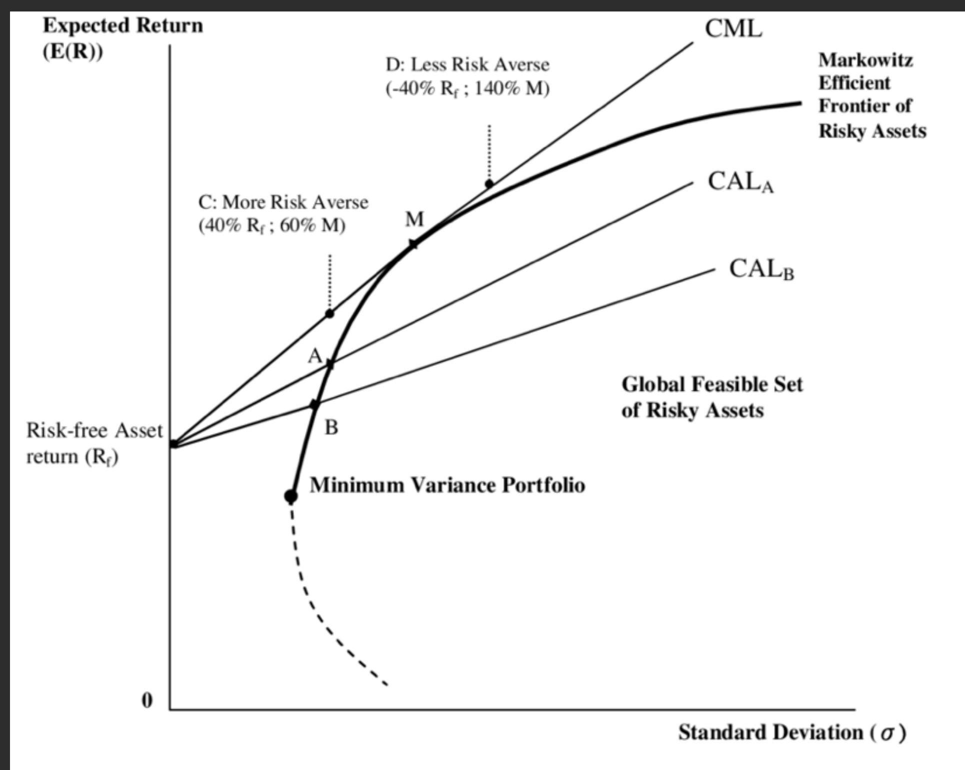
Building Diversified Portfolios That Outperform Out of Sample

A REVIEW OF HIERARCHICAL RISK PARITY (HRP)

Paper by: Marcos López de Prado (2016)

Presented by: Bramastya Zaki

THE CORE PROBLEM



For over 60 years, Markowitz's Mean-Variance optimization has been the theoretical standard, yet it consistently fails in practice due to three critical issues:

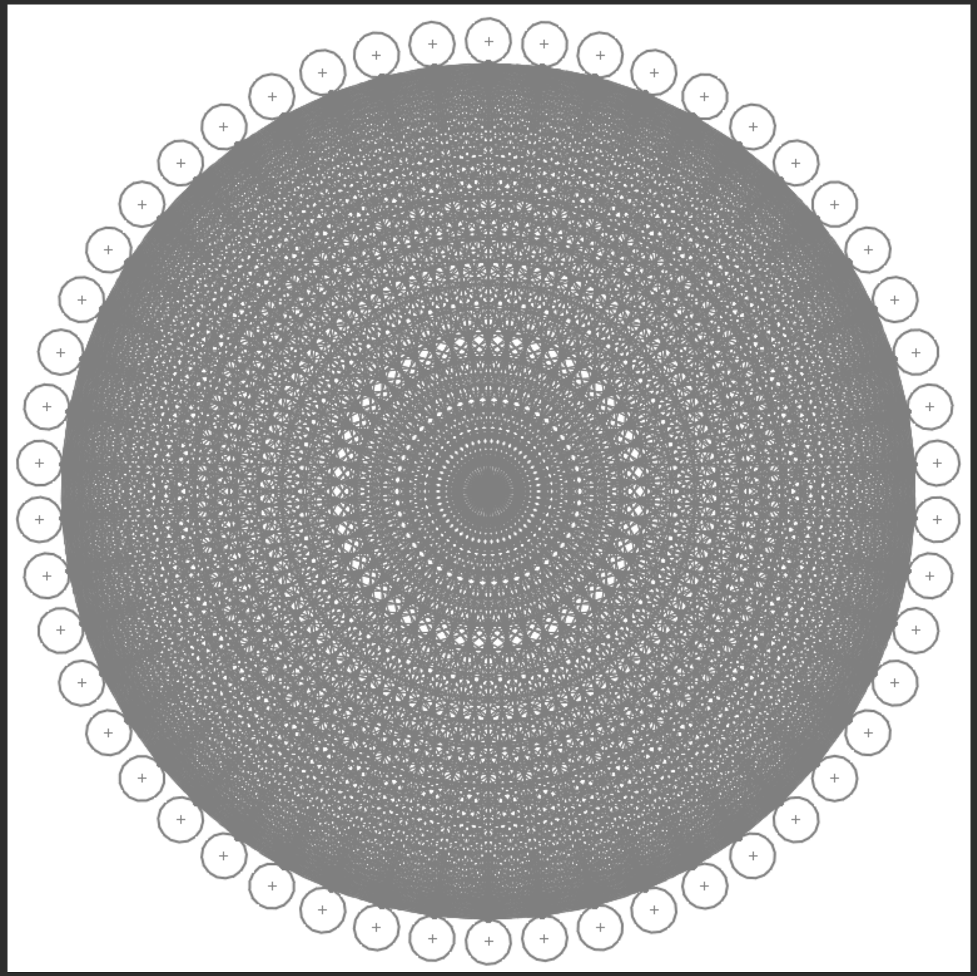
- **Instability:** Small changes to inputs lead to drastically different portfolios.
- **Concentration:** The method over-allocates to a few assets, defeating the purpose of diversification.
- **Underperformance:** It often loses to a naïve $1/N$ portfolio out-of-sample.

