Predicting Bond Prices Using PCA, KPCA and SVM

STA5069Z: Multivariate Analysis

Hlalumi Adams

University of Cape Town: Department of Statistical Sciences

March 20, 2024

Overview

Overview

- In the South African financial industry it has become very important to be able to model and predict changes in interest rates as this impacts inflation levels and individuals' spending abilities.
- One of the main drivers of interest rates are bond prices as exists an inverse relationship between bond prices and interest rates.

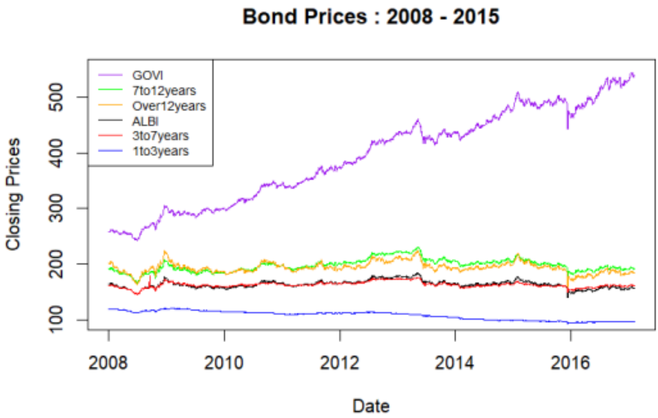
We will analyze how PCA, kPCA and SVM methods succeed in describing the dynamics of South African bond market prices at different points in time ,especially in the proximity of major marker events.

Dataset

The data series studied in this paper are daily-frequency observations from bond market data; the data is daily sampled closing price data and runs from 1994 to 2017. This data was purchased from INET-BFA for academic use . The data consists of bonds with different maturities and types namely ALBI, 1 to 3 years bonds, 3 to 7 years, 7 to 12 years, over 12 years and GOVI. Data for the Global Financial crisis (2008-2009) and Nene-gate (2015) period will be considered .
Variables are as follows:

- Duration
- Convexity
- Annualized Volatility Close
- TRI Average Yield
- Total Return Index
- Total Return Index Ytd

Dataset



Bond Prices : 2008 - 2015
GOVI
7to12years
Over12years
ALBI
3to7years
1to3years
Closing Prices
100
200
300
400
500
2008
2010
2012
2014
2016
Date

PCA: Introduction

Given a high-dimensional data set, i.e., information on multiple variables for a set of n observations, we wish to find a low-dimensional representation of the data that will retain most of the information in the original data. This "information" is captured by the variance-covariance matrix.

Definition

"Information" = "total variation" =variance-covariance matrix $= \sum_{j=1}^{r} var(X_j) = tr(\sum_{\mathrm{XX}})$

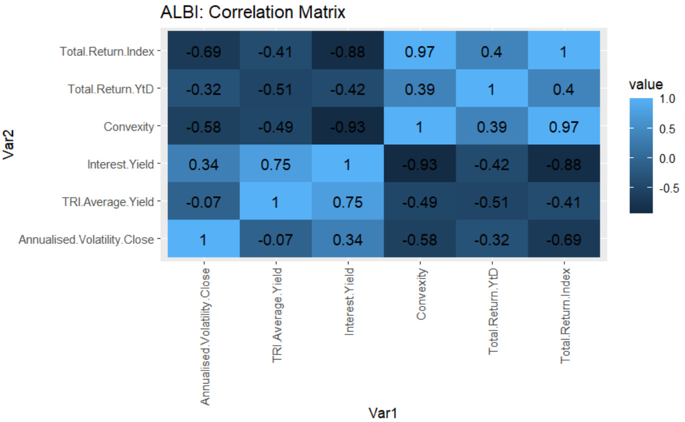
Using SVD :

Definition

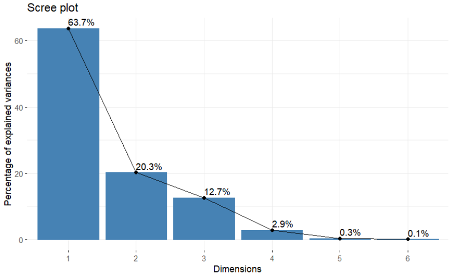
$\sum_{XX} = U \Lambda U^T - > \mathrm{U}^\mathrm{T}\mathrm{U} = \mathrm{I}_\mathrm{r} - > tr(\sum_{\mathrm{XX}}) = \sum_{\mathrm{j}=1}^{\mathrm{r}}(\lambda_\mathrm{j})$

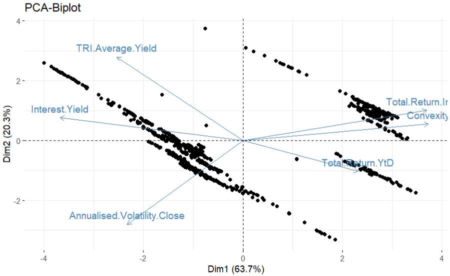
Why PCA?

PCA will help identify hidden patterns in the data sets by reducing the dimensionality of the data . This is useful when the variables within the dataset are highly correlated , indicating redundancy in the data.



PCA:Results


Scree plot
Percentage of explained variances
Dimensions
63.7%
20.3%
12.7%
2.9%
0.3%
0.1%


PCA-Biplot
TRI.Average.Yield
Interest.Yield
Total.Return.Ir
Convexity
Total.Return.YtD
Annualised.Volatility.Close
Dim1 (63.7%)
Dim2 (20.3%)

Kernel PCA: Introduction

- KPCA is an extension of traditional PCA. Unlike PCA, which is linear, KPCA can handle nonlinear relationships between data points.
- KPCA transforms the data onto a higher-dimensional space and uses standard PCA in this higher-dimensional space to project the data back onto a lower-dimensional space.

