PCA as a Tool for Analyzing the Market

Efe Yigitbasi

Content

- Modeling the Market
- Dimension Reduction and PCA
- An Example: Stock Market
- Another Example: Bond Market
- Conclusions

Modeling the Market

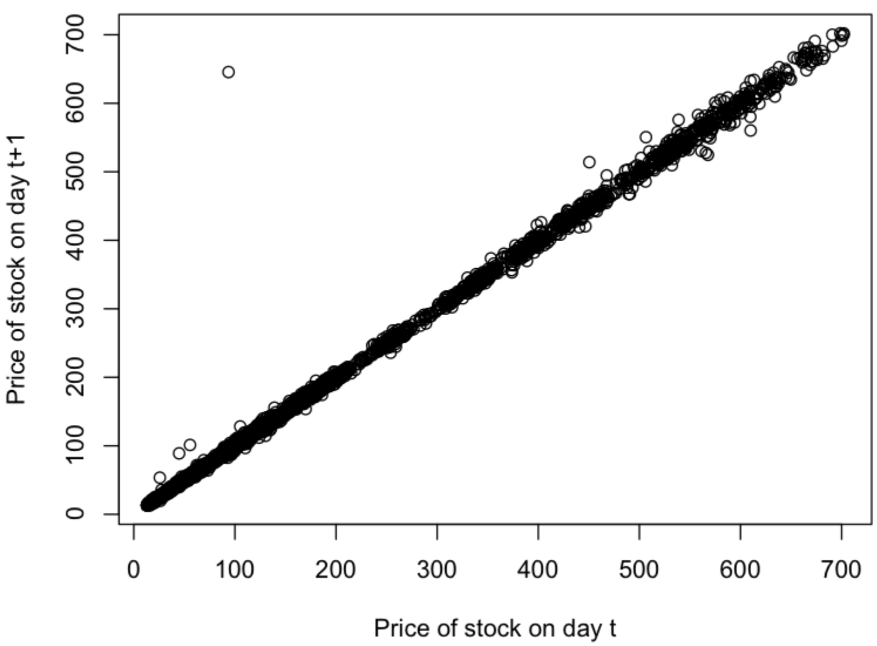
- We have to find a way to choose parameters for our models.
- Our parameters should be invariant in time.
- We should be able to reconstruct prices from invariants.

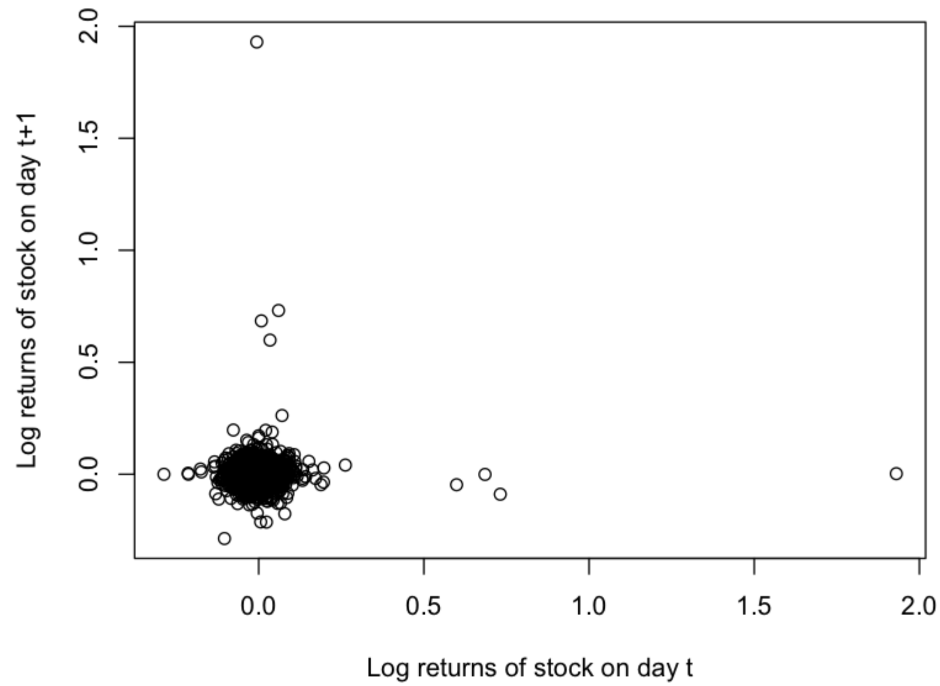
Invariants

- Stocks:
  - Price is not an invariant.
  - Logarithmic returns is.
- Bonds:
  - Price is not an invariant.
  - Yield to maturity is.

Invariants

- The difference between the invariants and prices:





Invariants

- Logarithmic (compounded) returns for a stock:

$$C_{t,\tau} \equiv \ln\left(\frac{P_t}{P_{t-\tau}}\right).$$

- Yield to maturity for bonds:

$$Y_t^{(v)} \equiv -\frac{1}{v}\ln\left(Z_t^{(t+v)}\right).$$

Dimension Reduction

- In general our invariants can be expressed as a single multivariate parameter:

$$X_{T \times K}(Y, C)$$

- Price will be a function of this parameter:

$$P_{T \times K} = f(X_{T \times K})$$

Dimension Reduction

- Usually the dimension of this parameter is larger than the actual degree of randomness in our market.

$$K > \text{Number of independent dimensions}$$

- Reasons:
  - Derivatives.
  - Hidden dependencies.

Dimension Reduction

- One method of reducing dimensions is Principal Component Analysis (PCA).

- We assume our parameter is a combination of some common factors and small perturbations:

$$\mathbf{X} \equiv \mathbf{q} + \mathbf{B}\mathbf{F}(\mathbf{X}) + \mathbf{U}.$$

Principal Component Analysis

- Construct the covariance matrix:

$$C_{K\times K} = \frac{X'X}{T}$$

- Find the eigenvalues and eigenvectors of the covariance matrix.

