An Analysis of Hierarchical Portfolios

IEDA Final Year Project

Jacky Lau Temirlan Amir

Supervisor: Prof. Daniel P. Palomar

HKUST

May 2021

Outline

Portfolio Allocation

- How to allocate wealth across a universe of assets?
- Modern portfolio theory [1]: maximize the expected return subject to a given risk level.

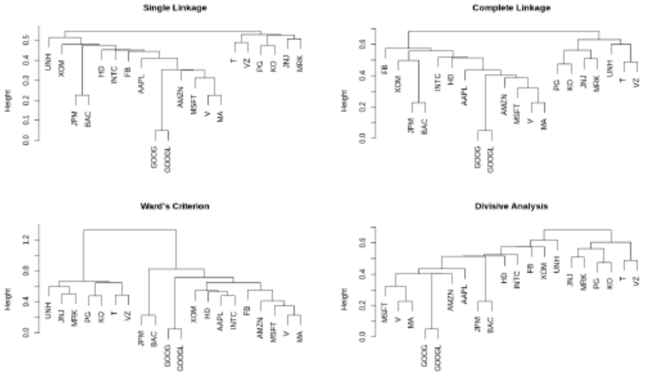
$$\begin{aligned} \underset{\mathbf{w}}{\text{maximize}} \quad & \mathbf{w}^T\boldsymbol{\mu} - \lambda \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} \\ \text{subject to} \quad & \mathbf{1}^T\mathbf{w} = 1,\ \mathbf{w} \geq \mathbf{0} \end{aligned}$$

  - Requires estimating an invertible covariance matrix of asset returns.
  - Sensitive to parameter estimation error.

Hierarchical Risk Parity (HRP)

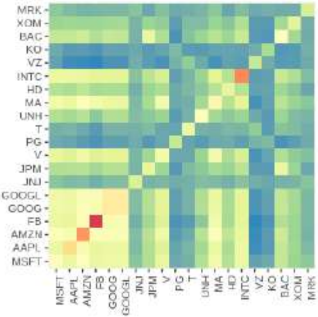
- Proposed by Lopez de Prado (2016) [2].
- Utilizes machine learning and graph theory to build diversified portfolios.
- Doesn't require an invertible covariance matrix estimate.
- Supposedly less risky out of sample.

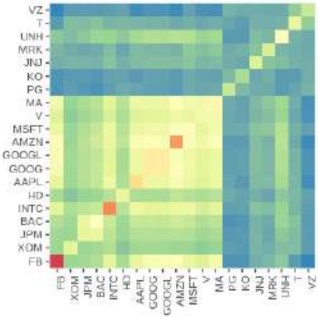
Hierarchical Clustering



Hierarchical Risk Parity (HRP)

1. Hierarchical clustering to reorder covariance matrix with distance matrix $D_{i,j} = \sqrt{\frac{1}{2}(1 - \rho_{i,j})}$.
2. Recursive bisection to allocate weights.



