6.0.1. Hierarchical Risk Parity (HRP)

Due to the multitude of robustness issues related to traditional portfolio optimisers, De Prado (2016) developed a new approach incorporating machine-learning methods and graph theory (**???**). De Prado (2016) argues that the "lack of hierarchical structure in a correlation matrix allows weights to vary freely in unintended ways" and that this contributes to the instability issues. His HRP algorithm requires only a singular co-variance matrix and can utilize the information within without the need for the positive definite property (De Prado, 2016). This procedure works in three stages:

De Prado (2016) carried out an in sample simulation study comparing the respective allocations of the long-only minimum variance, IVP and HRP portfolios using a co-variance matrix using a condition number that is "not unfavourable" to the minimum variance portfolio. The simulated data consisted of 10000 observations across 10 variables. The following findings were made: The minimum variance portfolio concentrated 92.66% of funds in the top 5 holdings and assigned a zero weight to 3 assets. *Conversly*, HRP only assigned 62.5% of its funds to the top 5 holdings (De Prado, 2016). The minimum variance portfolio's objective function causes it to build highly concentrated portfolio's in favor of a small reduction in volatility; the HRP portfolio had only a slightly higher volatility (De Prado, 2016). This apparent diversification advantage achieved by the minimum variance portfolio is rather deceptive as the portfolio remains highly susceptible to idiosyncratic risk incidents within its top holdings (De Prado, 2016). This claim was further validated by the finding that HRP achieved significantly lower out of sample variance compared to the minimum variance portfolio. ## Appendix A {-}

Some appendix information here

Appendix B