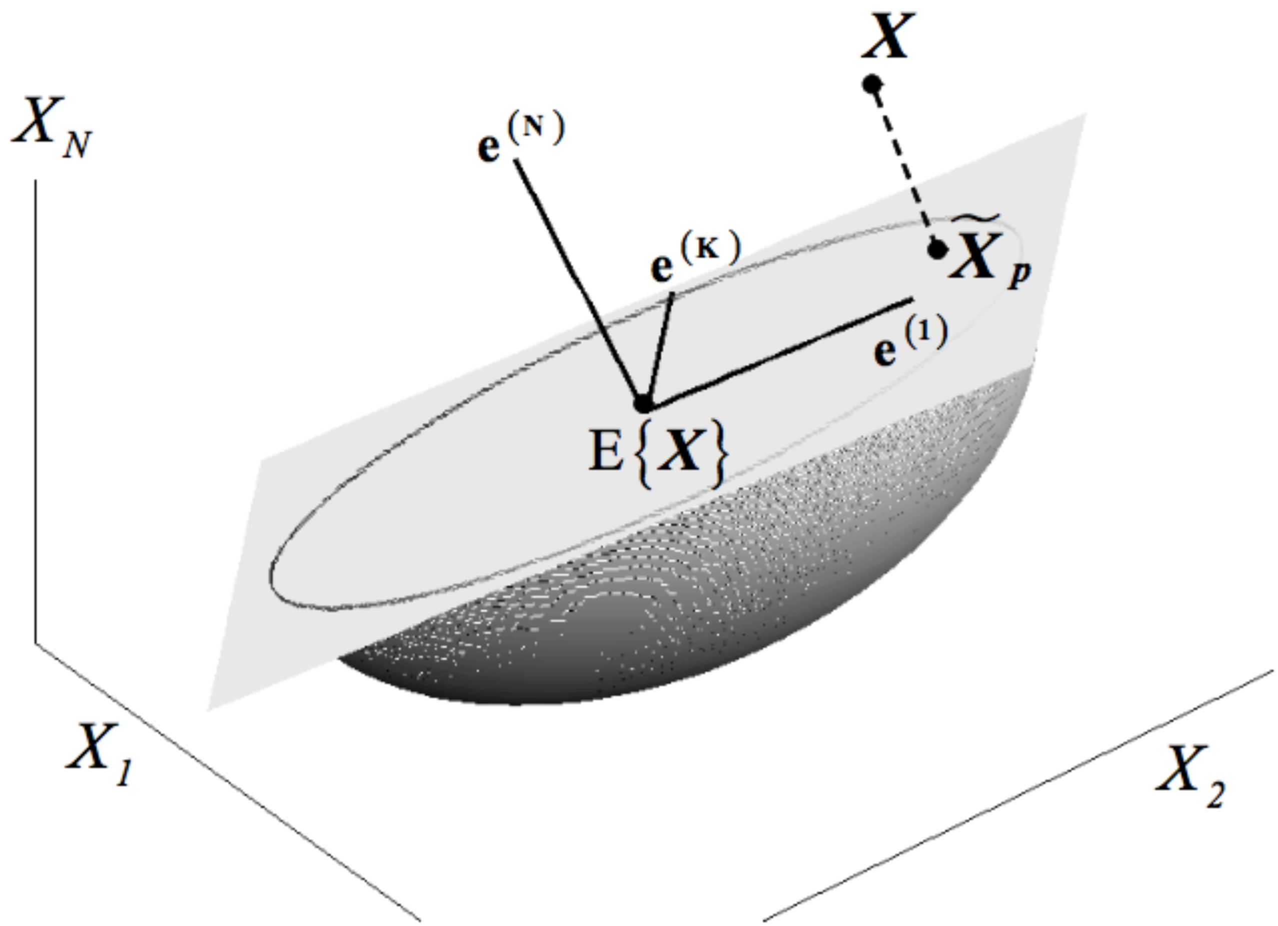
Principal Component Analysis

- The eigenvectors are the principal axes of the location-dispersion ellipsoid.

- They represent the directions of most variance.

- The dimension of the ellipsoid can be reduced depending on the eigenvalues.