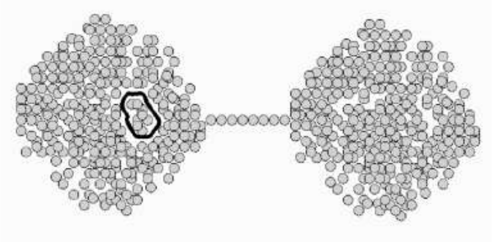


Recursive Bisection

1. Initialize weights $\mathbf{w} = (w_1, \cdots, w_n)^T = \mathbf{1}$, list of items $L = \{1, \cdots, n\}$.
2. While $|L| > 1$:
    1. Bisect L into 2 subsets $L^{(1)}, L^{(2)}$ such that $L^{(2)}$ consist of the first $\frac{n}{2}$ items (or $\frac{n+1}{2}$ if n is odd) and $L^{(2)}$ the remaining.
    2. Let $\Sigma^{(1)}, \Sigma^{(2)}$ denote the covariance matrices for assets in $L^{(1)}, L^{(2)}$ respectively. Define $w^{(1)}, w^{(2)}$ to be the inverse weighted portfolios corresponding to $\Sigma^{(1)}, \Sigma^{(2)}$ respectively, and $V^{(1)} = (w^{(1)})^T\, \Sigma^{(1)} w^{(1)}, \quad V^{(2)} = (w^{(2)})^T\, \Sigma^{(2)} w^{(2)}$ to be the variances of the portfolios.
    3. Compute the multiplier $\alpha = 1 - \frac{V^{(1)}}{V^{(1)} + V^{(2)}}$.
    4. Rescale weights. $w_i \leftarrow \alpha w_i, \quad \forall i \in L^{(1)}$
    5. Rescale weights. $w_j \leftarrow (1 - \alpha) w_j, \quad \forall j \in L^{(2)}$.
3. Repeat (2) recursively for $L \leftarrow L^{(1)}$ and $L \leftarrow L^{(2)}$.

Hierarchical Risk Parity (HRP)

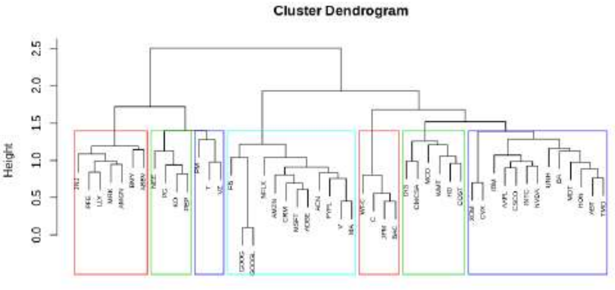
- Performs well out of sample.
- Only uses hierarchical clustering to obtain order of assets.
- Single linkage clustering can result in the chaining effect.



Hierarchical Equal Risk Contribution (HERC)

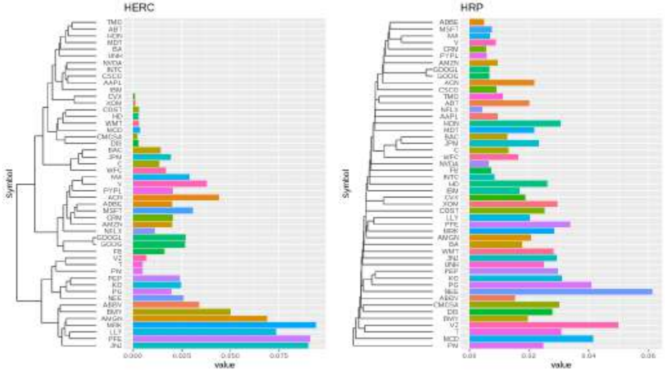
Proposed by Raffinot (2018) [3].

1. Perform hierarchical clustering with Ward's method.
2. Select the optimal number of clusters by computing the gap statistic [4].
3. Top-down recursive division.
4. Naive risk parity/inverse variance portfolio (IVP) within clusters.



Hierarchical Equal Risk Contribution (HERC)

HERC is less diversified compared to HRP.



Outline

Motivation

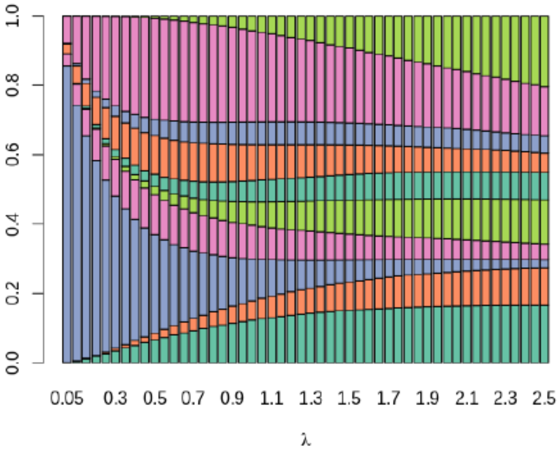
- Explore some extensions to HRP and HERC.
- Improve performance of these methods.
- Implement the existing methods in R for public use.