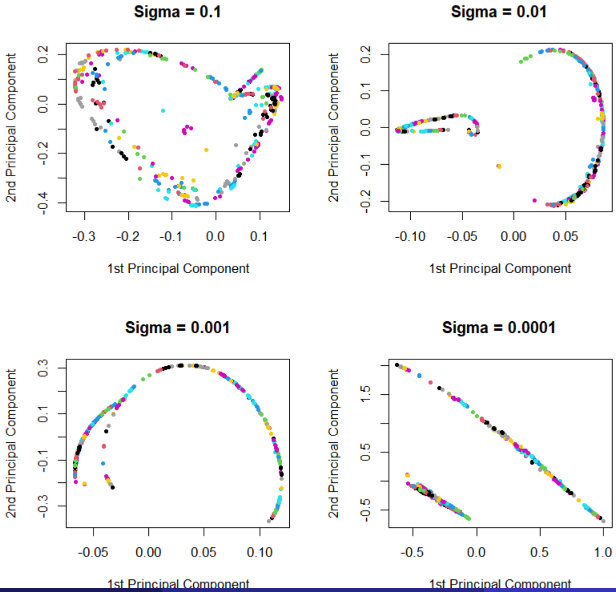
We use the Gaussian (radial-basis function network) Kernel:

$$\kappa(x_i, x_{i'}) = exp(-\frac{||x_i - x_{i'}||}{2\sigma^2})$$

Compute the kernel principal components using :

$$y(x) = \phi(\mathbf{x})^T \mathbf{v} = \Sigma_{i=1}^{n} \alpha_i \kappa(x_i, x_{i'})$$

KPCA: Results

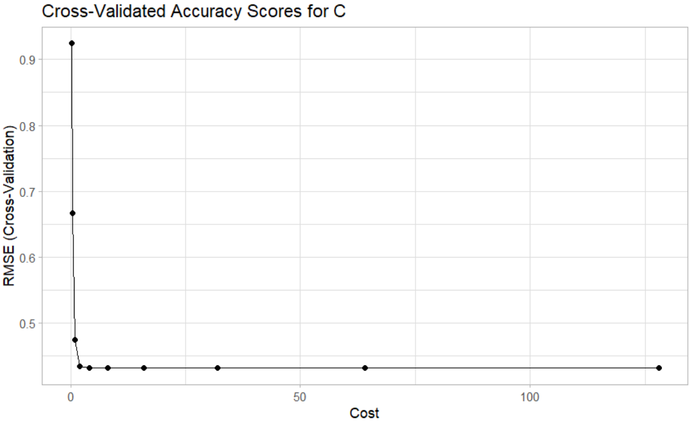


SVM

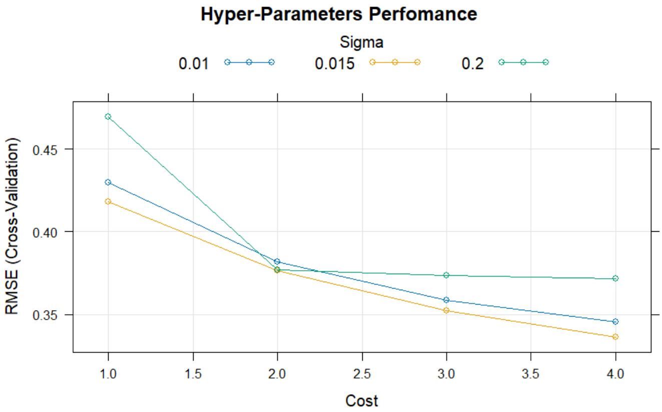
- Support vector machines (SVMs) are a group of supervised machine-learning models that can be used for classification and regression. It will seek an optimal hyperplane for separating two classes in a multidimensional space.
- Equally the regression problem (SVR) is a generalization of the classification problem we seek to find a good fitting hyperplane in a kernel-induced feature space that will have good generalization performance using the original features.

SVM: Hyper-Parameters

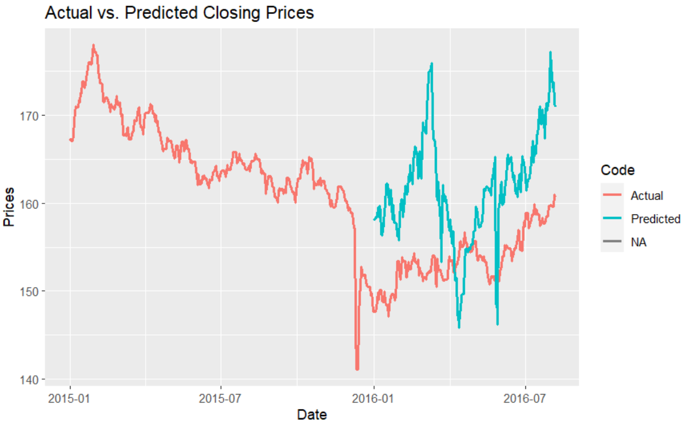
- The radial basis kernel function has two hyperparameters: σ and C.
- To find the optimal cost function we tune and train an SVM using the radial basis kernel function with auto-tuning for the σ parameter and 10-fold CV.
- Plotting the results , we see that smaller values of the cost parameter ($C = 4$) provide better cross-validated scored for the training data.



SVM: Hyper-Parameters



SVM: Predictions



The End