Portfolio Optimization

Graph-Based Portfolios

Daniel P. Palomar (2024). *Portfolio Optimization: Theory and Application.*
Cambridge University Press.

portfoliooptimizationbook.com

Latest update: October 11, 2024

Outline

Abstract

Graphs offer a compact representation of big data, enabling analysis of large networks and pattern extraction. For financial data, asset graphs provide crucial information for modern portfolio design, potentially enhancing the mean-variance portfolio formulation. However, the optimal incorporation of graph information in portfolio optimization remains an open question. These slides explore some attempts in the literature (Palomar 2024, chap. 12).

Outline

Introduction

- **Markowitz's mean-variance portfolio:**
  - Balances expected return and risk.
  - Optimization problem:

$$\begin{array}{ll} \underset{\boldsymbol{w}}{\text{maximize}} & \boldsymbol{w}^\mathsf{T}\boldsymbol{\mu} - \frac{\lambda}{2}\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w} \\ \text{subject to} & \boldsymbol{w} \in \mathcal{W}, \end{array}$$

  - λ: risk-aversion hyper-parameter.
  - $\mathcal{W}$: constraint set (e.g., $\mathcal{W} = \{\boldsymbol{w} \mid \mathbf{1}^\mathsf{T}\boldsymbol{w} = 1, \boldsymbol{w} \geq \mathbf{0}\}$).

- **Challenges with estimation errors:**
  - Mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ are prone to errors.
  - Estimation errors significantly impact portfolio performance.

- **Improvement using graph of assets:**
  - Potential for enhancement by incorporating asset graph connectivity.
  - Graph connectivity patterns may reveal key investment insights.

Graphs and distance matrices

- **Graph-based portfolio construction:**
  - Utilizes graph information encoded as a distance matrix $\boldsymbol{D}$.
  - Distance reflects the relationship between asset pairs.
- **Correlation-based distance matrix:**

$$D_{ij} = \sqrt{\frac{1}{2}(1 - \rho_{ij})}$$

  where ρ_{ij} is the correlation between assets i and j.
- **Connection with Euclidean distance:**
  - Standardized data columns: $\tilde{\boldsymbol{x}}_i = (\boldsymbol{x}_i - \mu_i)/\sigma_i$
  - Empirical correlation:

$$\rho_{ij} = \frac{1}{T}\tilde{\boldsymbol{x}}_i^\mathsf{T}\tilde{\boldsymbol{x}}_j$$

  - Normalized squared Euclidean distance:

$$\frac{1}{T}\|\tilde{\boldsymbol{x}}_i - \tilde{\boldsymbol{x}}_j\|_2^2 = 2(1 - \rho_{ij})$$

Graphs and distance matrices

- **Other distance functions:**
  - Minkowski metric based on p-norm:
  $$D_{ij} = \|\tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_j\|_p$$
  where $\|\boldsymbol{a}\|_p = \left(\sum_{t=1}^{T} |a_i|^p\right)^{1/p}$
  - Manhattan distance ($p = 1$)
  - Euclidean distance ($p = 2$).

- **Holistic distance matrix approach:** Euclidean distance between distance vectors:
  $$\tilde{D}_{ij} = \|\boldsymbol{d}_i - \boldsymbol{d}_j\|_2$$
  - $\boldsymbol{d}_i$: ith column of $\boldsymbol{D}$.
  - Reflects similarity of assets with the entire asset universe.
  - Overcomes the limitation of pairwise-only information.

Toy example

- **Correlation matrix $\boldsymbol{C}$:**

$$\boldsymbol{C} = \begin{bmatrix} 1 & 0.7 & 0.2 \\ 0.7 & 1 & -0.2 \\ 0.2 & -0.2 & 1 \end{bmatrix}$$

- **Correlation-based distance matrix $\boldsymbol{D}$:**

$$\boldsymbol{D} = \begin{bmatrix} 0 & 0.3873 & 0.6325 \\ 0.3873 & 0 & 0.7746 \\ 0.6325 & 0.7746 & 0 \end{bmatrix}$$

- **Euclidean distance matrix of correlation distances $\tilde{\boldsymbol{D}}$:**

$$\tilde{\boldsymbol{D}} = \begin{bmatrix} 0 & 0.5659 & 0.9747 \\ 0.5659 & 0 & 1.1225 \\ 0.9747 & 1.1225 & 0 \end{bmatrix}.$$

Graphs and distance matrices

Other advanced graph estimation methods (Palomar 2024, chap. 5):

- **Heavy-tailed Markov random field (MRF) with degree control:**

$$\begin{array}{ll} \underset{\boldsymbol{w} \geq \boldsymbol{0}}{\text{maximize}} & \log \text{gdet}(\mathcal{L}(\boldsymbol{w})) - \dfrac{N+\nu}{T} \displaystyle\sum_{t=1}^{T} \log \left(\nu + (\boldsymbol{x}^{(t)})^\mathsf{T} \mathcal{L}(\boldsymbol{w}) \boldsymbol{x}^{(t)}\right) \\ \text{subject to} & \mathfrak{d}(\boldsymbol{w}) = \boldsymbol{1}, \end{array}$$

  where

  - $\text{gdet}(\cdot)$: generalized determinant
  - $\boldsymbol{w}$: graph weight vector
  - $\mathcal{L}(\boldsymbol{w})$: Laplacian operator
  - $\mathfrak{d}(\boldsymbol{w})$: degree operator
  - ν: controls heavy-tailness.

Graphs and distance matrices

- ***k*-component heavy-tailed MRF with degree control:** Aims for a k-component graph (graph with k clusters)

$$
\begin{array}{ll}
\underset{\boldsymbol{w}\geq\boldsymbol{0},\boldsymbol{F}\in\mathbb{R}^{N\times k}}{\text{maximize}} & \log \text{gdet}(\mathcal{L}(\boldsymbol{w})) - \dfrac{N+\nu}{T}\displaystyle\sum_{t=1}^{T}\log\left(\nu + (\boldsymbol{x}^{(t)})^{\mathsf{T}}\mathcal{L}(\boldsymbol{w})\boldsymbol{x}^{(t)}\right)\\
& \qquad\qquad\qquad\qquad + \gamma\,\text{Tr}\left(\boldsymbol{F}^{\mathsf{T}}\mathcal{L}(\boldsymbol{w})\boldsymbol{F}\right)\\
\text{subject to} & \mathfrak{d}(\boldsymbol{w}) = \boldsymbol{1}, \quad \boldsymbol{F}^{\mathsf{T}}\boldsymbol{F} = \boldsymbol{I},
\end{array}
$$

where

- γ: regularization hyper-parameter
- $\boldsymbol{F}$: enforces low-rank property.

Outline

Hierarchical clustering and dendrograms

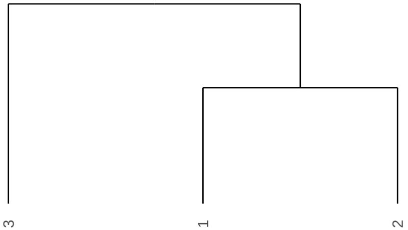
- **Clustering overview:**
  - Multivariate statistical analysis technique.
  - Used in machine learning, data mining, pattern recognition, bioinformatics, finance, etc.
  - Groups elements into clusters based on similar characteristics.
  - Unsupervised classification method.
- **Hierarchical clustering:**
  - Forms a recursive nested clustering.
  - Builds a binary tree of data points representing nested groups.
  - Allows data exploration at different levels of granularity.
  - Contrasts with partitional clustering, which finds all clusters simultaneously without a hierarchical structure.
- **Dendrogram:**
  - Visual representation of the hierarchical clustering tree.
  - Encodes the successive or hierarchical clustering process.
  - Provides a complete, interpretable description of the clustering in graphical format.
  - Popular due to its high interpretability.