Modified Recursive Bisection 1

1. Initialize weights $\mathbf{w} = (w_1, \cdots, w_n)^T = \mathbf{1}$, $L = \{1, \cdots, n\}$.
2. While $|L| > 1$:
   1. Bisect L into 2 (roughly equal) subsets $L^{(1)},\ L^{(2)}$.
   2. Let $\Sigma^{(1)}, \Sigma^{(2)}$ denote the covariance matrices for assets in $L^{(1)},\ L^{(2)}$ respectively. Compute the inverse weighted portfolios $w^{(1)}, w^{(2)}$, corresponding to $\Sigma^{(1)}, \Sigma^{(2)}$ respectively, and compute the variances $V^{(1)} = (w^{(1)})^T\ \Sigma^{(1)} w^{(1)}, \quad V^{(2)} = (w^{(2)})^T\ \Sigma^{(2)} w^{(2)}$.
   3. Compute the multiplier α.
      1. Initialize $\alpha = 1 - \frac{\lambda V^{(1)}}{V^{(1)} + V^{(2)}}$, let w_j denote the weight for some $j \in L$.
      2. $\alpha \leftarrow \min(w_j^{-1} \sum_{i \in L^{(1)}} b_i,\ \max(w_j^{-1} \sum_{i \in L^{(1)}} a_i,\ \alpha))$
      3. $\alpha \leftarrow 1 - \min(w_j^{-1} \sum_{i \in L^{(2)}} b_i,\ \max(w_j^{-1} \sum_{i \in L^{(2)}} a_i,\ 1 - \alpha))$
   4. Rescale weights. $w_i \leftarrow \alpha w_i, \quad \forall i \in L^{(1)}$
   5. Rescale weights. $w_j \leftarrow (1 - \alpha) w_j, \quad \forall j \in L^{(2)}$.
3. Repeat (2) recursively for $L \leftarrow L^{(1)}$ and $L \leftarrow L^{(2)}$.

Modified Recursive Bisection 1

- Allows for a tuning parameter and weight constraints.



Modified Recursive Bisection 1

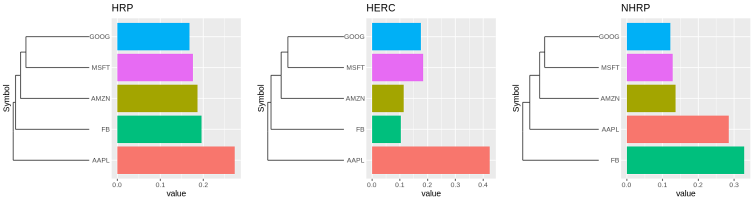
Parameter λ controls the level of diversification.

Modified Recursive Bisection 2

1. Initialize weights $\mathbf{w} = (w_1, \cdots, w_n)^T = \mathbf{1}$, minimum spanning tree $\mathcal{T}$ obtained from hierarchical clustering.
2. While $\mathcal{T}$ has subtrees:
    1. Let $\mathcal{T}_1$ and $\mathcal{T}_2$ denote the left and right subtrees respectively.
    2. Let $\Sigma^{(1)}, \Sigma^{(2)}$ denote the covariance matrices for assets in $\mathcal{T}_1$ and $\mathcal{T}_2$ respectively. Define $w^{(1)}, w^{(2)}$ to be the inverse weighted portfolios corresponding to $\Sigma^{(1)}, \Sigma^{(2)}$ respectively, and $V^{(1)} = (w^{(1)})^T \, \Sigma^{(1)} w^{(1)}, \quad V^{(2)} = (w^{(2)})^T \, \Sigma^{(2)} w^{(2)}$ to be the variances of the portfolios.
    3. Compute the multiplier α.
        1. Initialize $\alpha = 1 - \frac{\lambda V^{(1)}}{V^{(1)} + V^{(2)}}$, let w_j denote the weight for some $j \in \mathcal{T}$.
        2. $\alpha \leftarrow \min(w_j^{-1} \sum_{i \in L^{(1)}} b_i, \; \max(w_j^{-1} \sum_{i \in L^{(1)}} a_i, \; \alpha))$
        3. $\alpha \leftarrow 1 - \min(w_j^{-1} \sum_{i \in L^{(2)}} b_i, \; \max(w_j^{-1} \sum_{i \in L^{(2)}} a_i, \; 1 - \alpha))$
    4. Rescale weights. $w_i \leftarrow \alpha w_i, \quad \forall i \in \mathcal{T}^{(1)}$
    5. Rescale weights. $w_j \leftarrow (1 - \alpha) w_j, \quad \forall j \in \mathcal{T}^{(2)}$.
3. Repeat (2) recursively for $\mathcal{T} \leftarrow \mathcal{T}^{(1)}$ and $\mathcal{T} \leftarrow \mathcal{T}^{(2)}$.