Principal Component Analysis



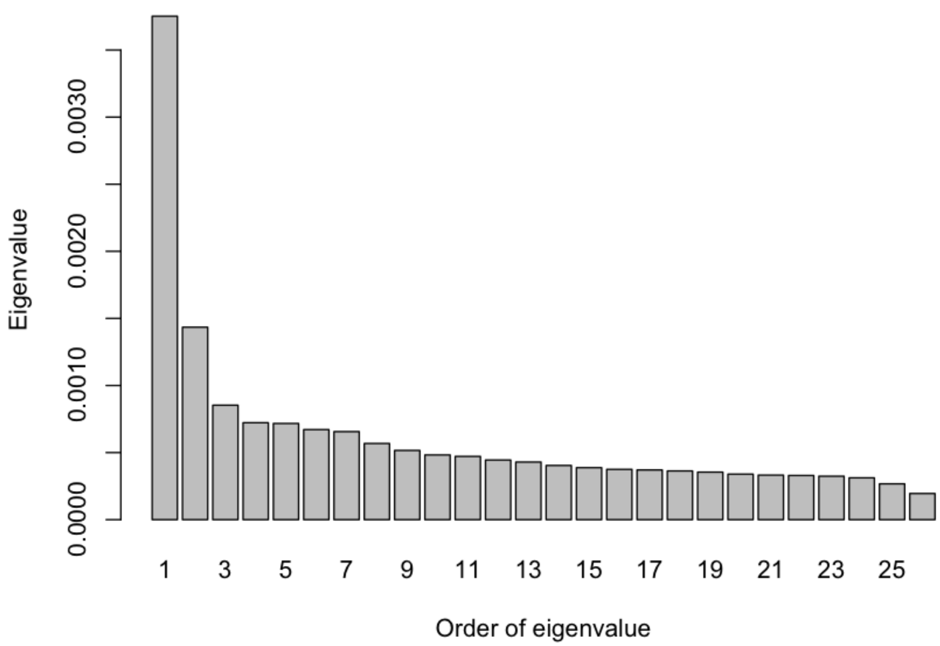
An Example: Stock Market

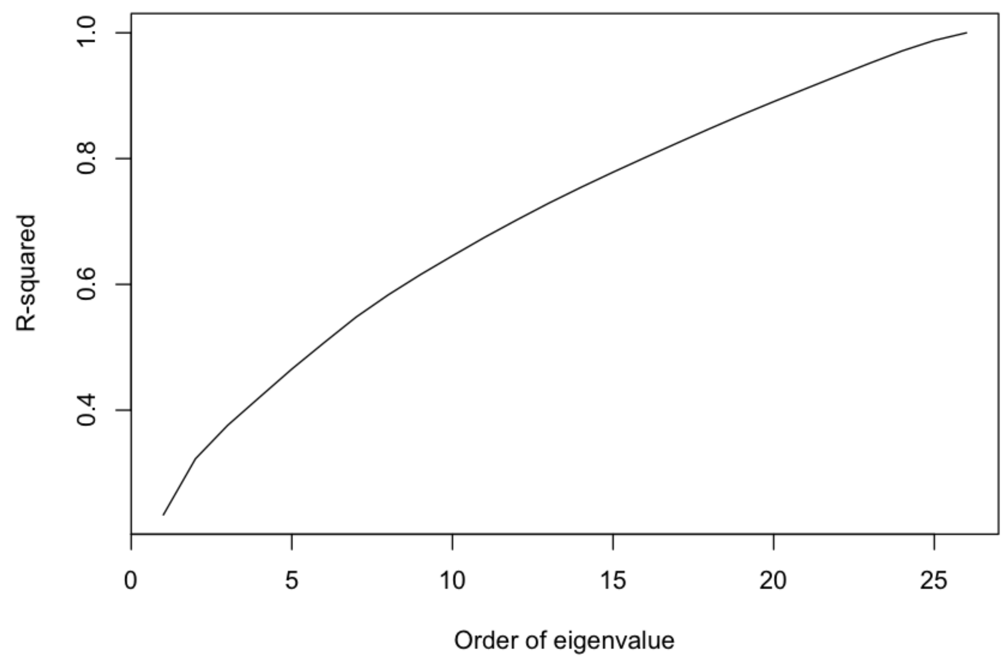
- 26 stocks from the Dow-Jones Index.
- Daily close prices from: 1/1/1990 to: 4/20/2015
- Use R-squared for measuring the explanatory power of the first N eigenvalues:

$$R^2 = \frac{\sum_{i=1}^{N} \lambda_i}{\sum_{i=1}^{K} \lambda_i}$$

An Example: Stock Market

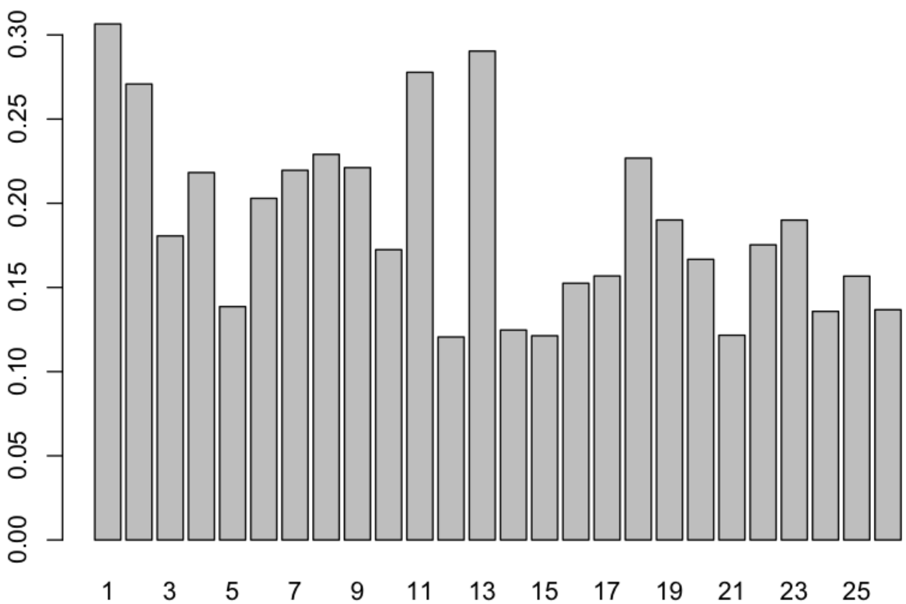
A look at eigenvalues



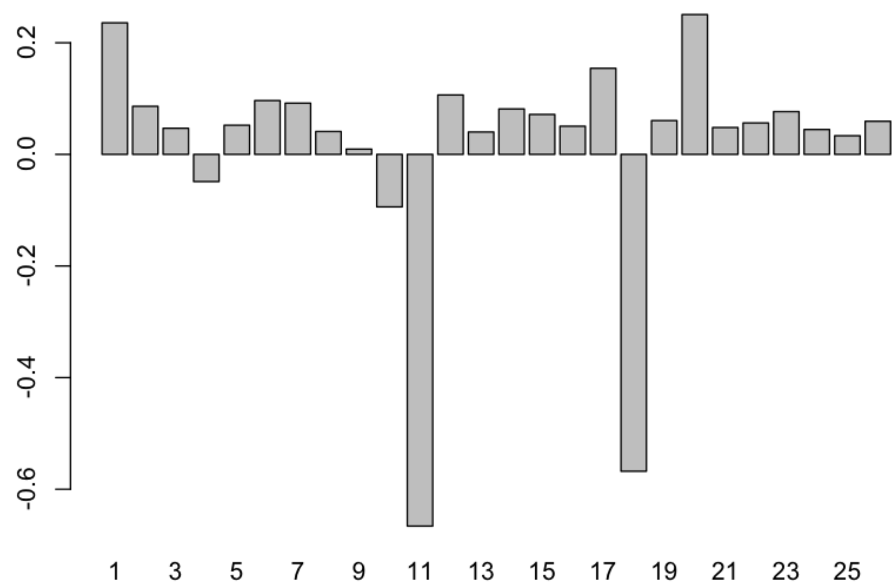


An Example: Stock Market

A look at eigenvectors



Bar plot of the first eigenvector



Bar plot of the third eigenvector

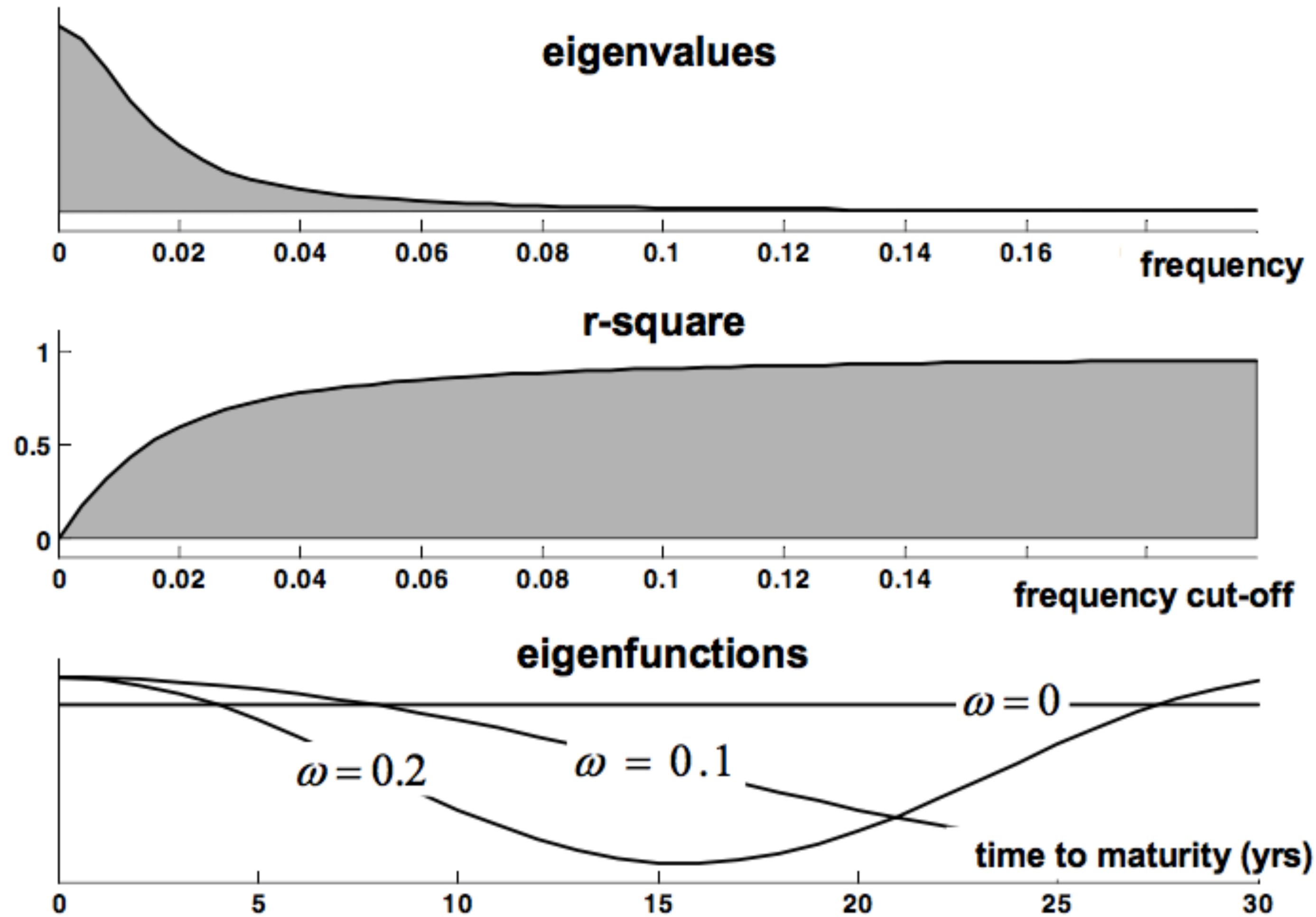
Another Example: Bond Market

- Daily US Treasury Bond yield rates from: 2/9/2006 to: 4/20/2015.
- Maturities: 1 Mo, 3 Mo, 6 Mo, 1 yr, 2 yr, 3 yr, 5 yr, 7 yr, 10 yr, 20 yr, 30 yr.
- Using the same R-squared as stock market.

Another Example: Bond Market

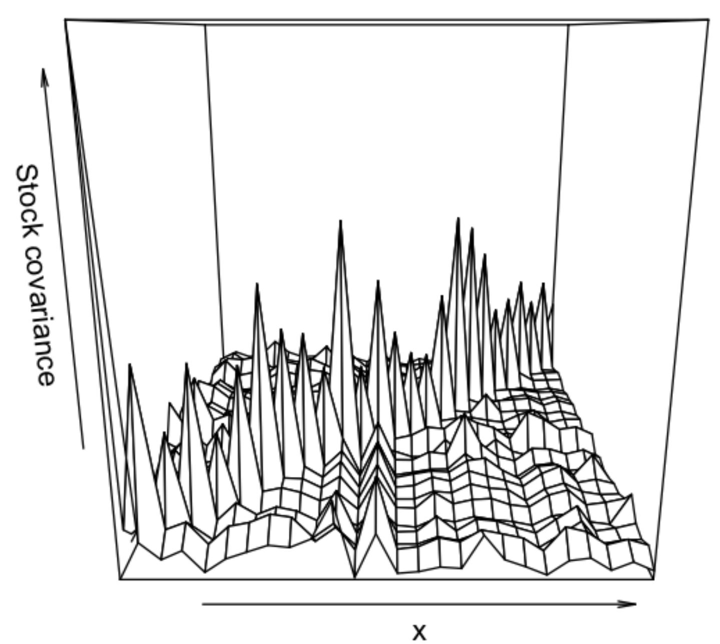
- The covariance matrix has some nice properties:
    - Smooth in both of its arguments.
    - Nearly constant diagonal terms.
    - Only depends on one parameter.
- Infinite dimensional case -> Toeplitz Operator.