Hierarchical clustering and dendrograms

Consider a toy example with distance matrix:

$$\tilde{\boldsymbol{D}} = \begin{bmatrix} 0 & 0.5659 & 0.9747 \\ 0.5659 & 0 & 1.1225 \\ 0.9747 & 1.1225 & 0 \end{bmatrix}$$

The dendrogram groups first the first and second elements since they have the smallest distance:



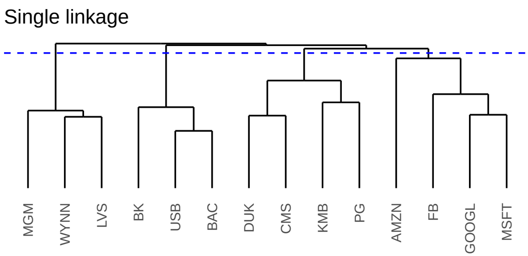
Basic procedure

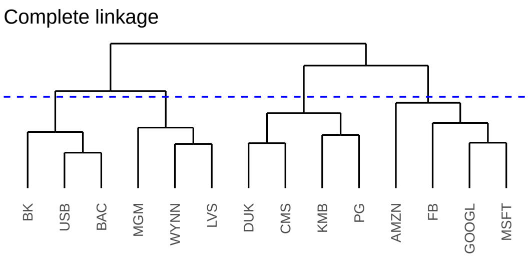
- **Hierarchical clustering process:**
  - Requires a distance matrix $\boldsymbol{D}$.
  - Sequentially clusters items based on distance.
- **Methods for hierarchical clustering:**
  - **Agglomerative (bottom-up):**
    - Starts with each item as a singleton cluster.
    - Merges the closest clusters sequentially.
    - Continues until one cluster remains.
  - **Divisive (top-down):**
    - Starts with all items in one cluster.
    - Recursively divides each cluster into smaller ones.
- **Levels of hierarchy:**
  - Each level represents a grouping into disjoint clusters.
  - The entire hierarchy is an ordered sequence of groupings.

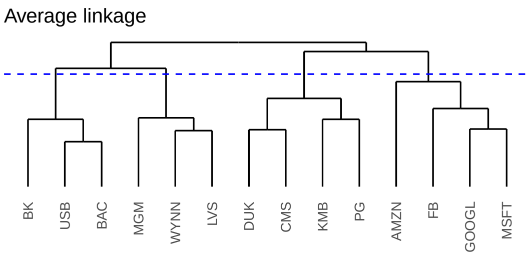
Basic procedure

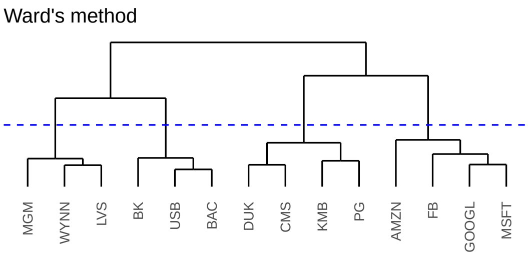
- **Linkage clustering methods:** measure of dissimilarity between clusters:
  - **Single linkage:**
    - Distance is the minimum distance between any two points in the clusters.
    - Related to the minimum spanning tree (MST).
  - **Complete linkage:**
    - Distance is the maximum distance between any two points in the clusters.
  - **Average linkage:**
    - Distance is the average distance between any two points in the clusters.
  - **Ward's method:**
    - Distance is the increase in squared error when merging clusters.
    - Related to distances between cluster centroids.
- **Effects of Linkage Method:**
  - Significantly impacts the resulting hierarchical clustering.
  - Single linkage may cause a "chaining" effect and imbalanced groups.
  - Complete linkage tends to produce more balanced groups.
  - Average linkage is an intermediate case.
  - Ward's method often yields results similar to average linkage.

Dendrograms of S&P 500 stocks









Number of clusters

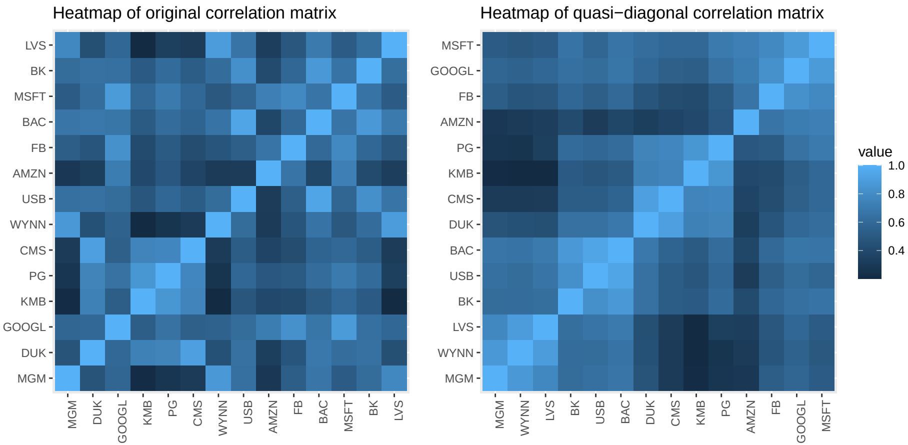
- **Determining the number of clusters:**
  - Traversing the dendrogram from top to bottom transitions from one giant cluster to N singleton clusters.
  - In practice, dealing with N singleton clusters may lead to overfitting.
- **Simplification vs. detail:**
  - Fewer clusters simplify the data but lose fine details, too many clusters might identify spurious patterns.
  - The challenge lies in choosing the optimal number of clusters.
- **Automatic detection of optimal clusters:**
  - Essential to avoid overfitting.
  - Aids in identifying the most appropriate number of clusters.
- **Gap statistic:**
  - Determines the optimal number of clusters.
  - Compares empirical within-cluster dissimilarity to uniformly distributed data.
  - Identifies the balance between simplification and preserving significant patterns.

Quasi-diagonalization of correlation matrix

- **Quasi-diagonalization of correlation matrix:**
  - Hierarchical clustering reorders items in the correlation matrix.
  - Groups similar assets closer and dissimilar assets farther apart.
  - Known as *matrix seriation* or *matrix quasi-diagonalization*.
  - An old statistical technique for revealing inherent clusters.
- **Benefits of quasi-diagonalization:**
  - Rearranges the correlation matrix into a quasi-diagonal form.
  - Reveals similar assets as blocks along the main diagonal.
  - Enhances visual pattern recognition compared to a randomly ordered matrix.
- **Visualization through heatmaps:**
  - Heatmaps can display the original and quasi-diagonal correlation matrices.
  - Original matrix with randomly ordered stocks shows no clear pattern.
  - Quasi-diagonal matrix, after reordering, clearly shows correlated stocks in diagonal blocks.
- **Identification of clusters:**
  - Quasi-diagonal matrix allows for easy identification of asset clusters.
  - Corresponding dendrograms can confirm the number and composition of these clusters.