Nodal Hierarchical Risk Parity (NHRP)

- Complete linkage instead of single linkage.
- Using the modified recursive bisection 2 which follows the tree structure.



hierarchicalPortfolios

Code is available for public use at
https://github.com/jackylauu/hierarchicalPortfolios.

Outline

Asset Universes

Three asset universes:

- 30 stocks from the Dow Jones Industrial Average (DJIA) over a period of 20 years from 2000 to 2020;
- 455 stocks from Standard and Poor's 500 (SP500) over a period of 20 years from 2000 to 2020;
- ETF Multi-Asset dataset with 40 assets over a period of 10 years from 2010 to 2020.

Backtesting Methods

- Backtest on historical data from 3 asset universes with monthly rebalancing.
- Use bootstrap [5] to resample data and compare observed performance metrics using error plots.
- Perform 1000 multivariate simulations with random shocks (additive and multiplicative noise).

Bootstrap Algorithm

1. Identify portfolio rebalancing date for the rebalancing window of 1 month
2. For each rebalancing date lookback window is 1 year (252 days) log-returns.
3. Resample log-returns for each lookback window 100 times, and generate 100 different time-series.
4. Obtain a set portfolio weights from portfolio allocation algorithms for each bootstrapped sample set.
5. Compute portfolio returns for the period from the current rebalancing date to the next rebalancing date (1 month).
6. Repeat the process until we obtain a set of portfolio returns for the entire observation period.

Simulated Data

- Use mean and covariance matrix from DJIA dataset.
- Generate multivariate normal to obtain 10-years simulated log-returns data
- Apply additive and multiplicative noise factor (random shocks)
- The starting period for shocks is random and ranges several weeks to several month
- 1000 iterations to obtain 1000 simulated sets

Benchmark Portfolios

Five benchmark portfolios:

- Global minimum variance portfolio (GMVP)
- Inverse variance portfolio (IVP)
- Equally weighted portfolio (uniform)
- Risk parity portfolio (RPP)
- Maximum Sharpe ratio portfolio (MSRP)

Performance Measures

- Sharpe Ratio
- Maximum drawdown
- Value at risk (VaR) with $\alpha = 95\%$
- Conditional value at risk (CVaR) with $\alpha = 95\%$

Outline

Real Data

	Sharpe Ratio	Annual Return	Annual Volatility	Max Drawdown	VaR (0.95)	CVaR (0.95)
NHRP	1.3178	0.1633	0.1239	0.1349	0.0127	0.0188
HRP λ=0.9	1.2861	0.1569	0.1220	0.1336	0.0125	0.0189
HRP	1.2814	0.1569	0.1224	0.1339	0.0127	0.0190
RPP	1.2798	0.1663	0.1300	0.1411	0.0134	0.0201
HRP λ=1.1	1.2780	0.1571	0.1229	0.1339	0.0127	0.0190
HERC-EW	1.2625	0.1780	0.1410	0.1581	0.0138	0.0215
IVP	1.2565	0.1579	0.1257	0.1382	0.0131	0.0195
Uniform	1.2406	0.1715	0.1383	0.1535	0.0145	0.0214
GMVP	1.2231	0.1527	0.1249	0.1382	0.0128	0.0193
MSRP	0.7372	0.1163	0.1578	0.1769	0.0163	0.0235

Table: Backtest results on historical data of 30 stocks composing the DJIA between 2010-2020