Another Example: Bond Market

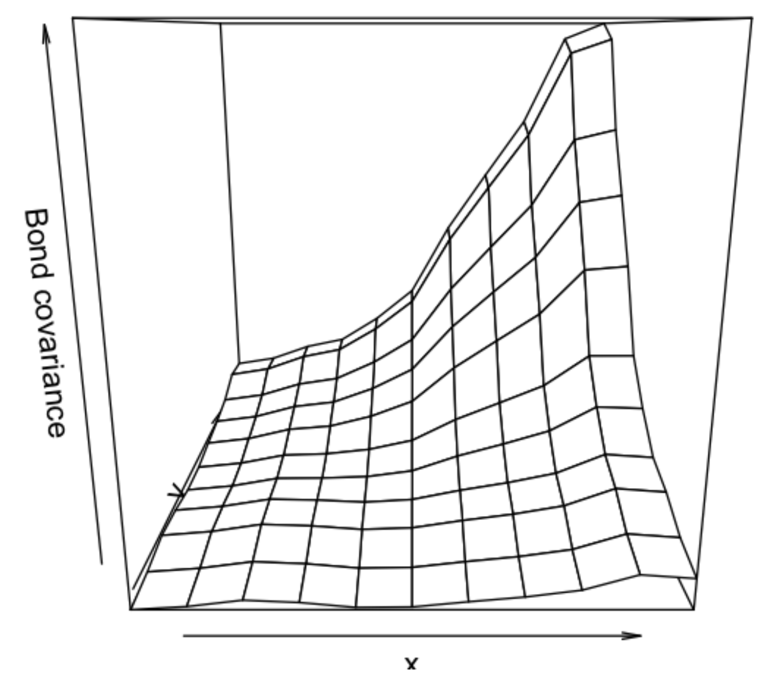


Another Example: Bond Market

Representation of covariance matrix of stocks

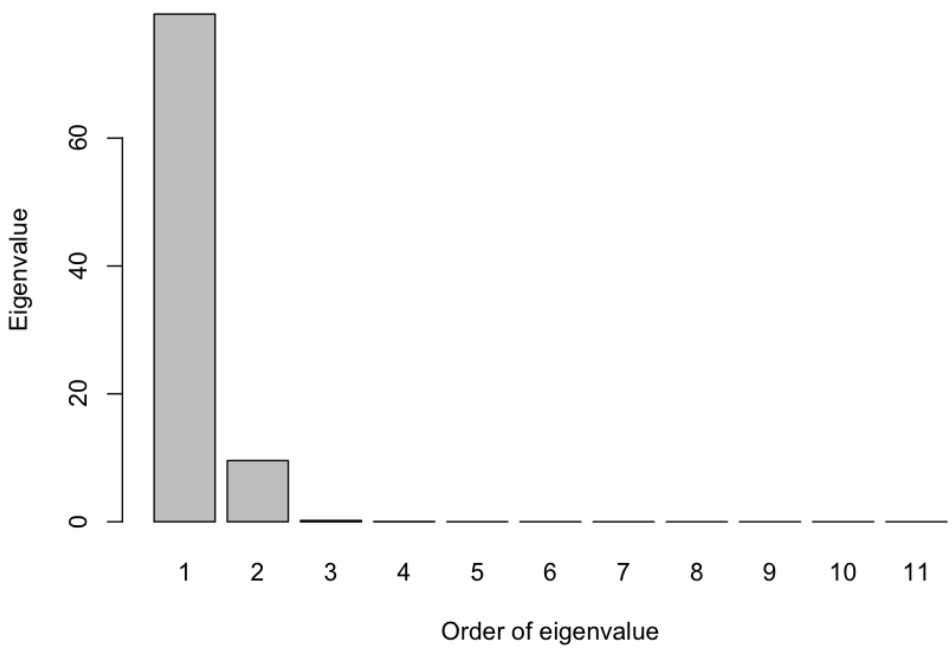


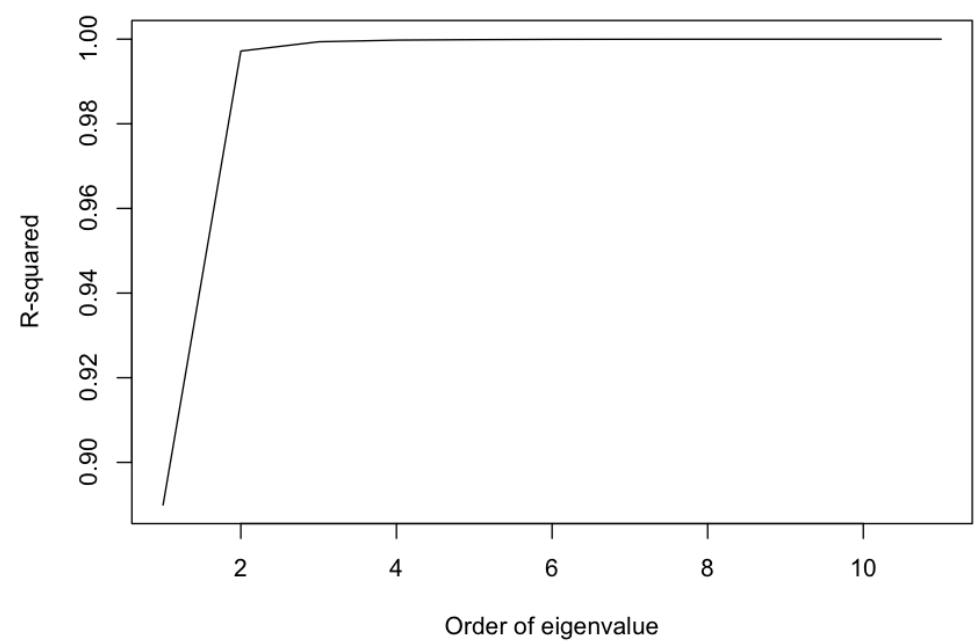
Representation of covariance matrix of bonds