Quasi-diagonalization of correlation matrix

Effect of seriation in the correlation matrix of S&P 500 stocks:



Outline

Hierarchical clustering-based portfolios

- **Portfolio design based on graph of assets:**
  - Aims to create robust, diversified portfolios with better risk-adjusted performance.
  - Less reliance on noisy estimates of mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

- **Hierarchical clustering for diversification:**
  - Distributes capital weights across hierarchically nested clusters.
  - Identifies isolated stocks contributing to diversification.
  - Visualized using the hierarchical tree layout.

- **Capital allocation in hierarchical clustering-based portfolios:**
  - Total capital starts at the top of the dendrogram.
  - Capital is allocated top-down through the hierarchy.
  - Each division of a cluster into sub-clusters splits the capital accordingly.
  - Portfolios for sub-clusters are designed at each split.

Hierarchical clustering-based portfolios

- **Characteristics of hierarchical clustering-based portfolios:**
  1. **Distance matrix:** Defines the graph (e.g., correlation-based, distance matrix of columns, sophisticated graph learning).
  2. **Linkage method:** Employed in the clustering process (e.g., single, complete, average, Ward).
  3. **Clustering stopping criterion:** Determines when to stop clustering (e.g., single-item clusters, gap statistic).
  4. **Splitting criterion:** Recursively splits the assets (e.g., bisection, dendrogram-based).
  5. **Intra-weight allocation:** Allocation of weights within clusters.
  6. **Inter-weight allocation:** Allocation of weights across clusters.

- **We will explore:**
  - Hierarchical $1/N$ Portfolio.
  - Hierarchical Risk Parity Portfolio.
  - Hierarchical Equal Risk Contribution Portfolio.

Outline

1. Introduction
2. Hierarchical clustering and dendrograms
3. Hierarchical clustering-based portfolios
   - Hierarchical $1/N$ portfolio
   - Hierarchical risk parity (HRP) portfolio
   - Hierarchical equal risk contribution (HERC) portfolio
4. Numerical experiments
5. Summary

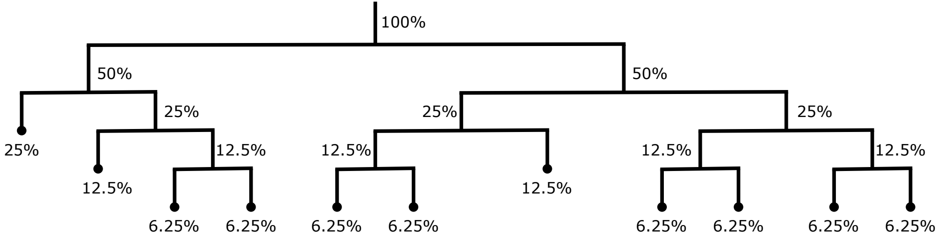
Hierarchical $1/N$ portfolio

- **Cluster-based waterfall portfolio overview:**
  - Introduced by (Papenbrock 2011) in his PhD thesis.
  - Utilizes hierarchical tree from correlation-based distance matrix.
  - Allocation process splits weights equally at each dendrogram split.

- **Allocation process:**
  - Proceeds in a top-down manner through the dendrogram.
  - Splits weights equally at each splitting point.
  - Illustration provided in the next figure.

Hierarchical $1/N$ portfolio

Illustration of the hierarchical $1/N$ portfolio construction in a top-down manner:



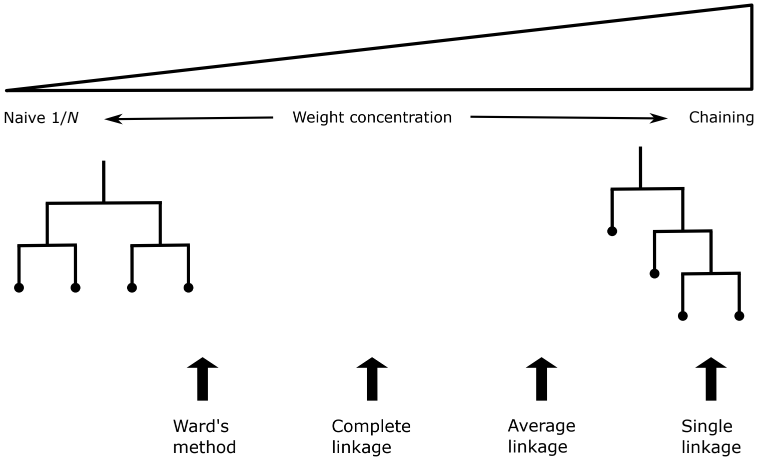
Hierarchical $1/N$ portfolio

Impact of linkage method on weight allocation:

- **Single linkage:**
  - Suffers from "chaining" effect, leading to high weights on some stocks.
- **Complete linkage:**
  - Produces more even groups and weights.
- **Average linkage:**
  - Intermediate between single and complete linkage.
- **Ward's method:**
  - Similar to complete linkage but with even more balanced groups and weights.
- **Regular $1/N$ portfolio:**
  - Represents equalized weights without using graph information, termed *naive* $1/N$ portfolio.

Hierarchical $1/N$ portfolio

Chaining effect of different linkage methods on the hierarchical $1/N$ allocation:



Hierarchical $1/N$ portfolio

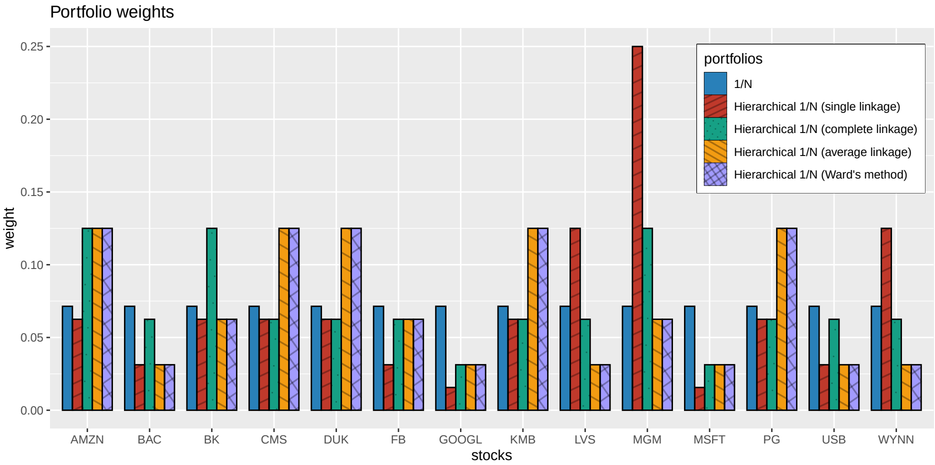
Summary:

1. **Distance matrix:**
   - Correlation-based: $D_{ij} = \sqrt{\frac{1}{2}(1-\rho_{ij})}$.
2. **Linkage method:**
   - Single linkage for high-risk investors.
   - Ward's method for risk-averse investors.
3. **Clustering stopping criterion:**
   - Continues to single-item clusters.
4. **Splitting criterion:**
   - Follows the dendrogram.
5. **Intra-weight allocation:**
   - $1/N$ portfolio strategy.
6. **Inter-weight allocation:**
   - $1/N$ portfolio with $N=2$, i.e., 50% - 50% split at each branching.