Real Data

	Sharpe Ratio	Annual Return	Annual Volatility	Max Drawdown	VaR (0.95)	CVaR (0.95)
NHRP	1.1762	0.1703	0.1448	0.3025	0.0131	0.0215
HRP lam=0.9	1.0940	0.1681	0.1537	0.3437	0.0136	0.0231
RPP	1.0746	0.1741	0.1620	0.3575	0.0143	0.0242
HRP	1.0736	0.1664	0.1550	0.3476	0.0137	0.0233
HRP lam=1.1	1.0559	0.1650	0.1563	0.3506	0.0138	0.0235
uniform	1.0423	0.1850	0.1775	0.3859	0.0153	0.0266
IVP	1.0160	0.1622	0.1597	0.3654	0.0141	0.0239
HERC-EW	0.8172	0.1427	0.1746	0.4233	0.0140	0.0256

Table: Backtest results on historical data of 455 stocks composing the S&P500 index between 2010-2020.

Real Data

	Sharpe Ratio	Annual Return	Annual Volatility	Max Drawdown	VaR (0.95)	CVaR (0.95)
NHRP	1.0194	0.0133	0.0122	0.0113	0.0012	0.0016
HRP λ=0.9	0.9712	0.0180	0.0144	0.0142	0.0014	0.0021
HRP	0.9671	0.0178	0.0145	0.0142	0.0014	0.0021
HRP λ=1.1	0.9578	0.0179	0.0147	0.0141	0.0014	0.0021
GMVP	0.9227	0.0271	0.0290	0.0250	0.0028	0.0041
IVP	0.9033	0.0199	0.0165	0.0152	0.0015	0.0023
MSRP	0.8782	0.0179	0.0207	0.0214	0.0019	0.0032
RPP	0.8494	0.0138	0.0128	0.0116	0.0012	0.0017
Uniform	0.4280	0.0350	0.0766	0.0815	0.0080	0.0111
HERC-EW	0.3146	0.0202	0.0715	0.0701	0.0073	0.0104

Table: Backtest results on historical data of ETF multi-asset dataset between 2010-2020.

Real Data

	Sharpe Ratio	Annual Return	Annual Volatility	Max Drawdown	VaR (0.95)	CVaR (0.95)
NHRP	0.1615	0.0526	0.3254	0.3692	0.0305	0.0441
HERC-EW	0.1390	0.0629	0.4523	0.3745	0.0353	0.0433
Uniform	0.1300	0.0573	0.4408	0.4613	0.0436	0.0624
RPP	0.1256	0.0518	0.4123	0.4385	0.0405	0.0584
HRP lam=0.9	0.1132	0.0423	0.3740	0.4129	0.0365	0.0528
HRP	0.1105	0.0420	0.3797	0.4184	0.0370	0.0535
HRP lam=1.1	0.1078	0.0415	0.3851	0.4235	0.0375	0.0542
IVP	0.0954	0.0399	0.4180	0.4512	0.0411	0.0586
GMVP	0.0951	0.0429	0.4512	0.4698	0.0446	0.0635

Table: Backtest results on SP500 stocks between September 2008-December 2009

Simulated Data

	Sharpe Ratio	Annual Return	Annual Volatility	Max Drawdown	VaR (0.95)	CVaR (0.95)
HRP λ=0.9	0.3027	0.0452	0.1470	0.1265	0.0149	0.0188
HRP λ=1.1	0.3009	0.0466	0.1493	0.1323	0.0148	0.0190
HRP	0.2991	0.0459	0.1481	0.1293	0.0149	0.0189
NHRP	0.2960	0.0455	0.1472	0.1225	0.0145	0.0185
IVP	0.2926	0.0455	0.1522	0.1334	0.0149	0.0194
GMVP	0.2874	0.0430	0.1431	0.1253	0.0144	0.0182
RPP	0.2867	0.0442	0.1551	0.1393	0.0153	0.0199
uniform	0.1902	0.0318	0.1637	0.1501	0.0160	0.0211
MSRP	0.1309	0.0227	0.1733	0.1672	0.0184	0.0234
HERC-EW	0.1077	0.0163	0.1580	0.1389	0.0158	0.0204

Table: Backtest results on simulated data with random shock.