Another Example: Bond Market

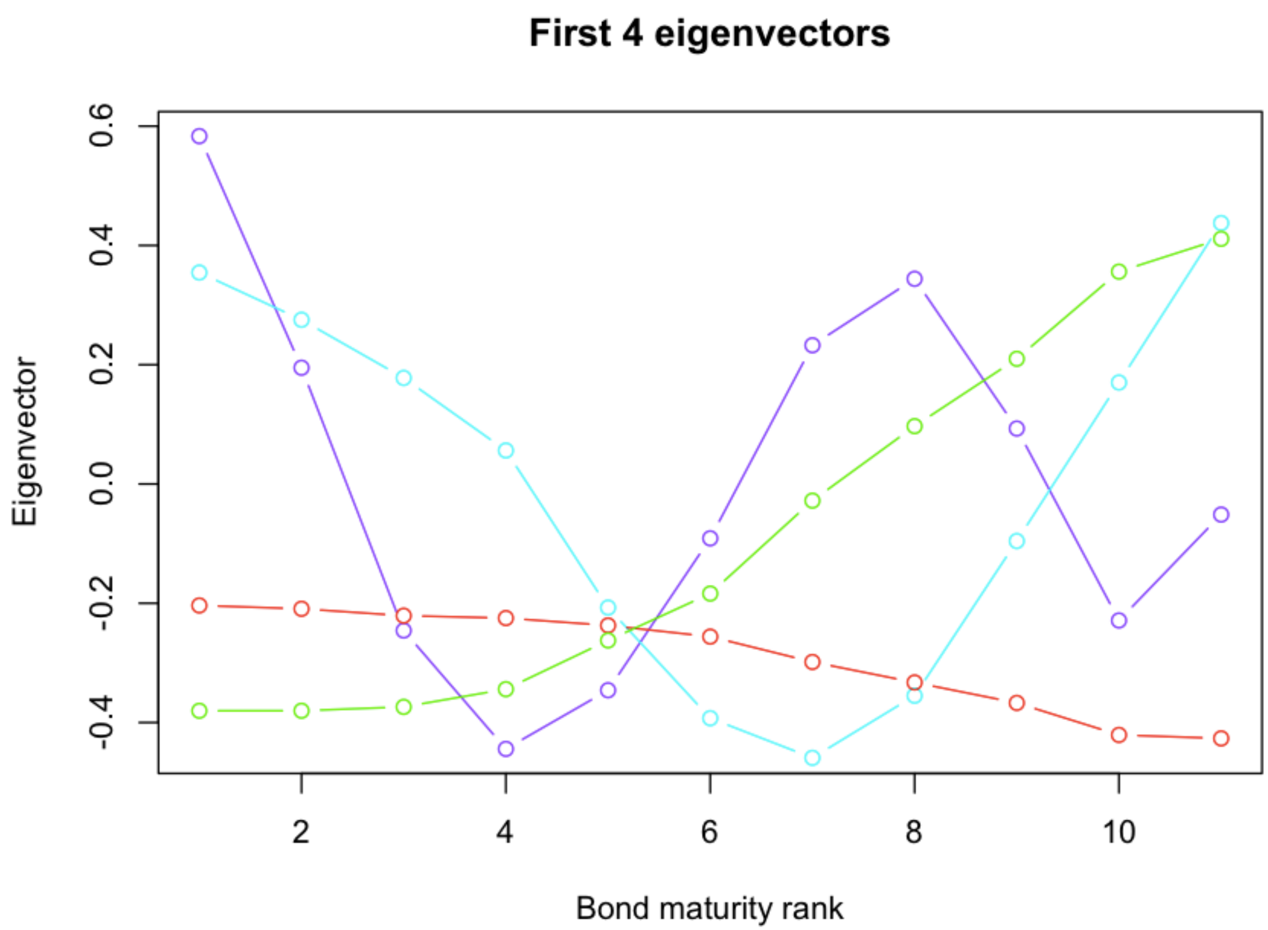
A look at eigenvalues





Another Example: Bond Market

A look at eigenvectors:

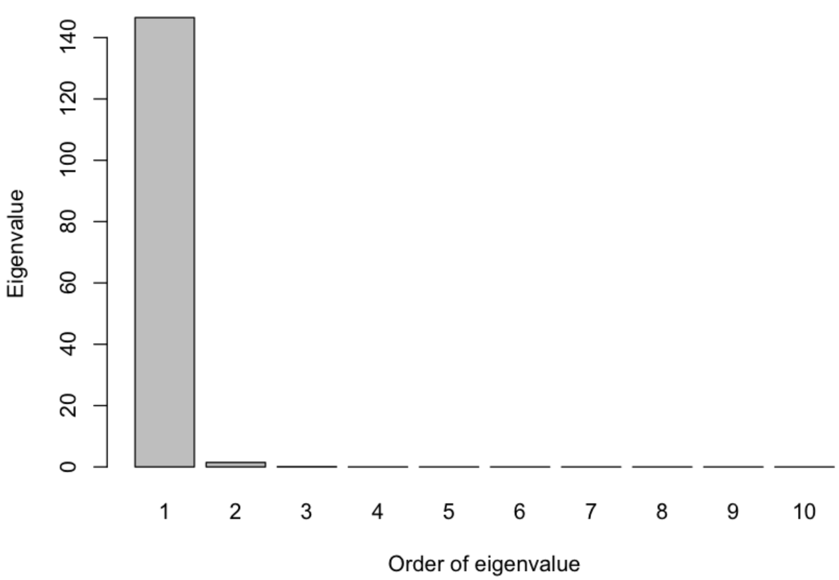


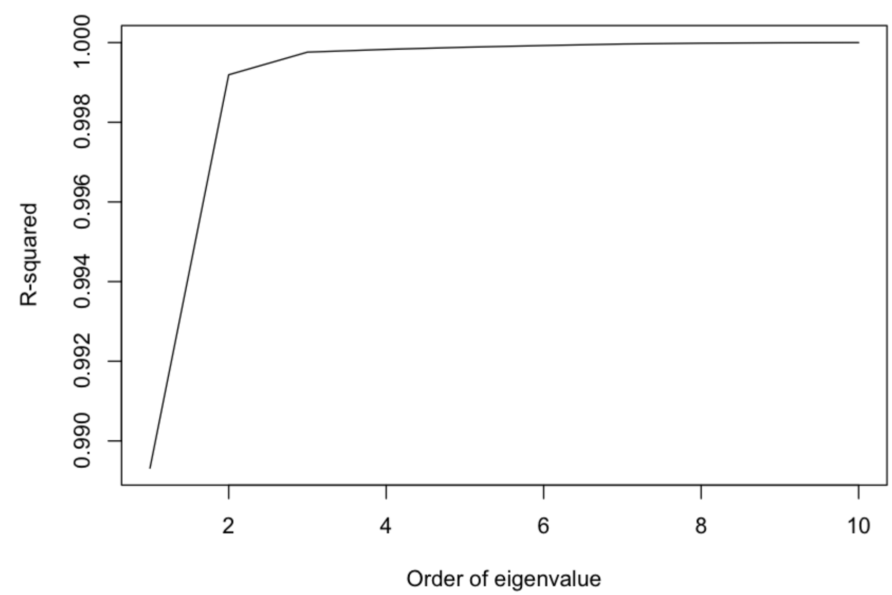
Another Example: Bond Market

- How about different countries?
- Monthly Spanish Government Bond yields
  - from: 10/1/2004 to: 4/1/2015