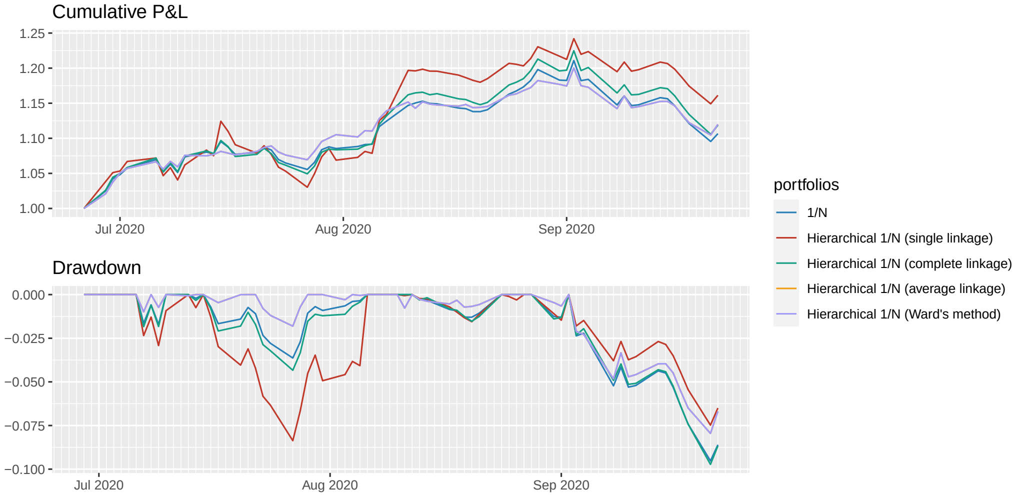
Numerical results

- **Comparing hierarchical $1/N$ portfolios with different linkage methods:**
  - Hierarchical $1/N$ portfolios are compared using single, complete, average, and Ward's linkage methods.
  - Naive $1/N$ portfolio serves as a benchmark.
- **Backtest results:**
  - Ward's method seems to be a good choice for hierarchical $1/N$ portfolio construction.
  - The original publication (Papenbrock 2011) supports the use of Ward's method for subsequent analysis.

Portfolio allocation of hierarchical $1/N$ portfolios with different linkage methods:



Backtest performance of hierarchical $1/N$ portfolios with different linkage methods:



Numerical results

- **Comparison of hierarchical $1/N$ portfolio using different distance matrices:**
  1. **Correlation-based distance matrix:**
     - As per (Papenbrock 2011):

$$D_{ij} = \sqrt{\frac{1}{2}(1-\rho_{ij})}$$

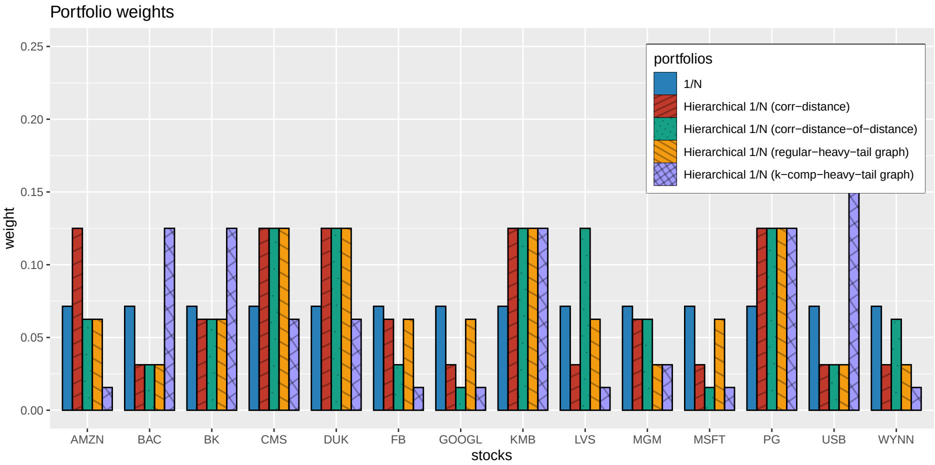
  2. **Correlation-based distance-of-distance matrix:**
     - As used in (López de Prado 2016):

$$\tilde{D}_{ij} = \|\boldsymbol{d}_i - \boldsymbol{d}_j\|_2$$

  3. **Graphs estimated via heavy-tailed MRF:**
     - Regular heavy-tailed MRF.
     - k-component heavy-tailed MRF.

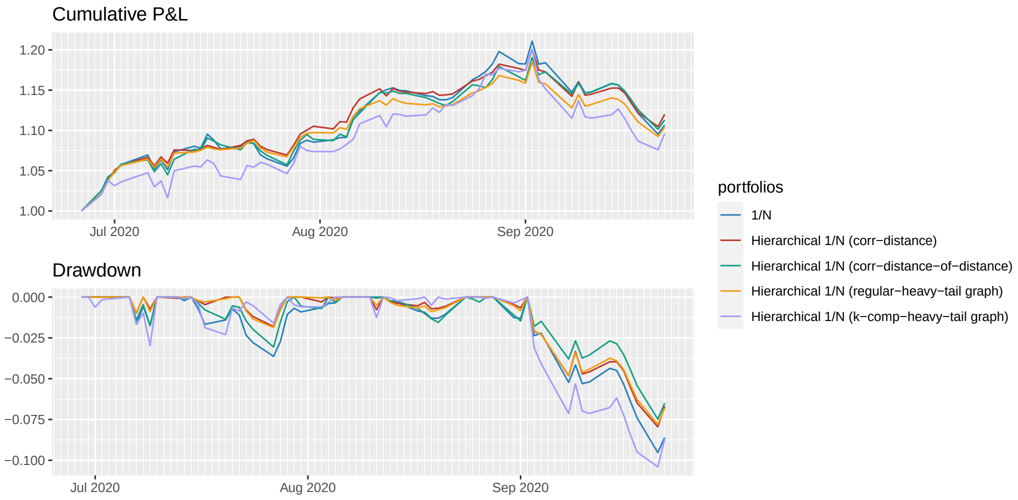
- **Backtest observations:**
  - The simple correlation-based distance-of-distance matrix appears to have a better drawdown profile.
  - More exhaustive backtests are recommended to draw definitive conclusions.

