**Assignment 1.2 PCA on Treasury Zero Coupon Bond Yield**

# **Introduction**

In terms structure of interest rate curve, it is well known the first 3 factors can explain over 99% of total variance. Meanwhile we can interpret the factors with economic meanings. First factor is level; second factor is slope; third factor is curvature.

We download zero coupon bond yield historical data: 5-year history, 9 tenors 6M, 1Y, 2Y, 3Y, 5Y, 7Y, 10Y, 20Y, 30Y.

# **task 1, download data**

The script is provided for this step, which is to parse the raw xml data into dataframe format. You only need to call the function below to process.

The source file DailyTreasuryYieldCurveRateData.xml is downloaded from federal reserve website

Hint: place testdownload.py in the same directory of your script; put “import” in the header of program, *import testdownload.py*, then you can use the download function

# start\_date = '2014-01-01'

# end\_date = '2018-12-31'

# xml\_filename = 'DailyTreasuryYieldCurveRateData.xml'

# hist\_data = testdownload.ParseTreasuryYields(start\_date, end\_date, xml\_filename)

# **task 2, calculate correlation matrix**

PCA is commonly performed on standardized data (mean 0, standard deviation 1), this is to remove the noise of volatility and different unit measures from raw data.

Standardized data means we need to apply eigen decomposition on correlation matrix of simple yield return. In interest rate, yield return is y(t) – y(y-1)

# **task 3, calculate modified duration, yield volatility and bond price return volatility**

We used **daily return**, but we want to calculate **monthly volatility**, scaled by **sqrt(22).**

# **task 4, calculate eigen value, eigen vector**

Calculate eigen value, eigen vector of correlation matrix. Sort the eigen value and eigen vector in descending order. Although sklearn has the PCA package, it is more flexible to use standard eigen decomposition method np.linalg.eigh

Calculate the variance explained by first 3 factors respectively.

Reconstruct correlation matrix by the first 2 factors.

# **task 5, calculate monthly 95% portfolio VaR**

A bond portfolio is constructed in this way:

40m invest in 1y, -100m in 7y, 50m invest in 20y

Construct factor model by PCA component, calculate portfolio beta of first 3 factors. Calculate portfolio VaR of 1 factor model, and 2 factor model respectively

# **task 6, save all result in a csv file**

# **SUBMISSION**

Write a word document, simply write down result with a few simple equations to explain the calculation.

In task 6, in the script, save CSV file in folder C:\temp\Assignment2\GrouptX.csv, X is your team group no.

In the output file, save sequentially, volatility and VaR scaled refers to monthly risk horizon.

* Dataframe, column would be yield volatility%, price return volatility%, modified duration, current yield level
* Eigen vector, eigen value, sorted by descending order
* Variance explained by 1st factor, 2nd factor and 3rd factor
* Correlation matrix reconstructed by first and second factor
* Betas of the portfolio and portfolio VaR