Another Example: Bond Market

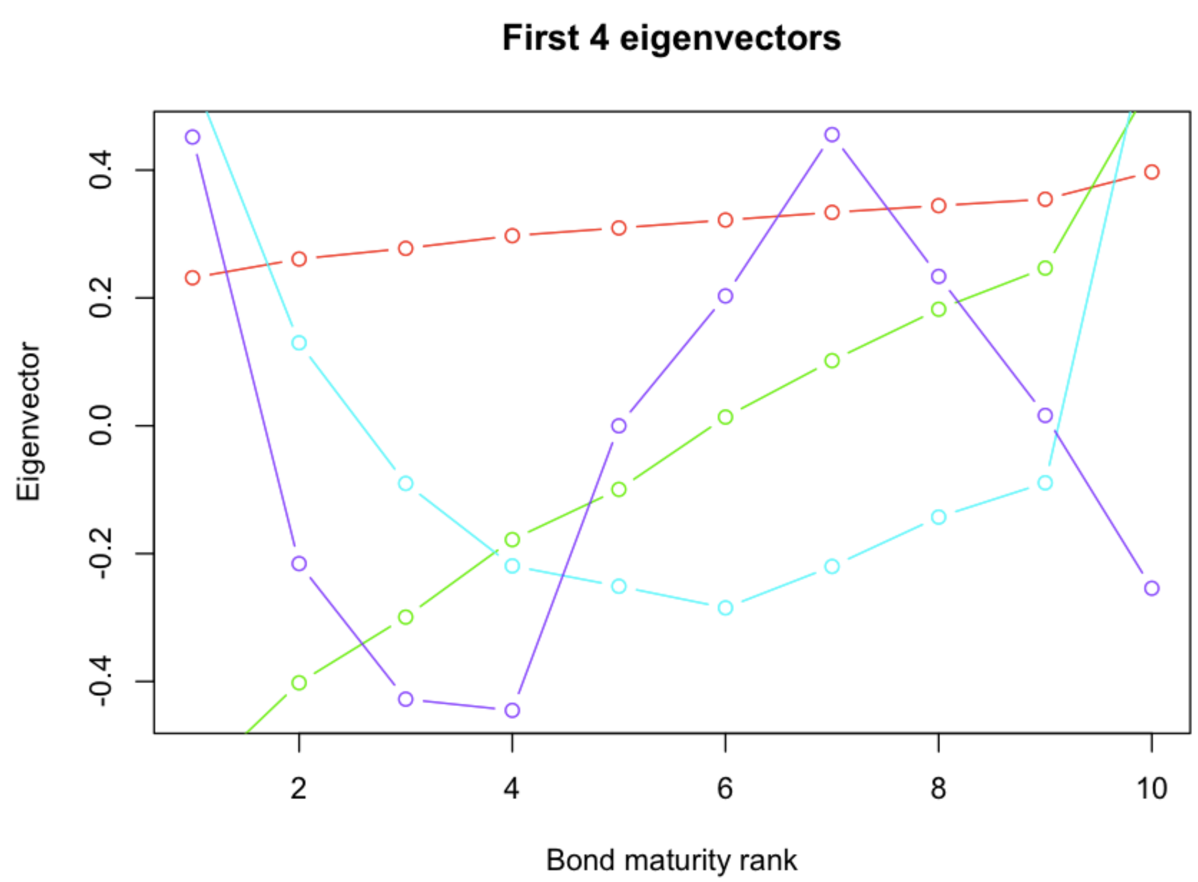
A look at eigenvalues:





Another Example: Bond Market

A look at eigenvectors:

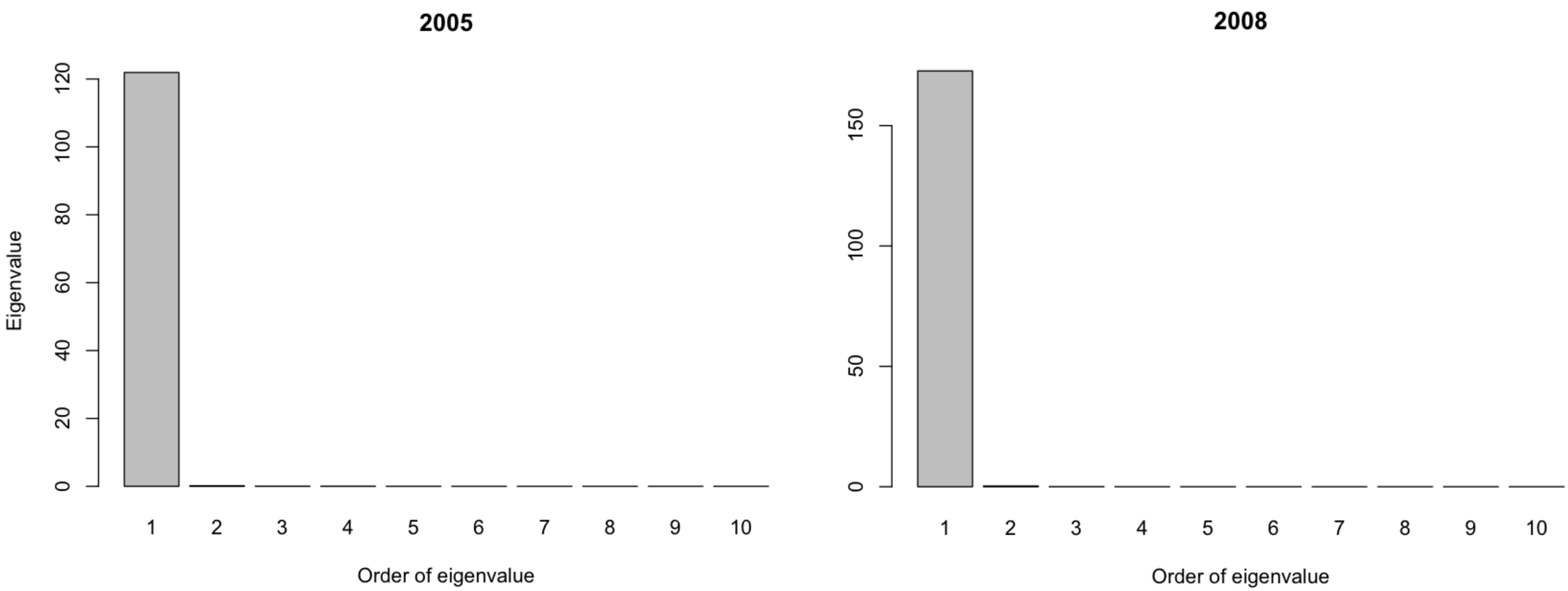


Another Example: Bond Market

- They look similar.
- Maybe we can look for specific years:
    - Before 2007.
    - After 2007.