Portfolio allocation of hierarchical $1/N$ portfolios with different distance matrices:



Numerical results

Backtest performance of hierarchical $1/N$ portfolios with different distance matrices:

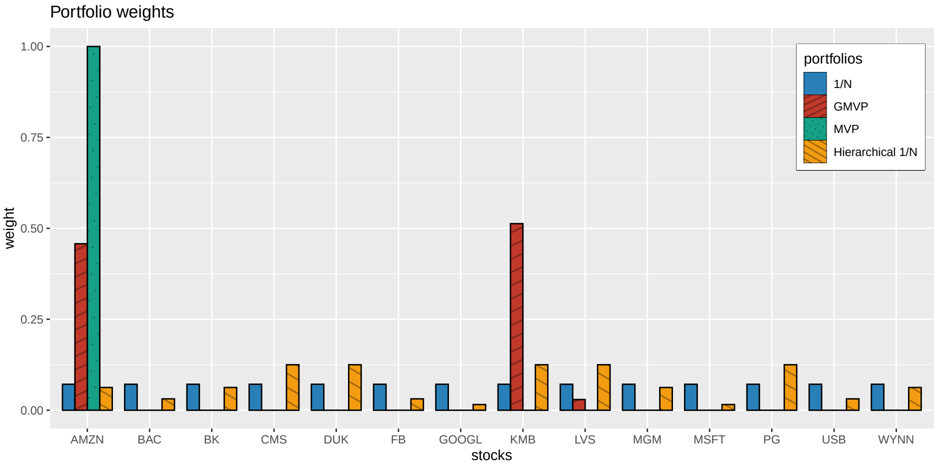


Numerical results

- **Final comparison of hierarchical $1/N$ portfolio:**
  - Selected version uses Ward's method for linkage and correlation-based distance-of-distance matrix.
  - Benchmarks: naive $1/N$ portfolio, global minimum variance portfolio (GMVP), and Markowitz mean-variance portfolio (MVP).
- **Observations:**
  - Hierarchical $1/N$ portfolio shows a distinct allocation pattern, emphasizing diversification.
  - MVP exhibits the worst drawdown due to sensitivity in estimating $\boldsymbol{\mu}$.
  - GMVP and naive $1/N$ portfolio show better performance than MVP.
  - Hierarchical $1/N$ portfolio demonstrates the mildest drawdown, indicating superior risk management.
- **Considerations and further evaluation:**
  - The presented backtest is anecdotal and not sufficient for definitive conclusions.
  - A proper empirical evaluation requires multiple randomized backtests.
  - Further analysis is necessary to robustly assess the performance of the hierarchical $1/N$ portfolio against benchmarks.

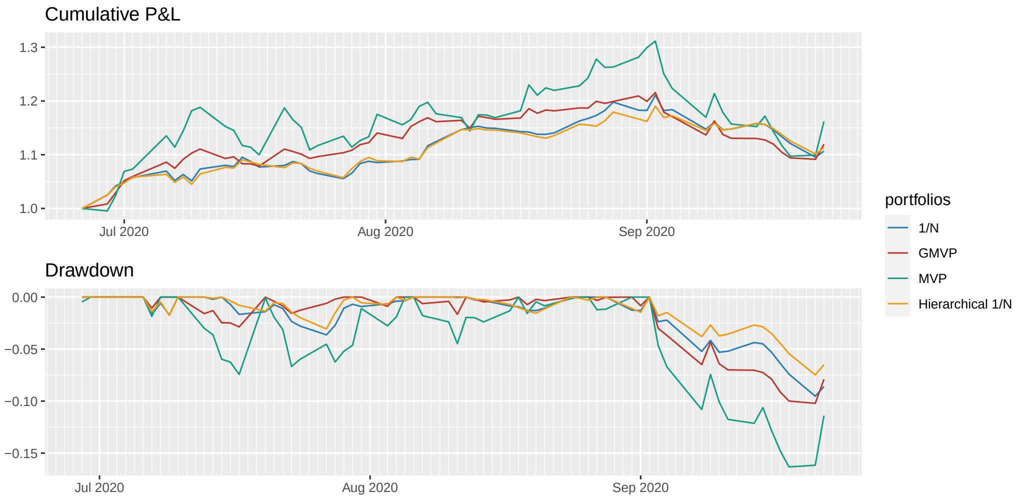
Numerical results

Portfolio allocation of hierarchical $1/N$ portfolio along benchmarks:



Numerical results

Backtest performance of hierarchical $1/N$ portfolio along benchmarks:



Outline

Hierarchical risk parity (HRP) portfolio

- **Hierarchical risk parity (HRP) portfolio overview:**
  - Introduced by (López de Prado 2016).
  - Based on hierarchical tree from correlation-based distance-of-distance matrix.
  - Utilizes single linkage method for clustering.
  - Allocation process uses inverse-variance portfolio (IVarP) for weight splitting.
- **Global minimum variance portfolio (GMVP) recap:**
  - Minimizes portfolio variance subject to budget constraint.
  - Solution simplifies to IVarP if covariance matrix $\boldsymbol{\Sigma}$ is diagonal:

$$\boldsymbol{w} = \frac{\boldsymbol{\sigma}^{-2}}{\mathbf{1}^\mathsf{T}\boldsymbol{\sigma}^{-2}}.$$

- **Inverse-variance portfolio (IVarP) for $N = 2$ assets:**
  - Weight allocation based on inverse of variances:

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \sigma_2^2/(\sigma_1^2 + \sigma_2^2) \\ \sigma_1^2/(\sigma_1^2 + \sigma_2^2) \end{bmatrix}.$$

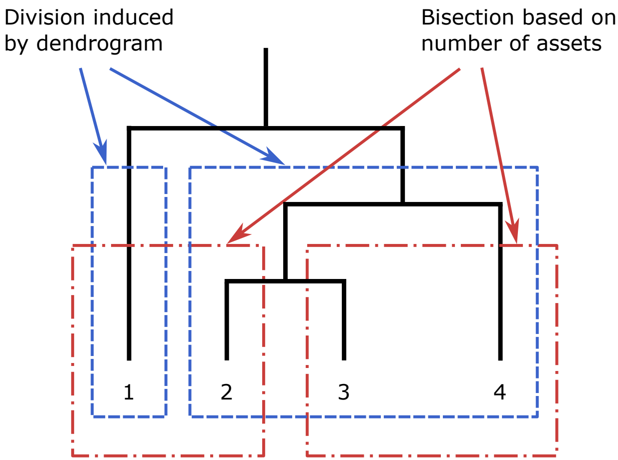
Hierarchical risk parity (HRP) portfolio

- **HRP portfolio design process:**
  - Similar to hierarchical $1/N$ portfolio, allocation proceeds top-down through the dendrogram.
  - **Differences:**
    - Uses bisection for splitting, not the dendrogram's natural structure.
    - Weight splitting based on IVarP for $N = 2$ assets.
  - Empirical evaluation required to compare performance with hierarchical $1/N$ portfolio.

- **Interpretation and connection to GMVP:**
  - HRP can be seen as a refined version of the IVarP.
  - At each step, weights are scaled based on inverse variances, ignoring correlations between subsets.
  - Correlations considered only in variance computation of subsets.
  - When covariance matrix is diagonal, IVarP, GMVP, and HRP coincide.
  - Connection between HRP and GMVP explored further in subsequent sections.