Bootstrapped Data (DJIA)

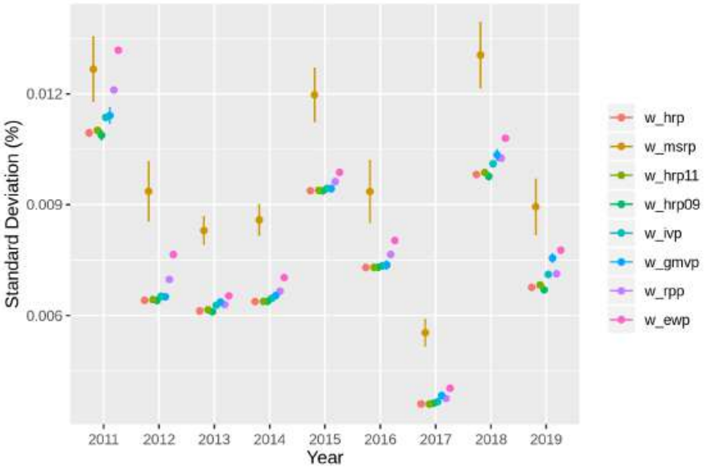


Figure: The annualized volatility of DJIA returns over time.

Bootstrapped Data (DJIA)

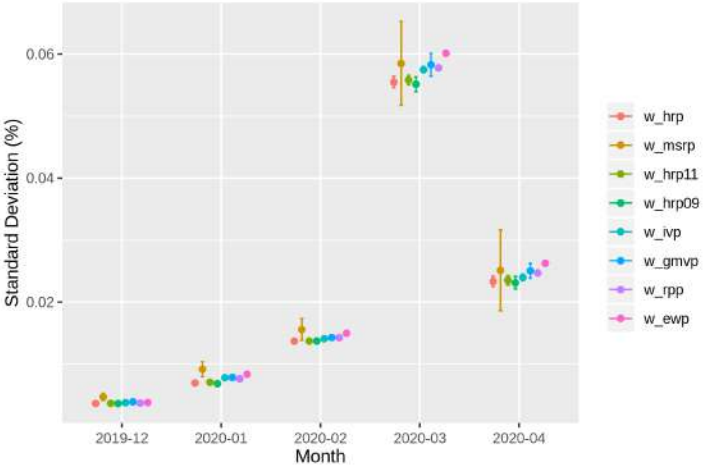


Figure: The annualized volatility of DJIA returns during January to May 2020.

Bootstrapped Data (S&P500)

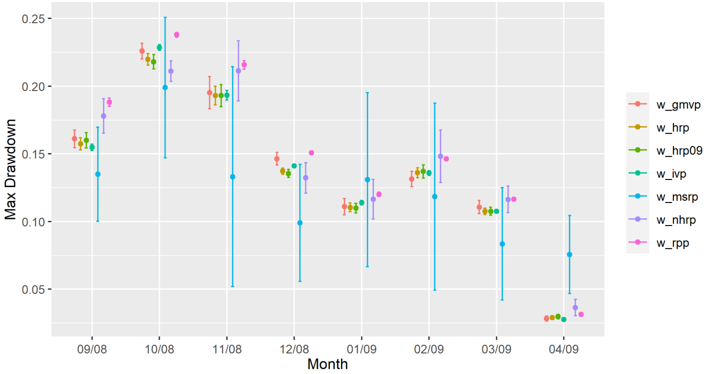


Figure: The maximum drawdown of DJIA returns between September 2008-April 2009.

Outline

Takeaways

- We proposed a slight modification to HRP as well as our method NHRP which combines HRP and HERC to outperform out of sample.

References I

- Markowitz, H. Portfolio Selection. *The Journal of Finance* **7,** 77–91. ISSN: 00221082, 15406261 (1952).
- Lopez de Prado, M. Building Diversified Portfolios that Outperform Out of Sample. *The Journal of Portfolio Management* **42,** 59–69 (July 2016).
- Raffinot, T. The Hierarchical Equal Risk Contribution Portfolio. *SSRN Electronic Journal* (Jan. 2018).
- Tibshirani, R., Walther, G. & Hastie, T. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **63,** 411–423 (2001).
- Efron, B. & Tibshirani, R. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science* **1,** 54–75 (1986).