The **root cause** is its reliance on **matrix inversion.** The stability of this operation is measured by the condition number of the covariance matrix, which explodes as assets become more correlated.

$$\text{Condition Number} = \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} \to \infty \text{ as asset correlation } \uparrow$$

This leads to **Markowitz's Curse:** the more you need diversification, the more unstable the solution becomes.

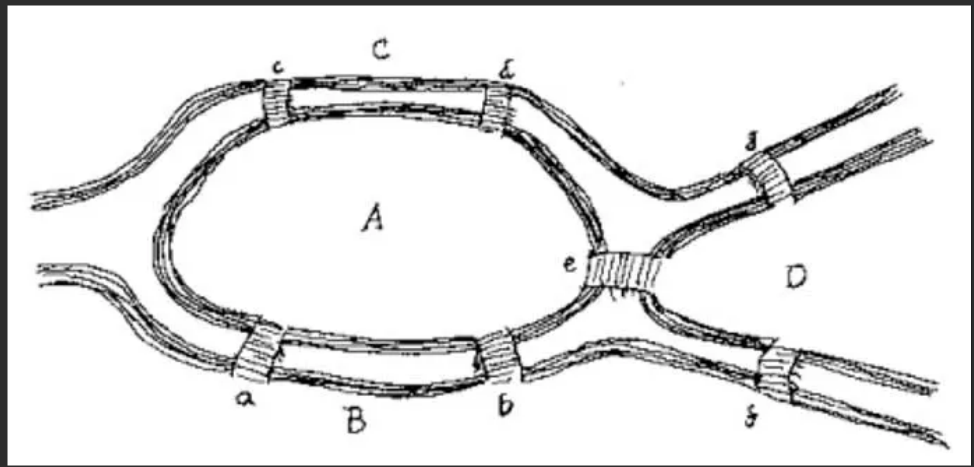
WHY DO TRADITIONAL OPTIMIZERS FAIL?



The instability from matrix inversion stems from two deep seated problems in financial data:

1. **Noise Induced Instability** Financial data has an extremely low signal to noise ratio. Most information in a covariance matrix is statistically indistinguishable from random noise, making its inversion highly unreliable.
2. **Signal Induced Instability** Even with perfect information, the signal itself high correlation between assets causes instability. As correlation increases, the covariance matrix becomes ill conditioned, which magnifies the impact of any input guesses on the final portfolio weights.

A NEW PERSPECTIVE: FROM GEOMETRY TO TOPOLOGY



The paper's key insight is to change the mathematical framework used to represent the relationships between assets.

- **Traditional Approach: Complete Graph** A covariance matrix treats the universe as a fully connected graph where every asset is a potential substitute for every other. Small estimation errors can propagate across the entire system, causing instability.
- **HRP Approach: Hierarchical Tree** HRP simplifies the structure to a tree, where assets are only compared to their closest peers within a hierarchy. This structure naturally contains estimation errors and is more robust.
- This is like choosing a simple **subway map (topology)** over a complex **geographic map (geometry)** to navigate a city.

The Proposed Solution: Hierarchical Risk Parity (HRP)

HRP is a novel portfolio construction method that uses graph theory and machine learning to build portfolios **without requiring matrix inversion.**

It operates in three distinct stages:

1. **Tree Clustering:** Group assets into a hierarchy based on their correlation.
2. **Quasi-Diagonalization:** Reorder the covariance matrix according to this hierarchy.
3. **Recursive Bisection:** Allocate capital down the tree, splitting weights based on cluster risk.

