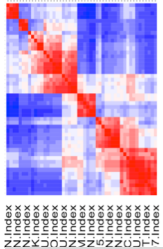
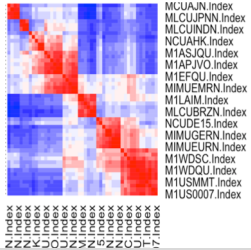
Market Regimes Using Corr Matrix & Clustering

Measure 1: Find the correlation matrix between a pair of window

corr matrix @ **1st** window

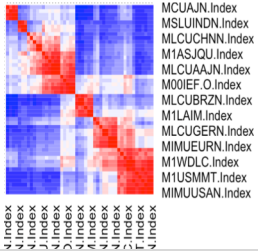


corr matrix @ **2nd** window



...

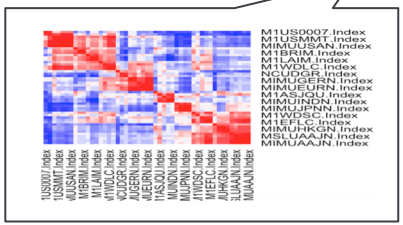
corr matrix @ **last** window



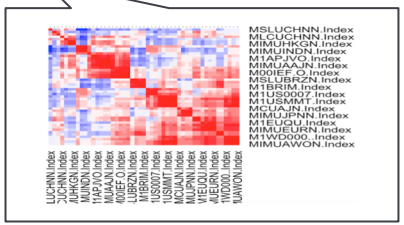
Red: positively correlated **Blue**: negatively correlated **White**: no correlations

Step 2: Calculate Difference Score between a pair of matrices

$$\zeta(t_1, t_2) \equiv \left\langle \left| C_{ij}(t_1) - C_{ij}(t_2) \right| \right\rangle_{ij}$$



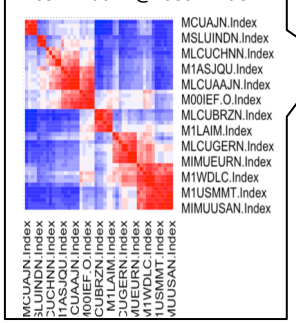
corr matrix @ 1st window



corr matrix @ 2nd window

Measure 2: Find the largest 5 eigenvalues of each correlation matrix

corr matrix @ **last** window



Explains >80% information

Largest 5 Eigenvalues:

30.70
21.51
7.14
4.54
3.40

In vector format:

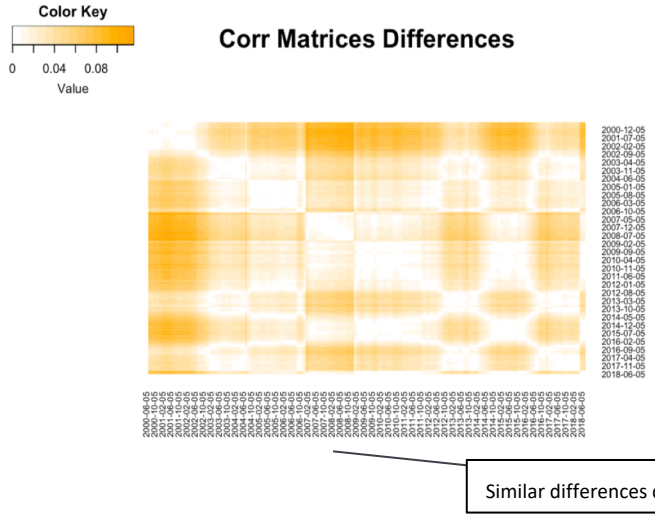
$$\begin{pmatrix} 30.70 \\ 21.51 \\ 7.14 \\ 4.54 \\ 3.40 \end{pmatrix}$$