Hierarchical risk parity (HRP) portfolio

Comparison of bisection splitting and dendrogram-based splitting:



Hierarchical risk parity (HRP) portfolio

Summary:

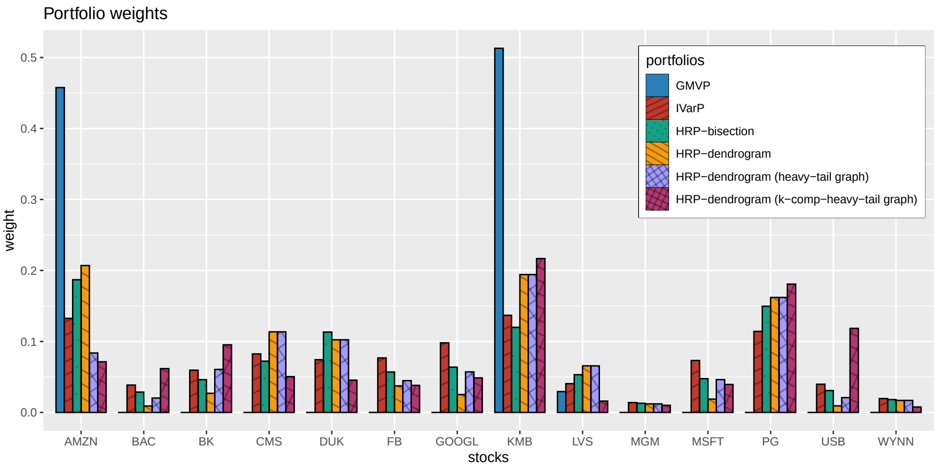
1. **Distance matrix:**
   - Correlation-based distance-of-distance matrix.
2. **Linkage method:**
   - Single linkage.
3. **Clustering stopping criterion:**
   - Continues to single-item clusters.
4. **Splitting criterion:**
   - Bisection, ignoring dendrogram grouping sizes.
5. **Intra-weight allocation:**
   - IVarP.
6. **Inter-weight allocation:**
   - IVarP for $N = 2$.

Numerical results

- **Comparison of HRP portfolios with benchmarks:**
  - HRP portfolios compared with global minimum variance portfolio (GMVP) and inverse-variance portfolio (IVarP).
  - Two versions of HRP: one with bisection split and another with dendrogram split.
- **Observations:**
  - GMVP shows concentration in two assets, while others are more diversified.
  - HRP portfolios show similar diversification to IVarP.
  - HRP portfolios exhibit slight improvement over IVarP.
  - Graphs estimated via heavy-tailed MRF methods may offer better performance than correlation-based methods.
  - HRP portfolios aim to balance diversification and risk management.
  - The choice of splitting method in HRP (bisection vs. dendrogram) may not significantly alter the diversification profile compared to IVarP.
  - The performance of HRP portfolios in terms of drawdown and P&L suggests potential advantages over traditional IVarP, especially when using advanced graph estimation methods.

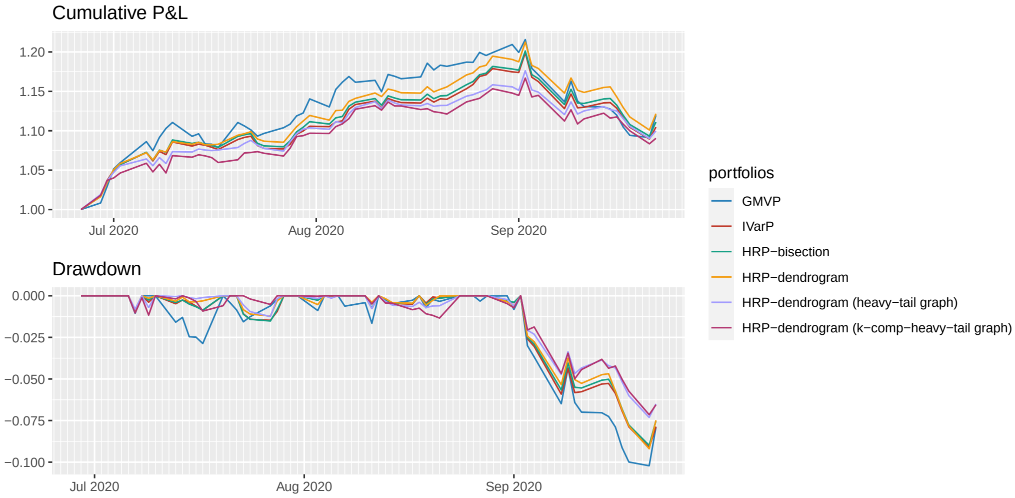
Numerical results

Portfolio allocation of HRP portfolios and benchmarks:



Numerical results

Backtest performance of HRP portfolios and benchmarks:



Outline

Hierarchical equal risk contribution (HERC) portfolio

- **Hierarchical equal risk contribution (HERC) portfolio overview:**
  - Introduced by (Raffinot 2018).
  - Refines and extends the hierarchical $1/N$ and HRP portfolios.
  - Incorporates early stopping based on the gap statistic for cluster selection.
  - Utilizes equal risk contribution (ERC) for weight allocation among clusters.

- **Key differences from previous approaches:**
  - **Early stopping with Gap statistic:**
    - Automatically selects the appropriate number of clusters.
    - Avoids clustering down to single assets.
  - **General equal risk contribution:**
    - Splits weights based on alternative risk measures (e.g., standard deviation, conditional value-at-risk).
    - Formula for two clusters:
$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \mathrm{RC}_1/(\mathrm{RC}_1 + \mathrm{RC}_2) \\ \mathrm{RC}_2/(\mathrm{RC}_1 + \mathrm{RC}_2) \end{bmatrix},$$
      where RC_i is the risk contribution of the ith cluster.

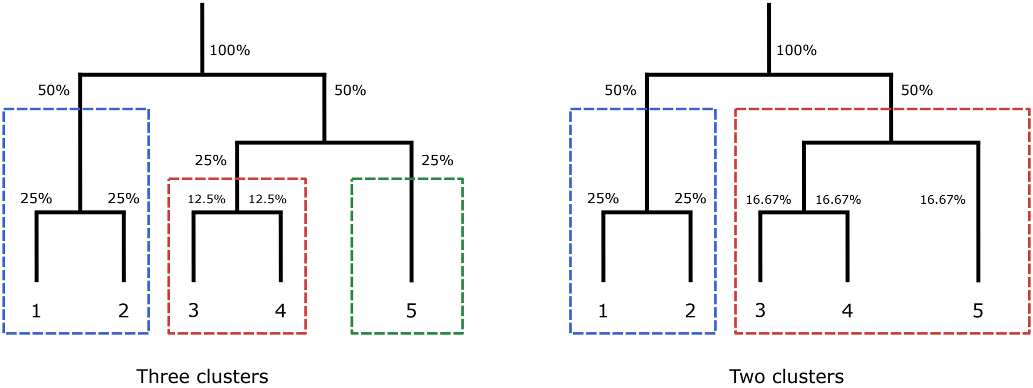
Hierarchical equal risk contribution (HERC) portfolio

- **Main findings:**
  - Hierarchical $1/N$ portfolio is a strong baseline.
  - HERC portfolios based on downside risk measures (especially conditional drawdown-at-risk) show statistically better risk-adjusted performances.

- **Illustration of early stopping:**
  - The next figure demonstrates early stopping in hierarchical clustering with a toy dendrogram.
  - Groups assets into clusters based on the gap statistic, avoiding overly granular clustering.

Hierarchical equal risk contribution (HERC) portfolio

Effect of early stopping in the hierarchical clustering process:



Hierarchical equal risk contribution (HERC) portfolio

Summary:

1. **Distance matrix:**
   - Correlation-based distance-of-distance matrix.
2. **Linkage method:**
   - Ward's method.
3. **Clustering stopping criterion:**
   - Gap statistic for optimal cluster selection.
4. **Splitting criterion:**
   - Follows the dendrogram structure.
5. **Intra-weight allocation:**
   - $1/N$ portfolio strategy.
6. **Inter-weight allocation:**
   - Equal risk contribution based on various risk measures.