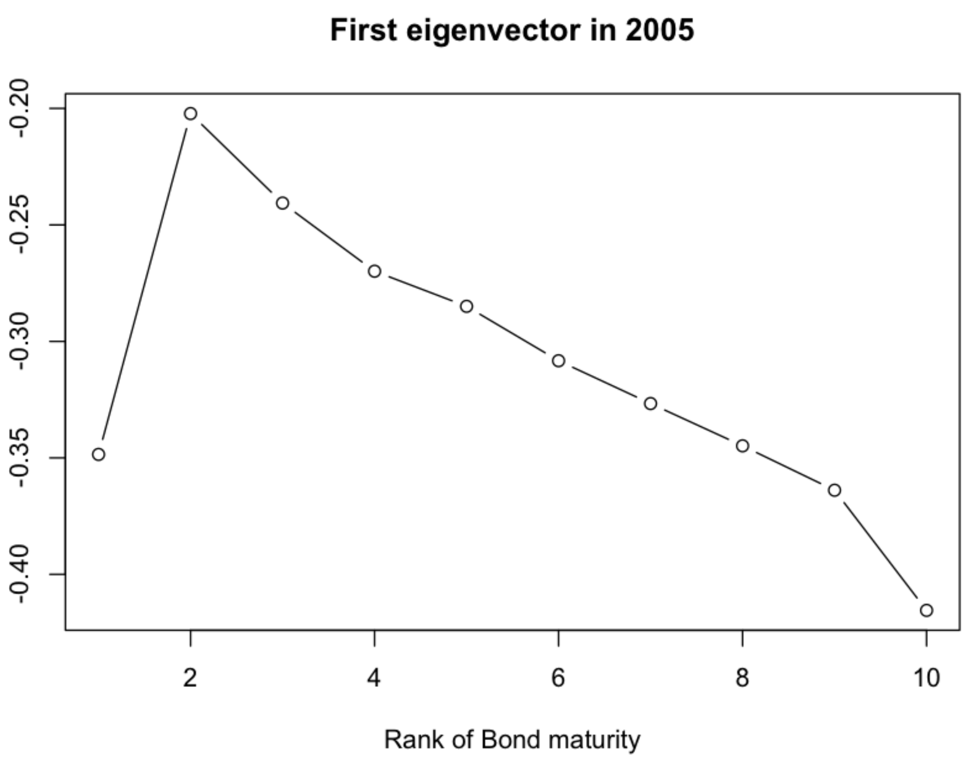
Another Example: Bond Market

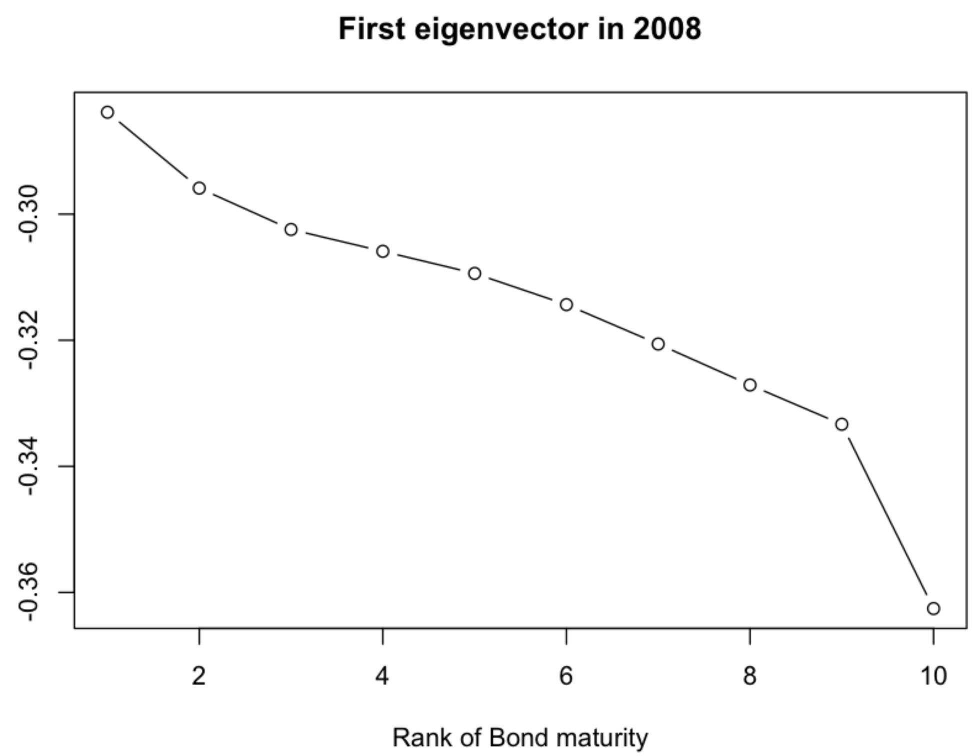
Eigenvalues for two different years:



Another Example: Bond Market

First eigenvectors for two different years:





Conclusions

- Stocks:
    - Dimension of the stock market is not as easily reducible as the bond market.
    - Gives valuable information about common factors, and correlation between different stocks (sectors?)
    - Might give better results for larger number of stocks.
- Bonds:
    - Number of important dimensions is usually less than 3.
    - Different markets and different years tend to give similar results.
    - Relation to the continuous case can make it easier to analyze.

Thank you!