Step 2: Calculate Difference Score between a pair of vectors

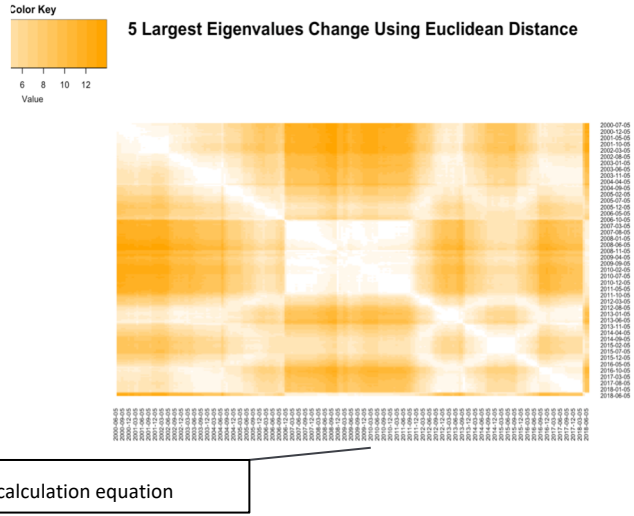
- Euclidean Distance $\sqrt{\sum |x_i - y_i|^2}$
- Cosine Distance $\frac{\sum x_i y_i}{\sqrt{\sum x_i^2}\sqrt{\sum y_i^2}}$

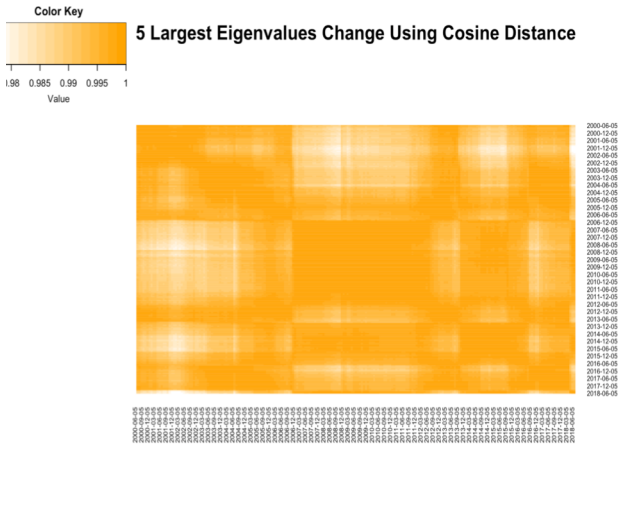
Two Measures Heat Map Comparison

Correlation Matrix Differences



5 Largest Eigenvalue Differences





Similar differences calculation equation

- Their scales are different
  - 5 Largest Eigenvalues Change Using Cosine Distance - **tiny**
  - Correlation matrix differences- **small**
  - 5 Largest Eigenvalues Change Using Euclidean Distance - **big**
- These three graphs look different, though the “Correlation Matrix Differences” graph and the “5 Largest Eigenvalues Change Using Euclidean Distance” graph look similar

PAM Method Clustering

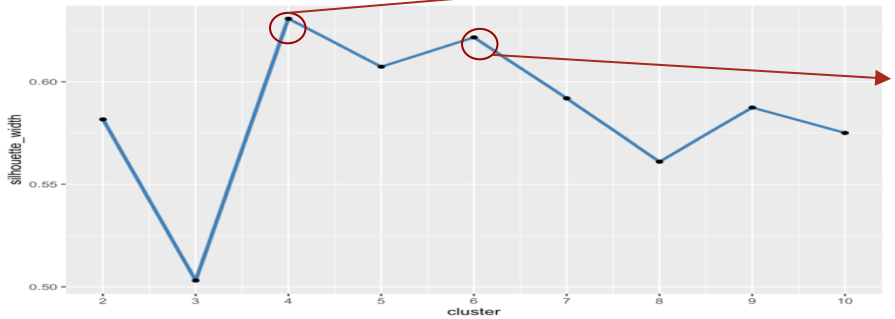
PAM method

- PAM stands for “partition around medoids”. The algorithm is intended to find a sequence of objects called medoids that are centrally located in clusters.
- Input: Dissimilarity matrix
  - correlation matrix differences
  - 5 Largest Eigenvalues Cosine Distance
  - 5 Largest Eigenvalues Euclidean Distance