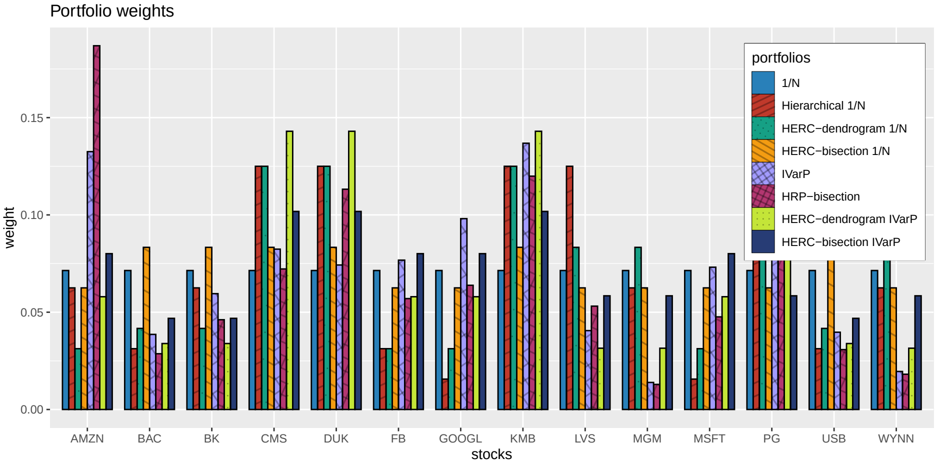
Conclusion: HERC portfolio represents a sophisticated approach to portfolio construction, balancing risk across clusters for improved risk-adjusted returns.

Numerical results

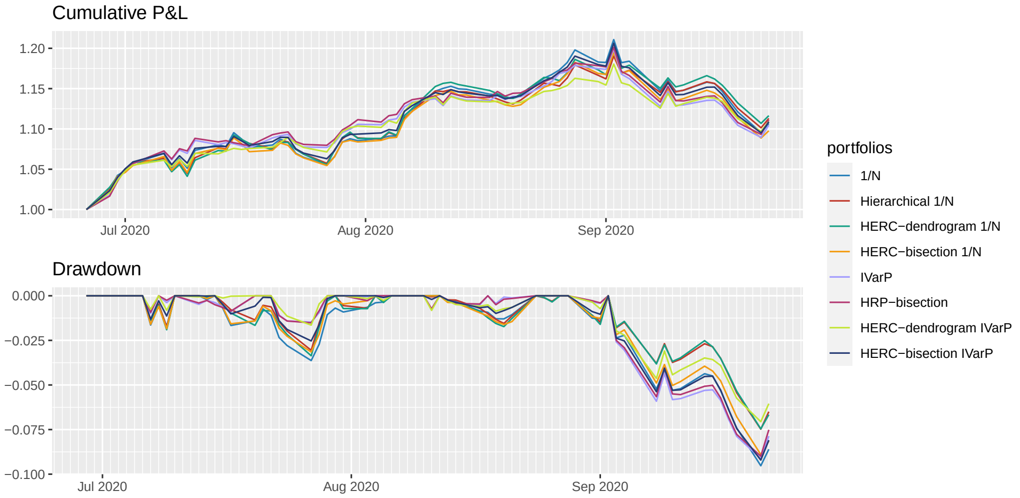
- **Simplification in weight splitting for HERC portfolios:**
  - For weight allocation, two risk contribution measures are considered:
    - $\text{RC}_i = 1$: Leads to a 50% - 50% split, similar to the hierarchical $1/N$ portfolio.
    - $\text{RC}_i = 1/\sigma_i^2$: Aligns with the inverse-variance portfolio (IVarP) formula, akin to the HRP portfolio.

- **Comparison of HERC portfolios with benchmarks:**
  - Benchmarks include: $1/N$ portfolio, hierarchical $1/N$ portfolio, inverse-variance portfolio (IVarP), and HRP portfolio.
  - Two versions of HERC portfolios are evaluated: one with bisection split and another with dendrogram split.

Portfolio allocation of HERC portfolios and benchmarks:



Backtest performance of HERC portfolios and benchmarks:



Numerical results

- **Observations and further evaluation:**
  - Difficult to draw definitive conclusions from a single backtest.
  - More exhaustive backtests are necessary to robustly assess the performance of HERC portfolios.
  - Future analysis should aim to evaluate the risk-adjusted returns and drawdown characteristics of HERC portfolios in various market conditions.

- **Conclusion:**
  - The HERC portfolio introduces a nuanced approach to portfolio construction by incorporating risk contributions and early stopping based on the gap statistic.
  - Its performance relative to traditional and hierarchical portfolio strategies warrants further empirical investigation to fully understand its benefits and limitations.

From portfolio risk minimization to hierarchical portfolios

- The basic structure of hierarchical portfolios is heuristic and suboptimal, which is understandable since the motivation was not optimality but stability against estimation errors.

- On the other hand, portfolios designed based on the minimization of some properly chosen measure of risk are not heuristic by definition but optimal according to the design criterion.

- Can we make an explicit connection between the two paradigms?

- Indeed, it is possible to design a continuum between hierarchical portfolios and optimally designed portfolios (Palomar 2024, sec. 12.3.4).

Outline

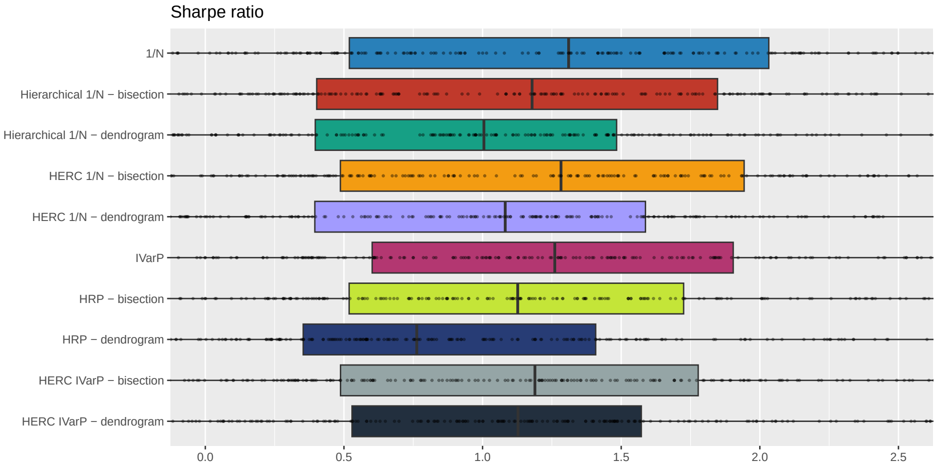
Numerical experiments

- **Overview:**
  - Conducted multiple randomized backtests using S&P 500 stocks from 2015-2020.
  - Generated 200 resamples with $N = 50$ stocks and a random two-year period.
  - Walk-forward backtest with a 1-year lookback, reoptimizing monthly.
  - Caution: Results are indicative and should be supplemented with more exhaustive backtests.

