DSA302 Financial Data Analysis

Group Project, Term 1, 2020 – 2021

Introduction

This group project consists of two parts. Part 1 is quite structured. It consists of some analysis which should be performed with accuracy. Part 2 is quite open. You are given some broad guidelines under which you will develop your own analysis and draw conclusions/recommendations. Both parts will involve analysis of zero coupon bond prices and spot rates of interest.

Part 1: Estimating the Yield Curve

1.1 Interest Rate Theory

Zero coupon bonds are priced based on spot rates of interest. Let the price of a zero coupon bond which matures at time t with face value 100 be $P(t)$. Investors who purchase the bond today (time 0) at price $P(t)$ will receive 100 at time t and nothing else prior to that. If the spot rate of interest for maturity at time t is $r(t)$, then the bond price is given by

$$P(t) = 100\, e^{-r(t)t} \tag{1}$$

Spot rates of interest are unobservable quantities, while bond prices are observable. If an analyst has a good set of data of bond prices (with or without coupons) she may construct a good set of zero coupon bond (with given maturity) prices from which to compute the spot rates of interest.

In continuous time context, the spot rate function $r(t)$ is related to the instantaneous forward rate at time t, denoted by $f(t)$, as follows:

$$r(t) = \frac{1}{t}\int_0^t f(s)ds. \tag{2}$$

Like spot rate of interest, forward rate of interest is unobservable. Given bond prices, we can estimate the forward rate of interest and then compute the spot rate of interest. A plot of the spot rate against time to maturity is called the yield curve.

Given a set of zero coupon bond prices P_i with maturity at time t_i, for $i = 1, \cdots, n$ and $0 < t_1 < t_2 < \cdots < t_n$, there are two possible ways to compute the spot rates.

Method 1: We can compute $r(t_i)$ from equation (1) as follows:

$$r(t_i) = -\frac{1}{t_i}\log(0.01P_i) \tag{3}$$

The $(t_i,\ r(t_i))$ data points may then be smoothed by some numerical methods such as polynomial or spline smoothing.

Method 2: We first compute the forward rate applicable to the interval $(t_j,\ t_{j+1})$, denoted by f_j, by

$$f_j = -\frac{\log P_{j+1} - \log P_j}{t_{j+1} - t_j}, \qquad \text{for } j = 1, \cdots n-1. \tag{4}$$

The spot rate function $r(t)$ is then computed as

$$r(t) = \frac{1}{t}\left[\sum_{i=2}^{j} f_{i-1}(t_i - t_{i-1}) + f_j(t - t_j)\right], \qquad \text{for } t_j < t \le t_{j+1};\ j = 2, \cdots, n-1, \tag{5}$$

with $r(t) = -(\log 0.01P_1)/t_1$ for $t \le t_1$ and $r(t) = f_1$ for $t_1 < t \le t_2$.

It can be seen that for Method 1, the spot rate at time t_i depends on the price of the zero coupon bond P_i and not other data points. In contrast, for Method 2, the spot rate at time t_i depends on all data with maturity prior to t_i.

1.2 Guideline to Part 1

The file `ZCBP.txt` contains a set of zero coupon bond prices (per 100 face value) with their corresponding maturities (in years) up to 30 years. You are required to do the following:

(1) Plot the bond prices versus their maturities.

(2) Estimate the empirical spot rates of interest using Method 1.

(3) Smooth the empirical spot rates in (2) using a second order polynomial. You may apply the `optim( )` function to minimize the sum of squares of errors. Plot the smoothed curve versus the empirical spot rates.

(4) Estimate the empirical spot rates of interest using Method 2 for any maturities up to 30 years. You may apply the function `cut( )` (or other methods of your choice to determine which interval t falls in) and `cumsum( )` to evaluate equation (5).

(5) Plot the yield curve computed in (4) against the smoothed yield curve computed in (3).

(6) Comment on your results.

Part 2: Fitting the Yield Curve

2.1. Nelson-Siegel Model and Nelson-Siegel-Svensson Model

Nelson and Siegel (1985) propose a model (NS Model) to capture the shape of the yield curve. They propose the following spot rate $r(t)$ as a function of maturity t:

$$r(t) = \theta_0 + \left(\theta_1 + \frac{\theta_2}{\theta_3}\right)\left(\frac{1 - e^{-\theta_3 t}}{\theta_3 t}\right) - \frac{\theta_2}{\theta_3} e^{-\theta_3 t}, \qquad \theta_0,\ \theta_3 > 0. \tag{6}$$

The NS Model was subsequently extended in Svensson (1994) to the Nelson-Siegel-Svensson (NSS) Model with an additional term as follows:

$$r(t) = \theta_0 + \left(\theta_1 + \frac{\theta_2}{\theta_3}\right)\left(\frac{1 - e^{-\theta_3 t}}{\theta_3 t}\right) - \frac{\theta_2}{\theta_3} e^{-\theta_3 t} + \frac{\theta_4}{\theta_5}\left[\frac{1 - e^{-\theta_5 t}}{\theta_5 t} - e^{-\theta_5 t}\right], \quad \theta_0,\ \theta_3,\ \theta_5 > 0. \tag{7}$$

The NS and NSS Models have been used extensively in the literature to study the shapes of various yield curves and their changes over time.

Given a set of spot rates $r(t_i)$ at various maturities t_i, the parameters $\boldsymbol{\theta} = (\theta_0, \theta_1, \theta_2, \theta_3)$ in the NS Model and $\boldsymbol{\theta} = (\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$ in the NSS Model can be estimated using a nonlinear regression model. The model parameters can be estimated by minimizing the sum of squares of the error to provide the **nonlinear least squares** (NLS) estimates.

2.2 Guideline to Part 2

The Excel file `ZCBYF86.csv` contains the yield curves (spot rates of interest in percent per annum) from 1986-01-02 to 2020-08-28, downloaded from `quandl.com`. The maturities are from 1 year to 30 years in steps of 1 year.

Below are some research questions you may consider:

(1) Present the data in a suitable way.

tell the fit of these models, and choice of model

(2) Fit the NS and NSS Models to the yield data using the NLS method. Compare the two models using some suitable diagnostics and model selection criteria.

(3) The changing patterns of the yield curve can be studied through the parameters $\boldsymbol{\theta}$. What information can you extract from the estimates of $\boldsymbol{\theta}$?

(4) How may the data tell you about the response of the spot rates at the long end with respect to the spot rates at the short end?

maturity at year 25,30 vs maturity at year 1,2
how does long end depend on short end?

(5) Fitting the NS and NSS Models on each day is a static exercise. Can the NS and NSS Models be used to *predict* future interest rate movements?

Report and Evaluation

(1) This project constitutes 25% of course assessment, with Part 1 taking up 10% and Part 2 taking up 15%.

(2) Each group will submit a report of no more than 10 pages (A4 single line spacing), and deliver a presentation of no more than 20 min, inclusive of 5 min of Q&A. `R` codes will be submitted separately (not inclusive in the 10-page report). The presentation will be in Week 13 (Week 8 is recess).

(3) The report must have the following declaration, signed by all group members: *"Each member of this group contributes honestly and fairly to the completion of this report."*.