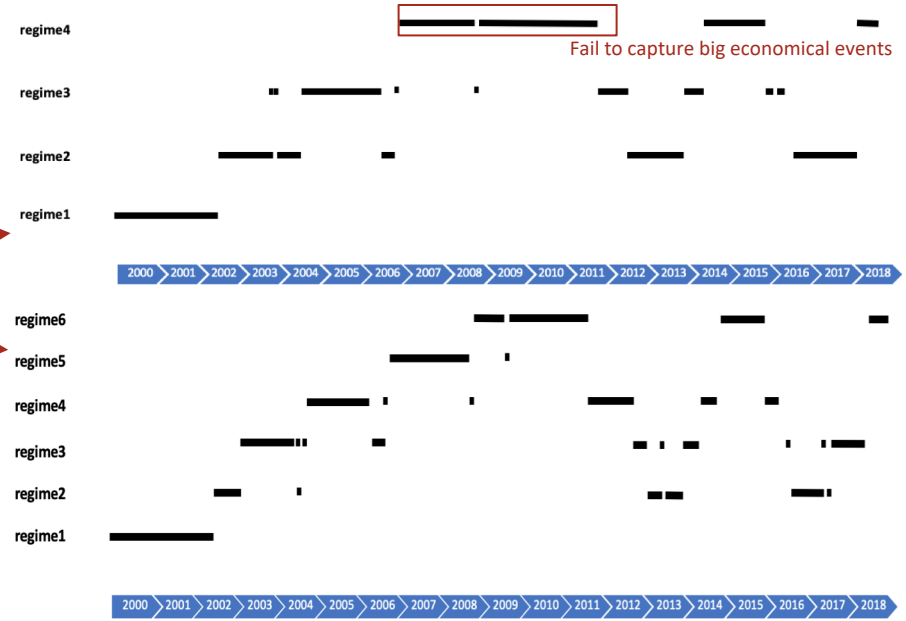
Note: use 2 yrs data first because it’s preferred

Silhouette analysis

- reflects how well points fit in their respective clusters.
- number of optimal clusters are 4 and 6



Eg: Silhouette graph of Corr Matrix Diff



Optimization: A New Clustering Method Construction

Use windows as observations, and assets & factors as features

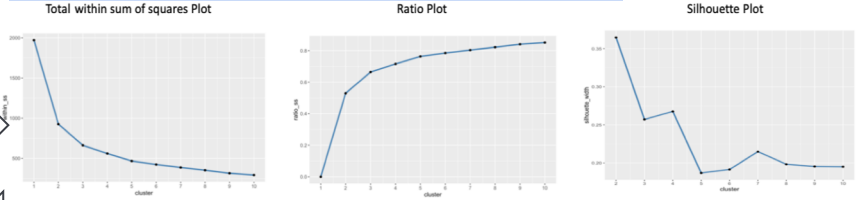
1. Within each window, find the sum of each asset&factor returns

Window1:
Row1 (factor1, factor2, ..., factor49, asset1, asset2, … , asset 51)
Row2 (factor1, factor2, ..., factor49, asset1, asset2, … , asset 51)
...
Row515 (factor1, factor2, ..., factor49, asset1, asset2, … , asset 51)
---
colSum sum factor1, … sum factor49, sum asset1, … sum asset51

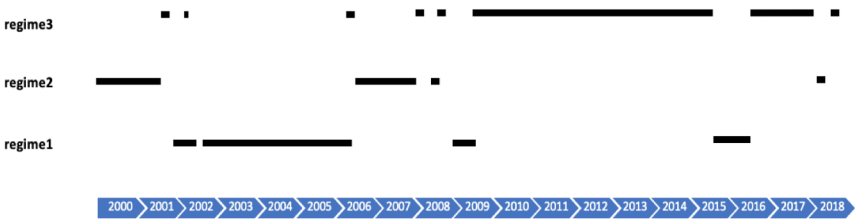
2. Cluster on observations (windows)