- **Observations:**
  - **Splitting**: Natural dendrogram splits might be expected to outperform, but it seems that bisection might provide more balanced clusters (while utilizing dendrogram ordering).
  - **Graph learning**: Sophisticated methods do not clearly outperform simple graph-based approaches.

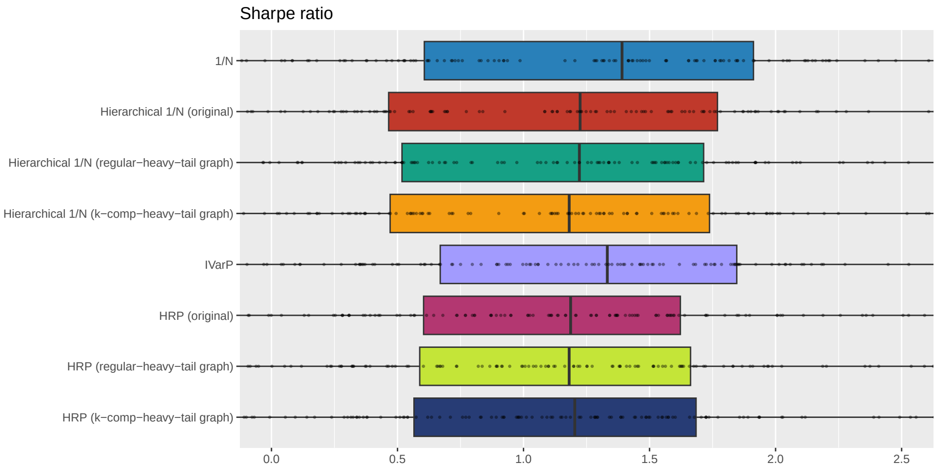
Numerical experiments - Splitting: bisection versus dendrogram

Comparison of graph-based portfolios: bisection versus dendrogram splitting:



Numerical experiments - Graph estimation: simple versus sophisticated

Comparison of graph-based portfolios: simple versus sophisticated graph learning methods:



Numerical experiments - Final comparison

- **Portfolios compared:**
  - Hierarchical $1/N$ portfolio.
  - HERC $1/N$ portfolio.
  - HRP portfolio.
  - HERC IVarP.
- **Benchmarks:**
  - $1/N$ portfolio.
  - IVarP.
- **Empirical results:**
  - The following table and figures show no significant differences among methods.
- **Conclusion:**
  - Further exhaustive comparison needed to draw clear conclusions.
  - Current analysis does not favor one graph-based portfolio method over others.

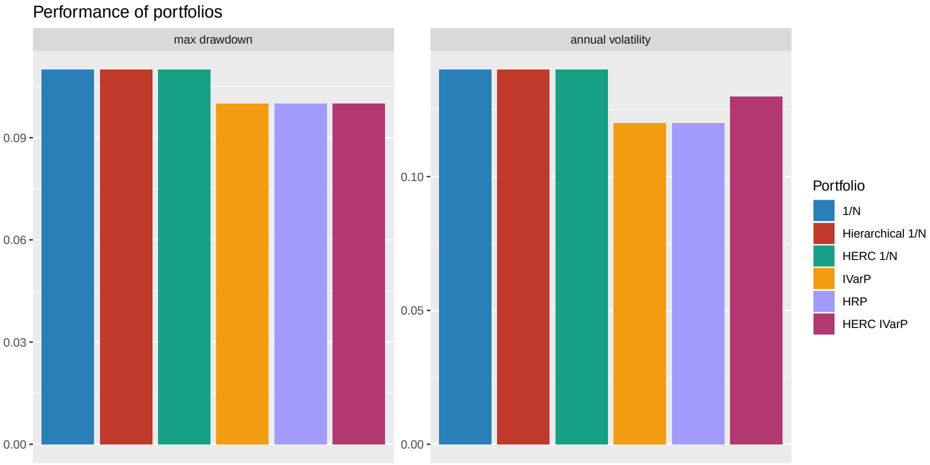
Numerical experiments - Final comparison

Comparison of selected graph-based portfolios: performance measures:

Portfolio	Sharpe ratio	annual return	annual volatility	max drawdown
1/N	1.01	14%	14%	11%
Hierarchical 1/N	0.81	12%	14%	11%
HERC 1/N	0.99	14%	14%	11%
IVarP	1.04	13%	12%	10%
HRP	0.91	11%	12%	10%
HERC IVarP	0.89	12%	13%	10%

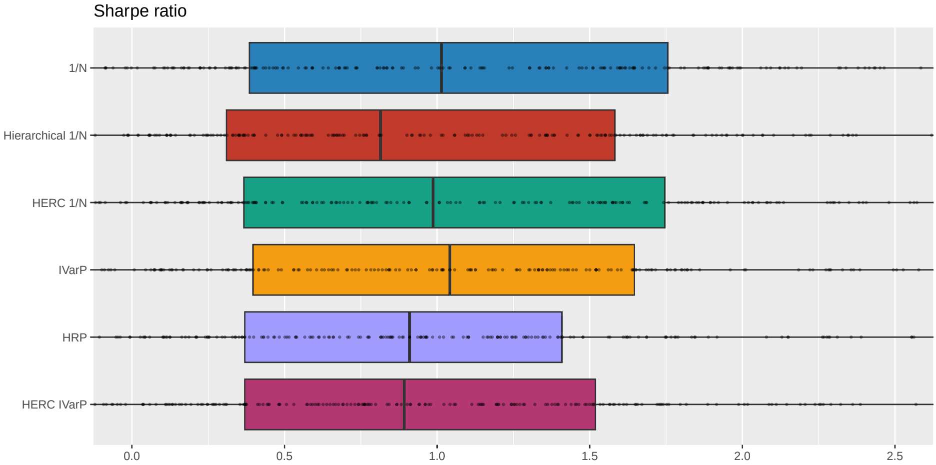
Numerical experiments - Final comparison

Comparison of selected graph-based portfolios: barplots of maximum drawdown and annualized volatility:



Numerical experiments - Final comparison

Comparison of selected graph-based portfolios: boxplots of Sharpe ratio:



Outline

Summary

Graphs compactly represent big data, revealing underlying structure and patterns.

Key takeaways for portfolio design using graphs:

- Graphs represent asset relationships: nodes are assets, edges are pairwise relationships.
- Financial graphs can be learned from data, e.g., based on heavy-tailed Markov random fields or k-component versions for clustered graphs (Palomar 2024, chap. 5).
- Hierarchical clustering partitions assets into clusters at different levels of detail.
- Graph information should be incorporated into portfolio formulation, with notable examples being hierarchical $1/N$, risk parity, and equal risk contribution portfolios.

References I

López de Prado, M. 2016. "Building Diversified Portfolios That Outperform Out of Sample." *Journal of Portfolio Management* 42 (4): 59–69.

Palomar, D. P. 2024. *Portfolio Optimization: Theory and Application*. Cambridge University Press.

Papenbrock, J. 2011. "Asset Clusters and Asset Networks in Financial Risk Management and Portfolio Optimization." PhD thesis, Karlsruher Institute für Technologie.

Raffinot, T. 2018. "The Hierarchical Equal Risk Contribution Portfolio." *SSRN Electronic Journal*. https://dx.doi.org/10.2139/ssrn.3237540.