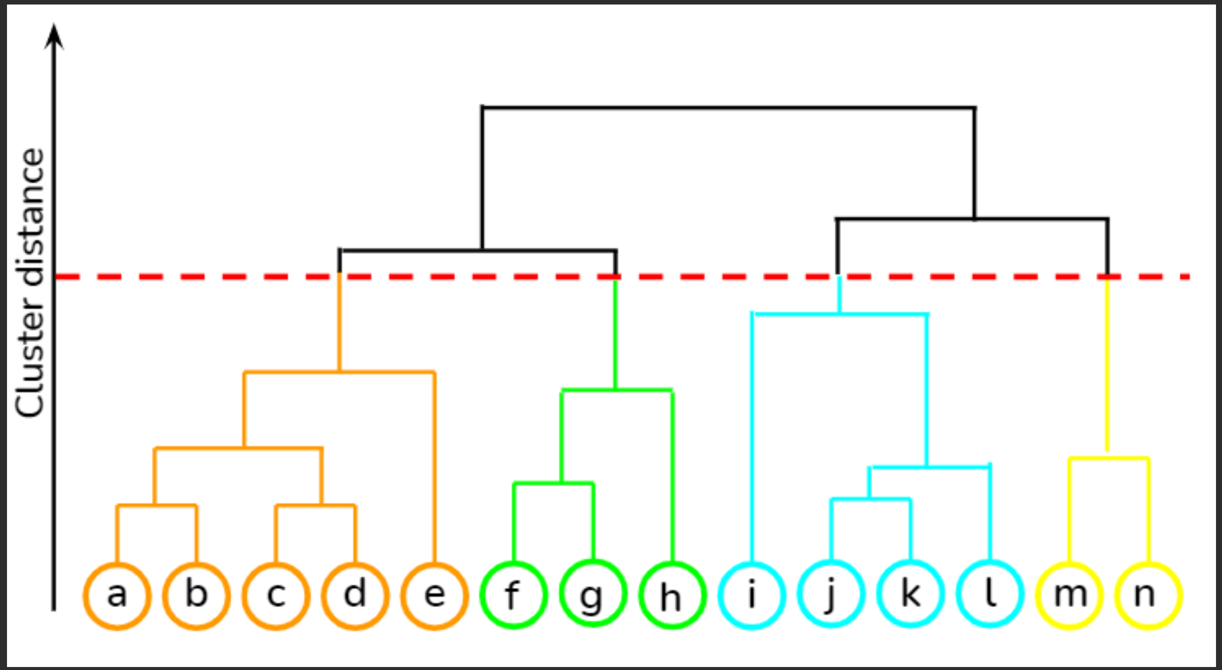
This approach is also far more data efficient. A standard 50 asset Markowitz portfolio needs about five years of data to estimate a stable covariance matrix. HRP can achieve a robust allocation with only two months of data, making its assumptions more realistic.

HRP STEP 1: TREE CLUSTERING

The goal is to discover the natural hierarchical structure of the assets.

First, the correlation matrix is converted into a distance matrix, which is a proper metric space. A common metric is:

$$d_{i,j} = \sqrt{\frac{1}{2}(1 - \rho_{i,j})}$$

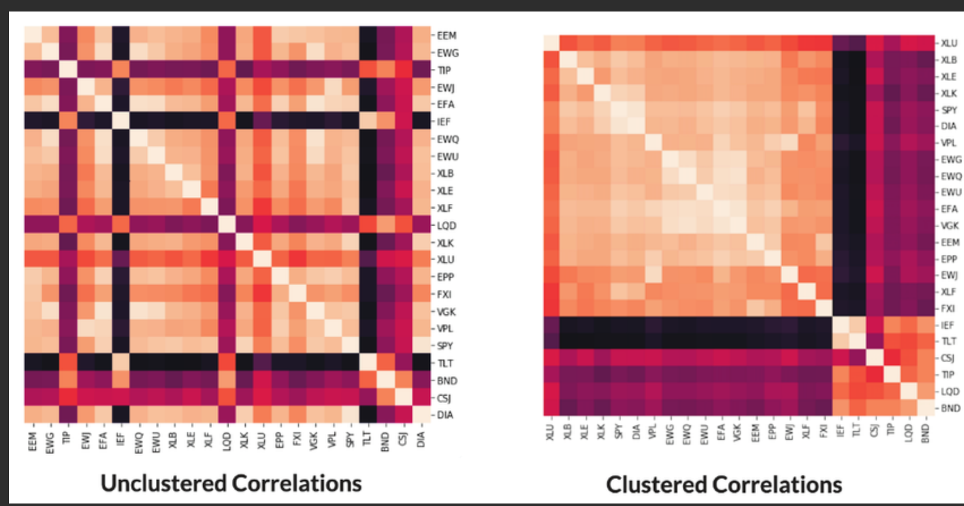


Next, a hierarchical clustering algorithm groups the assets. It iteratively pairs the closest assets or clusters until a single tree containing all assets is formed.