Window1 (sum factor1,..., sum factor49, sum asset1, … , sum asset 51)
Window2 (sum factor1,..., sum factor49, sum asset1, … , sum asset 51)
....
Last Window (sum factor1,...,sum factor49,sum asset1, … ,sum asset 51)

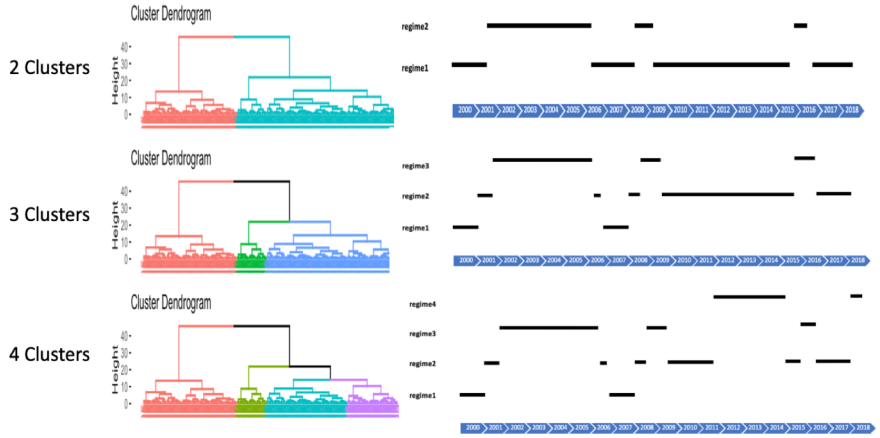
K-means: determining optimal regime numbers by graph



Optimal: 3 Clusters



Hierarchical Clustering



Optimization & Regime Selection

➔ **20** total market regime combinations:
4 time period (0.5 year, 1 year, 1.5 year, 2year) * 3 difference measures (corr matrix, 5 Largest Eigenvalues Cosine & Euclidean)
+ 8 (new method - k-means & hierarchical clustering for 0.5 year, 1 year, 1.5 year, 2year)

Which one is the best?

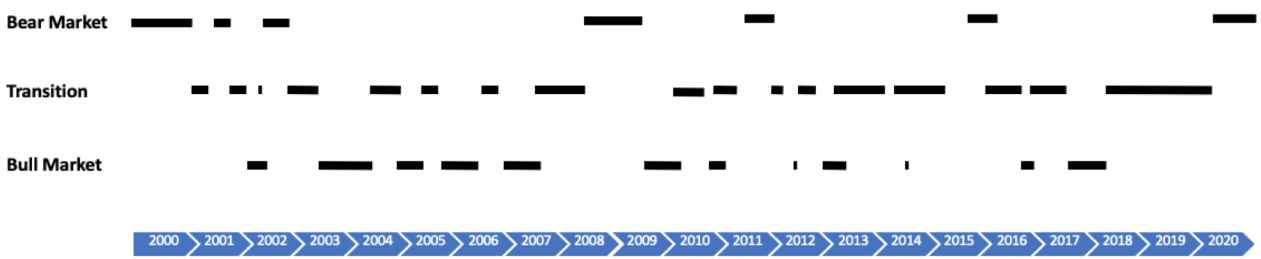
Criteria 1: Define Market Regime by Economic Cycle

- Regime should capture bull market (2002-2007 & 2009-2020) and major bear market (2000-2002, 2007-2009, 2020)

Criteria 2: Define Market Regime by volatility

- Step1: find the volatility of each assets under the same regime
- Step2: Within one regime, average all assets' volatilities (each regime has one averaged volatility value)
- **Bull markets**: slow and steady, low volatility **Bear markets**: assets lose value and prices become volatile

Preferred - Optimization method, Half year, 3 Clusters



Why it's good?

- Economic life cycle: captured bear & bull
- volatility
  Bear: 0.01741643 high
  Bull: 0.0101847 low
  Transition:0.01045842