The output is a dendrogram that visually represents the relationships, with closely related assets on the same branches.

HRP STEP 2: QUASI DIAGONALIZATION

This stage reorganizes the covariance matrix according to the tree structure from Step 1.



- The rows and columns of the matrix are reordered so that assets that are close in the tree structure are placed next to each other in the matrix.
- This creates a **quasi-diagonal** structure, where the largest values (highest covariance) are concentrated in blocks along the main diagonal.
- This is achieved without changing the underlying basis of the investments, unlike methods like PCA.

HRP Step 3: Recursive Bisection

This is where capital allocation happens, flowing top-down through the tree structure.

1. **Initialize:** Start with a single cluster containing all assets and 100% of the portfolio weight.
2. **Bisect:** Split the cluster into two sub-clusters based on the reordered matrix.
3. **Calculate Cluster Risk:** For each sub-cluster, compute its variance using an inverse-variance allocation *within* that sub-cluster.
4. **Allocate:** Distribute the parent cluster's weight between the two sub-clusters in inverse proportion to their respective variances.
5. **Recurse:** Repeat this splitting and allocation process down the tree until each individual asset has received its final weight.

A Comparison of Allocations

HRP produces allocations that provide a structured compromise between the extremes of other methods.

Method	Characteristics
CLA (Min-Variance)	Highly **concentrated.** Tends to give zero weight to many assets, making it vulnerable to idiosyncratic shocks affecting its few large positions.
IVP (Risk Parity)	**Ignores correlation.** Spreads weight based only on individual asset variance, making it vulnerable to systemic shocks that affect correlated assets.
HRP	**Structured Diversification.** Diversifies across both clusters and individual items, providing a balance that is more resilient to both types of shocks.

The Key Result: Superior Out of Sample Performance

The paper's central claim is validated through 10,000 Monte Carlo simulations designed to test real-world, out-of-sample performance.

Out-of-Sample Variance (Lower is Better)

$$\sigma^2_{\mathrm{CLA}} = 0.1157 \quad \text{(Worst - despite being "optimal" in-sample)}$$
$$\sigma^2_{\mathrm{IVP}} = 0.0928$$
$$\sigma^2_{\mathrm{HRP}} = 0.0671 \quad \text{(Best)}$$

The performance improvement is significant:

- HRP delivered **72% lower variance** than CLA, translating to a potential **31% Sharpe ratio boost.**
- HRP delivered **38% lower variance** than traditional risk parity (IVP).

The Big Picture: Beyond HRP

López de Prado emphasizes that machine learning's true value in asset management extends far beyond this single algorithm.

The goal of ML is often **not to ~~predict prices~~**, which he calls the "least interesting" application due to high noise and risk of overfitting.

Instead, its real power lies in solving other critical problems:

- **Optimal "Bet Sizing":** Using ML to dynamically adjust position sizes based on an algorithm's confidence, much like a professional poker player.
- **Detecting Structural Breaks:** Identifying regime changes to know when a strategy may no longer be valid.
- **Diversifying the *Methods*:** The most robust approach is not to rely on a single allocator, but to use an **ensemble of methods.**

Key Advantages of HRP

HRP is a powerful modern alternative because it is:

- **Robust:** It can handle singular or ill-conditioned covariance matrices, an impossible task for traditional quadratic optimizers.
- **Stable:** The hierarchical tree structure contains and limits the propagation of estimation errors.
- **Interpretable:** The top-down allocation mimics how many portfolio managers think (e.g., Asset Class → Sector → Security), making decisions transparent.
- **Flexible:** The framework can be easily adapted to include constraints or investor views.
- **Scalable:** With a deterministic complexity of $O(N^2 \log_2 N)$, the algorithm is fast and can be applied to massive datasets.

Critical Assessment & Limitations

A balanced view of the paper reveals both its strengths and areas for further exploration.

Strengths

- Addresses a well-documented, 60-year-old problem in finance.
- Built on a rigorous mathematical foundation (proof of metric space included).
- Validated with comprehensive Monte Carlo simulations, not just a single backtest.
- Practical and implementable.

Open Questions & Limitations

- **Choice of Hyperparameters:** How sensitive are the results to the choice of distance metric or clustering linkage method?
- **Return Forecasts:** The base version is purely risk based. How does it perform when return forecasts are incorporated?
- **Regime Changes:** The paper does not explicitly test performance during major structural breaks or financial crises.

Most Interesting Insight: The problem isn't just the optimization objective—it's the mathematical representation of the system. Changing from a fully connected graph to a hierarchical tree solves the same economic problem in a more robust way.

Thank You